A Bayesian Nonparametric Approach for Functional Data Classification with Application to Hepatic Tissue Characterization

Kassandra M. Fronczyk[1], Michele Guindani[2], Brian P. Hobbs[2], Chaan S. Ng[3] and Marina Vannucci[4]

[1]Research Staff Member, Operational Evaluation Division, Institute for Defense Analyses, Alexandria, VA, USA. [2]Assistant Professor, Department of Biostatistics, University of Texas MD Anderson Cancer Center, Houston, TX, USA. [3]Professor, Department of Diagnostic Radiology, University of Texas MD Anderson Cancer Center, Houston, TX, USA. [4]Professor, Department of Statistics, Rice University, Houston, TX, USA.

Supplementary Issue: Statistical Systems Theory in Cancer Modeling, Diagnosis, and Therapy

ABSTRACT: Computed tomography perfusion (CTp) is an emerging functional imaging technology that provides a quantitative assessment of the passage of fluid through blood vessels. Tissue perfusion plays a critical role in oncology due to the proliferation of networks of new blood vessels typical of cancer angiogenesis, which triggers modifications to the vasculature of the surrounding host tissue. In this article, we consider a Bayesian semiparametric model for the analysis of functional data. This method is applied to a study of four interdependent hepatic perfusion CT characteristics that were acquired under the administration of contrast using a sequence of repeated scans over a period of 590 seconds. More specifically, our modeling framework facilitates borrowing of information across patients and tissues. Additionally, the approach enables flexible estimation of temporal correlation structures exhibited by mappings of the correlated perfusion biomarkers and thus accounts for the heteroskedasticity typically observed in those measurements, by incorporating change-points in the covariance estimation. This method is applied to measurements obtained from regions of liver surrounding malignant and benign tissues, for each perfusion biomarker. We demonstrate how to cluster the liver regions on the basis of their CTp profiles, which can be used in a prediction context to classify regions of interest provided by future patients, and thereby assist in discriminating malignant from healthy tissue regions in diagnostic settings.

KEYWORDS: functional data analysis, Bayesian analysis, Bayesian nonparametrics, computed tomography perfusion

Introduction

Functional imaging techniques have been progressively used in oncology to improve tumor detection, characterization, staging, prognosis, and treatment of cancer patients.[1] In particular, computed tomography perfusion (CTp) has recently emerged as a widely available, inexpensive and noninvasive, imaging technique to evaluate changes in tumor vascular physiology and tumor biology. CTp can potentially lead to biomarkers for tumor characterization, as well as for early assessment of therapeutic response in multiple cancers.[2,3] Any modern CT scanner system can perform CTp: a small bolus of iodine-based intravenous contrast is injected rapidly over a target region, and repeated images are taken at regular frequent time points before, during, and after the passage of the contrast agent through the tumor vasculature. Hence, this dynamic acquisition allows to measure temporal changes in tissue density after injection with the contrast agent. Tissue perfusion can then be estimated in the target region through the use of different kinetic models.[4]

The importance of assessing tumor perfusion and vascular permeability is related to their association with tumor angiogenesis, ie, the formation of new blood vessels, in several tumors.[5–8] Neovascularization is considered an important process in cancer progression and tumor growth. For example, high tumor angiogenesis activity has been often associated with distant metastases and adverse clinical outcomes.[9–11]

The angiogenic vasculature of malignant solid tumors is usually characterized by dense, dilated, and tortuous microvessels. In particular, blood volume (BV) and blood flow (BF) are higher in malignant tumors due to the proliferation of new blood vessels through the process of tumor angiogenesis.[12,13] Therefore, it has been suggested that perfusion imaging can yield biomarkers of angiogenesis and tumor growth and thereby greatly enhance the clinical development of antiangiogenic therapies. This is because changes in tumor perfusion appear to occur soon after therapy initiation with antiangiogenic drugs.[14–16]

CTp parameters are commonly calculated after the scan session by using commercially available postprocessing software platforms (CT Perfusion 4, version 4.3.1, Advantage Windows 4.4; GE Healthcare). Typical CTp parameters include BF (mL/100 g/minute), BV (mL/100 g), mean transit time (MTT, seconds), and permeability surface area product (PS, mL/100 g/minute), which are obtained using standard deconvolution physiological modeling, based on the central volume principle (BF = BV/MTT) first described in the context of cerebral perfusion.[17,18] PS is a measure of capillary permeability and it reflects the flux of solutes from blood plasma to the interstitial space.[19] As an illustration, Figure 1 reports the time course of the PS parameter observed during the imaging session for two representative normal and tumor regions of interest (ROIs) in a single patient. The disruption and decline in PS values within the first 100 seconds of the scan in Figure 1B are usually the result of neoangiogenesis and microvascular attenuation.

The trajectory of the perfusion parameters over the scan period can be affected by many factors, which may pertain to patient and tumor characteristics. For example, in squamous cell carcinoma, BF and PS values have been shown to be significantly higher in subjects with longer local tumor control than in subjects with local recurrence.[20] This consideration highlights the heterogeneity of the perfusion parameter–time curves both between ROIs within a patient as well as between patients, in addition to the fact that the temporal resolution and duration of the CT acquisition can also impact the CT perfusion parameter values.[21,22] Yet, heretofore, the clinical implications that have been *inferred* from perfusion CT studies have predominantly relied on simple hypothesis testing approaches, which discard much of the acquired information. By way of contrast, more advanced statistical approaches that facilitate characterization of distinct perfusion signatures that are attributable to different types of tissues (eg, normal tissues vs different stages of tumor growth), as well as similar cluster patterns of responses within and between patients, could provide new insights for enhancing the use of this technology in diagnostic settings. Moreover, analytical tools that facilitate the study of CTp heterogeneity could be critical for the development and effective application of perfusion-based surrogate markers of therapeutic response.[23]

In this manuscript, we consider a functional data analytic approach[24] and assume that the observations collected on each CTp parameter during the scan period are a realization, up to a measurement error, of an underlying perfusion process. The corresponding *perfusion curves* are estimated in a flexible way through the use of Gaussian processes (GP),[25] which are characterized by possibly different parameters for the normal and tumor tissues. In order to take into account the heterogeneity of the functional CTp responses, we further propose to borrow strength in the estimation of the inferred time course patterns by using a Bayesian nonparametric mixture model, more specifically by employing a *functional* Dirichlet process prior.[26–28] This modeling choice also allows us to achieve dimension reduction by clustering the observed dynamics of the CTp parameters into a smaller set of *canonical* response curves. On the contrary, estimation of the CTp parameters using any parametric higher order polynomial basis function would suffer from nonlocal influence, and thus would be unable to accommodate the inherently local temporal trends that are exhibited in the data.

Bayesian hierarchical models and Dirichlet Process (DP) mixtures have been successfully exploited to represent functional curves, eg, clustering spline coefficients in Bayesian multivariate adaptive regression splines models.[29] In computer modeling[30] and machine learning,[25] Gaussian process realizations are often used as a basis to model random functions.[31] Finite mixtures and Dirichlet process mixtures of Gaussian processes have also been proposed to model a sample of curves, possibly encoding complex spatiotemporal or covariate dependencies.[32–37]

With respect to those contributions, the method we propose for the analysis of functional CTp responses takes explicitly into account prior information about the perfusion imaging experiment and the vascular physiology of normal
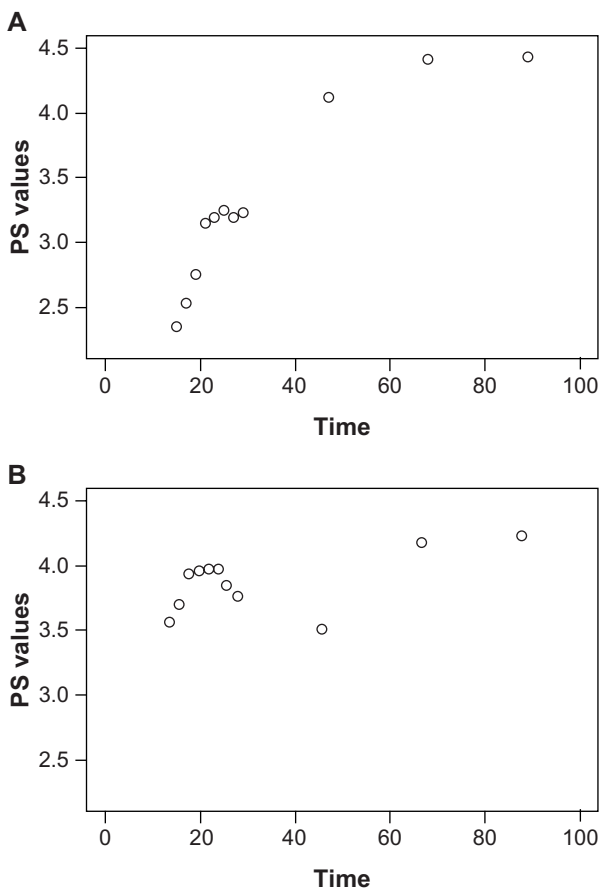

Figure 1. Illustrative plot of the observed measurements on the PS perfusion characteristic in two ROIs for a representative patient, as a function of scan time, zoomed in the first 100 seconds of the scan: (**A**) normal tissue and (**B**) tumor tissue.

and tumor tissues. Because the time course mappings of CT perfusion parameters are expected to stabilize some time after injection of the contrast agent, in order to obtain a reliable quantification of the perfusion characteristics,[21] our proposal accommodates change-points in the distribution of the functional responses over time, namely, in their temporal covariance structure. Moreover, since the perfusion time courses may vary across tissue types, we also allow for distinct change-points in the curves characterizing the normal and tumor tissues.

In addition, we demonstrate how to use the results from the posterior inference obtained from our Bayesian modeling framework to characterize ROIs from future patients as malignant tumor tissues to reflect diagnostic clinical settings. The classification is based on posterior predictive computations and the use of the Bayes factor, which weights the evidence in the data for the tested hypotheses.[38] More specifically, we describe how the classification can be conducted by compounding the contribution of time courses from multiple perfusion parameters. Figure 2 summarizes the main features of our contribution.

The remainder of the manuscript is organized as follows. The "Methods" section describes the details of Bayesian nonparametric approach. Methods for posterior inference and classification are succinctly described in "Gaussian processes with varying autocorrelation" and "Posterior inference" sections. In the "Results" section, we describe the posterior and clustering results for analysis of tissue permeability in a sample of 16 patients with neuroendocrine liver metastases. We also provide an illustration of the classification performance of our algorithm on the basis of two studies: one aimed at comparing the predictive ability of our model with respect to commonly used classifiers in our small dataset, and a more comprehensive simulation study that further illustrates its general properties. We provide some concluding remarks in the "Discussion" section.

Methods

Bayesian mixture for modeling the heterogeneity of CTp characteristics. Without loss of generality in this section, we detail our modeling strategy for a single perfusion

characteristic. More specifically, let $\boldsymbol{Y}_{ij} = (Y_{ij}(t_{1ij}), ..., Y_{ij}(T_{ij}))^T$ denote the vector of perfusion values collected on each ROI $j = 1, ..., n_i$, and on each patient $i = 1, ..., n$. Notice that we allow for scans obtained at different temporal frequencies across regions and patients. Furthermore, let z_{ij} be a binary variable, such that $z_{ij} = 1$ if the jth ROI corresponds to a tumor region, and $z_{ij} = 0$ if the jth ROI is from normal tissue. In this section, we assume to know the tissue type (ie, tumor or normal) of each sampled region. However, this might not always be the case. In the "Classification performance" section, we provide a way to predict from the perfusion data if a tissue is normal or tumor based on the knowledge accumulated from previous data within the two conditions.

Following the typical functional data analysis framework,[24] we envision that for either the normal or tumor tissue, the CTp responses can be formally described by the following model,

$$Y_{ij} = \theta_{ij} + \epsilon_{ij}, \quad i = 1, ..., n,$$

where the ϵ_{ij} are independent realizations of a Gaussian white noise process with variance $\sigma^2_{z_{ij}}$ dependent on the tissue type, ie, $\epsilon_{ij} \sim N\left(0, \sigma^2_{z_{ij}} I_{d_{ij}}\right)$, with d_{ij} indicating the number of elements in the observation vector, Y_{ij}, and I_{dij} a d_{ij}-dimensional identity matrix. Equivalently, given the θ_{ij}, we can write

$$Y_{ij} \mid \theta_{ij} \overset{\mathrm{ind}}{\sim} N_{d_{ij}}\left(\theta_{ij}, \sigma^2_{z_{ij}} I_{d_{ij}}\right), \quad z_{ij} = 0, 1. \quad (1)$$

Arguably, the inferential interest often pertains to the modeling of the mean vector θ_{ij}, which can be seen as the realization of a random function $\theta(t)$ on the observation points, ie, the data are supposed to be noisy realization of the underlying *perfusion curves*, $y_{ij}(t) = \theta_{ij}(t) + \epsilon_{ij}(t)$, with $\epsilon_{ij}(t) \overset{\mathrm{i.i.d.}}{\sim} N\left(0, \sigma^2_{z_{ij}}\right), t \in \mathbb{R}^+$. A typical assumption specifies $\theta_{ij}(t)$'s as independent realizations of a Gaussian process. Instead, we propose to borrow strength in the estimation by introducing probabilistic dependence across the θ_{ij}'s. More
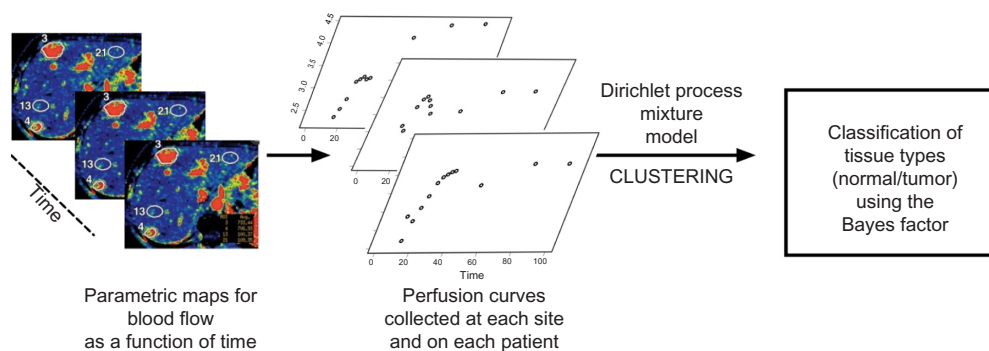

Figure 2. General scheme of our modeling framework and inferential objectives.

specifically, we assume that the θ_{ij}'s are sampled from a prior (unknown) probability measure $G_{z_{ij}}$ on $\mathbb{R}^+$, which is also tissue-dependent, that is, $\theta_{ij} \mid G_{z_{ij}} \overset{\text{i.i.d.}}{\sim} G_{z_{ij}}, z_{ij} = 0, 1$. The vector of observed CTp parameter values can then be described by the following location mixture of Gaussian densities,

$$Y_{ij} \mid G_{z_{ij}} \overset{\text{i.i.d.}}{\sim} \int N_{d_{ij}}\left(\cdot \mid \theta_{ij}, \sigma^2_{z_{ij}} I_{d_{ij}}\right) G_{z_{ij}}\left(d\theta_{ij}\right), z_{ij} = 0, 1, \quad (2)$$

where, for notational simplicity, we have used the symbol $G_{z_{ij}}(\cdot)$ to also indicate the finite-dimensional distribution of the random function $\theta^{z_{ij}}_{ij}(t)$ observed at time points $t_{1ij}, \ldots, T_{ij}$.

Functional Dirichlet process priors. We follow a Bayesian nonparametric approach and further assign a prior probability to the mixing distributions $G_{z_{ij}}(\cdot), z_{ij} = 0, 1$. More specifically, we assume a *functional* Dirichlet process prior,[27] by taking $G_{z_{ij}}$ as the realization of a Dirichlet process.[26] The Dirichlet process has been widely employed in Bayesian nonparametric (BNP) models due to its various properties. For example, the use of a Dirichlet process prior allows borrowing of information across observations in the estimation of the parameters of the model by automatically providing an unsupervised clustering of the data. More formally, the Dirichlet process formulation implies that $G_{z_{ij}}$ is almost surely a discrete probability measure of the form,

$$G_{z_{ij}}(\cdot) = \sum_{l=1}^{\infty} p_l \, \delta_{\theta_l^*}(\cdot), \quad (3)$$

where δ_θ denotes a probability measure degenerate on the atom θ_l^*. In our framework, this implies that $G_{z_{ij}}$ concentrates probability masses p_l on a set of *canonical* perfusion curves θ_l^* in $\mathbb{R}^+$, so that curves characterized by similar trajectories can be clustered together. Both the weights p_l's and the curves θ_l^* are random functions, with probability laws specified as follows. For the sequence of weights $(p_1, p_2, \ldots)$, we assume a stick-breaking prior with parameter a, ie, $p_1 = V_1, p_l = V_l \prod_{r=1}^{l-1}(1 - V_r), V_r \overset{\text{i.i.d.}}{\sim}$ *Beta* $(1, \alpha)$.[39] The θ_l^*'s are i.i.d. from a nonatomic probability measure G_0 on Θ, independent of the p_l's. In particular, G_0 is commonly regarded as a parametric centering (base) distribution, since $E(G) = G_0$, whereas the parameter $\alpha > 0$ is a precision parameter, since it controls the variability of G around G_0, with larger values of a resulting in realizations of G that are closer to G_0. By assuming a random mixing distribution G with a DP prior, we do not restrict the model to a specific parametric form, and we increase its flexibility to capture different types of trajectories of the perfusion values. In the following, we denote our functional prior by $G \sim f\text{DP}(\alpha, G_0)$.

Gaussian processes with varying autocorrelation. We further take G_0 as a stationary Gaussian process, that is, we assume that the distribution of the *canonical curves* θ_l^* on any finite set of time points in $\mathbb{R}^+$ is Gaussian.[25] Just as a multivariate Gaussian distribution is fully specified by its mean

and covariance matrix, a Gaussian process is also specified by the mean and a covariance function. In symbols, we write $\theta_l^* \overset{\text{i.i.d.}}{\sim} G_0 = \text{GP}\left(\theta_0, R(\phi)\right)$, where θ_0 may be either known (eg, a constant function) or itself assigned a prior. In the following, we assume $\theta_0 \sim N(0, \eta^2)$, with η^2 large so to encode vague information on the marginal expectation of the process. $R(\phi)$ represents the covariance function of the process, as a function of a set of parameters ϕ. The covariance function encodes the smoothness properties of the process (eg, mean square continuity and differentiability). A class of covariance functions that has proven to be attractive in various respects is the so-called *Matérn* covariance function.[40] Let $\phi = (\psi, v, \tau)$, with $\psi > 0$, $v > 0$, and $\tau > 0$, and let $\Delta_t > 0$ denote a time interval of length Δ_t from t. Then, the covariance function can be written as

$$R_{v,\psi}(\Delta_t) = \frac{2^{-v+1}\tau^2}{\Gamma(v)} (\psi \Delta_t)^v H_v(\psi \Delta_t)$$

where $H_v()$ is the modified Bessel function of order v,[41 ch. 9]. The parameter ψ can be viewed as a decay parameter, since it governs the rate at which the covariance drops as a function of the distance, and v is a parameter that controls the degree of smoothness of the process. For $v = \frac{1}{2}$, we get the exponential covariance function, $R_{v,\psi}(\Delta_t) = \tau^2 \exp\left(-\psi \Delta_t\right)$. The resulting stochastic process is continuous but not differentiable at the origin; thus, it may be appropriate for modeling curves that can possibly vary abruptly in their gradients. Finally, τ represents a variance term. In Bayesian inference for Gaussian processes, it is often common to assume $\tau = 1$, due to the poor identifiability of the model parameters, especially when the inferential interest is on estimating the correlation structure of the data.[42]

In order to obtain a reliable quantification of the perfusion characteristics, since the perfusion values are expected to show low autocorrelation soon after injection of the contrast agent, but subsequently they become more stable and highly correlated,[21] we allow for changes in the correlation structure of the process over time. More specifically, we assume that the decay parameter in the correlation function may vary with time. Mathematically, we could describe these changes as different states s_k of the process, for $k = 1, \ldots, K$ (K finite). Each state would be characterized by a specific correlation function $R_{s_k}(\phi)$, and we would allow the process to move across the different *correlation* states by employing a general hidden Markov model framework.[43,44] However, based on our knowledge of the perfusion characteristics, we only expect one change point during the scan period. Let $\bar{t}$ indicate such (unknown) change point. Therefore, we consider two states s_1 and s_2 and assume that, for $t > \bar{t}$, the process cannot revert anymore to the preceding state. We further assume that the change point may be typically different for the normal and tumor tissues

due to the different environment the contrast agent faces. Let $T_1 = [0, \bar{t}]$, $T_2 = [\bar{t}, T]$, where we set $T = \max_{ij} T_{ij}$ for notational simplicity. Then, we assume that the base distribution of the functional Dirichlet prior (3) is characterized as

$$G_0^{z_{ij}} = GP\left(\theta_0^{z_{ij}}, R^{z_{ij}}\left(\phi_{z_{ij}} \right) \right), \tag{4}$$

where $R\left(\phi_{z_{ij}} \right)$ is a block diagonal matrix of the form

$$R\left(\phi_{z_{ij}} \right) = \begin{bmatrix} R\left(\phi_{1z_{ij}} \right) & 0 \\ 0 & R\left(\phi_{2z_{ij}} \right) \end{bmatrix},$$

with $R\left(\phi_{1z_{ij}} \right)$ defined on T_1 and $R\left(\phi_{2z_{ij}} \right)$ defined on T_2. Note that our formulation implicitly assumes conditional independence of the observations before and after the change point. For $z_{ij} = 0, 1$, we can summarize our model as follows

$$Y_{ij} \mid \theta_{ij}, \sigma_{z_{ij}} \sim N\left(\theta_{ij}, \sigma_{z_{ij}}^2 I \right) \quad j = 1, \dots n_i; i = 1, \dots, \imath$$
$$\theta_{ij} \mid G_{z_{ij}} \sim G_{z_{ij}}$$
$$G_{z_{ij}} \sim f\,DP\left(\alpha_{z_{ij}}, G_0^{z_{ij}} \right), \tag{5}$$
$$p\left(\sigma_{z_{ij}}, \phi_{z_{ij}}, \alpha_{z_{ij}}, \bar{t}_{z_{ij}} \right) \sim p\left(\sigma_{z_{ij}} \right) p\left(\phi_{z_{ij}} \right) p\left(\alpha_{z_{ij}} \right) p\left(\bar{t}_{z_{ij}} \right),$$

where $G_0^{z_{ij}} = GP\left(\theta_0^{z_{ij}}, R^{z_{ij}}\left(\phi_{z_{ij}} \right) \right)$, and $p\left(\sigma_{z_{ij}} \right)$ is typically an inverse gamma prior, $p\left(\alpha_{z_{ij}} \right)$ is a gamma distribution, and $p\left(\phi_{z_{ij}} \right)$ and $p\left(\bar{t}_{z_{ij}} \right)$ are both (independent) uniform distributions.

Posterior inference. Lai and Xiang[45] have recently considered a simple Bayesian model for multiple parameter changes in a multiparameter exponential family, developing explicit formulas for the estimators of the change-points. Our Dirichlet process formulation does not allow the use of the formulas of Lai and Xiang. Thus, we rely on a Markov chain Monte Carlo (MCMC) algorithm to obtain samples from the posterior distribution of the parameters of interest and conduct posterior inference. In this section, we provide a very brief description of the basic steps of the MCMC algorithm. Since perfusion parameters can be acquired at different time points across patients, we start by considering the union set of all observed time points. We then treat the values of the perfusion parameters at the times not observed as missing at random.[46,47] Following the standard Bayesian approach,[48] we impute the missing values within each MCMC iteration by drawing from the relevant posterior predictive distributions conditionally on the observed data and the currently sampled values of the parameters. Due to the normality and conjugacy of the model, the corresponding posterior predictive distribution is multivariate normal. More precisely, let Y_{ij}^{obs} denote the set of observed values for ROI j and patient i, and let Y_{ij}^{mis} indicate

the missing values, which need to be imputed at each iteration. Let $H = \sigma^2 I + \left(R^{z_i}\left(\phi_{z_{ij}} \right) \right)$, with $H^{\text{mis,mis}}$, $H^{\text{obs,obs}}$, and $H^{\text{mis,obs}}$ denoting the corresponding submatrices. Then, the distribution of Y_{ij}^{mis} conditional on Y_{ij}^{obs} is multivariate normal with mean $\tilde{\mu}_{ij} = \theta_0^{z_{ij}, \text{miss}} + H^{\text{miss,obs}} \left[H^{\text{obs,obs}} \right]^{-1} \left(Y_{ij}^{\text{obs}} - \theta_0^{z_{ij}, \text{obs}} \right)$ and covariance matrix $\tilde{\Sigma}_{ij} = \sigma_{z_{ij}}^2 \left\{ H^{\text{obs,obs}} - H^{\text{miss,obs}} \left[H^{\text{obs,obs}} \right]^{-1} \left[H^{\text{miss,obs}} \right]^{\text{T}} \right\}$. The remainder of the model parameters is updated as described below, where, for notational simplicity, we omit the subscript z_{ij}.

a. Sampling the canonical curves θ_l^*. We update the canonical curves using the algorithm in Ref. 49, which is based on the Pólya urn representation of the Dirichlet process.[26,50] The canonical curves are then resampled within each cluster to improve the mixing of the chain, similar to the algorithm originally described in Ref. 51.

b. Updating the change point $\bar{t}$. We use a Metropolis–Hastings step to update t. Since the set of time points is discrete, we consider a multinomial proposal, such that the proposed change point is sampled with probability proportional to the values of the unnormalized likelihoods at each time t, calculated assuming $t = \bar{t}$.

c. Updating the covariance parameters $\left\{ \phi_{1z_{ij}}, \phi_{2z_{ij}} \right\}$. The update of $\left\{ \phi_{1z_{ij}}, \phi_{2z_{ij}} \right\}$ depends on the particular form of the covariance function we specify. For the data analysis described in "Discussion" section, we use a Matérn with parameter $v = 1/2$. Then, the resulting exponential correlations are parameterized by the decay parameters $\psi_{1z_{ij}}$ and $\psi_{2z_{ij}}$ for the process in states s_1 and s_2, respectively. Conditional on the current sampled value of $\bar{t}$ at the MCMC iteration, $\left\{ \psi_{1z_{ij}}, \psi_{2z_{ij}} \right\}$ are updated with standard Metropolis–Hastings techniques.

The updates of the other model parameters are standard.[52] For example, the full conditional of θ_0 is multivariate normal and that of σ^2 is an inverse gamma. The update of a follows the procedure described in Ref. 49. We refer to the Supplementary File for a more detailed description of the MCMC algorithm.

The posterior samples so obtained can be used to provide posterior estimates of the parameters of interest, as well as highest posterior credible intervals. For example, the MCMC sample average provides an ergodic estimate of the posterior expected value of a parameter. Let $\theta_{ij}^{(b)}(t)$, $b = 1, \dots, B$, indicate the posterior draws of θ_{ij} at time t after burn-in. Then $E\left(\theta_{ij}(t) \mid \text{data} \right) \approx \frac{1}{B} \sum_{b=1}^{B} \theta_{ij}^{(b)}(t)$. The accuracy of the estimate will increase as B, the length of the chain, increases. We found

that in our applications, assuming $B = 10,000$ was generally enough to produce good estimates. Similarly, we can acquire 95% posterior credible intervals of the parameters by estimating the posterior density of the MCMC draws and determining the corresponding 2.5% and 97.5% quantiles.

Classification via Bayes factors. We can use the information obtained from posterior inference to classify new observations, eg, from ROIs for which the tissue class is heretofore unknown. Let Y_m^{new} indicate the CT perfusion time courses, measured for a characteristic $m = 1, \ldots, M$ on a given ROI. A clinician may be interested to classify the corresponding tissue based on the values of Y_m^{new} and the knowledge gathered from previous datasets. This can be accomplished by taking advantage of the updating scheme typical of the Bayesian framework. More specifically, we can compute the Bayes factor [BayesF, 38] to discriminate between normal and tumor tissues. The Bayes factor is related to the posterior odds of one hypothesis relative to another (eg, tumor vs normal tissue), and it actually coincides with them if the prior probability of each hypothesis is the same. The $\log_{10}$ of the Bayes factor has been traditionally referred to as the *weight of evidence* in the data for the tested hypothesis.[53] Then, if $\log_{10}(\text{BayesF}) > 0$, the data indicate some evidence against the hypothesis, with values $\log_{10}(\text{BayesF}) > 0.5$ and higher denoting substantial and increasing evidence. Let z^{new} be a binary indicator, such that $z^{new} = 1$ if the newly examined tissue is classified as tumor, and $z^{new} = 0$ if normal. Then, we can compute the Bayes factor for each characteristic $m = 1, \ldots, M$ as follows,

$$
\begin{aligned}
\text{BayesF}(Y_m^{new}) &= \frac{\Pr\left(Y_m^{new} \mid z^{new} = 1, \text{data}\right)}{\Pr\left(Y_m^{new} \mid z^{new} = 0, \text{data}\right)} \\
&= \frac{\int \Pr\left(Y_m^{new} \mid z^{new} = 1, \theta_{z=1}\right) \Pr\left(\theta_{z=1} \mid \text{data}\right) d\theta_{z=1}}{\int \Pr\left(Y_m^{new} \mid z^{new} = 0, \theta_{z=0}\right) \Pr\left(\theta_{z=0} \mid \text{data}\right) d\theta_{z=0}},
\end{aligned}
\tag{6}
$$

where the quantities $\Pr\left(Y_m^{new} \mid z, \text{data}\right)$ and $\Pr(\theta_z \mid \text{data})$ denote, respectively, the posterior predictive density of Y_m^{new} and the posterior probability of the model parameters if we assume that the data are from tissue type z, based on the already available CT perfusion data.

For clinical purposes, the classification of targeted ROIs shall rely on the Bayes factor computed either on a single perfusion characteristic (eg, blood flow) or from a combination of the perfusion characteristics, eg, only those which have been previously found in the literature to be associated with angiogenesis and tumor growth. For M different characteristics, the decision will be based on the evaluation of the following compound log-Bayes factor,

$$
\log \text{BayesF} = \sum_{m=1}^{M} w_m \log \text{BayesF}\left(Y_m^{new}\right),
\tag{7}
$$

where $w_m \geq 0$ represent weights that can be chosen to reflect the relative importance that a clinician might attribute to the different perfusion parameters in the decision process. If $w_m = 0$, then the mth perfusion characteristic will not contribute to the classification. Due to the lack of available prior information, the choice $w_1 = w_2 = \ldots = w_m = 1$ effectuates equivalence among the perfusion parameters. The choice of the weights could also be informed by existing knowledge about the dependence between perfusion characteristics, eg, by choosing weights inversely proportional to the strength of association of each CTp parameter with the others.

Results

A case study from patients with neuroendocrine liver metastases. In this section, we apply our Bayesian nonparametric approach to a retrospective clinical study. Both our study and the retrospective clinical study were approved by the MD Anderson Institutional Review Board (IRB). The study included 16 patients (6 men and 10 women) with metastatic neuroendocrine tumors who received optional CT perfusion imaging between April 2007 and September 2009 as part of two IRB approved clinical trials. CT perfusion was performed for a target lesion in the liver, which was clinically or radiologically determined to be malignant. Images were obtained with a 64-row multidetector CT scanner (VCT; GE Healthcare) and with acquisition time of 12–590 seconds.[21] We consider the time courses of four perfusion characteristics, as follows: BF, BV, MTT, and PS. The measurements were obtained on 27 separate ROIs in liver metastases and 25 separate ROIs in normal liver tissue. More specifically, for each of the eight axial slice locations of each dataset, a liver tumor ROI was drawn free hand around the periphery of the primary target lesion, using an electronic cursor and mouse, with reference to the source cine CT images and perfusion parametric maps, displaying the images at soft tissue windows (width = 350 HU, level = 40 HU). Wherever possible, a second tumor ROI was delineated, provided it was greater than 1.5 cm in diameter. ROIs were placed in the abdominal aorta and in the portal vein on the source images to provide these vascular inputs. Refer to Ref. 21 for more details. In this section, we illustrate the results of our analysis for the time courses of the PS characteristic, since PS values have been previously associated with the markers of angiogenesis.[54] More specifically, we apply model Equation (5) to a normalizing log-transformation of the PS perfusion values, and then perform posterior inference as described in the "Posterior inference" section. The log scale is frequently used in CT perfusion analysis to adjust for conditionally asymmetric residual error at a given acquisition time and to alleviate the effect of heteroskedasticity as a function of time. In terms of prior choices, we assume that the mean of the base G_0 $\theta_{0z_{ij}}$, is the null vector 0 with diagonal variance–covariance matrix $4I$. Given the range of values observed in CT perfusion data, this setting corresponds to a vague (noninformative) prior on the elements of $\theta_{0z_{ij}}$. The prior on the range parameter of the Matérn covariance function, $\psi_{z_{ij}}$, is taken to be uniform over the values from 0 to 10. The prior specification is completed by assuming

a gamma prior $\alpha \sim \text{Ga}(1, 1)$ on the precision parameter of the functional Dirichlet process, a vague inverse gamma prior on the sampling variance, $\sigma_{z_{ij}}^2 \sim \text{Inv-Ga}(5,5)$, and a uniform prior across the full range observed time domain $(0, 600)$ for the change point, $\bar{t}_{z_{ij}}$.

Figure 3 displays the posterior distributions of the number of clusters for the normal (left panel) and the tumor liver tissue (right panel). In both cases, the posterior is characterized by high variance and long tails. However, the modal number of clusters is estimated around 5 in the normal tissue and 6 in the tumor tissue type. Hence, we may conclude that there is evidence of heterogeneous clustering patterns among the CT perfusion time courses. The MCMC algorithm typically samples possibly different cluster configurations at each iteration. Several methods have been used in the literature to obtain a single-point estimate of clustering draws from the posterior distribution.[55,56] In the following, we consider the maximum a posteriori (MAP) estimate. The MAP clustering is the clustering estimate that maximizes the posterior density.

Figure 4 (top) displays a cubic spline interpolation of the original PS values, with different colors to highlight the different MAP clusters. In both normal and tumor samples, the clusters appear to represent different time course trajectories, either because of their behavior within the first 300 seconds of acquisition or because of the resultant level that they achieve at the end of the scanning period. Figure 5 shows the posterior distribution of the change point in the covariance structure of the CT perfusion values. It is well known that for many of the correlation functions typically used for Gaussian processes, the decay parameters are only weakly identifiable.[57,58] Accordingly, the posterior distribution of the change-points is quite diffuse. Nevertheless, one can recognize a mode within the first 100 seconds of the acquisition time. More specifically, the posterior expectation is 89.6 seconds (67.8 SD) for the normal tissue and 93.6 seconds (68.4 SD) for the tumor tissue. These values reflect the heteroskedasticity patterns typically observed in CT perfusion values.

Classification performance. In order to assess the classification performance of our Bayesian nonparametric modeling approach, we devise two different strategies. First, we consider the dataset analyzed in "A case study from patients with neuroendocrine liver metastases" section and we select one normal and one tumor curve for each of the 16 patients. We then use formula (6) to classify the selected curves either as normal or tumor, separately for each perfusion parameter. Ng et al.[21] have recently discussed the minimal acquisition duration that is necessary to attain stability and good quantification of the CT perfusion parameter values. There is clear motivation to reduce the overall duration of CT acquisition to the shortest time possible to reduce radiation exposure without compromising quantification of the CT perfusion parameter values. A few studies have suggested that acquisition times of 30–60 seconds might be satisfactory for some of the CT perfusion parameters.[59,60] However, Ng et al.[21] suggest that for most parameters stabilization with moderate confidence is attained only between 220 and 360 seconds of acquisition. Those studies did not address the predictive ability of the measurements collected over the different acquisition times to discriminate between tissue types. For this reason, in the following, we will evaluate the predictive ability of observations collected over 30–100 seconds in addition to the full-time courses over the complete 590 seconds of acquisition duration.
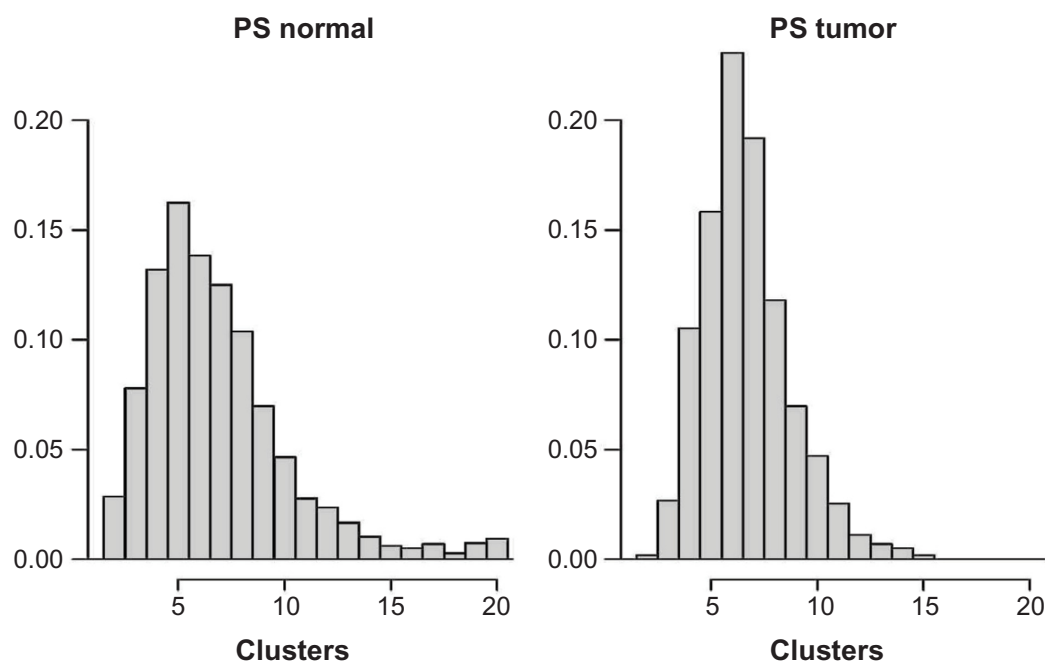

Figure 3. Posterior of the number of clusters of PS time courses for the normal (left) and tumor (right) liver ROIs.
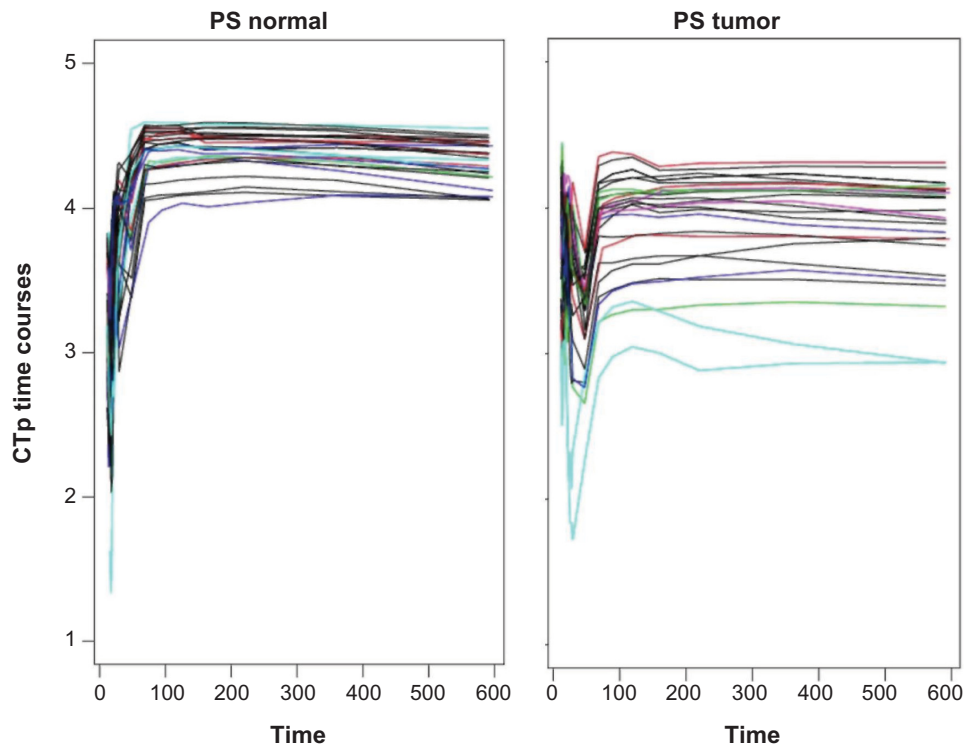
Figure 4. Posterior clustering: cubic spline interpolation of the observed log-PS values, color coded according to the MAP estimate for normal tissue type (left) and tumor tissue type (right).

For comparison, we also implement the commonly used classifiers, linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and support vector machine (SVM) algorithms. LDA assumes that the measurements from each class are normally distributed with the same variance–covariance matrix for both tumor and normal tissues, whereas QDA relaxes the homoskedasticity assumption. As these classifiers require the same number of time points for each observation, we first fit cubic spline interpolations to the data in order to be able to evaluate all the classifiers
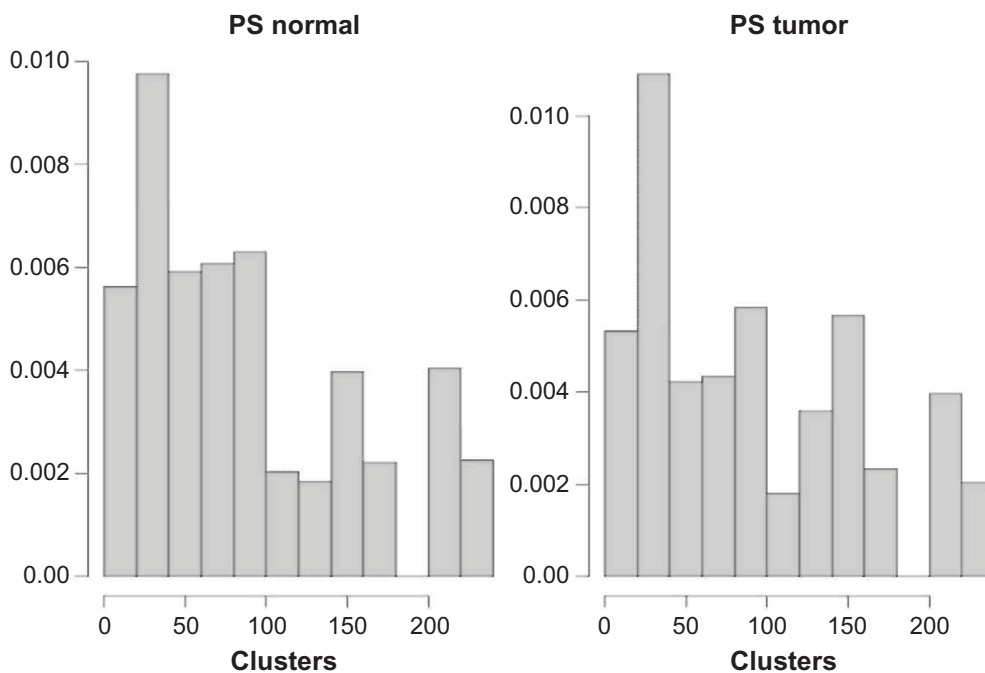

Figure 5. Posterior distribution of the change point in the correlation structure of log PS values for the normal (left) and tumor (right) liver tissues.

Table 1. Prediction true negative (normal) and true discover (tumor) rates using LDA, QDA, and the proposed BNP between 30–100 seconds and 0–590 seconds.

	LDA	LDA	QDA	QDA	SVM	SVM	BNP	BNP
Normal	(30–100 sec)	(0–590 sec)	(30–100 sec)	(0–590 sec)	(30–100 sec)	(0–590 sec)	(30–100 sec)	(0–590 sec)
logBF	75%	94%	81%	94%	75%	63%	100%	100%
logBV	69%	88%	63%	88%	75%	50%	94%	75%
logMTT	75%	88%	69%	94%	50%	56%	100%	100%
logPS	69%	88%	88%	94%	69%	69%	100%	100%
Tumor								
logBF	50%	88%	56%	94%	63%	56%	75%	56%
logBV	81%	88%	81%	88%	75%	50%	69%	56%
logMTT	69%	88%	88%	82%	50%	56%	82%	63%
logPS	94%	100%	94%	94%	69%	56%	100%	82%

on the same acquisition intervals. This is particularly true for the PS and the BF parameter, which is consistent with previous literature. The discriminant analysis techniques (run after curve interpolation) do not appear to perform consistently well, especially when using the first 30–100 seconds of the acquisition scans. This is particularly relevant considering that higher acquisition duration would require high overall radiation exposures.

Table 1 shows the true negative and true discover rates for the prediction of normal and tumor curves, separately for the four CT perfusion parameters. The BNP method we propose generally seems to perform better than the alternative methods in most cases.

To further test our classification method, we also devise a simulation strategy where we consider a larger dataset, obtained by generating 1,000 distinct time courses for each of the four CT perfusion parameters, as follows. First, we obtain spline interpolations of the observed time courses in the liver metastases case study considered in "A case study from patients with neuroendocrine liver metastases" section. The spline interpolations are fit using the R package spline on each of the 52 curves and for each of the four characteristics.[61] Then, we select one of the 52 original observations by sampling with repetition, separately for each characteristics. A new time course is generated as a draw from a multivariate normal distribution with the selected interpolation as the mean and a diagonal covariance matrix, with variance η^2. This simulation scheme allows us to evaluate the effect of the noise η^2 on the prediction and also to more accurately assess the asymptotic performance of our classifier under multiple realizations of the generating models.

Tables 2 and 3 report the results of the classification for varying degrees of the noise η^2 ($\eta^2 = 0.1, 0.2, 0.5$) using the first time frame of 30–100 seconds of acquisition duration and the full scanning period, respectively. The column corresponding to *Combination* refers to the compound log-Bayes factor in equation (7), where we assume all weights

$w_1 = w_2 = w_3 = w_4 = 1$. Such a choice can be considered as default, unless prior information leads to alternative choices. As expected, the prediction accuracy decreases with higher values of noise. The best predictor appears to be represented by the PS values, whose performance also appears more robust to higher levels of noise. Interestingly enough, the accuracy of the prediction with PS values increases slightly when using the full acquisition duration from 0 to 590 seconds, whereas for all other perfusion characteristics, a relatively better performance is obtained by considering the values between 30 and 100 seconds. This is consistent with the findings in the study by Ng et al.[21], where it is shown that the PS values require longer acquisition times than do BF, BV, and MTT to reach comparable levels of stabilization in CT perfusion values. Among the other parameters, the BF and the MTT also appear to be relatively robust to higher values of noise. BV appears to be the less informative of the parameters in terms of prediction performance. Finally, the compound Bayes factor (7) appears to be the best classifier only when considering the first 30–100 seconds of acquisition duration. If the full-time course is considered, the performance is notably reduced. This result appears to be mainly due to the negative accuracy of the BV, and therefore, one might consider assigning a weight w_m close to zero to this perfusion parameter.

Discussion

With the advances of the contrast-enhanced functional imaging technology, the development of noninvasive perfusion imaging biomarkers for tissue characterization, cancer prognostication, and detection has emerged as an area of recent focus in clinical cancer research. Moreover, these imaging features have the potential to enhance quantitative evaluations for measuring therapeutic response to antiangiogenic treatment strategies. Future endeavors to further develop and translate this technology should rely on appropriate analytical models for characterizing the multiple sources of variance that are inherent to the acquisition and measurement

Table 2. Classification results for the large simulation with 1,000 simulated time courses described in the "Classification performance" section, for varying degree of noise $\eta^2 = 0.1, 0.2, 0.5$ and considering the first 30–100 seconds of acquisition duration.

30–100 SECS	BF	BV	MTT	PS	COMBINATION
$\eta^2 = 0.1$					
Normal	342/503 (68%)	332/503 (66%)	397/503 (79%)	412/503 (82%)	423/503 (84%)
	<0.001	<0.001	0.03	0.36	
Tumor	357/497 (72%)	333/497 (67%)	378/497 (76%)	402/497 (81%)	442/497 (89%)
	<0.001	<0.001	<0.001	0.07	
$\eta^2 = 0.2$					
Normal	337/503 (67%)	322/503 (64%)	377/503 (75%)	357/503 (71%)	392/503 (78%)
	<0.001	<0.001	<0.001	0.01	
Tumor	338/497 (68%)	323/497 (65%)	347/497 (70%)	400/497 (80%)	407/497 (82%)
	<0.001	<0.001	<0.001	0.57	
$\eta^2 = 0.5$					
Normal	287/503 (57%)	206/503 (41%)	292/503 (58%)	312/503 (62%)	347/503 (69%)
	<0.001	<0.001	<0.001	0.02	
Tumor	293/497 (59%)	193/497 (39%)	293/497 (59%)	318/497 (64%)	359/497 72%)
	<0.001	<0.001	<0.001	0.005	

Note: The *P*-values of the comparison of each individual CT characteristics and the combination (two-sample test for equality of proportions) are reported under each result.

of temporal changes in contrast enhancement obtained from dynamic CT.

In this article, we presented a Bayesian nonparametric functional analytic approach for the analysis of CT perfusion time courses, which allows for heterogeneity observed across patients and tissues, by clustering the measurements of perfusion characteristics into groups characterized by similar temporal behavior. Our approach takes explicitly into account prior information about the perfusion imaging experiment and the vascular physiology of normal and tumor tissues, by accommodating change-points in the temporal covariance structure of responses over time in order to appropriately describe the

Table 3. Classification results for the large simulation with 1,000 simulated time courses described in the "Classification performance" section, for varying degree of noise $\eta^2 = 0.1, 0.2, 0.5$ and considering the full acquisition duration.

0–590 SECS	BF	BV	MTT	PS	COMBINATION
$\eta^2 = 0.1$					
Normal	331/503 (66%)	236/503 (47%)	377/503 (75%)	448/503 (89%)	247/503 (49%)
	<0.001	0.49	<0.001	<0.001	
Tumor	338/497 (68%)	203/497 (41%)	362/497 (73%)	457/497 (92%)	268/497 (54%)
	<0.001	<0.001	<0.001	<0.001	
$\eta^2 = 0.2$					
Normal	312/503 (62%)	221/503 (44%)	352/503 (70%)	411/503 (82%)	226/503 (45%)
	<0.001	0.75	<0.001	<0.001	
Tumor	313/497 (63%)	189/497 (38%)	337/497 (68%)	442/497 (89%)	238/497 (54%)
	<0.001	0.002	<0.001	<0.001	
$\eta^2 = 0.5$					
Normal	261/503 (52%)	150/503 (30%)	277/503 (55%)	352/503 (70%)	211/503 (42%)
	0.002	<0.001	<0.001	<0.001	
Tumor	278/497 (56%)	129/497 (26%)	278/497 (56%)	358/497 (72%)	249/497 (50%)
	0.07	<0.001	0.07	<0.001	

Note: The *P*-values of the comparison of each individual CT characteristics and the combination (two-sample test for equality of proportions) are reported under each result.

temporal heteroskedasticity observed in such measurements. We demonstrate how our approach can lead to improved performance with respect to commonly used classification methods for classifying ROIs in future patients, especially when considering the typical length of acquisition times. Therefore, our approach can assist in discriminating malignant from healthy tissue regions in diagnostic settings.

The clinical study was limited only to metastases to the liver from a specific tumor (metastases arising from neuroendocrine tumors) and consisted of a relatively small number of patients. The extent to which the conclusions of our case study might be generalizable to other tissues and tumors requires future exploration. However, we expect that the proposed methodology for functional data classification is generalizable to any perfusion study. Our data were obtained using a specific acquisition protocol based on relatively high temporal sampling initially (in the first 30s), and more sparsely sampled data in its second phase (out to 590s). While one would not expect more widely spaced temporal sampling in the second phase to have substantial impact on analyses, an ideal dataset might be the one that was acquired at a high temporal sampling in the second phase as well. However, obtaining such data would require high overall radiation exposures and present practical challenges in acquisition because of the constraints of breathing motion and registration.

Furthermore, we should note that our study does not characterize treatment-induced changes in the perfusion values. Using data from a larger randomized clinical trial, the clustering detection could be informed by treatment information, eg, by using an ANOVA-dependent Dirichlet process.[62] Finally, our modeling formulation is more computationally demanding than competing approaches. While our extensive simulation studies suggest that computations do not slow down substantially for datasets comprising a few hundred patients' samples, for larger datasets, computational speed could be improved through the use of variational Bayes approaches.[63] The precision of the classification, when considering the full acquisition duration, is dramatically decreased, as presented in Table 3. Scarpa and Dunson[64] have recently proposed enriched stick-breaking processes for functional data, which enable incorporation of prior information about attributes of the curves in the classification. A similar feature selection-based strategy could be used to guide the classification of the curves so that specific characteristics of each perfusion parameter are taken into account and/or relevant portions of the time courses can contribute to the inference more than others.

Additionally, while our framework considered ROI-level inference, which is perhaps most prevalent in clinical practice, the methodology could be adapted to accommodate voxelwise analysis following image registration. The methods, when implemented in this context, could be leveraged to create spatial–temporal posterior probability maps, providing diagnostic tools that could potentially be used to enhance existing tumor segmentation approaches.

Author Contributions

Conceived and designed the experiments: CSN. Analyzed the data: KMF. Wrote the first draft of the manuscript: MG. Contributed to the writing of the manuscript: KMF, MG, BH, CSN, MV. Agree with manuscript results and conclusions: KMF, MG, BH, CSN, MV. Jointly developed the structure and arguments for the paper: MG, BH, CSN, MV. Made critical revisions and approved final version: MG, BH, CSN, MV. All authors reviewed and approved of the final manuscript.

Supplementary Material

Supplementary File. MCMC details.

REFERENCES

1. Torigian DA, Huang SS, Houseni M, Alavi A. Functional imaging of cancer with emphasis on molecular techniques. *CA Cancer J Clin*. 2007;57:206–24.
2. Miles KA, Griffiths MR. Perfusion ct: a worthwhile enhancement? *Br J Radiol*. 2003;76:220–31.
3. Hansen ML, Norling R, Lauridsen C, et al. Computed tomography (ct) perfusion in abdominal cancer: technical aspects. *Diagnostics (Basel)*. 2013;3:261.
4. García-Figueiras R, Goh VJ, Padhani AR, et al. Ct perfusion in oncologic imaging: a useful tool? *AJR Am J Roentgenol*. 2013;200:8–19.
5. Tateishi U, Kusumoto M, Nishihara H, Nagashima K, Morikawa T, Moriyama N. Contrast-enhanced dynamic computed tomography for the evaluation of tumor angiogenesis in patients with lung carcinoma. *Cancer*. 2002;95:835–42.
6. Ng Q-S, Goh V, Fichte H, et al. Lung cancer perfusion at multi–detector row ct: reproducibility of whole tumor quantitative measurements. *Radiology*. 2006;239:547–53.
7. Ash L, Teknos TN, Gandhi D, Patel S, Mukherji SK. Head and neck squamous cell carcinoma: Ct perfusion can help noninvasively predict intratumoral microvessel density. *Radiology*. 2009;251:422–8.
8. Chen Y, Zhang J, Dai J, Feng X, Lu H, Zhou C. Angiogenesis of renal cell carcinoma: perfusion ct findings. *Abdom Imaging*. 2010;35:622–8.
9. Fidler IJ, Ellis LM. The implications of angiogenesis for the biology and therapy of cancer metastasis. *Cell*. 1994;79:185–8.
10. Ebos JML, Kerbel RS. Antiangiogenic therapy: impact on invasion, disease progression, and metastasis. *Nat Rev Clin Oncol*. 2011;8:316–316.
11. Moserle L, Casanovas O. Anti-angiogenesis and metastasis: a tumour and stromal cell alliance. *J Intern Med*. 2013;273:128–37.
12. Folkman J, Kalluri R. Tumor angiogenesis. *Cancer Medicine*. 6 ed. PC Decker; Hamilton (ON): 2003.
13. Wang Y, Hobbs BP, Ng CS. CT perfusion characteristics identify metastatic sites in liver. *Biomed Res Int*. 2015;2015:6.
14. Jain RK, Duda DG, Willett CG, et al. Biomarkers of response and resistance to antiangiogenic therapy. *Nat Rev Clin Oncol*. 2009;6:327–38.
15. Pope W. Assessing radiographic response in glioblastoma following avastin treatment. In: Chen TC, Chamberlain M, eds. *Controversies in Neuro-Oncology (Avastin and Malignant Gliomas)*. Bentham; 2010:59–73.
16. Moserle L, Jiménez-Valerio G, Casanovas O. Antiangiogenic therapies: going beyond their limits. *Cancer Discov*. 2014;4:31–41.
17. Axel L. Tissue mean transit time from dynamic computed tomography by a simple deconvolution technique. *Invest Radiol*. 1983;18:94–9.
18. Ng CS, Chandler AG, Wei W, et al. Reproducibility of ct perfusion parameters in liver tumors and normal liver. *Radiology*. 2011;260:762–70.
19. Luczyńska E, Aniol J. Neoangiogenesis in prostate cancer. *Contemp Oncol*. 2013;17:229–33.
20. Bisdas S, Rumboldt Z, Šurlan-Popovič K, et al. Perfusion CT in squamous cell carcinoma of the upper aerodigestive tract: long-term predictive value of baseline perfusion ct measurements. *AJNR Am J Neuroradiol*. 2010;31:576–81.
21. Ng CS, Hobbs BP, Chandler AG, et al. Metastases to the liver from neuroendocrine tumors: effect of duration of scan acquisition on CT perfusion values. *Radiology*. 2013;269:758–67.
22. Ng CS, Chandler AG, Wei W, et al. Effect of sampling frequency on perfusion values in perfusion ct of lung tumors. *AJR Am J Roentgenol*. 2013;200(2):W155–62.
23. Petralia G, Summers P, Viotti S, Montefrancesco R, Raimondi S, Bellomi M. Quantification of variability in breath-hold perfusion ct of hepatocellular carcinoma: a step toward clinical use. *Radiology*. 2012;265:448–56.
24. Ramsay J, Silverman B. *Functional Data Analysis*. 2nd ed. New York: Springer; 2005.

25. Rasmussen C, Williams C. *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, Massachusetts; 2006.
26. Ferguson TS. A Bayesian analysis of some nonparametric problems. *Ann Stat*. 1973;1:209–30.
27. Petrone S, Guindani M, Gelfand AE. Hybrid Dirichlet mixture models for functional data. *J R Stat Soc Series B Stat Methodol*. 2009;71:755–82.
28. Müller P, Mitra R. Bayesian nonparametric inference – why and how. *Bayesian Anal*. 2013;8:269–302.
29. Bigelow JL, Dunson DB. Bayesian adaptive regression splines for hierarchical data. *Biometrics*. 2007;63:724–32.
30. Oakley J, O'Hagan A. Bayesian inference for the uncertainty distribution of computer model outputs. *Biometrika*. 2002;89:769–84.
31. Zhu H, Vannucci M, Cox DD. A Bayesian hierarchical model for classification with selection of functional predictors. *Biometrics*. 2010;66:463–73.
32. Shi J, Wang B. Curve prediction and clustering with mixtures of Gaussian process functional regression models. *Stat Comput*. 2008;18:267–83.
33. Gelfand AE, Kottas A, MacEachern SN. Bayesian nonparametric spatial modeling with Dirichlet process mixing. *J Am Stat Assoc*. 2005;100:1021–35.
34. Duan JA, Guindani M, Gelfand AE. Generalized spatial Dirichlet process models. *Biometrika*. 2007;94:809–25.
35. Wu J, Wang X, Walker S. Bayesian nonparametric inference for a multivariate copula function. *Methodol Comput Appl Probab*. 2014;16:747–63.
36. Wade S, Dunson DB, Petrone S, Trippa L. Improving prediction from Dirichlet process mixtures via enrichment. *J Mach Learn Res*. 2015;15:1041–71.
37. Wade S, Walker SG, Petrone S. A predictive study of Dirichlet process mixture models for curve fitting. *Scand Stat Theory Appl*. 2014;41:580–605.
38. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc*. 1995;90:773–95.
39. Sethuraman J. A constructive definition of Dirichlet priors. *Stat Sin*. 1994;4:639–50.
40. Matern B. *Spatial Variation*. 2nd ed. Berlin: Springer Verlag; 1986.
41. Abramowitz M, Stegun I. *Handbook of Mathematical Functions*. New York: Dover; 1965.
42. Banerjee S, Carlin B, Gelfand A. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC, Boca Raton; 2004.
43. Eddy SR. What is a hidden Markov model? *Nat Biotech*. 2004;22:1315–6.
44. Scott SL. Bayesian methods for hidden Markov models: recursive computing in the 21st century. *J Am Stat Assoc*. 2002;97:337–51.
45. Lai T, Xing H. A simple Bayesian approach to multiple change points. *Stat Sin*. 2011;21:539–10.
46. Little R, Rubin D. *Statistical Analysis with Missing Data*. New York: J. Wiley and Sons; 1987.
47. Rubin D. Multiple imputation after 18+ years (with discussion). *J Am Stat Assoc*. 1996;91:473–89.
48. Daniels MJ, Hogan J. *Missing Data in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Boca Raton; 2008.
49. Escobar M. Estimating normal means with a Dirichlet process prior. *J Am Stat Assoc*. 1994;89:268–77.
50. Blackwell D, MacQueen J. Ferguson distributions via pólya urn schemes. *Ann Stat*. 1973;1:353–5.
51. Bush C, MacEachern S. A semiparametric Bayesian model for randomised block designs. *Biometrika*. 1996;83:275–85.
52. Hoff PD. A First Course in Bayesian Statistical Methods. Springer, New York; 2009.
53. Good I. Studies in the history of probability and statistics. XXXVII A. M. Turing's statistical work in World War II. *Biometrika*. 1979;66:393–6.
54. Osimani M, Bellini D, Di Cristofano C, et al. Perfusion MDCT of prostate cancer: correlation of perfusion CT parameters and immunohistochemical markers of angiogenesis. *AJR Am J Roentgenol*. 2012;199:1042–8.
55. Kim S, Tadesse MG, Vannucci M. Variable selection in clustering via Dirichlet process mixture models. *Biometrika*. 2006;93:877–93.
56. Dahl DB. Modal clustering in a class of product partition models. *Bayesian Anal*. 2009;4:243–64.
57. Sahu S, Gelfand A, Holland D. Spatio-temporal modeling of fine particulate matter. *J Agric Biol Environ Stat*. 2006;11:61–86.
58. Berrocal V, Gelfand A, Holland D. A spatio-temporal downscaler for output from numerical models. *J Agric Biol Environ Stat*. 2010;15:176–97.
59. Kambadakone A, Sharma A, Catalano O, Hahn P, Sahani D. Protocol modifications for CT perfusion (CTp) examinations of abdomen-pelvic tumors: impact on radiation dose and data processing time. *Eur Radiol*. 2011;21:1293–300.
60. Yeung TPC, Yartsev S, Bauman G, He W, Fainardi E, Lee T-Y. The effect of scan duration on the measurement of perfusion parameters in CT perfusion studies of brain tumors. *Acad Radiol*. 2013;20:59–65.
61. Core Team R. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2015.
62. Iorio MD, Müller P, Rosner GL, MacEachern SN. An ANOVA model for dependent random measures. *J Am Stat Assoc*. 2004;99:205–15.
63. Blei D, Jordan M. Variational inference for Dirichlet process mixtures. *Bayesian Anal*. 2006;1:121–43.
64. Scarpa, B. and Dunson, D.B. (2014). Enriched stick breaking processes for functional data. Journal of the American Statistical Association, 109, 647–660.