University of Michigan School of Public Health

The University of Michigan Department of Biostatistics Working Paper Series

Longitudinal Image Analysis of Tumor/Brain Change in Contrast Uptake Induced by Radiation

Xiaoxi Zhang[*]   Tim Johnson[†]

Rod Little[‡]   Yue Cao[**]

[*]Pfizer Inc., xiaoxi.zhang@pfizer.com

[†]University of Michigan Biostatistics, tdjtdj@umich.edu

[‡]rlittle@umich.edu

[**]University of Michigan, yuecao@umich.edu

Longitudinal Image Analysis of Tumor/Brain Change in Contrast Uptake Induced by Radiation

Xiaoxi Zhang, Tim Johnson, Rod Little, and Yue Cao

Abstract

This work is motivated by a quantitative Magnetic Resonance Imaging study of the differential tumor/healthy tissue change in contrast uptake induced by radiation. The goal is to determine the time in which there is maximal contrast uptake, a surrogate for permeability, in the tumor relative to healthy tissue. A notable feature of the data is its spatial heterogeneity. Zhang, Johnson, Little, and Cao (2008a and 2008b) discuss two parallel approaches to "denoise" a single image of change in contrast uptake from baseline to a single follow-up visit of interest. In this work we explore the longitudinal profile of the tumor/healthy tissue change in contrast uptake. In addition to the spatial correlation, we account for temporal correlation by jointly modeling multiple images on the individual subjects over time. We fit a two-stage model. First, we propose a longitudinal image model for each subject. This model simultaneously accounts for the spatial and temporal correlation and denoises the observed images by borrowing strength both across neighboring pixels and over time. We propose to use the area under the receiver operating characteristics (ROC) curve (AUC) to summarize the differential contrast uptake between tumor and healthy tissue. In the second stage, we fit a population model on the AUC values and estimate when it achieves the maximum.

Longitudinal Image Analysis of Tumor/Brain Change in Contrast Uptake Induced by Radiation

Xiaoxi Zhang[*][1]**, Timothy D. Johnson**[2]**, Roderick J. A. Little**[2]**,**

and

Yue Cao[2]

[1] Pfizer Inc., New York, NY 10017

[2] University of Michigan, Ann Arbor, MI 48109

[*]*email:* xiaoxi.zhang@pfizer.com

SUMMARY: This work is motivated by a quantitative Magnetic Resonance Imaging study of the differential tumor/healthy tissue change in contrast uptake induced by radiation. The goal is to determine the time in which there is maximal contrast uptake, a surrogate for permeability, in the tumor relative to healthy tissue. A notable feature of the data is its spatial heterogeneity. Zhang, Johnson, Little, and Cao (2008a and 2008b) discuss two parallel approaches to "denoise" a single image of change in contrast uptake from baseline to a single follow-up visit of interest. In this work we explore the longitudinal profile of the tumor/healthy tissue change in contrast uptake. In addition to the spatial correlation, we account for temporal correlation by jointly modeling multiple images on the individual subjects over time. We fit a two-stage model. First, we propose a longitudinal image model for each subject. This model simultaneously accounts for the spatial and temporal correlation and denoises the observed images by borrowing strength both across neighboring pixels and over time. We propose to use the area under the receiver operating characteristics (ROC) curve (AUC) to summarize the differential contrast uptake between tumor and healthy tissue. In the second stage, we fit a population model on the AUC values and estimate when it achieves the maximum.

KEY WORDS: area under curve; Markov random field; population model; quantitative MRI; receiver operating characteristics curves; reversible jump MCMC; spatial-temporal model.

1. Introduction

This work is motivated by a pilot imaging study investigating the effects of radiation therapy on the vascular permeability of high-grade gliomas, a particularly virulent type of brain cancer. Despite advances in both radiotherapy and chemotherapy, the overall survival time of a glioma patient has not significantly increased from its one year median survival since diagnosis (Curran et al. 1993).

Researchers hypothesize that tumor vasculature damage due to radiation would disrupt the blood tumor barrier (BTB), at least transiently, allowing larger molecules to cross the BTB. If true, this suggests that chemotherapy should commence during the time when the BTB damage is at its peak. For instance, if the BTB damage peaks early during radiotherapy, greater tumor control could be achieved by starting chemotherapy during this time, in contrast to current practice of sequentially administering radiotherapy followed by chemotherapy.

Eleven subjects, diagnosed with high-grade gliomas, volunteered to take part in this study conducted at the University of Michigan School of Medicine in the Department of Radiation Oncology. The subjects underwent, standard-of-care radiation therapy over a course of four to six weeks. Six imaging sessions were scheduled for each subject. The first image was obtained approximately one week prior to the onset of radiotherapy. The other five images were acquired at weeks one and three during therapy, and at one, three and six months post therapy. Two T1-weighted magnetic resonance images (MRI), one with and one without contrast enhancement agent—Gadolinium diethylenetriaminepentaacetic acid (Gd-DTPA)—were obtained at each exam. All images were registered to the baseline computed tomography (CT) image used for treatment planning. The molecular weight of Gd-DTPA is approximately that of many chemotherapeutic molecules. Therefore, the contrast uptake (the volume of Gd-DTPA that crosses the BTB) is a good surrogate of vascular permeability. It

is representative of the volume of chemotherapeutic drug that crosses the BTB. The contrast uptake is quantified as the log ratio of the contrast enhanced T1-weighted MRI image to the non-contrast enhanced T1-weighted MRI image (e.g. Figure 1). The resulting contrast uptake image is an example of quantitative MRI (qMRI). It quantitates a physiological process (in this case vascular permeability). This is in contrast to an anatomical MRI exams that only allow visualization of anatomical structures and abnormalities and whose values have no absolute, quantitative interpretation, only relative meaning.

Despite a large body of literature on medical image analysis, especially functional MRI, proper statistical analysis of qMRI data is limited. Two important features of qMRI data are spatial correlation and heterogeneity. Although they seem to be discordant features, examination of Figure 1 reveals that contrast uptake in healthy tissue is rather homogeneous and one would expect that there is high spatial correlation (at least locally). The tumor, visible in the upper left quadrant, is much more heterogeneous with respect to its contrast uptake. This heterogeneity can, more or less, be divided into two relatively homogeneous regions: the core of the tumor and its periphery, or annulus. The core of a tumor is typically hypoxic (low oxygen content due to a lack of blood supply) and appears dark in Figure 1, for example, indicating low contrast uptake. Whereas, the annulus is typically rich in neovasculature (newly formed blood vessels that are typically disorganized and leaky) and appears light in shades indicating high contrast uptake (see Figure 1 for example). It is known that hypoxia is protective against damage due to radiation and chemotherapy. Therefore, the core is usually a source of tumor regrowth after therapy. The focus of this imaging study is on demonstrating that radiation therapy can transiently increase vascular permeability in the tumor core.

In many imaging studies, spatial information is ignored (e.g. Cao et al. 2005; Moffat et al. 2005; Hamstra et al. 2005). Zhang, Johnson, Little, and Cao (2008a and 2008b) develop an

appropriate statistical model (a Gaussian hidden Markov random field model) for qMRI data that smoothes the data while respecting the heterogeneity not only within the tumor but also between the tumor and healthy tissue. They propose both a maximum likelihood and a Bayesian estimation method for image smoothing with edge preservation. However, their work focuses on a single contrast change—the change in contrast uptake from the baseline image to the three-week image, and thus ignores the temporal aspects in the data. In this manuscript we extend their Bayesian approach to account for the temporal profile of the change in contrast uptake in addition to the spatial correlation of the data. In doing so, we can answer the main question of interest: at what time, if any, does the differential contrast uptake achieve a maximum.

In this article we first extend the Bayesian hierarchical model proposed in Zhang et al. 2008b by building a spatial-temporal model, accounting for temporal, as well as spatial, correlation. The model smoothes the observed images, while preserving edges, by borrowing strength across pixels and over time. We then summarize the differential tumor/healthy tissue "true" contrast uptake for each subject using the area under the receiver operating characteristics (ROC) curve (AUC, Pepe, 2003). The AUC is a robust measure of the difference in the distributions of the "true" contrast uptake of diseased and healthy tissue (Pepe, 2003). We then combine information across subjects in several population models treating the AUC as a response variable and investigate the population level temporal profile of the tumor/healthy tissue contrast uptake. The best population model is chosen based on the predictive distribution criterion proposed by Gelfand, Dey, and Chang (1992).

In the next section we describe the details of the proposed model. In Section 3 we outline the model implementation. Results from simulation studies and the motivating study are presented in Section 4. A small sensitivity analysis is also conducted in that section. We conclude the paper with a discussion of the model and ideas for future research.

[Figure 1 about here.]

2. Model Specification

In this section, we first describe a longitudinal image model at the subject level that borrows strength both across pixels and over time. It smoothes the observed contrast uptake while preserving edges between homogeneous regions. Later, we summarize the differential tumor/healthy tissue response to radiation using AUCs, and then build a population model treating AUCs as the outcome of interest.

2.1 *Longitudinal Image Model*

We build a longitudinal image model for each subject using the following notation. Let j $(j = 1, 2, \cdots, J)$ index the visits. Since we assume the images are registered over time for each subject, we use a common index i $(i = 1, 2, \cdots, N)$ to denote pixels regardless of visits (pixel i at visit j and pixel i at visit j' are the same pixel). We denote the observed contrast uptake of pixel i at visit j by y_{ij}, and write the image at visit j in vector form as $\mathbf{y}_j = (y_{1j}, \cdots, y_{Nj})^{\mathrm{T}}$ and all J images as $\mathbf{y} = (\mathbf{y}_1^{\mathrm{T}}, \cdots, \mathbf{y}_J^{\mathrm{T}})^{\mathrm{T}}$.

2.1.1 *Distribution of the Data.* We decompose the observed contrast uptake, $\mathbf{y}$, into the "true" contrast uptake and observational noise. We specify the observed contrast uptake as an unobserved "true" value plus a measurement error, i.e. $y_{ij} = \nu_{ij} + e_{ij}$ where $e_{ij} \sim \mathrm{N}(0, \sigma_e^2)$, or $[y_{ij} \mid \nu_{ij}, \sigma_e^2] \sim \mathrm{N}(\nu_{ij}, \sigma_e^2)$, for all i, j. The ν_{ij} is the unobserved "true" contrast uptake at pixel i, visit j, the temporal and spatial structure of which will be specified in the following section. We write in vector form $\boldsymbol{\nu}_j = (\nu_{1j}, \cdots, \nu_{Nj})^{\mathrm{T}}$ and $\boldsymbol{\nu} = (\boldsymbol{\nu}_1^{\mathrm{T}}, \cdots, \boldsymbol{\nu}_J^{\mathrm{T}})^{\mathrm{T}}$.

2.1.2 *Prior Distributions.*

(1) We specify flat (improper) prior distribution for the standard deviation of the measurement error, i.e. $P(\sigma_e^2) \propto 1$.

(2) The decomposition of the observed contrast uptake into the "truth" and conditionally independent noises in the above section is similar to Zhang et al. (2008a and 2008b). They impose a spatial structure (a Markov Random Fields prior, MRF, Besag 1974) directly on the "true" change in contrast uptake. Their model is sufficient when analyzing one single change image, i.e. subtracting the baseline observed contrast uptake image from the observed image at three-week. However, we need additional modeling of the longitudinal perspective when we analyze multiple qMRI images taken over time.

Considering that the "true" contrast uptake evolves over time during and after radio-therapy, this motivates us to conceptualize the "true" contrast uptake at current visit as the "true" contrast uptake from the last visit plus a change in contrast uptake and an evolution error.

For $j > 1$, this is formalized as $\nu_{ij} = \nu_{i,j-1} + \delta_{ij} + \varepsilon_{ij}$, $\varepsilon_{ij} \sim \mathrm{N}(0, \tau_{ij}^2)$, where δ_{ij} is the structural "change" in contrast uptake and ε_{ij} is the evolution error of pixel i at visit j, respectively. To keep the notation consistent, we define $\nu_{i0} \equiv 0$. Hence, $[\nu_{ij} \mid \nu_{i,j-1}, \delta_{ij}, \tau_{ij}^2] \sim \mathrm{N}(\nu_{i,j-1} + \delta_{ij}, \tau_{ij}^2)$ for $i = 1, 2, \cdots, N$ and $j = 1, 2, \cdots, J$. We write in vector form $\boldsymbol{\delta}_j = (\delta_{1j}, \cdots, \delta_{Nj})^{\mathrm{T}}$, $\boldsymbol{\delta} = (\boldsymbol{\delta}_1^{\mathrm{T}}, \cdots, \boldsymbol{\delta}_J^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{\tau}_j^2 = (\tau_{1j}^2, \cdots, \tau_{Nj}^2)^{\mathrm{T}}$, and $\boldsymbol{\tau}^2 = (\boldsymbol{\tau}_1^{2\,\mathrm{T}}, \cdots, \boldsymbol{\tau}_J^{2\,\mathrm{T}})^{\mathrm{T}}$. The above prior distribution of $\boldsymbol{\nu}$ implies $E(\boldsymbol{\nu}_j \mid \boldsymbol{\nu}_{j-1}, \boldsymbol{\delta}_j, \boldsymbol{\tau}_j^2) = \boldsymbol{\nu}_{j-1} + \boldsymbol{\delta}_j$. Therefore, $\boldsymbol{\delta}_1$ is interpreted as the expectation of the true contrast uptake at the first visit (i.e. $E(\boldsymbol{\nu}_1 \mid \boldsymbol{\nu}_0, \boldsymbol{\delta}_1, \boldsymbol{\tau}_j^2) = \boldsymbol{\delta}_1$), and the interpretation of $\boldsymbol{\delta}_j$ $(j > 1)$ is the structural change in contrast uptake from visit $(j-1)$ to visit j. Henceforth, we refer to $\boldsymbol{\delta}_j$ $(j = 1, \cdots, J)$ as the true "change" in contrast uptake with change in quotes for simplicity.

(3) Since the radiation is administered in a gradient fashion, i.e. the dose is intensified in the tumor area and tapered off in healthy tissue, it is reasonable to assume its impact ("change" in contrast uptake) is spatially correlated. In the meanwhile, there exists

spatial heterogeneity within and surrounding the tumor. This motivates us to model the "change" in contrast uptake with a MRF prior for each visit, which only only accounts for the spatial correlation in the "change" but also respects the distinct boundaries between tumor and healthy tissue. Zhang et al. (2008a and 2008b) have successfully applied MRF in modeling the observed change in contrast uptake images $(y_{ij} - y_{ij'})$, which can be considered as the "prototype" of the proposed model if we force $\varepsilon_{ij} \equiv 0$.

Specifically, let δ_{ij} $(i = 1, 2, \cdots, N_j)$ assume one of M_j possible values, $(\mu_{1j}, \cdots, \mu_{M_j j})$ for each visit j. We write in vector form $\boldsymbol{\mu}_j = (\mu_{1j}, \cdots, \mu_{M_j j})^{\mathrm{T}}$ and $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^{\mathrm{T}}, \cdots, \boldsymbol{\mu}_J^{\mathrm{T}})^{\mathrm{T}}$. We specify the prior distributions of the M_j possible values as $\mu_{kj} \sim \mathrm{Uniform}(\mu_{\min}, \mu_{\max})$ for all $1 \leqslant k \leqslant M_j$, where $\mu_{\min}$ and $\mu_{\max}$ are the minimum and maximum of the observed "change" in contrast uptake.

(4) Similar to (3), for each visit j, let τ_{ij}^2 $(i = 1, 2, \cdots, N_j)$ assume one of M_j possible values denoted by $(\sigma_{1j}^2, \cdots, \sigma_{M_j j}^2)$. We assume, a priori, $\sigma_{kj}^2 \sim \mathrm{Inv\text{-}Gamma}(\alpha_\sigma, \beta_\sigma)$ for all $1 \leqslant k \leqslant M_j$. We write in vector form $\boldsymbol{\sigma}_j^2 = (\sigma_{1j}^2, \cdots, \sigma_{M_j j}^2)^{\mathrm{T}}$, and $\boldsymbol{\sigma}^2 = (\boldsymbol{\sigma}_1^{2\mathrm{T}}, \cdots, \boldsymbol{\sigma}_J^{2\mathrm{T}})^{\mathrm{T}}$. We set $\alpha_\sigma = 2.5$ and $\beta_\sigma = 5$ as constants, such that the prior mode of σ_{kj}^2 is 0.3. Later in Section 4.1, we investigate the sensitivity of the results to other choices of α_σ.

(5) We introduce hidden MRF labels Z_{ij} on the state space $\mathcal{S}_j = \{1, \cdots, M_j\}$. Let Z_{ij} index the value of δ_{ij} and τ_{ij}^2 out of the M_j choices, i.e. $\delta_{ij} = \mu_{kj}$ and $\tau_{ij}^2 = \sigma_{kj}^2$ when $Z_{ij} = k$.

The MRF prior on the hidden labels takes the form

$$\mathrm{P}(\mathbf{Z}_j = \mathbf{z}_j \mid \beta_j, M_j) = g^{-1}(\beta_j, M_j)\exp\Big\{\beta_j \sum_{i \sim i'} \mathrm{I}[z_{ij} = z_{i'j}]\Big\},$$

where $\mathrm{I}[\cdot]$ is the indicator function, $i \sim i'$ denote pixel i and i' are neighbors (i.e. immediately adjacent), and $g^{-1}(\beta_j, M_j) = \sum_{\mathbf{Z}_j} \exp\Big\{\beta_j \sum_{i \sim i'} \mathrm{I}[z_{ij} = z_{i'j}]\Big\}$ is the normalizing constant to be estimated via thermal integration (Ogata, 1989; Gelman and Meng, 1998).

This model is also known as the Potts model in statistical physics (Potts, 1952). The collection of hidden labels, $\mathbf{z}_j = (Z_{1j} = z_{1j}, \cdots, Z_{Nj} = z_{Nj})^{\mathrm{T}}$, is called a configuration.

We also write $\mathbf{z} = (\mathbf{z}_1^{\mathrm{T}}, \cdots, \mathbf{z}_J^{\mathrm{T}})^{\mathrm{T}}$. The spatial regularization parameter β_j controls the strength of the spatial correlation of the labels, e.g. large β_j encourages smoother configurations.

(6) We assume $M_j \sim \mathrm{Uniform}(M_{\min}, M_{\max})$, where $M_{\min} = 2$ and $M_{\max} = 20$, such that the number of components is allowed to vary on a reasonable range. We write in vector form $\mathbf{M} = (M_1, \cdots, M_J)^{\mathrm{T}}$.

(7) Let $\beta_j \sim \mathrm{Uniform}(0, \beta_{\max})$, where $\beta_{\max} = 3$. Similarly, we write in vector form $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_J)^{\mathrm{T}}$.

Note that we compute the normalizing constant $g(\beta_j, M_j)$ off-line on the grid values of $\beta_j = 0, 0.1, \cdots, 3.0$ and $M = 2, 3, \cdots, 20$ (Zhang et al. 2008b). Its value on non-grid points will be computed via interpolation.

2.1.3 *Joint Distribution and Conditional Posterior Distributions.* The joint distribution is

$$
\mathrm{P}(\sigma_e^2) \prod_{j=1}^{J} \left[\mathrm{P}(M_j)\mathrm{P}(\beta_j)\mathrm{P}(\mathbf{Z}_j \mid \beta_j, M_j) \times \prod_{k=1}^{M_j} \left[\mathrm{P}(\mu_{kj})\mathrm{P}(\sigma_{kj}^2) \right.\right.
$$
$$
\left.\left. \times \prod_{i \in D_{kj}} \left\{ \mathrm{P}(\nu_{ij} \mid \nu_{i,j-1}, \mu_{kj}, \sigma_{kj}^2)\mathrm{P}(y_{ij} \mid \nu_{ij}, \sigma_e^2) \right\} \right] \right],
$$

where $D_{kj} = \{i : Z_{ij} = k\}$ is the collection of pixels with the same label k at visit j, i.e. a component. By Bayes Theorem the conditional posterior distributions of μ_{kj}, σ_{kj}^2, ν_{ij}, and σ_e^2 have closed form. The conditional posterior distributions of $\mathbf{Z}_j$ and β_j satisfy

$$
\mathrm{P}(\mathbf{Z}_j \mid \cdot) \propto g^{-1}(\beta_j, M_j)\exp\left\{ \sum_{i \sim i'} \beta_j \mathrm{I}[z_{ij} = z_{i'j}] \right\}
$$
$$
\times \prod_{k=1}^{M_j} \prod_{i \in D_{kj}} \sigma_{kj}^{-1} \exp\left\{ -.5\sigma_{kj}^{-2}(\nu_{ij} - \nu_{i,j-1} - \mu_{kj})^2 \right\},
$$
$$
\mathrm{P}(\beta_j \mid \cdot) \propto g^{-1}(\beta_j, M_j)\exp\left\{ \sum_{i \sim i'} \beta_j \mathrm{I}[z_{ij} = z_{i'j}] \right\},
$$

for $j = 1, \cdots, J$.

2.1.4 *Marginal Posterior Mean of Contrast Uptake.* Following Zhang et al. (2008b), we characterize the contrast uptake by its posterior mean, i.e. $\nu_{ij} = \nu_{i,j-1} + \sum_{k=1}^{M_j} \mu_{kj} \mathrm{P}(Z_{ij} = k \mid \mathbf{y})$. It can be estimated via Markov chain Monte Carlo (MCMC) draws, $\boldsymbol{\delta}_{ij}^{(t)}$ and $\boldsymbol{\nu}_{i,j-1}^{(t)}$, at iteration t $(1 \leqslant t \leqslant T)$: $\hat{\nu}_{ij} = T^{-1} \sum_{t=1}^{T} \left[\nu_{i,j-1}^{(t)} + \delta_{ij}^{(t)} \right]$, where $\delta_{ij}^{(t)} = \mu_{kj}^{(t)}$ when $z_{ij}^{(t)} = k$. Similarly, we summarize the standard deviation of pixel intensity with its posterior mean, ϕ_i. The corresponding Monte Carlo estimate is $\hat{\phi}_{ij} = T^{-1} \sum_{t=1}^{T} \tau_{ij}^{(t)}$, where $\tau_{ij}^{(t)} = \sigma_{kj}^{(t)}$ when $z_{ij}^{(t)} = k$.

2.2 *Summarizing Differential Tumor/Healthy Tissue Response Using AUC*

Although the longitudinal image model detailed in Section 2.1 estimates the "true" contrast uptake by smoothing the images, it does not directly answer the question of interest: When is there is maximal contrast uptake in the tumor relative to healthy tissue? We propose using the AUC as a nonparametric and robust measure of the differences in the distributions of the "true" contrast uptake of diseased and healthy pixels (Pepe, 2003).

As a descriptive device, the ROC curve is most commonly used for evaluating diagnostic tests, providing false positive and true positive rates over a range of thresholds. We use it to provide a distribution-free description of the separation between two distributions. Let Y_D and $Y_{\bar{D}}$ denote the distribution of diseased and healthy populations (e.g. the "true" contrast uptake of tumor and healthy pixels). The ROC curve plots the false positive fraction ($\mathrm{FPF}(c) = \mathrm{P}(Y_{\bar{D}} > c)$) against the true positive fraction ($\mathrm{TPF}(c) = \mathrm{P}(Y_D > c)$) for a range of threshold values c (assuming large values indicate disease). The AUC is a non-parametric and robust measure of the separation of Y_D and $Y_{\bar{D}}$. It can be interpreted as the probability that a randomly drawn value from Y_D is greater than a random draw from $Y_{\bar{D}}$ (Bamber, 1975; Hanley and McNeil, 1982). When the distributions are discrete, the AUC based on linear interpolation is $\mathrm{AUC} = \mathrm{P}(Y_D > Y_{\bar{D}}) + .5\mathrm{P}(Y_D = Y_{\bar{D}})$. Large values of AUC suggest

good separation of the two distributions. In the ideal case, when the two distributions of Y_D and $Y_{\bar{D}}$ are perfectly separated, AUC reaches its maximum value, 1.

One benefit of using the AUC as a summary statistic of the differential tumor/healthy tissue response is that the healthy tissue of the same subject is serving as the control for the tumor region. As we have observed in some situations, the contrast uptake images from one visit can be overall brighter/darker than another visit due to MRI machine receiver gain (e.g. $\mathbf{y}_4$ in Figure 1 is brighter than the rest of the visits). Although in such situations the image intensities of all pixels are shifted, the AUC values are invariant, i.e. $P(Y_D > Y_{\bar{D}}) = P(Y_D + d > Y_{\bar{D}} + d)$ for an arbitrary shift in images intensities d.

As we stated in the introduction, the contrast uptake profile of tumor core and annulus are pathologically different and need to be examined separately. We divide the tumor into core and annulus by applying the edge preserving smoothing discussed in Zhang et al. (2008b) to the baseline contrast uptake image and taking the 95% percentile of the posterior mean image as the threshold between core and annulus. We use superscripts to denote tumor core and annulus respectively, i.e. A^c for core and A^a for annulus. The next section discusses building population model on the AUCs (A^c and A^a). Since the model for A^c and A^a will be considered separately and the superscript will be omitted for simplicity thereafter.

2.3 *Population Model of the AUC*

In building the population model, we introduce a subject index h $(1 \leqslant h \leqslant H)$. We denote the posterior mean of AUCs for subject h at visit j $(1 \leqslant j \leqslant J_h)$ by A_{hj} (gray profiles in Figures 5a and 5c). Since the AUC is bounded between 0 and 1, we apply a logit transformation to A_{hj} before fitting the population model, i.e. $\text{logit}(A_{hj}) = \log A_{hj} - \log(1 - A_{hj})$.

The subject AUC profiles in Figures 5a and 5c indicate the need for separate models for tumor core and annulus. Considering the limited number of visits for each individual, we opt for a linear function of time (days since the initiation of radiation, t_{hj}) with a variable

change point as the population mean model. The change point location (τ) is treated as a model parameter. As considerable inter-subject variability is seen in the profiles, we introduce individual random effects to capture the inter-subject variability in the patient population, e.g. random intercept, random slope for the second linear segment, or both. We define $[t]_+ = \max(0, t)$. Explicitly, the models we consider are

$$M_1: \quad \text{logit}(A_{hj}) = \alpha + \beta_1 t_{hj} + \beta_2 [t_{hj} - \tau]_+ + b_{0h} + e_{hj}.$$

$$M_2: \quad \text{logit}(A_{hj}) = \alpha + \beta_1 t_{hj} + (\beta_2 + b_{1h})[t_{hj} - \tau]_+ + e_{hj}.$$

$$M_3: \quad \text{logit}(A_{hj}) = \alpha + \beta_1 t_{hj} + (\beta_2 + b_{1h})[t_{hj} - \tau]_+ + b_{0h} + e_{hj}.$$

where $b_{0h} \sim \text{N}(0, \phi^2)$ and $b_{1h} \sim \text{N}(0, \psi^2)$ are the individual-specific random intercept and slope, respectively, and $e_{hj} \sim \text{N}(0, \sigma^2)$ is the error term. We specify the following prior distributions for model parameters, $\phi^2 \sim \text{Inv-Gamma}(a_\phi, b_\phi)$, $\psi^2 \sim \text{Inv-Gamma}(a_\psi, b_\psi)$, $\sigma^2 \sim \text{Inv-Gamma}(a_\sigma, b_\sigma)$, $\tau \sim \text{Uniform}(0, 183)$ and an improper flat prior density for $\boldsymbol{\theta} = (\alpha, \beta_1, \beta_2)^{\text{T}}$, i.e. $\text{P}(\boldsymbol{\theta}) \propto 1$. These result in a proper joint posterior distribution. We set $a_\phi = a_\psi = a_\sigma = 1$, $b_\phi = b_\psi = 0.1$, and $b_\sigma = 0.5$ such that ϕ^2, ψ^2, and σ^2 are on reasonable scales, a priori.

We evaluate the adequacy of the population models via a cross-validation approach (Gelfand, Dey, and Chang, 1992), which aims to validate the conditional predictive distributions when each observation is left out. We denote the remaining data by $\mathbf{A}_{\backslash hj}$ when the jth visit of subject h is left out and use a checking function $d_{hj} = \text{logit}(A_{hj}) - E[\text{logit}(\mathcal{A}_{hj}) \mid \mathbf{A}_{\backslash hj}]$, where $E[\text{logit}(\mathcal{A}_{hj}) \mid \mathbf{A}_{\backslash hj}]$ is the posterior predictive mean of $\text{logit}(\mathcal{A}_{hj})$.

The overall fit of a candidate model is evaluated as the sum of squares of the checking functions, i.e. $D = \sum_{h,j} d_{hj}^2$. The best-fitting model among the candidates, Model (M1), (M2), and (M3), is the one with the smallest D (Gelfand, Dey, and Chang, 1992).

A straightforward approach is to fit the population model multiple times with one observation left out each time. Instead, we use the method described by Chen, Shao and Ibrahim

(2000) to simplify the posterior predictive check. We denote the posterior distribution by $P(\boldsymbol{\theta} \mid \cdot)$ and the distribution of the data by $P(A_{hj} \mid \boldsymbol{\theta})$. It follows from $P(\boldsymbol{\theta} \mid \mathbf{A}) \propto P(\boldsymbol{\theta} \mid \mathbf{A}_{\setminus hj})P(A_{hj} \mid \boldsymbol{\theta})$, where $\mathbf{A} = (A_{hj}, \mathbf{A}_{\setminus hj})$, that

$$
\begin{aligned}
d_{hj} &= \operatorname{logit}(A_{hj}) - \int \int \operatorname{logit}(\mathcal{A}_{hj})P(\mathcal{A}_{hj} \mid \boldsymbol{\theta})P(\boldsymbol{\theta} \mid \mathbf{A}_{\setminus hj})d\boldsymbol{\theta}d\mathcal{A}_{hj} \\
&= \operatorname{logit}(A_{hj}) - \frac{\int \int \operatorname{logit}(\mathcal{A}_{hj})P(\mathcal{A}_{hj} \mid \boldsymbol{\theta})P(\boldsymbol{\theta} \mid \mathbf{A})/P(A_{hj} \mid \boldsymbol{\theta})d\boldsymbol{\theta}d\mathcal{A}_{hj}}{\int P(\boldsymbol{\theta} \mid \mathbf{A})/P(A_{hj} \mid \boldsymbol{\theta})d\boldsymbol{\theta}}.
\end{aligned}
$$

The above equation only depends on $P(\boldsymbol{\theta} \mid \mathbf{A})$ not $P(\boldsymbol{\theta} \mid \mathbf{A}_{hj})$. Therefore, the population model only needs to be fitted once with all the observations.

3. Sketch of Implementation

We implement the subject-level longitudinal image model by running separate Monte Carlo chains for each subject. At every iteration, we update the parameters for the J_h visits sequentially. For each visit j $(1 \leqslant j \leqslant J_h)$, the conditional posterior distribution of $\boldsymbol{\mu}_j$, $\boldsymbol{\sigma}_j^2$, and $\boldsymbol{\nu}_j$ have closed form and are updated via standard Gibbs sampling steps. The spatial regularization parameter, β_j, requires a Metropolis-Hastings step. We use the Swendsen-Wang algorithm (Swendsen and Wang, 1987; Higdon, 1998; Zhang et al. 2008b) to update the hidden configurations. The most challenging part of the sampler is to update the number of components for each visit via RJMCMC. The trans-dimensional proposal follows Zhang et al. (2008b) with some modifications (see Appendix for details). We then compute the tumor core and annulus AUCs for each visit. After all parameters indexed by visits are updated, we finish the current iteration by drawing σ_e^2 via a Gibbs step and proceed to the next iteration.

In the population AUC model, the conditional posterior distribution of $\boldsymbol{\theta}$, ϕ^2, ψ^2, and σ^2 have closed form by Bayes Theorem and are updated via Gibbs steps. The change point location, τ, is updated via a Metropolis-Hastings step. The best model is chosen via cross-

validation (Gelfand et al., 1992). The MCMC estimate of d_{hj} is as follows

$$\hat{d}_{hj} = \mathrm{logit}(A_{hj}) - \frac{S^{-1}\sum_{s=1}^{S}\mathrm{logit}(\mathcal{A}_{hj}^{(s)})/P\big(A_{hj}\mid\boldsymbol{\theta}^{(s)}\big)}{S^{-1}\sum_{s=1}^{S}1/P\big(A_{hj}\mid\boldsymbol{\theta}^{(s)}\big)},$$

where $\big(\mathcal{A}^{(s)},\boldsymbol{\theta}^{(s)}\big)$ are posterior draws at iteration $s(1\leqslant s\leqslant S)$. We fit the candidate Models (M1)-(M3) and select the one with the smallest $\hat{D}=\sum_{h,j}\hat{d}_{hj}^{2}$.

4. Results

To evaluate the performance of the longitudinal image model, we conduct a series of simulation studies. In each set, we generate multiple simulations, indexed by $l=1,2,\cdots,L$. We compute the average Mean Squared Error (aMSE) and average bias (abias) for the "true" contrast uptake estimate for simulation l at visit j, both averaged over pixels within an image, $\mathrm{aMSE}(\boldsymbol{\nu}_{j}^{l})=N^{-1}\sum_{i=1}^{N}(\hat{\nu}_{ij}^{l}-\nu_{ij}^{l,\mathrm{true}})^{2}$ and $\mathrm{abias}(\boldsymbol{\nu}_{j}^{l})=N^{-1}\sum_{i=1}^{N}(\hat{\nu}_{ij}^{l}-\nu_{ij}^{l,\mathrm{true}})$, where $\nu_{ij}^{l,\mathrm{true}}$ is the true contrast uptake used to generate the lth simulation and $\hat{\nu}_{ij}^{l}$ is its estimate (i.e. posterior mean). We also define the average aMSE and abias across the L simulations as $\overline{\mathrm{MSE}}(\boldsymbol{\nu}_{j})=L^{-1}\sum_{l=1}^{L}\mathrm{aMSE}(\boldsymbol{\nu}_{j}^{l})$ and $\overline{\mathrm{bias}}(\boldsymbol{\nu}_{j})=L^{-1}\sum_{l=1}^{L}\mathrm{abias}(\boldsymbol{\nu}_{j}^{l})$. To illustrate the benefits of modeling the temporal correlation, we compare this method with a set of cross-sectional models on the simulated data, i.e. fitting an independent model for the observed contrast uptake image at each visit as Zhang et al. (2008b). Essentially, the cross-sectional model assumes $y_{ij}=\eta_{ij}+e_{ij}$, $e_{ij}\sim\mathrm{N}(0,\omega_{ij}^{2})$ where the value of η_{ij} and ω_{ij}^{2} are indexed by hidden label Z_{ij}, i.e. $\eta_{ij}=\xi_{kj}$ and $\omega_{ij}^{2}=\upsilon_{kj}^{2}$ when $z_{ij}=k$. We further run the algorithm on the real dataset and summarize the differential tumor/healthy tissue response with AUC, and select the best fitting population AUC model for the tumor core and annulus separately.

4.1 *Simulation Study of Random "Change" Images*

We assume there are six images $(J=6)$ on 128×128 lattice. We divide the lattice into a skeleton of regions in various shapes (Figure 2). The same skeleton is used for all visits. For each visit, we assign the same value of true "change" to all pixels in the same region

($\delta_{ij} = \delta_{i'j}$ for all j and $i \sim i'$), which is independently and randomly drawn from eight candidate levels ($M_j = 8$ and $\boldsymbol{\mu}_j = (-5.95, -4.25, -2.55, -0.85, 0.85, 2.55, 4.25, 5.95)^{\mathrm{T}}$ for all j). We set the standard deviation of all evolution errors to 0.7, i.e. $\sigma_{kj} = 0.7$ for all j and k. The standard deviation of the measurement errors is $\sigma_e = 1$. The top row of Figure 3 shows an example of the true "change" images. We follow the model assumptions and generate the "true" contrast uptake images from $\nu_{ij} = \nu_{i,j-1} + \delta_{ij} + \varepsilon_{ij}$, where $\boldsymbol{\nu}_0 = 0$ and $\varepsilon_{ij} \sim \mathrm{N}(0, \tau_{ij}^2)$. The observed images are generated as $y_{ij} = \nu_{ij} + e_{ij}$, where $e_{ij} \sim \mathrm{N}(0, \sigma_e^2)$ (Figure 3 third row). Twenty sets of images are generated in the above fashion ($L = 20$).

[Figure 2 about here.]

We run the proposed algorithm on each of the twenty simulated datasets. The pixel-wise posterior means of $\boldsymbol{\delta}_j$ and $\boldsymbol{\nu}_j$ for the example dataset are displayed in Figure 3 second and fourth row. The posterior mean images of the true "change" are similar to the corresponding true values. However, we note that the errors are more likely to occur on the boundaries of two adjacent regions when they have means that are close to one another. By modeling the temporal correlation, the posterior mean images of the "true" contrast uptake ($\hat{\boldsymbol{\nu}}_j$) are less noisy than the observed images ($\mathbf{y}_j$) while the edges between regions with different "true" values are preserved.

We also compare the proposed longitudinal model with J independent cross-sectional models, one for each visit. The cross-sectional model assumes that each observed image is a hidden MRF deteriorated by Gaussian noise. At each visit j, the true pixel mean (η_{ij}) is the counterpart of the "true" contrast uptake (ν_{ij}) in the proposed model. We compare the average MSE of the two approaches averaged over pixels within an image. Both methods have minimal bias. However, the proposed longitudinal approach produces uniformly smaller average MSEs than the cross-sectional model by borrowing strength over time in addition to across neighboring pixels. Because the noises (both evolution and measurement errors)

add up over visits, the MSE is non-decreasing as the signal to noise ratio decreases for both approaches. The average MSEs of the cross-sectional model increase roughly as multiples of its baseline value while the longitudinal model only increases moderately for the follow-up visits ($j > 1$) after borrowing strength across visits. By way of data generation, the "true" contrast uptake is essentially continuous. We argue that the posterior draws of ν from the longitudinal model are on a continuous scale for $j > 1$, whereas the posterior draws of η in the cross-sectional model are discrete. This explains why the proposed method fits the data better with smaller average MSEs.

We investigated the sensitivity of the results to alternative choices of the hyperprior $\beta_\sigma = 1$ and 5, instead of $\beta_\sigma = 3$. The $\overline{\text{MSE}}(\boldsymbol{\nu}_j)$ is not sensitive to the choice of λ — the change in average MSE is less than 5%.

[Figure 3 about here.]

[Table 1 about here.]

4.2 *Simulation Study of No Treatment Effect*

We generate a second set of simulations assuming no treatment effect. We assume the regions within the outer ring are the "tumor" with relatively low baseline contrast uptake, while the rest of the regions are "healthy tissue". At baseline ($j = 1$), we assign the "tumor" regions with means randomly drawn from $(-5.95, -4.25, -2.55, -0.85)$, and the means of the "healthy" regions drawn from $(-2.55, -0.85, 0.85, 2.55, 4.25, 5.95)$. The other parameters remain the same as in Section 4.1. The regions within the outer ring have smaller means (darker colors) than the surrounding regions, on average. The rest of the true "change" in contrast uptake is set to zero, i.e. $\boldsymbol{\delta}_j = 0$ for $j > 1$. We then generate $\boldsymbol{\nu}_j$ and $\mathbf{y}_j$. We generate twenty sets of such datasets and run the proposed algorithm on each set (not shown). Similarly, the average MSEs of the proposed longitudinal model are significantly smaller than

the MSEs of the cross-sectional model (Table 1) with the comparative advantage increasing over time.

4.3 *Simulation Study of Treatment Effect*

We also simulate a third scenario assuming a treatment effect. At baseline, $\boldsymbol{\delta}_1$ is generated in the same way as Section 4.2. During the hypothetical radiotherapy period (the second and third visits, $j = 2$ and 3), we assume the "diseased" regions are more responsive to radiation than the "healthy" regions — the true "changes" of the "diseased" regions are randomly drawn from $(2.55, 4.25)$, while the "change" of the "healthy" regions are drawn from $(-0.85, 0.85)$. Hence, the center regions are brighter than the surrounding regions during the treatment period. We assume no long-term effect after the completion of treatment, i.e. $\boldsymbol{\delta}_j = 0$ for $j = 4, 5, 6$. We set the rest of the parameter values in the same way as Section 4.2. Again, twenty sets of such simulations are generated ($L = 20$). The MSEs of the proposed longitudinal model are once again smaller than the MSEs of the cross-sectional model (Table 1) with the difference increasing over time.

4.4 *Application*

With satisfactory results from the simulation studies, we fit our proposed model to the real dataset. Taking subject 1 for example, the posterior mean of "true" contrast uptake (Figure 4) smoothes the observed images (Figure 1). Further, the posterior mean image of true "change" in contrast uptake (Figure 4) reveals a differential pattern for the diseased and healthy tissue.

[Figure 4 about here.]

4.4.1 *Population Model of the AUC.* We use the 95th percentile of baseline healthy tissue contrast uptake to segment the tumor core from the annulus and fit separate population models for them (see Figure 5a for core and Figure 5c for annulus). The population models

are fitted after logit transformation of the posterior means of the AUCs, and then transferred back to its original scale (Section 2.2).

For the tumor core, the best fitting model is Model (M3) with both a random intercept and a random slope for the second linear segment. It profiles the steep increase in the population AUC earlier on after the initiation of radiotherapy, and then the moderate decrease afterwards (Table 2). The change point location is around ten days after the initiation of radiotherapy (mode is 10 days, first and third quartiles are 9 and 14, Figure 5b). The solid and dashed black curves in Figure 5a display the posterior mean and its point-wise 95% credible interval of the population AUC. The population mean reaches its maximum at around three weeks after the initiation of radiation.

We notice that the AUC profile of one subject (the bottom one in Figure 5a) is notably different from the rest of the subjects. We refit the model excluding this subject. The parameter estimates do not change substantially (Table 2).

[Figure 5 about here.]

[Table 2 about here.]

Model (M2) is the best fitting model for the tumor annulus (Table 2). Figure 5c shows that the population AUC first decreases and then remains rather flat. The posterior distribution of the change point location centers around one week after the initiation of radiotherapy (mode is 7 days, first and third quartiles are 6 and 10, Figure 5d). We also notice the AUC of one subject fluctuates more than usual in Figure 5c and refit the model excluding this subject, the parameter estimates are listed in Table 2. The change in the parameter values is small.

4.4.2 *Sensitivity Analysis.* In the above analysis, the tumor core and annulus are segmented based on the 95th percentile of baseline healthy tissue contrast uptake. We conduct

a sensitivity analysis on other choices of threshold values of the tumor core versus annulus, i.e. 90th and 97.5^{th} percentile of healthy tissue contrast uptake at baseline. The parameter estimates of the tumor core are given in Table 2. Using a higher threshold, a larger portion of the tumor with high baseline contrast uptake compared to other core pixels is included in the tumor core. This leads to higher AUCs, and hence a larger intercept term in the population model α^c. Since the slopes of the two linear pieces do not change much, the temporal profile of the differential tumor/healthy tissue response to radiation is not sensitive to the threshold in the range of threshold value investigated.

The parameters of the tumor annulus population model are more sensitive to the choice of the core/annulus threshold (Table 2). As the threshold increases (e.g. 97.5^{th} percentile), a larger portion of the tumor with low baseline contrast uptake (compared to other tumor annulus pixels) is excluded from the tumor annulus, which leads to larger AUCs (the probability of randomly picking a pixel in the tumor annulus with higher baseline contrast uptake than a healthy pixel is higher), and hence the intercept term in the population model is larger. Furthermore, the change point location is shifted towards zero, which reflects a steeper decline in AUCs early on after the initiation of radiotherapy.

5. Discussion

A notable feature of the longitudinal image model is that it simultaneously models spatial and temporal correlations. In many cases, the two correlations are modeled separately or sequentially. In this manuscript, we use hidden MRFs to model the spatial correlation in the true "change" of contrast uptake. Furthermore, we build the current "true" contrast uptake on the "true" contrast uptake at the previous visit, to model temporal correlation in a manner that is computationally tractable.

In many statistical problems, a population model that borrows strength across individuals is more desirable than modeling each individual separately. The proposed longitudinal image

model is fitted for each subject independently. We argue that the number of subjects in the motivating study is rather limited compared to the heterogeneity of tumors. Hence, a pixel-level population model is not necessarily more beneficial than the current one, e.g. imposing a hyperprior on the pixel "true" change, μ. It adds one more level on the model hierarchy, and is computationally more difficult. We believe that the disadvantages outweigh the benefits. Therefore, we only fit individual longitudinal image models in this work.

Although the longitudinal image model discussed in this manuscript helps to "denoise" the images by borrowing strength both across neighboring pixels and over time, the tens of thousands of pixel-wise contrast uptake profiles do not provide useful information individually. We summarize the differential tumor/healthy tissue responses with AUC, which is intuitive and interpretable.

The analysis in this manuscript only deals with the area under the entire ROC curve. The partial AUC ($\mathrm{pAUC}(t_0) = \int_0^{t_0} \mathrm{ROC}(t)dt$) is of possible interest, especially when there is special interest in the area under the ROC curve with a low false positive fraction. The partial AUC is considered a compromise between the AUC and a single point $\mathrm{ROC}(t_0)$ at a given threshold value t_0 (McClish 1989; Thompson and Zucchini 1989).

The models and computer algorithms have been developed for either 2D or 3D data. For the purpose of illustrating the model and algorithm, we analyze the MRI slice with the largest tumor volume. It is possible that the results could change if the entire brain volume is used.

Extensions of the proposed model include adding covariate information in the longitudinal image model, for example radiation dose maps, and specifying more flexible functional forms for the mean model, such as a cubic-spline, which better represents the nature of biological response to radiotherapy.

Acknowledgments

References

Besag J (1974), Spatial Interaction and the Statistical Analysis of Lattice Systems (with discussions), *Journal of the Royal Statistical Society. Series B*, **36**:192-236.

Cao Y, Tsien CI, Shen Z, Tatro DS, Ten Haken R, Kessler ML, Chenevert TL, and Lawrence TS (2005), Use of Magnetic Resonance Imaging to Assess Blood-Brain/Blood-Glioma Barrier Opening During Conformal Radiotherapy, *Journal of Clinical Oncology* **23**, 4127-4136.

Chen HM, Shao QM, Ibrahim JG (2000), Monte Carlo Methods in Bayesian Computation, Springer.

Curran WJ, Scott CB, Horton J, et al. (1993), Recursive Partitioning Analysis of Prognostic Factors in three Radiation Therapy Oncology Group Malignant Glioma Trials, Journal of National Cancer Insistitue, **85**:704-710.

Egan JP (1975), Signal Detection Theory and ROC Analysis, *Academic Press, New York*.

Gelfand AE, Dey DK, and Chang H (1992), Model Determination using Predictive Distributions with Implmentation via Sampling-Based Methods, *Bayesian Statistics*, **4** 147-167.

Gelman A and Meng XL (1998), Simulating Normalizing Constants: from Importance Sampling to Bridge Sampling to Path Sampling, *Statistical Science*, **13**(2), 163-185.

Green DM and Swets JA (1966), Signal Detection Theory and Psychophysics, *John Wiley & Sons*.

Green PJ (1995), Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination, *Biometrika*, **82**(4):711-732.

Green PJ and Richardson S (1997), On Bayesian Analysis of Mixtures with an Unknown Number of Components, *Journal of the Royal Statistical Society. Series B*, **59**(4):731-792.

Hamstra DA, Chenevert TL, Moffat BA, Johnson TD, Meyer CR, Mukherji SK, Quint DJ, Gebarski SS, Fan X, Tsien CI, Lawrence TS, Junck L, Rehemtulla A, and Ross BD (2005), Evaluation of the functional diffusion map as an early biomarker of time-to-progression and overall survival in high-grade glioma, *Proceedings of the National Academy of Sciences* **46**, 16759-16764;

Higdon D (1998), Auxiliary Variable Methods for Markov Chain Monte Carlo with Applications, *Journal of the American Statistical Association*, **93**.

Medical Research Council (2001), Randomized Trial of Procarbazine, Lomustine, and Vincristine in the Adjuvant Treatment of High-Grade Astrocytoma: A Medical Research Council Trial, *J. Clin. Oncol.*, 19(2):509-518.

Moffat BA, Chenevert TL, Lawrence TS, Meyer CR, Johnson TD, Dong Q, Tsien C, Mukherji S, Quint DJ, Gebarski SS, Robertson PL, Junck LR, Rehemtulla A, and Ross BD (2005), Functional diffusion map: A noninvasive MRI biomarker for early stratification of clinical brain tumor response, *Proceedings of the National Academy of Sciences* **102**, 5524-5529.

Ogata Y (1989), A Monte Carlo Method for High Dimensional Integration, *Numerische Mathematik*, **55**, 137-157.

Pepe MS (2003), The Statistical Evaluation of Medical Tests for Classification and Prediction, *Oxford University Press*.

Zhang X, Johnson TD, Little RJA, and Cao Y (2008a), Quantitative Magnetic Resonance Image Analysis via the EM Algorithm with Stochastic Variation, *Annals of Applied Statistics*, **2(2)**, 736-755.

Zhang X, Johnson TD, Little RJA, and Cao Y (2008b), A Bayesian Image Analysis of the Change in Tumor/Brain Contrast Uptake, *The University of Michigan Department of Biostatistics Working Paper Series*, Working Paper **74**.
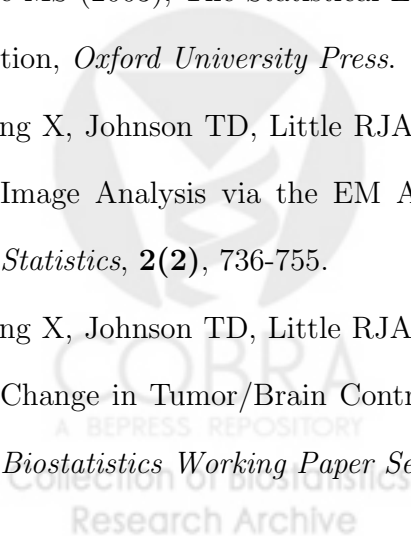
Appendix

We treat the number of components at each visit as an independent parameter, and estimate it via RJMCMC (Green, 1995; Green and Richardson, 1997). We follow Zhang et al. (2008b) in trans-dimensional proposals.

We first randomly choose between a split and a merge proposal with

$$
\mathrm{P_{split}}(M_j) = \begin{cases} 0 & \text{if } M_j = M_{\max} \\ 1 & \text{if } M_j = M_{\min} \\ 0.5 & \text{otherwise} \end{cases} \quad \text{and } \mathrm{P_{merge}}(M_j) = \begin{cases} 0 & \text{if } M_j = M_{\min} \\ 1 & \text{if } M_j = M_{\max} \\ 0.5 & \text{otherwise} \end{cases} .
$$

When a trans-dimensional move for subject h, visit j, is proposed (with M_j^* components), the new configuration is denoted by $\mathbf{z}_j^* = (z_{1j}^*, \cdots, z_{Nj}^*)^{\mathrm{T}}$ with $z_{ij} \in \{1, 2, \cdots, M_j^*\}$, the new component parameters are $\boldsymbol{\mu}_j^* = (\mu_{1j}^*, \cdots, \mu_{M_j^* j}^*)^{\mathrm{T}}$ and $\boldsymbol{\sigma}_j^* = (\sigma_{1j}^*, \cdots, \sigma_{M_j^* j}^*)^{\mathrm{T}}$. We also update the "true" contrast uptake at visit j with new values $\boldsymbol{\nu}_j^*$.

Split Proposal. If a split is chosen, we randomly pick a component k ($1 \leqslant k \leqslant M_j$) to split, i.e. $\mathrm{P_{select}^{split}}(k) = 1/M_j$. We need two extra parameters for the Gaussian mean and variance. To match the increase of dimension, we introduce two independent random variables, $u_1, u_2 \sim$ Beta$(2, 2)$, and define a bijective transformation $(\mu_{k_1}^*, \mu_{k_2}^*, {\sigma_{k_1}^*}^2, {\sigma_{k_2}^*}^2) = \psi(\mu_k, \sigma_k^2, u_1, u_2)$ that matches the first two moments,

$$
\mu_{k_1}^* = \mu_k - u_1 \sigma_k, \qquad \mu_{k_2}^* = \mu_k + u_1 \sigma_k,
$$

$$
{\sigma_{k_1}^*}^2 = 2u_2(1 - u_1^2)\sigma_k^2, \qquad {\sigma_{k_2}^*}^2 = 2(1 - u_2)(1 - u_1^2)\sigma_k^2. \tag{A.1}
$$

The acceptance rate of a split proposal is the product of prior ratio, likelihood ratio, proposal ratio, and the Jacobian of the bijective transformation.

The prior ratio in a split move (from k in M_j components to k_1 and k_2 in $M_j^* = M_j + 1$

components) is

$$\frac{\mathrm{P}(M_j^*)\mathrm{P}(\beta_j^*)\mathrm{P}(\mathbf{Z}_j^* \mid \beta_j, M_j) \prod_{k=1}^{M_j^*} \left\{\mathrm{P}(\mu_{kj}^*)\mathrm{P}(\sigma_{kj}^2{}^*)\right\}}{\mathrm{P}(M_j)\mathrm{P}(\beta_j)\mathrm{P}(\mathbf{Z}_j \mid \beta_j, M_j) \prod_{k=1}^{M_j} \left\{\mathrm{P}(\mu_{kj})\mathrm{P}(\sigma_{kj}^2)\right\}}$$

$$= \frac{\beta_j^{*\,\alpha_\beta-1} \exp\{-\beta_\beta \beta_j^*\}}{\beta_j^{\alpha_\beta-1} \exp\{-\beta_\beta \beta_j\}} \times \frac{g^{-1}(\beta_j^*, M_j^*) \exp\left\{\sum_{i \sim i'} \beta_j^* \mathrm{I}[z_{ij}^* = z_{i'j}^*]\right\}}{g^{-1}(\beta_j, M_j) \exp\left\{\sum_{i \sim i'} \beta_j \mathrm{I}[z_{ij} = z_{i'j}]\right\}}$$

$$\times \frac{\sigma_\mu^{-1} \exp\left\{\sigma_\mu^{-2}(\mu_{k_1 j}^* - \mu_\mu)^2\right\} \sigma_\mu^{-1} \exp\left\{-.5\sigma_\mu^{-2}(\mu_{k_2 j}^* - \mu_\mu)^2\right\}}{\sigma_\mu^{-1} \exp\left\{-.5\sigma_\mu^{-2}(\mu_{kj} - \mu_\mu)^2\right\}}$$

$$\times \frac{\left(\sigma_{k_1 j}^{-2}{}^*\right)^{\alpha_\sigma-1} \exp\left(-\beta_\beta \sigma_{k_1 j}^{-2}{}^*\right) \left(\sigma_{k_2 j}^{-2}{}^*\right)^{\alpha_\sigma-1} \exp\left(-\beta_\beta \sigma_{k_2 j}^{-2}{}^*\right)}{\left(\sigma_{kj}^{-2}\right)^{\alpha_\sigma-1} \exp\left(-\beta_\beta \sigma_{kj}^{-2}\right)}.$$

The likelihood ratio term in the split proposal is

$$\prod_{i=1}^{N} \exp\left[-.5\sigma_e^{-2}\left\{(y_{ij} - \nu_{ij}^*)^2 - (y_{ij} - \nu_{ij})^2\right\}\right].$$

We construct the reversible split and merge move on the equivalence classes, i.e. the lattice is partitioned into same-labeled equivalence classes. In order to compute the proposal ratio, we need to calculate the allocation probabilities associated with both moves.

We start with a split move. Let $\epsilon_1, \epsilon_2, \cdots, \epsilon_L$ denote the L equivalence classes with label $E_1, E_2, \cdots, E_L$. The allocation probability of class ϵ_l ($l = 1, 2, \cdots, L$) assuming label $k^*(k^* = 1, 2, \cdots, M_j + 1)$ in the proposed configuration $\mathbf{z}_j^*$ is

$$p_{k^*}^{l}{}^* = \mathrm{P}(Z_{ij}^* = k^*, \text{ for all } i \in \epsilon_l)$$

$$\propto \prod_{i \in \epsilon_l} \exp\left\{-.5\sigma_e^{-2}(y_{ij} - \nu_{ij}^*)^2\right\}.$$

The allocation probability from $\mathbf{z}_j^*$ is $\mathrm{P}_{\mathrm{alloc}}(\mathbf{z}_j^*) = \prod_{l=1}^{L} p_{E_l}^{l}$. The probability of class ϵ_l assuming label $k(k = 1, 2, \cdots, M_j)$ satisfies

$$p_k^l = \mathrm{P}(Z_{ij} = k, \text{ for all } i \in \epsilon_l)$$

$$\propto \prod_{i \in \epsilon_l} \exp\left\{-.5\sigma_e^{-2}(y_{ij} - \nu_{ij})^2\right\}.$$

The $p_k^{l}{}^*$ satisfy $\sum_{k=1}^{M_j+1} p_k^{l}{}^*$ for all $l = 1, 2, \cdots, L$. Let $E_l^*(E_l^* = 1, 2, \cdots, M_j+1)$ denote the new label of class ϵ_l drawn based on the above probabilities, i.e. $p_{E_l^*}^{l}{}^* = \mathrm{P}(Z_{ij}^* = E_l^*, \forall i \in \epsilon_l)$.

Hence, the allocation probability from configuration $\mathbf{z}_j$ conditional on the equivalence classes is $\mathrm{P}_{\mathrm{alloc}}(\mathbf{z}_j) = \prod_{l=1}^{L} p_{E_l^*}^l{}^*$.

The proposal ratio of a split move is

$$\frac{\mathrm{P}_{\mathrm{split}}(M_j - 1)\mathrm{P}_{\mathrm{select}}^{\mathrm{split}}(k)\mathrm{P}_{\mathrm{alloc}}(\mathbf{z}_j^*)}{\mathrm{P}_{\mathrm{merge}}(M_j)\mathrm{P}_{\mathrm{select}}^{\mathrm{merge}}(k_1, k_2)\mathrm{P}_{\mathrm{alloc}}(\mathbf{z}_j)}.$$

Merge Proposal. The merge proposal is determined by the above split move. We randomly pick a pair of components with adjacent means, say k_1 and k_2 with $\mathrm{P}_{\mathrm{select}}^{\mathrm{merge}}(k_1, k_2) = 1/(M_j - 1)$. Inverting Equation (A.1), the parameters for the new component are

$$\mu_k^* = 0.5\left(\mu_{k_1} + \mu_{k_2}\right),$$

$$\mu_k^{*2} + \sigma_k^{*2} = 0.5(\mu_{k_1}^2 + \sigma_{k_1}^2 + \mu_{k_2}^2 + \sigma_{k_2}^2).$$

We denote the new set of parameters by $\boldsymbol{\mu}^* = (\mu_1^*, \mu_2^*, \cdots, \mu_{M-1}^*)^{\mathrm{T}}$ and $\boldsymbol{\sigma}^{*2} = (\sigma_1^{*2}, \sigma_2^{*2}, \cdots, \sigma_{M-1}^{*2})^{\mathrm{T}}$. The proposal ratio of a merge move is the reciprocal of the split move.

24

Baseline, $\mathbf{y}_1$

One week after treatment initiation, $\mathbf{y}_2$

Three weeks after treatment initiation, $\mathbf{y}_3$

One month after treatment completion, $\mathbf{y}_4$

Three months after treatment completion, $\mathbf{y}_5$

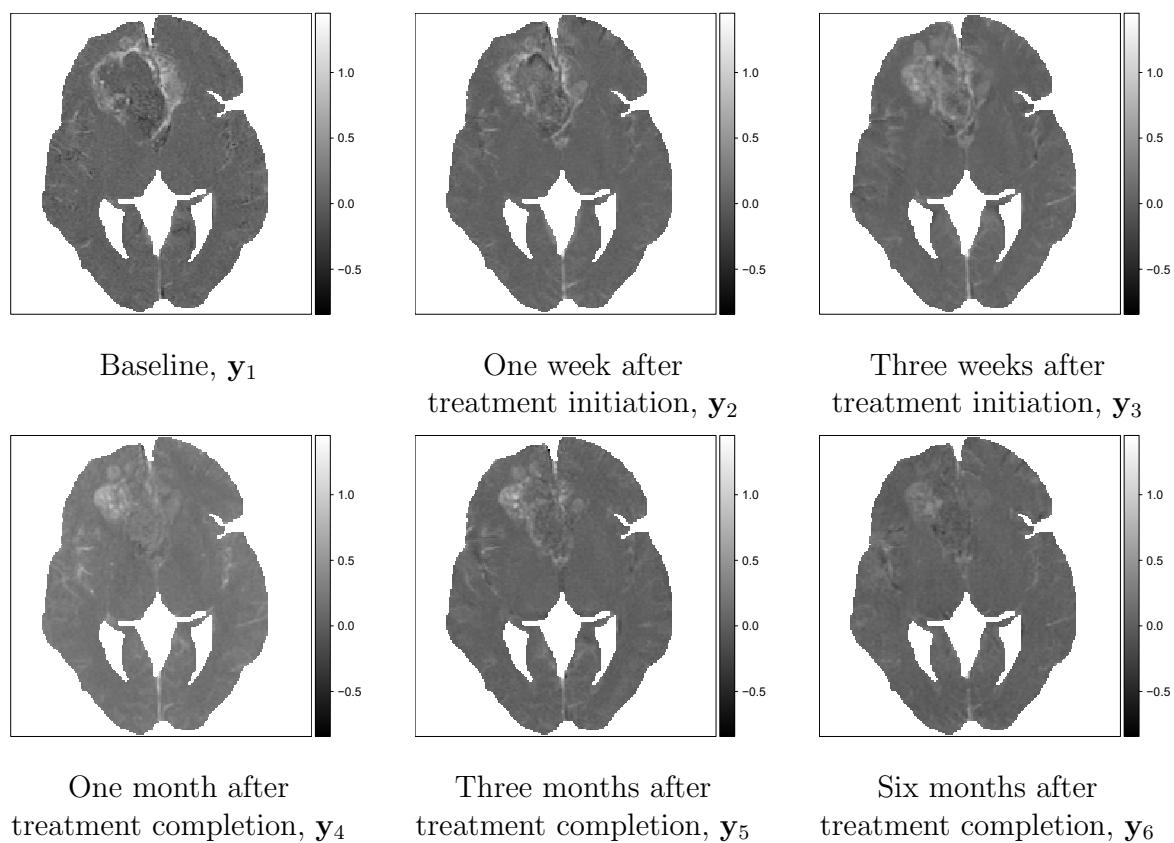Six months after treatment completion, $\mathbf{y}_6$

Figure 1. Observed contrast uptake $(\mathbf{y}_j)$ for subject 1.

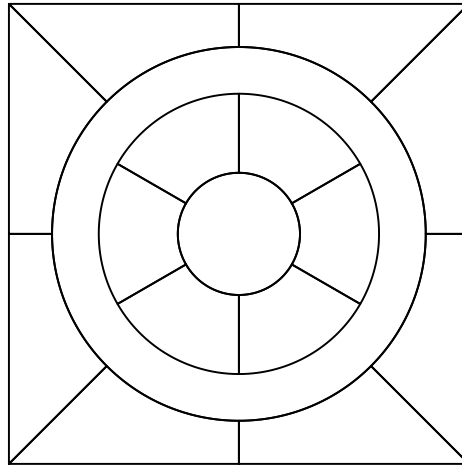**Figure 2.** The skeleton of "true" change image.
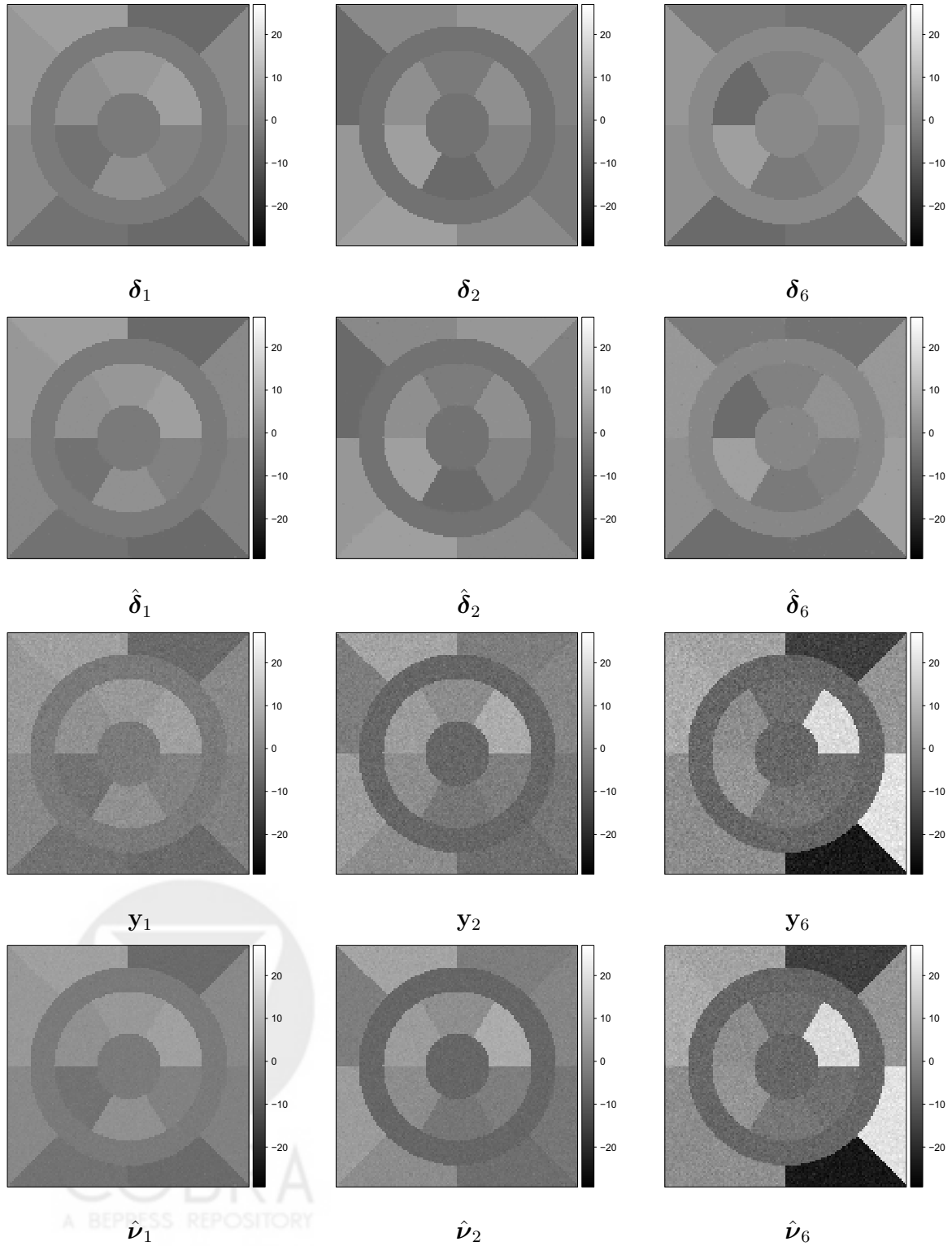
Figure 3. Simulation study of random true "change" images: true "change" in contrast uptake ($\boldsymbol{\delta}_j$), posterior mean ($\hat{\boldsymbol{\delta}}_j$) observed contrast uptake ($\mathbf{y}_j$), and posterior mean of "true" contrast uptake ($\hat{\boldsymbol{\nu}}_j$).

$$\hat{\boldsymbol{\delta}}_1 \qquad\qquad\qquad \hat{\boldsymbol{\delta}}_2 \qquad\qquad\qquad \hat{\boldsymbol{\delta}}_6$$

$$\hat{\boldsymbol{\nu}}_1 \qquad\qquad\qquad \hat{\boldsymbol{\nu}}_2 \qquad\qquad\qquad \hat{\boldsymbol{\nu}}_6$$
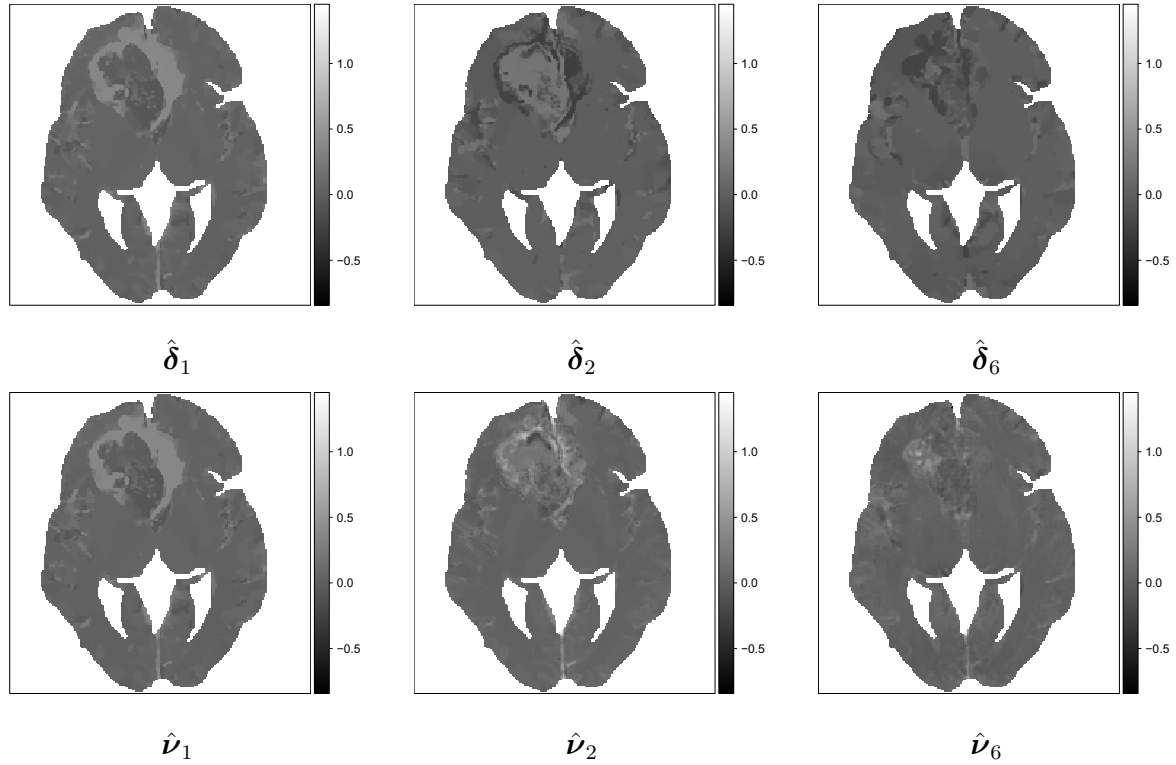
Figure 4. Results for subject 1: posterior mean of $\boldsymbol{\delta}_j$ and posterior mean of $(\boldsymbol{\nu}_j)$ for subject 1.

(a)

(b)

(c)

(d)

Figure 5. Results on tumor core versus healthy tissue: (a) posterior mean of AUC for all subjects and point-wise posterior mean and 95% credible interval of population AUC; (b) posterior distribution of the change point location (τ); results on tumor annulus versus healthy tissue: (c) posterior mean of AUC for all subjects and point-wise posterior mean and 95% credible interval of population AUC; (d) posterior distribution of the change point location (τ).
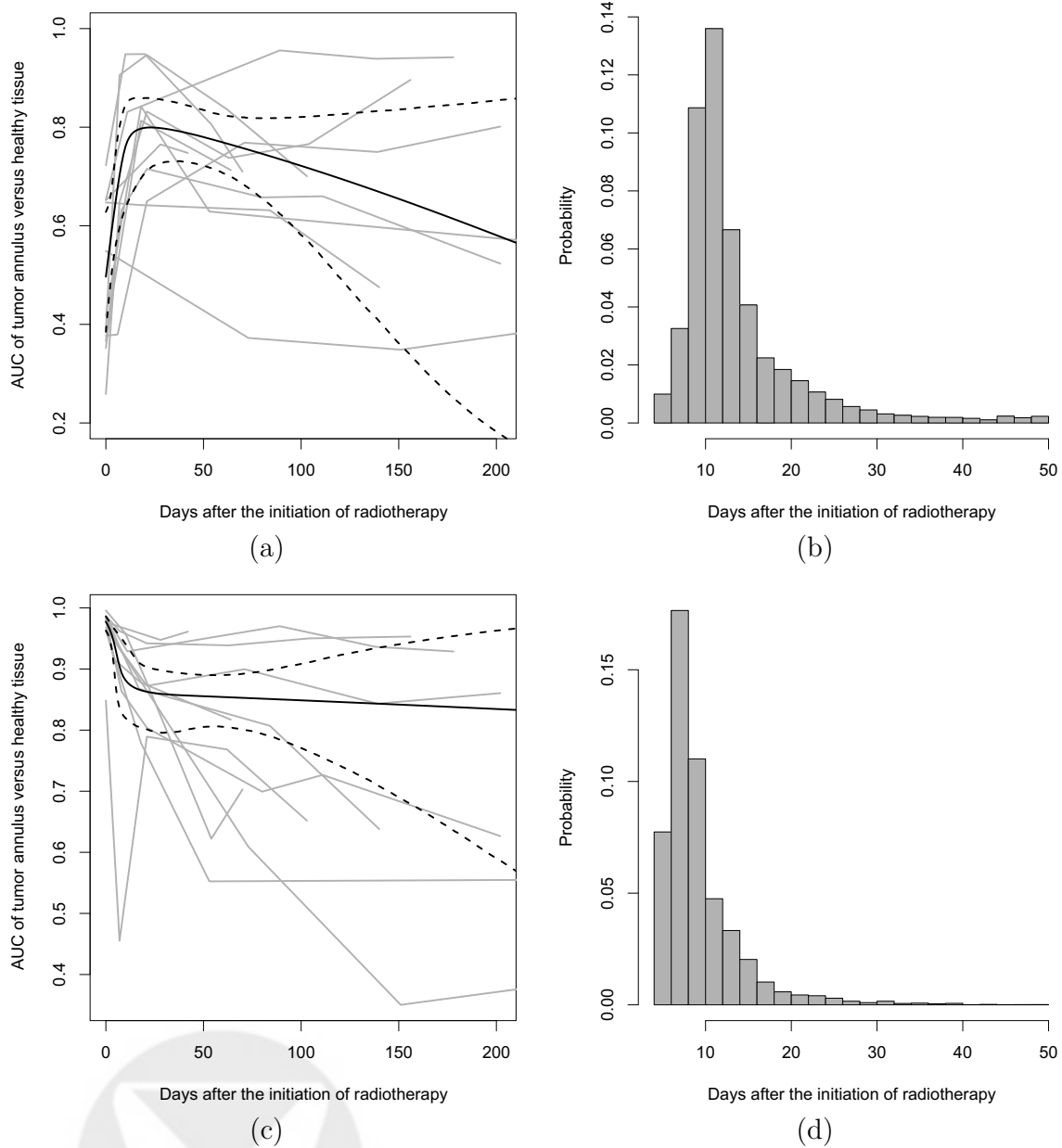
Table 1
Average MSE and bias of the proposed method vs. a cross-sectional approach — simulation studies of random
"change" images, no treatment effect (NTE), and treatment effect (TE).

		$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
Rand.	Long.	2.60 ± 0.03	3.28 ± 0.07	3.45 ± 0.05	3.54 ± 0.04	3.89 ± 0.08	5.20 ± 0.08
$(\times 10^{-1})$	C.S.	5.22 ± 0.47	9.81 ± 0.40	14.32 ± 0.73	19.90 ± 1.18	24.27 ± 1.09	29.2 ± 1.18
NTE	Long.	3.04 ± 0.10	3.75 ± 0.05	3.74 ± 0.05	3.85 ± 0.04	3.92 ± 0.03	5.67 ± 0.07
$(\times 10^{-1})$	C.S.	5.64 ± 0.16	11.45 ± 0.79	17.01 ± 1.07	21.35 ± 0.95	25.59 ± 0.72	29.62 ± 0.78
TE	Long.	2.71 ± 0.15	3.48 ± 0.19	3.80 ± 0.07	3.83 ± 0.04	3.93 ± 0.03	5.62 ± 0.07
$(\times 10^{-1})$	C.S.	5.63 ± 0.16	9.83 ± 0.65	14.66 ± 0.81	20.30 ± 1.51	25.26 ± 1.70	30.36 ± 1.99

Table 2

Sensitivity analysis of posterior parameter estimates for the population AUC model for tumor core and annulus versus healthy tissue.

	90th		95th		97.5^{th}	
	All subj.	Excl. one	All subj.	Excl. one	All subj.	Excl. one
Core						
τ	12.6 ± 4.4	12.3 ± 3.5	13.7 ± 7.2	12.9 ± 5.2	16.2 ± 10.1	15.4 ± 8.8
α	-0.21 ± 0.21	-0.26 ± 0.19	-0.03 ± 0.24	-0.07 ± 0.22	0.15 ± 0.26	0.10 ± 0.25
β_1	0.14 ± 0.05	0.15 ± 0.04	0.13 ± 0.05	0.14 ± 0.05	0.11 ± 0.06	0.12 ± 0.06
$\beta_2(\times 10^{-2})$	-0.65 ± 0.42	-0.75 ± 0.42	-0.60 ± 0.52	-0.70 ± 0.49	-0.61 ± 0.64	-0.74 ± 0.65
ϕ	0.69 ± 0.09	0.62 ± 0.08	0.75 ± 0.10	0.67 ± 0.09	0.78 ± 0.10	0.71 ± 0.09
ψ	0.14 ± 0.03	0.15 ± 0.04	0.14 ± 0.03	0.15 ± 0.04	0.14 ± 0.03	0.15 ± 0.04
σ	0.13 ± 0.03	0.14 ± 0.03	0.13 ± 0.03	0.14 ± 0.03	0.13 ± 0.03	0.14 ± 0.03
Annulus						
τ	14.1 ± 10.0	15.0 ± 5.3	10.1 ± 5.2	13.6 ± 4.2	8.3 ± 3.0	11.7 ± 3.2
α	3.36 ± 0.26	3.57 ± 0.19	3.77 ± 0.26	3.97 ± 0.18	4.15 ± 0.26	4.38 ± 0.18
β_1	-0.17 ± 0.09	-0.13 ± 0.05	-0.23 ± 0.09	-0.17 ± 0.05	-0.29 ± 0.09	-0.22 ± 0.06
$\beta_2(\times 10^{-2})$	-0.02 ± 0.50	-0.01 ± 0.50	-0.04 ± 0.43	-0.02 ± 0.43	-0.10 ± 0.46	0.01 ± 0.39
ψ	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01
σ	0.65 ± 0.17	0.37 ± 0.09	0.67 ± 0.16	0.36 ± 0.08	0.71 ± 0.17	0.35 ± 0.08