HHS Public Access

Author manuscript

Stat Med. Author manuscript; available in PMC 2019 September 30.

Published in final edited form as:

Stat Med. 2018 September 30; 37(22): 3214–3229. doi:10.1002/sim.7810.

Detection of Prostate Cancer with Multi-Parametric MRI Utilizing the Anatomic Structure of the Prostate

Jin Jin[1], **Lin Zhang**[1], **Ethan Leng**[2], **Gregory J. Metzger**[2], and **Joseph S. Koopmeiners**[1,*]

[1]Division of Biostatistics, University of Minnesota, Minneapolis, MN, 55455, U.S

[2]Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota, Minneapolis, MN, 55455, U.S

Summary

Multi-parametric magnetic resonance imaging (mpMRI), which combines traditional anatomic and newer quantitative MRI methods, has been shown to result in improved voxel-wise classification of prostate cancer as compared to any single MRI parameter. While these results are promising, substantial heterogeneity in the mpMRI parameter values and voxel-wise prostate cancer risk has been observed both between and within regions of the prostate. This suggests that classification of prostate cancer can potentially be improved by incorporating structural information into the classifier. In this paper, we propose a novel voxel-wise classifier of prostate cancer that accounts for the anatomic structure of the prostate by Bayesian hierarchical modeling, which can be combined with post-hoc spatial Gaussian kernel smoothing to account for residual spatial correlation. Our proposed classifier results in significantly improved area under the ROC curve (0.822 vs. 0.729, $p < 0.001$) and sensitivity corresponding to 90% specificity (0.599 vs. 0.429, $p < 0.001$), compared to a baseline model that does not account for the anatomic structure of the prostate. Furthermore, the classifier can also be applied on voxels with missing mpMRI parameters, resulting in similar performance, which is an important practical consideration that cannot be easily accommodated using regression-based classifiers. In addition, our classifier achieved high computational efficiency with a closed-form solution for the posterior predictive cancer probability.

Keywords

Bayesian classifier; Multi-parametric magnetic resonance imaging; Prostate cancer; Spatial classifier; Voxel-wise classification

1 | INTRODUCTION

Multi-parametric magnetic resonance imaging (mpMRI), a combination of traditional anatomic and newer quantitative MRI methods, is continuing to evolve as a valuable tool for the detection of prostate cancer.[1,2] Although the potential has been shown in the literature, its application in clinical practice is limited due to the substantial variability of the standard

*Correspondence: Joseph Koopmeiners, 420 Delaware St. SE, Minneapolis, MN55455. koopm007@umn.edu.

radiological assessment.[3] This variability persists even with the development of the Prostate Imaging Reporting and Data System (PI-RADS) standard,[4,5] as the standard is qualitative and does not provide sufficient guidance on the combination of the results comprising an mpMRI study. In contrast, quantitative predictive models provide user independent predictions combining all parameters, and thus addressing the limitations of direct radiologic interpretation. In the literature, various Computer Aided Detection (CAD) algorithms for the voxel-wise prostate cancer classification utilizing mpMRI have been developed and evaluated, including linear and nonlinear regressions, Clustering Methods, Support Vector Machine, Ensemble learning and Naive Bayes.[6, 7, 8, 9, 10, 11, 12,13] CAD algorithms utilizing textural feature models (radiomics-based approaches) have also been discussed.[14, 15,16]

Based on certain anatomic features, the prostate can be segmented into two main regions, the peripheral zone (PZ), which refers to the sub-capsular portion of the posterior aspect of the prostate gland that surrounds the distal urethra, and central gland (CG), which is comprised of multiple structures including the transition, central, anterior fibromuscular, and periurethral zones.[17,18] The whole gland (WG) refers to the combination of the PZ and CG.

Previously, Metzger et al.[19] developed an mpMRI model which generated a voxel-wise composite biomarker score (CBS) for the classification of prostate cancer. The CBS model linearly combined multiple mpMRI parameters and showed significantly improved prostate cancer classification compared to using any single mpMRI parameter. An important finding of their work is that, while combining multiple parameters improves classification, there is substantial heterogeneity in the mpMRI parameters both within and between the PZ and CG and, in addition, substantial variability exists in the cancer prevalence throughout the whole prostate gland with a much higher prevalence observed in the PZ. In this sense, the region (i.e. PZ or CG) meets the classic definition of a confounder, in that the region is associated both with the predictors (mpMRI parameters) and the outcome (voxel-wise cancer occurrence). This finding motivated the investigation of region-specific voxel-wise classifiers, which demonstrated a benefit over a single WG model, but requires manually guided region segmentation.

Previous authors have noted that the accuracy of voxel-wise classifiers for prostate cancer could be improved by developing region-specific CAD algorithms, as compared to a single, WG CAD algorithm that ignores differences between regions of the prostate.[13, 20,21] In region-specific CAD, classifiers are built separately for different regions of the prostate. However, this requires an additional step of region segmentation, which typically requires manual guidance from clinicians with expertise. Although techniques for automatic segmentation have been described,[22,23] few publications have discussed how to generate a fully automated classifier to model the anatomic structure of the prostate without the need for an additional manually guided, region segmentation step.

In this paper, we propose a fully automated, voxel-wise classifier, which aims to improve voxel-wise classification accuracy of prostate cancer to account for regional heterogeneity in the mpMRI parameters and cancer risk without the need for manual segmentation. We will present two approaches to accounting for the structure of the prostate in a voxel-wise classifier: a region-specific model, which models the mpMRI parameters and cancer risk as

a function of region and uses the 2-D coordinates to segment the prostate, and a model that ignores the regions and directly models the mpMRI parameters and voxel-wise cancer risk as a function of the 2-D coordinates. Simulation results and the application of our models to the data from Metzger et al.[19] indicate that both approaches improve the area under the ROC curve (AUC) and sensitivity corresponding to 90% specificity (S90) relative to a baseline model that does not account for the structure of the prostate. In addition, we applied a Spatial Gaussian Kernel smoothing technique to the posterior predictive cancer probabilities to account for potential residual spatial correlation, which also significantly improved AUC and S90. Our proposed classifier is available in a closed-form, thus avoiding the need for Markov chain Monte Carlo (MCMC), resulting in a computationally efficient classifier.

The remaining sections of the paper are organized as follows. In Section 2, we briefly describe the data available for developing our multi-parametric classifier. In Section 3, we present our analytic approach to developing a voxel-wise classifier that accounts for the anatomic structure of the prostate. In Section 4, we present a simulation study evaluating the impact of modeling regional heterogeneity on the performance of our classifier and, in Section 5, we apply our classifier to the data discussed in Section 2. We conclude with a discussion of the properties of our classifier and potential extensions in Section 6.

2 | DATAOVERVIEW ANDNOTATIONS

We begin by providing an overview of the data available for developing our classifier and introduce the notation that will be used throughout the paper. The collection procedure of the quantitative multi-parametric MRI and histology data used to train and test our classifier have been previously described.[19] Our data consists of 46 prostate slices obtained from 34 patients, with 2098 to 5756 voxels per slice. For simplicity, we randomly select one slice from patients with multiple slices to avoid having to model correlation between slices from the same subject. Figure 1 shows the image of an example slice of prostate segmented in the CG and PZ.

Let N be the total number of subjects, i.e. prostate slices, and n_i the number of voxels in slice i, with $i = 1, \ldots, N$. For voxel j in slice i, we observe an $m \times 1$ vector of mpMRI parameters, which we denote as $\boldsymbol{y_{ij}} = (y_{ij,1}, \ldots, y_{ij,m})^T$. MpMRI parameters available in our data set include calculated T2 values, the apparent diffusion coefficient (ADC), and dynamic contrast enhanced MRI (DCE-MRI) parameters, including, the area under the gadolinium concentration time curve at 90 seconds (AUGC90), the forward volume transfer constant (K^{trans}), reflux rate constant (k_{ep}) and fractional extravascular extracellular space (V_e). Each voxel also has a binary indicator of its cancer status derived from co-registered pathology data,[24] which we denote as c_{ij}, where $c_{ij} = 1$ indicates cancer and $c_{ij} = 0$ indicates noncancer. This ground truth will be used for the training and evaluation of the cancer classifiers. In addition, each voxel is identified by a set of location information $k_{ij} = \{r_{ij}, (a_{ij}, b_{ij})\}$, including a binary indicator of region, r_{ij}, with $r_{ij} = 1$ indicating PZ and $r_{ij} = 0$ indicating CG, and a standardized 2-D coordinate, (a_{ij}, b_{ij}), that has been re-scaled to have support $(-1, 1) \times (-1, 1)$ with (0, 0) being the center of each slice.

We denote the set of observed mpMRI parameters of all voxels as $\boldsymbol{Y} = \{\boldsymbol{y_{ij}} | i = 1, \ldots, N, j = 1, \ldots, n_i\}$, and their corresponding cancer status as $C = \{c_{ij} | i = 1, \ldots, N, j = 1, \ldots, n_i\}$. Similarly, we denote the generic parameter set that provides the location information for the voxels as $K = \{R, (A,B)\}$, where $R = \{r_{ij} | i = 1, \ldots, N, j = 1, \ldots, n_i\}$, and $(A,B) = \{(a_{ij}, b_{ij}) | i = 1, \ldots, N, j = 1, \ldots, n_i\}$ denote the corresponding region indicators and standardized 2-D coordinates, respectively. Finally, we let $\boldsymbol{Y}^*$, C^* and $K^* = \{R^*, (A^*,B^*)\}$ denote the mpMRI parameters, cancer status and location information (including region indicators and coordinates), respectively, for a new prostate slice for which voxel-wise classifications of cancer status are desired. For each slice, missing mpMRI parameter values for individual voxels can result from uncorrected motion between the separately acquired data sets used to calculate T2, T1 or pharmacokinetics parameters.

3 | METHODS

3.1 | Overview

In this section, we discuss the development of an mpMRI prostate cancer classifier that accounts for the anatomic structure of the prostate. We first introduce a general Bayesian modeling framework, which models the mpMRI parameters and cancer risk as a function of the location in the prostate. Cancer predictions for future observations can be obtained using Bayesian posterior predictive probabilities. Two models that incorporate location information will be discussed: “Mregion”, which models the mpMRI parameters and cancer risk by region of the prostate, and “Mcoord”, which models the mpMRI parameters and cancer risk by voxel coordinates. Both approaches can be combined with post-hoc spatial smoothing to account for residual spatial correlation and improve classification. A baseline model will also be presented, which ignores the anatomic structure of the prostate and will serve as a baseline for evaluating the gain in classification accuracy resulting from modeling the mpMRI parameters and cancer risk as a function of location.

3.2 | General Model Framework

Here we present our general framework for developing a voxel-wise classifier that accounts for heterogeneity in the mpMRI parameters and cancer risk as a function of location. We developed our voxel-wise classifiers by modifying the framework proposed by Wang et al. (2015), which generated region-wise classification for metastatic liver cancer.[25] Our basic approach is to jointly model the mpMRI parameters $\boldsymbol{Y}$ and the corresponding cancer status C conditional on the location information, and use Bayesian posterior predictive probabilities conditional on the observed mpMRI parameters to predict the unknown cancer status for voxels in a new prostate slice. This basic approach to Bayesian inference in the discriminant analysis setting has been previously pursued by a number of authors.[26, 27, 28, 29,30]

First, note that the joint model for $\boldsymbol{Y}$ and C conditional on the location information can be defined hierarchically as $f(\boldsymbol{Y}, C|K) = f(\boldsymbol{Y}|C,K)p(C|K)$. The two components of the likelihood are defined as follows. First, we assume that the $\boldsymbol{y_{ij}}$’s are independently distributed given the c_{ij}’s and k_{ij}’s:

$$\boldsymbol{y_{ij}} \mid c_{ij}, k_{ij} \overset{iid}{\sim} f(\boldsymbol{y_{ij}} \mid \boldsymbol{\theta}(c_{ij}, k_{ij})), \quad (1)$$

where $\boldsymbol{\theta}(c_{ij}, k_{ij})$ represents the set of model parameters, which are dependent on the cancer status c_{ij} and also location information k_{ij}. The cancer status indicators, c_{ij}'s, are assumed to be independent Bernoulli random variables with the voxel-wise cancer probability $\rho(k_{ij})$, also a function of k_{ij}. That is: $p(c_{ij}|k_{ij}) = \rho(k_{ij})^{c_{ij}}(1 - \rho(k_{ij}))^{1-c_{ij}}$. The joint likelihood is therefore:

$$f(\boldsymbol{Y}, C \mid K, \boldsymbol{\theta}(C, K), \rho(K)) = \prod_{i=1}^{N} \prod_{j=1}^{n_i} \left\{ f(\boldsymbol{y_{ij}} \mid \boldsymbol{\theta}(c_{ij}, k_{ij})) p(c_{ij} \mid \rho(k_{ij})) \right\}. \quad (2)$$

Estimation is completed under the Bayesian paradigm. We specify a prior distribution for $\boldsymbol{\theta}(C,K)$, while $\rho(K)$ is estimated by an Empirical Bayesian (EB) approach. The cancer status for voxels in a new slice can be predicted by Bayesian predictive probabilities conditional on the observed mpMRI parameters. The voxel-wise posterior predictive cancer risk given $\boldsymbol{y}^*$ can be expressed as:

$$p(c^* = 1 \mid \boldsymbol{y}^*, k^*, \boldsymbol{Y}, C, K) = \frac{f(\boldsymbol{y}^*, c^* = 1 \mid k^*, \boldsymbol{Y}, C, K)}{\sum_{c=0}^{1} f(\boldsymbol{y}^*, c^* = c \mid k^*, \boldsymbol{Y}, C, K)}, \quad (3)$$

where $f(\boldsymbol{y}^*, c^*|k^*,\boldsymbol{Y}, \boldsymbol{C},K)$ is the joint posterior predictive distribution for $(\boldsymbol{y}^*, c^*)$ conditional on the training data and the location information for the new voxel, k^*, given by:

$$f(\boldsymbol{y}^*, c^* \mid k^*, \boldsymbol{Y}, C, K) = \int f(\boldsymbol{y}^* \mid \boldsymbol{\theta}(c^*, k^*)) f(\boldsymbol{\theta}(c^*, k^*) \mid \boldsymbol{Y}, C, K) d\boldsymbol{\theta}(c^*, k^*) \cdot p(c^* \mid k^*), \quad (4)$$

and $f(\boldsymbol{\theta}(c^*, k^*)|\mathbf{Y}, C,K)$ is the posterior distribution for $\boldsymbol{\theta}(c^*, k^*)$:

$$f(\boldsymbol{\theta}(c^*, k^*) \mid \boldsymbol{Y}, C, K) = \frac{f(\boldsymbol{Y} \mid \boldsymbol{\theta}(c^*, k^*), C, K) \pi(\boldsymbol{\theta}(c^*, k^*))}{f(\boldsymbol{Y} \mid C, K)}. \quad (5)$$

Here $\pi(\boldsymbol{\theta}(c^*, k^*)|C,K)$ denotes the prior distribution for $\boldsymbol{\theta}(c^*, k^*)$, and the denominator in (5), $f(\boldsymbol{Y}|C,K) = \int f(\boldsymbol{Y}|\boldsymbol{\theta}(c^*, k^*), C,K)\pi(\boldsymbol{\theta}(c^*, k^*))d\boldsymbol{\theta}(c^*, k^*)$, is the likelihood averaged over the prior for $\boldsymbol{\theta}(c^*, k^*)$.[31]

3.3 | Regional Model (Mregion)

We propose our first model, "Mregion", which models the mpMRI parameters and cancer risk as a function of the binary indicator for region, R. In this case, the location information K will refer to the regional information, R, and the two components of the joint likelihood

are region-specific: $f(\mathbf{Y}, C|R) = f(\mathbf{Y}|C,R)p(C|R)$. In addition, we model the probability that a voxel is in either the PZ or CG as a function of its 2-D coordinates to avoid manually guided segmentation of the prostate. In this case, the full hierarchical model can be expressed as $f(\mathbf{Y}, C|R,A,B) = f(\mathbf{Y}|C,R)p(C|R)p(R|A,B)$.

We first discuss the modeling framework assuming that R is known, and then the prediction of R using the 2-D coordinates later in this subsection. With R known, for $f(\mathbf{Y}|C,R)$, we assume that the $\mathbf{y_{ij}}$'s are independently distributed given the c_{ij}'s and follow a multivariate normal distribution:

$$\mathbf{y_{ij}} \mid c_{ij}, r_{ij} \overset{iid}{\sim} \mathcal{MVN}(\boldsymbol{\mu}_{\mathbf{c_{ij}, r_{ij}}}, \boldsymbol{\Sigma}_{\mathbf{c_{ij}, r_{ij}}}), \quad (6)$$

where the mean and covariance parameters $\boldsymbol{\mu}_{\mathbf{c,r}}$ and $\boldsymbol{\Sigma}_{\mathbf{c,r}}$ are functions of the cancer status, c, and the region, r. The set of model parameters in the multivariate normal distribution is then $\boldsymbol{\theta}_{\mathbf{c,r}} = \{\boldsymbol{\mu}_{\mathbf{c,r}}, \boldsymbol{\Sigma}_{\mathbf{c,r}} | c, r \in \{0, 1\}\}$. Similarly, for the specification of $p(C|R)$, the cancer status indicators, c_{ij}'s, are assumed to be independent Bernoulli random variables with the voxel-wise probability of cancer, ρ_r, which is region-specific.

For the model parameters $\boldsymbol{\theta}_{\mathbf{c,r}}$ in (6), we specify the prior distribution, $\pi(\boldsymbol{\theta}_{\mathbf{c,r}})$, as follows. We assume a priori independence of cancer status and region, and, conditional on the cancer status c and region indicator r, we assign conjugate priors for the mean $\boldsymbol{\mu}_{\mathbf{c,r}}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{c,r}}$. In particular, for each combination of c and r, we put a flat prior on the mean $\boldsymbol{\mu}_{\mathbf{c,r}}$, and an independent inverse Wishart prior, $\mathcal{W}^{-1}(\boldsymbol{\Sigma}_{\mathbf{c,r}}|\delta,\boldsymbol{\Omega}_{\mathbf{c,r}})$, on the covariance $\boldsymbol{\Sigma}_{\mathbf{c,r}}$. The hyperparameters, δ and $\boldsymbol{\Omega}_{\mathbf{c,r}}$'s, are specified by setting $\delta = m-1$ to provide maximum entropy, while $\boldsymbol{\Omega}_{\mathbf{c,r}}$ is determined by an EB approach, where $\boldsymbol{\Omega}_{\mathbf{c,r}}$ is set equal to the maximizer of the marginal likelihood, $f(\mathbf{Y}|C,R,\boldsymbol{\Omega}_{\mathbf{c,r}}, \delta = m - 1)$. Rather than placing a prior on ρ_r, we instead estimate ρ_r using an EB approach, as well, in which case $\hat{\rho}_r$ equals the sample mean of the c_{ij}'s averaging over all prostate slices and voxels within region r. This allows us to maintain a closed-form solution for the posterior, avoiding MCMC and decreasing the computational burden of our method.

Combining the likelihood and prior results in the following posterior distribution for $\boldsymbol{\theta}_{\mathbf{c,r}}$:

$$\boldsymbol{\mu}_{\mathbf{c,r}}, \boldsymbol{\Sigma}_{\mathbf{c,r}} \mid \mathbf{Y}^{\mathbf{c,r}} \sim \mathcal{MVN}(\boldsymbol{\mu}_{\mathbf{c,r}} \mid \tilde{\mathbf{y}}^{\mathbf{c,r}}, \frac{\boldsymbol{\Sigma}_{\mathbf{c,r}}}{\sum_{i=1}^{N} n_i^{c,r}}) \times \mathcal{W}^{-1}(\boldsymbol{\Sigma}_{\mathbf{c,r}} \mid \sum_{i=1}^{N} n_i^{c,r} + \delta, \tilde{\mathbf{S}}^{\mathbf{c,r}} + \hat{\boldsymbol{\Omega}}_{\mathbf{c,r}}),$$

(7)

where $c, z \in \{0, 1\}$, $\mathbf{Y}^{\mathbf{c,r}}$ denotes the set of mpMRI parameters with cancer status c in region z, $n_i^{c,r}$ denotes the number of voxels with cancer status c in region z of slice i, and the two additional terms, $\tilde{\mathbf{y}}^{\mathbf{c,r}}$ and $\tilde{\mathbf{S}}^{\mathbf{c,r}}$, are defined as:

$$\tilde{\boldsymbol{y}}^{\boldsymbol{c},\boldsymbol{r}} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{n_i^{c,r}} \boldsymbol{y}_{\boldsymbol{ij}}^{\boldsymbol{c},\boldsymbol{r}}}{\sum_{i=1}^{N} n_i^{c,r}}, \quad (8)$$

$$\tilde{\boldsymbol{S}}^{\boldsymbol{c},\boldsymbol{r}} = \sum_{i=1}^{N} \boldsymbol{Y}_{\boldsymbol{i}}^{\boldsymbol{c},\boldsymbol{r}} \left(\boldsymbol{Y}_{\boldsymbol{i}}^{\boldsymbol{c},\boldsymbol{r}}\right)^{T} - \sum_{i=1}^{N}\sum_{j=1}^{n_i^{c,r}} \bar{\boldsymbol{y}}^{\boldsymbol{c},\boldsymbol{r}} \bar{\boldsymbol{y}}^{\boldsymbol{c},\boldsymbol{r}T}. \quad (9)$$

The posterior predictive density of the mpMRI parameters $\boldsymbol{y}^*$ for a new voxel given cancer status c^* and region r^* can then be written as:

$$f(\boldsymbol{y}^* \mid c^*, r^*, \boldsymbol{Y}, C, R) = (2\pi)^{-\frac{m}{2}} \left(\frac{\sum_{i=1}^{N} n_i^{c^*,r^*} + 1}{\sum_{i=1}^{N} n_i^{c^*,r^*}}\right)^{-\frac{m}{2}} \quad (10)$$

$$\times \frac{h\left(\sum_{i=1}^{N} n_i^{c^*,r^*} + \delta, \tilde{\boldsymbol{S}}^{\boldsymbol{c}^*,\boldsymbol{r}^*} + \widehat{\boldsymbol{\Omega}}_{\boldsymbol{c}^*,\boldsymbol{r}^*}\right)}{h\left(\sum_{i=1}^{N} n_i^{c^*,r^*} + \delta + 1, \tilde{\boldsymbol{S}}_{*}^{\boldsymbol{c}^*,\boldsymbol{r}^*} + \widehat{\boldsymbol{\Omega}}_{\boldsymbol{c}^*,\boldsymbol{r}^*}\right)},$$

where $\tilde{\boldsymbol{S}}_{*}^{\boldsymbol{c}^*,\boldsymbol{r}^*}$ is defined as $\tilde{\boldsymbol{S}}^{\boldsymbol{c},\boldsymbol{r}}$ in (9) with the new voxel included, and $h(\delta,\boldsymbol{\Omega})$ is the normalizing constant in the probability density function of $\mathscr{W}^{-1}(\boldsymbol{\Sigma}|\delta,\boldsymbol{\Omega})$. Combining the estimated cancer risk and the posterior predictive density of the mpMRI parameters results in the predictive distribution of c^* conditional on $\boldsymbol{y}^*$, r^*, $\boldsymbol{Y}$, C and R:

$$p(c^* = 1 \mid \boldsymbol{y}^*, r^*, \boldsymbol{Y}, C, R) = \frac{p(c^* = 1 \mid r^*) f(\boldsymbol{y}^* \mid c^* = 1, r^*, \boldsymbol{Y}, C, R)}{\sum_{c=0}^{1} p(c^* = c \mid r^*) f(\boldsymbol{y}^* \mid c^* = c, r^*, \boldsymbol{Y}, C, R)}, \quad (11)$$

where $p(c^* = 1|r^*) = \hat{\rho}_{r^*}$ is the estimated cancer risk for region r^*.

The posterior predictive probability in (11) assumes that the region R is known. For our data, the region annotation was conducted on the anatomic T2-weighted images using a segmentation software (Segasist, Ontario, Canada) requiring manual guidance. Our objective is to create a fully automated classifier, in which case this step must be avoided. Given the consistency of the prostate anatomy across patients, a rough but automatic estimation of a voxel's regional location can be obtained by using standardized coordinates to estimate the probability that a voxel is in the PZ or CG. This prediction can be performed in multiple ways. Here we adopt an adaptive polynomial spline regression with a logistic likelihood.[32] We estimate the vector of coefficients in the spline $\boldsymbol{\beta}_{\text{coord}}$ by iterated weighted least squares assuming the following regression model for $p(r = 1|a, b)$:

$$\text{logit}(p(r = 1 \mid a, b)) = \boldsymbol{x}_{\text{coord}}^{T}(a, b)\boldsymbol{\beta}_{\text{coord}}, \quad (12)$$

where $\boldsymbol{x}_{\text{coord}}(a, b)$ is the $d\times 1$ vector of the basis functions for the coordinates (a, b) determined by a model selection procedure (the details will be discussed in Section 4). This provides the probability that a voxel is in the PZ or CG and is fully automated. Combining (12) with (11) results in the posterior predictive probability of cancer for a new voxel conditional on $\boldsymbol{y}^*$ and (a^*, b^*):

$$p(c^* = 1 \mid \boldsymbol{y}^*, a^*, b^*, \boldsymbol{Y}, C, R) = \sum_{r=0}^{1} p(c^* = 1 \mid \boldsymbol{y}^*, r^* = r, \boldsymbol{Y}, C, R)p(r^* = r \mid a^*, b^*), \quad (13)$$

where $p(r^* = 1 \mid a^*, b^*) = e^{\boldsymbol{x}_{\text{coord}}^{T}(a^*, b^*)\hat{\boldsymbol{\beta}}_{\text{coord}}}$, and $p(c^* = 1|\boldsymbol{y}^*, r^* = r,\boldsymbol{Y}, C,R)$ is defined as in (11). The resulting predictive probability is the weighted average of the posterior predictive cancer risk for PZ and CG.

3.4 | Coordinate Model (Mcoord)

The Mregion model assumes that the distribution of the mpMRI parameters and the voxel-wise cancer risk are heterogeneous across regions but constant within a region. However, this model requires that region be predicted by coordinates to avoid manually guided segmentation. An alternate approach is to ignore the region information and directly model the distribution of the mpMRI parameters and the voxel-wise cancer risk as a function of the 2-D coordinates. We refer to this model as the "Mcoord" model. Here the location information K is summarized by the standardized 2-D coordinates, and the joint likelihood becomes $f(\boldsymbol{Y}, C|A,B) = f(\boldsymbol{Y}|C,A,B)p(C|A,B)$.

We again assume that the $\boldsymbol{y}_{ij}$'s are independently distributed random variables following a multivariate normal distribution with mean and covariance parameters varying by cancer status. As with Mregion, we also allow the mean parameter to vary as a function of the location in the prostate, but, unlike Mregion, the mean in Mcoord is allowed to vary gradually as a function of location. A multivariate adaptive polynomial spline regression[33,32] is adopted to describe the coordinate based mean:

$$\begin{aligned} &\boldsymbol{y}_{ij} \mid c_{ij}, a_{ij}, b_{ij} \overset{iid}{\sim} \mathcal{MVN}(\boldsymbol{\mu}_{\boldsymbol{c_{ij}}}(a_{ij}, b_{ij}), \boldsymbol{\Sigma}_{\boldsymbol{c_{ij}}}), \quad (14)\\ &\boldsymbol{\mu}_{\boldsymbol{c_{ij}}}^{T}(a_{ij}, b_{ij}) = \boldsymbol{x}_{\boldsymbol{c_{ij}}}^{\boldsymbol{T}}(a_{ij}, b_{ij}), \boldsymbol{\beta}_{\boldsymbol{c_{ij}}}, \end{aligned}$$

where $\boldsymbol{x}_{\boldsymbol{c_{ij}}}(a_{ij}, b_{ij})$ denotes the $d_{c_{ij}}\times 1$ vector of basis functions for the coordinate (a_{ij}, b_{ij}),$\boldsymbol{\beta}_{\boldsymbol{c_{ij}}}$ denotes the $d_{c_{ij}} \times m$ matrix of the corresponding coefficients, with $d_{c_{ij}}$ the number of basis functions used for cancer status c_{ij}. Notice that, under this setting, the means of the multivariate normal distributions are no longer constant, and the set of model parameters

become $\boldsymbol{\theta_c} = (\boldsymbol{\beta_c}, \boldsymbol{\Sigma_c} | c \in \{0, 1\})$. The cancer status indicators, c_{ij}'s, are still assumed to be independent Bernoulli random variables. However, we now allow the cancer risk $\rho(a_{ij}, b_{ij})$ to vary as a function of the coordinates (a_{ij}, b_{ij}).

We assume a priori independence of $\boldsymbol{\theta_c}$ given the value of c, and specify its prior distribution as follows. We place a flat prior on the spline regression coefficient $\boldsymbol{\beta_c}$ and an inverse Wishart prior on $\boldsymbol{\Sigma_c}$. As in Section 3.3, the hyperparameter δ is set to $m - 1$ to provide maximum entropy, and $\boldsymbol{\Omega_c}$ is determined by an EB approach where it is set equal to the maximizer of the marginal likelihood $f(\boldsymbol{Y}|C, A, B, \boldsymbol{\Omega_c}, \delta = m - 1)$. As with the Mregion model, we estimate the voxel-wise cancer risk using an EB approach to avoid having to implement MCMC but, in this case, we use an adaptive polynomial spline, $\text{logit}(\rho(a_{ij}, b_{ij})) = \boldsymbol{x_p}^T(a_{ij}, b_{ij})\boldsymbol{\beta_p}$, where $\boldsymbol{x_p}(a_{ij}, b_{ij})$ is the $d_p \times 1$ vector of basis functions of the coordinate (a_{ij}, b_{ij}), and estimate the vector of regression parameters, $\boldsymbol{\beta_p}$, using iterated weighted least squares.

Combining the likelihood and prior results in the following updated posterior for $\boldsymbol{\theta_c}$:

$$\boldsymbol{\beta_c}^T, \boldsymbol{\Sigma_c} \mid \boldsymbol{Y^c} \sim \mathcal{MN}_{m, d_c}(\boldsymbol{\beta_c}^T \mid \widetilde{\boldsymbol{\beta}}_{\boldsymbol{c}}^T, \boldsymbol{\Sigma_c}, \sum_{i=1}^{N} \sum_{j=1}^{n_i^c} \boldsymbol{x_{ij}^c} \boldsymbol{x_{ij}^c}^T) \times \mathcal{W}^{-1}(\boldsymbol{\Sigma_c} \mid \sum_{i=1}^{N} n_i^c + \delta, \widetilde{\boldsymbol{S}}^{\boldsymbol{c}} + \widetilde{\boldsymbol{\Omega}}_{\boldsymbol{c}}), \quad (15)$$

where $\boldsymbol{Y^c}$ denotes the set of mpMRI parmeters with cancer status c, $\mathcal{MN}_{m,d_c}$ represents an $m \times d_c$ matrix normal random variable, and n_i^c the number of voxels with cancer status c in slice i. Here $\boldsymbol{x_{ij}^c}$ denotes the $\boldsymbol{x_c}(a_{ij}, b_{ij})$ in (14), which is the vector of basis functions in the spline regression. The two additional terms, $\widetilde{\boldsymbol{\beta}}_{\boldsymbol{c}}$ and $\widetilde{\boldsymbol{S}}_{\boldsymbol{c}}$, are defined as:

$$\widetilde{\boldsymbol{\beta}}_{\boldsymbol{c}} = \left[\sum_{i=1}^{N} \sum_{j=1}^{n_i^c} \boldsymbol{x_{ij}^c} \boldsymbol{x_{ij}^c}^T \right]^{-1} \left[\sum_{i=1}^{N} \sum_{j=1}^{n_i^c} \boldsymbol{x_{ij}^c} \boldsymbol{y_{ij}^c}^T \right], \quad (16)$$

$$\widetilde{\boldsymbol{S}}^{\boldsymbol{c}} = \sum_{i=1}^{N} \sum_{j=1}^{n_i^c} \boldsymbol{x_{ij}^c} \boldsymbol{x_{ij}^c}^T - \sum_{i=1}^{N} \sum_{j=1}^{n_i^c} \boldsymbol{y_{ij}^c} \boldsymbol{x_{ij}^c}^T \widetilde{\boldsymbol{\beta}}_{\boldsymbol{c}}. \quad (17)$$

The posterior predictive density of the mpMRI parameters for a new voxel $\boldsymbol{y}^*$ given its cancer status c^* and coordinate (a^*, b^*) can then be written as:

$$f(\boldsymbol{y}^* \mid c^*, a^*, b^*, \boldsymbol{Y}, C, A, B) = (2\pi)^{-\frac{m}{2}} \left| \left[\sum_{i=1}^{N} \sum_{j=1}^{n_i^{c^*}} \boldsymbol{x}_{ij}^{\boldsymbol{c}^*} \boldsymbol{x}_{ij}^{\boldsymbol{c}^{*T}} \right]^{-1} \boldsymbol{x}_*^{\boldsymbol{c}^*} \boldsymbol{x}_*^{\boldsymbol{c}^{*T}} + \boldsymbol{I} \right|^{-\frac{m}{2}} \times \frac{h(\sum_{i=1}^{N} n_i^{c^*} + \delta, \widetilde{\boldsymbol{S}}^{\boldsymbol{c}^*} + \widehat{\boldsymbol{\Omega}}_{\boldsymbol{c}^*})}{h(\sum_{i=1}^{N} n_i^{c^*} + \delta + 1, \widetilde{\boldsymbol{S}}_*^{\boldsymbol{c}^*} + \widehat{\boldsymbol{\Omega}}_{\boldsymbol{c}^*})}, \tag{18}$$

where $\boldsymbol{x}_*^{\boldsymbol{c}^*}$ denotes the vector of basis functions for the new voxel given its cancer status c^*, and $\widetilde{\boldsymbol{S}}_*^{\boldsymbol{c}^*}$ is defined as in (17) with the new voxel included. Combining the estimated cancer risk and the posterior predictive density of the mpMRI parameters, the posterior predictive cancer risk for a new voxel can be expressed as:

$$p(c^* = 1 \mid \boldsymbol{y}^*, a^*, b^*, \boldsymbol{Y}, C, A, B) = \frac{p(c^* = 1 \mid a^*, b^*) f(\boldsymbol{y}^* \mid c^* = 1, a^*, b^*, \boldsymbol{Y}, C, A, B)}{\sum_{c=0}^{1} p(c^* = c \mid a^*, b^*) f(\boldsymbol{y}^* \mid c^* = c, a^*, b^*, \boldsymbol{Y}, C, A, B)}, \tag{19}$$

where $p(c^* = 1 \mid a^*, b^*) = e^{\boldsymbol{x}_{\boldsymbol{p}}^T(a^*, b^*)\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}}$ is the estimated cancer risk at (a^*, b^*).

To evaluate the improvement of the Mregion and Mcoord models, we also considered an Mbase model which ignores the location information, K. Details of Mbase can be found in Section 1 of the Supplemental Material, but the resulting posterior distribution, posterior predictive probability, and voxel-wise cancer classifier are equivalent to (7), (10) and (11), respectively, except that the parameters in Mbase do not depend on the region information R. Comparing the classification accuracy of Mregion and Mcoord to Mbase allows us to evaluate the benefit of modeling the anatomic structure of the prostate.

Table 1 summarizes the data components required to train and to obtain predictions for new voxels using each model. For the Mbase model, which ignores the structure of the prostate, voxel location information is not required to train the model or obtain predictions for new voxels. In contrast, the Mregion and Mcoord models both require that standardized coordinates are available to train the classifier and obtain predictions for future voxels. The primary distinction is that the Mregion model requires that the prostate be manually segmented to train the model, while the Mcoord models the parameters and voxel-wise cancer risk as a function of the coordinates, directly, and does not. Finally, we note that

segmentation is not required to obtain predictions using either model, which was one of our primary objectives.

3.5 | Predictions for Voxels with Missing mpMRI Parameters

A common problem in mpMRI data is that some mpMRI parameters are missing due to failed MRI parametric mapping, most frequently a result of uncorrected motion between or within datasets. It is a practical problem that complicates the application of regression-based classifiers. Implementing regression-based classifiers in the presence of missing data requires that the missing mpMRI parameters be imputed. In the Bayesian case, however, we simply integrate out the missing parameters from the joint predictive density of the mpMRI parameters and cancer status:

$$f(\boldsymbol{y}^*_{\boldsymbol{obs}}, c^* \mid k^*, \boldsymbol{Y}, C, K) = \int f(\boldsymbol{y}^*_{\boldsymbol{obs}}, \boldsymbol{y}^*_{\boldsymbol{mis}}, c^* \mid k^*, \boldsymbol{Y}, C, K) d\boldsymbol{y}^*_{\boldsymbol{mis}}, \quad (20)$$

where $\boldsymbol{y}^* = (\boldsymbol{y}^*_{\boldsymbol{obs}}, \boldsymbol{y}^*_{\boldsymbol{mis}})$, and $\boldsymbol{y}^*_{\boldsymbol{obs}}$ and $\boldsymbol{y}^*_{\boldsymbol{mis}}$ denote the observed and missing mpMRI parameters, respectively. In this case, voxel-wise predictions are obtained by replacing $f(\boldsymbol{y}^* | c^*, k^*, \boldsymbol{Y}, C, K)$ with $f(\boldsymbol{y}^*_{\boldsymbol{obs}} \mid c^*, k^*, \boldsymbol{Y}, C, K)$ in (11) and (19). In other words, if a voxel is missing one or more mpMRI parameters, the voxel's cancer status can still be predicted using only the observed parameters. If a voxel is missing all mpMRI parameters, then its cancer risk is estimated as a function of location in the prostate. This approach assumes missing completely at random (MCAR),[34] which is a strong and likely untrue assumption, but we will show that this approach still results in reasonable predictions in our data.

3.6 | Spatial Gaussian Kernel Smoother and Msmooth

Throughout, we have assumed independent joint distributions for the mpMRI parameters and cancer status across voxels. However, there is potential residual spatial correlation in the data, and accounting for spatial correlation has the potential to improve our predictions. Ideally, we would formally model the spatial correlation between voxels. However, formally modeling the spatial correlation in our setting is not trivial due to the size of our data.[35,36] Instead, we combine our classifier with a spatial Gaussian-Kernal smoothing technique to maintain computational efficiency. Here we adopt the Nadaraya-Watson estimator with Gaussian kernel.[37,38] In this case, the smoothed voxel-wise predictive probability of cancer at coordinate (a_{*j}, b_{*j}) can be expressed as:

$$\widetilde{\widehat{p}}(a_{*j}, b_{*j}) = \frac{\sum_{k=1, k\neq j}^{n_*} K_\sigma(\|(a_{*j}, b_{*j}) - (a_{*k}, b_{*k})\|_2) \hat{p}(a_{*k}, b_{*k})}{\sum_{k=1, k\neq j}^{n_*} K_\sigma(\|(a_{*j}, b_{*j}) - (a_{*k}, b_{*k})\|_2)}, \quad (21)$$

where n_* denotes the number of voxels in the new prostate slice, $\|\cdot\|_2$ denotes L_2 norm, K_σ denotes the gaussian kernel with bandwidth σ, and $\hat{p}(a_{*k}, b_{*k})$ is the unsmoothed voxel-wise predictive probability of cancer for voxel k. The bandwidth σ controls the contributions of

the neighboring voxels determined by their distance from (a_{*j}, b_{*j}), and is chosen by cross-validation to maximize the average AUC. Notice that the predictive probabilities for the voxels on the edges of a slide need an edge correction. This is realized by replacing the denominator in (21) by the inverse of an edge correction term $e(a_{*j}, b_{*j})$, which is the reciprocal of the kernel mass within the new prostate slice:

$$e(a_{*j}, b_{*j}) = \left[\int_{v \in G} k(\|v - (a_{*j}, b_{*j})\|_2) dv \right]^{-1}, \quad (22)$$

where G denotes the area within the new prostate slice. This spatial smoothing step was implemented using the "spatstat" package in R.[39]

4 | SIMULATION STUDIES

4.1 | Simulation Settings

We conducted a simulation study to evaluate the performance of our proposed classifiers under different scenarios. For each realization of our simulation study, we simulated 34 training slices and 10 test slices. For each slice, a mask (i.e shape of the prostate slice, including voxel-wise region indicators and 2-D spatial coordinates) was sampled with replacement from the 46 slices in the real data set. Within a slice, voxel-wise mpMRI parameters and binary cancer status indicators were simulated from the following model:

$$\begin{aligned} &\boldsymbol{y_{ij}} \mid c_{ij}, r_{ij} \sim \mathcal{MVN}(\boldsymbol{y_{ij}} \mid \boldsymbol{\mu}_{c_{ij}, r_{ij}}, \boldsymbol{\Gamma}_{c_{ij}, r_{ij}}), \quad (23) \\ &c_{ij} = I(p_{ij} > 0.5) = I(q_{ij} > 0), \\ &q_{ij} = \Phi^{-1}(p_{ij}) = q_{0, r_{ij}} + w_{ij}, \\ &\boldsymbol{w_i} \sim \mathcal{MVN}(\mathbf{0}, \boldsymbol{C_i}(\phi, \sigma^2)). \end{aligned}$$

For the j-th voxel in the i-th slice, cancer probability p_{ij} was simulated by adding a spatial random effect, w_{ij}, on the probit (which we denote as $\Phi^{-1}(\cdot)$) of region-specific cancer probability $q_{0,rij}$. To allow the voxel-wise cancer probabilities p_{ij}'s, $j = 1, \ldots, n_i$ to be spatially correlated, we assume that $\boldsymbol{w_i} = (w_{i1}, w_{i2}, \ldots, w_{i,ni})^T$ follows a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{C_i}(\phi, \sigma^2)$, i.e. we allow w_{ij}'s to be spatially correlated, and the (j, k)-th entry of $\boldsymbol{C_i}(\phi, \sigma^2)$ (i.e. the spatial correlation between w_{ij} and w_{ik}) is $\boldsymbol{C_i}(\phi, \sigma^2)_{j,k} = \sigma^2 e^{-\phi\sqrt{(a_{ij} - a_{ik})^2 + (b_{ij} - b_{jk})^2}}$, where σ^2 is the spatial variance and ϕ is the range of spatial correlation (large ϕ indicates stronger spatial correlation). The binary cancer status, c_{ij}, was set to be equal to 1 if p_{ij} was greater than 0.5 and 0 if p_{ij} was less than or equal to 0.5. Finally, we assumed that the distribution of voxel-wise mpMRI parameters $\boldsymbol{y_{ij}}$ varies between the PZ and CG, and between cancer and noncancer voxels, and simulated $\boldsymbol{y_{ij}}$ conditional on the region indicator r_{ij} and cancer indicator c_{ij}.

To understand the impact of the magnitude of regional heterogeneity and spatial pattern on the classification accuracy of the proposed models, we conducted simulations under different settings of the magnitude of the regional heterogeneity, the spatial correlation and the spatial variance. To vary the regional heterogeneity in the mpMRI parameters, we assumed that the difference in the mean of mpMRI parameters between the PZ and CG was either large (same magnitude as $(\boldsymbol{\mu}_C - \boldsymbol{\mu}_{NC})$, i.e. the mean difference between cancer and noncancer voxels), small (same magnitude as $\frac{1}{2}(\boldsymbol{\mu}_C - \boldsymbol{\mu}_{NC})$, i.e. half of the mean difference between cancer and noncancer voxels), or none. For the regional heterogeneity in the voxel-wise cancer probabilities, we varied the difference in cancer probability between PZ and CG: we assumed that voxel-wise cancer probabilities in the CG were $p_{CG} = 0.092$, which is the sample mean from real data, and set the difference in cancer probability between PZ and CG to be 0.184, (large) 0.092 (small) or 0 (none). In addition, we also varied the spatial variance (large: 0.8, small: 0.4 or little: 0.2) and the spatial correlation (strong: $\phi = 5$, weaker: $\phi = 8$ or little: $\phi = 12$), which induced different levels of spatial correlation strength among the simulated data.

We completed 1000 simulations under each setting of parameters. For each simulation, we trained the classifiers using 34 training prostate slices, and evaluated the performance using 10 test slices. We summarized the classification accuracy of the various classifiers by averaging the AUC across the 1000 simulated data sets. For each simulated data set, the spatial smoothing kernel bandwidth was treated as a tuning parameter and selected by Leave-one-out Cross Validation (LOOCV) to maximize the average AUC. The basis functions for a spline regression were determined by stepwise model selection to minimize the penalized residual sum of squares using the polymars algorithm.[33]

4.2 | Simulation Results

Figure 2 presents the AUC for the Mbase, Mregion, Mcoord and Msmooth models as a function of the magnitude of regional heterogeneity under different patterns of spatial variance and correlation (σ^2 and ϕ). In the figure, the error bars represent the 95% confidence intervals for the AUC, and the dots in the center of the bars are point estimates of AUC. We first look at the impact of the magnitude of regional heterogeneity on the performance of Mregion and Mcoord models. From each sub figure, we see that Mregion and Mcoord have similar performance to the Mbase model when there is no regional heterogeneity (i.e. when voxel-wise cancer probabilities and distribution of the mpMRI parameters are the same in the PZ and CG), but that the Mregion and Mcoord models have higher AUC than the Mbase in the presence of regional heterogeneity, which increases as the difference between the regions increases. This agrees with our expectation that Mregion and Mcoord can improve the classification accuracy of prostate cancer by accounting for the regional heterogeneity in the mpMRI parameters and voxel-wise cancer risk. Next, we consider the impact of spatial dependence on the relative performance of Mcoord and Msmooth, which captures the improvement in AUC that results from the spatial Gaussian kernel smoothing step. In the figures on the top left, where spatial correlation is stronger ($\phi = 5$) and spatial variance is larger ($\sigma^2 = 0.8$), the improvement in AUC from Mcoord to Msmooth is large but decreases in the presence of weaker spatial correlation and smaller spatial variance. When there is little spatial correlation ($\phi = 12$), we can see that spatial

Gaussian kernel smoothing has little improvement in AUC but does not over-smooth to the point of reducing the AUC relative to the Mbase model. Detailed simulation results can be found in Table 1 of the Supplemental Material.

In addition, we also evaluated the spatial Gaussian kernel smoothing step under scenarios with more complex patterns of spatial dependence. Specifically, we considered two scenarios: (1) the data sets were simulated with ϕ, the range parameter of spatial correlation, varying by prostate slice, so that the spatial correlation pattern varies across prostate slices; and (2) there is non-stationarity in the spatial correlation within each slice. To simulate data under this second scenario, we varied the range of spatial correlation by the location in a prostate slice. A detailed description of the simulation settings and results can be found in Section 2 of the Supplemental Material. In summary, similar performance was observed in these scenarios as was described above, except the improvement in classification accuracy by the spatial smoothing step is lower than when data were generated assuming a constant, stationary, spatial correlation structure for all slices.

5 | APPLICATION

5.1 | Methods

We will evaluate the classification accuracy of the proposed models using the data described in Section 2 to illustrate the benefit of accounting for the anatomic structure of the prostate in our classifier. Metzger et al.[19] developed two mpMRI classifiers: a whole gland (WG) classifier, which included ADC, K^{trans}, k_{ep} and AUGC90, and a PZ-specific classifier, which included T2 in place of K^{trans}. We evaluated different sets of parameters with the models proposed in Section 3 and found that the combination of ADC, K^{trans}, k_{ep} andAUGC90 resulted in the optimal classification accuracy and, therefore, were used in all subsequent modeling. Among the four mpMRI parameters, ADC values are approximately normally distributed, while K^{trans}, k_{ep} and AUGC90 values were skewed and were log transformed before analysis to achieve an approximately normal distribution.

We will evaluate the classification accuracy of the various classifiers by the ROC curve and its summaries, including the AUC and S90. Summaries of the ROC curve were estimated using LOOCV to account for overfitting due to building and evaluating the model on the same data set.[40,13] In our case, LOOCV was completed on the subject/slice level, rather than the voxel level due to correlation between voxels from the same subject. Estimates of the AUC and S90 were obtained for each subject by applying a classifier trained using data from all other subjects. Global estimates of AUC and S90 were obtained by averaging the subject-specific estimates of AUC and S90. For each model, the number of basis functions in the adaptive polynomial spline regression and the spatial smoothing kernel bandwidth were treated as potential tuning parameters and selected by LOOCV to maximize the average AUC. The basis functions for spline regression were determined by stepwise model selection to minimize the penalized residual sum of squares using the polymars algorithm.[33] 95% Confidence intervals and p-values for all model comparisons were calculated using a bootstrapping procedure that resampled cases, rather than voxels. Voxels were not resampled within a slice in order to retain the spatial structure of the data. Confidence intervals were

calculated using the percentile method. All results in Section 4.2 were obtained using 1000 resampled data sets.

5.2 | Results

Summaries of the classification accuracy of the Mbase, Mregion, Mcoord and Msmooth models are presented in Table 2 and Figure 3. Table 2 presents the estimated AUC and S90 for the complete voxels (i.e. voxels with no missing mpMRI parameters) and for all voxels (i.e. including voxels with missing mpMRI parameters). Figure 3 presents the ROC curves for the complete voxels and for all voxels.

We first consider the results for the complete voxels and will then consider the performance of our classifier for voxels with missing data. Results from the real data application support our simulation results. Both Mregion and Mcoord significantly improved the AUC (p = 0.013 and p = 0.011) and S90 (p = 0.015 and p = 0.012) relative to Mbase. The average AUC and S90 for Mregion were 0.038 and 0.046 higher than Mbase, respectively. We note that Mregion uses an automated procedure to segment the prostate into the PZ and CG. The average AUC by LOOCV for the automated segmentation in Mregion was 0.972, which performed similar to Mregion applied assuming known, manually guided segmentation (AUC: 0.767 vs. 0.763; S90: 0.475 vs. 0.476). Mcoord achieved similar performance, resulting in an AUC and S90 that were 0.043 and 0.065 higher than Mbase, respectively. Compared to Mregion, Mcoord achieved higher AUC (0.772 vs. 0.767) and S90 (0.494 vs. 0.475) but the differences were not significant (AUC, p = 0.413; S90, p = 0.320).

The additional step of spatial smoothing demonstrated significant improvement in classification accuracy, as well. Here, Msmooth is obtained by combining Mcoord, which performed better than Mcoord, with the spatial smoother discussed in Section 3.6. The smoothing step resulted in significantly higher AUC and S90 compared to Mcoord (p = 0.001 and p < 0.001). The AUC and S90 for MSmooth were 0.050 and 0.105 higher than Mcoord, respectively. This gain in AUC and S90 due to post-hoc spatial Gaussian kernel smoothing is similar in magnitude to the improvement observed due to modeling the anatomic structure of the prostate (i.e. comparing both Mregion and Mcoord to Mbase). Overall, after modeling the anatomic structure of the prostate and conducting spatial smoothing on the predicted cancer risk, the AUC and S90 for Msmooth were 0.093 (12.8%) and 0.170 (39.6%) higher than Mbase, respectively (p < 0.001 and p = 0.001).

From Table 2, we can compare the performance of our classifiers when applied to all voxels, including voxels with missing parameters, to the complete case results. 8.553% of voxels have at least one missing mpMRI parameter, most of which have one (3.915%) or three (4.347%) parameters missing. Even with this substantial amount of missing data, the estimated AUC and S90 when the classifier was applied to all voxels were similar to the results observed for the complete case analysis (AUC, 0.818 vs. 0.822, p = 0.326; S90, 0.596 vs. 0.599, p = 0.393).

Figure 4 presents heatmaps of the voxel-wise posterior predictive cancer probabilities for three randomly selected prostate slices in our data. In general, we observe improved classification moving from left to right. The Mbase predicted cancer probabilities are diffuse

and sometimes fail to cover the tumor, while the Mregion and Mcoord have identified areas with a high probability of cancer, and the overestimated cancer probabilities for the remaining areas have been reduced. In addition, while the Mregion and Mcoord predicted cancer probabilities are close, the Mcoord probability maps have more areas of deep red, indicating a high probability of cancer, which overlap with the cancer regions in the ground truth. Our final model, Msmooth, has further improved the classification by reducing the sharp difference among neighboring voxels and clearly identifying the regions of cancer, as compared to diffuse voxels with high probability of cancer.

Figure 5 presents a binary representation of classification accuracy and categorizes voxels as a true positive, false positive, false negative and true negative using the probability cut-off corresponding to 90% specificity. The area of true positives (i.e. the red regions) clearly increases in size moving from left to right, with a corresponding decrease in the false negatives (i.e. the blue regions). We also observe a decrease in false positives from left to right (i.e. the yellow regions) and the random false-positive chunks are eliminated in the figures for Msmooth. However, the maps do highlight the drawback of spatial smoothing, as it has the potential to over smooth some areas in a prostate slice and create continuous false positive areas adjacent to the predicted tumors. This is likely due to the use of a constant spatial smoothing bandwidth across all prostate slices or local non-stationarity in the underlying spatial process in some images. However, our simulation results presented in Section 4.2 illustrate that, under both scenarios, the average voxel-wise classification accuracy is improved by the spatial smoothing step.

For all models, we implement an EB approach to obtain a closed form for the posterior cancer probabilities, which potentially comes at the price of reduced classification accuracy relative to a fully Bayesian model that better captures the uncertainty of all model parameters. To ensure that there is minimal loss in the classification accuracy of prostate cancer, we also compared our proposed models, Mregion and Mcoord, with their fully Bayesian versions. Specifically, for the fully Bayesian version of Mregion, we assume a Unif(0, 1) prior on the prior cancer risk $p(c^*|r^*)$ in (11), and a flat prior on $\boldsymbol{\beta}_{\text{coord}}$ in (12). In addition, instead of estimating $\boldsymbol{\Omega}_{\boldsymbol{c,r}}$ in the inverse Wishart prior for $\boldsymbol{\Sigma}_{\boldsymbol{c,r}}$ in (6) by an EB approach, we instead considered the commonly used specification of setting$\boldsymbol{\Omega}_{\boldsymbol{c,r}} = \boldsymbol{I}_{\boldsymbol{m}}$, where $\boldsymbol{I}_{\boldsymbol{m}}$ is a $m \times m$ identity matrix. For the fully Bayesian modeling of Mcoord, we assume a flat prior on $\boldsymbol{\beta}_{\boldsymbol{p}}$ in Section 3.4, the model coefficient of the logistic spline regression that assumes that the prior cancer risk varies by 2-D coordinate. Similar to Mregion, we also assume an inverse Wishart prior for$\boldsymbol{\Sigma}_{\boldsymbol{c}}$, with $\delta = m - 1$ and$\boldsymbol{\Omega}_{\boldsymbol{c}}$ equal to $\boldsymbol{I}_{\boldsymbol{m}}$.

We conducted the fully Bayesian approaches in JAGS using the R package “Rjags” with 5000 MCMC iterations after 2000 iterations for initial burn-in, and estimated the AUC and S90 by LOOCV. For the Mregion model, the mean AUC is 0.766 for both empirical Bayesian model and fully Bayesian model, while the mean S90 is 0.475 and 0.474, respectively, for empirical Bayesian model and fully Bayesian model. Similarly, for the Mcoord model, the mean AUC by LOOCV is 0.770 and 0.772, and the mean S90 is 0.494 and 0.497, respectively, for empirical Bayesian model and fully Bayesian model. Our results suggest that, at least for the data set considered, there is little cost in classification accuracy associated with our EB approach relative to the fully Bayesian approach.

6 | DISCUSSION

We proposed a Bayesian mpMRI classifier that improves the voxel-wise classification of prostate cancer by accounting for the anatomic structure of the prostate. There is substantial heterogeneity between regions of the prostate in both the mpMRI parameters and the voxel-wise cancer risk, and exploiting this information improves the classification of prostate cancer. Additionally, we account for the residual spatial correlation in the voxel-wise predictive cancer probabilities by a spatial Gaussian kernel smoothing technique. Our simulation results indicate that our proposed classifiers result in improved classification accuracy relative to a baseline model that does not account for the structure of the prostate. Similarly, in an application of our proposed models to the data described in Section 2, our final model, Msmooth, resulted in a 0.093 (12.8%) improvement in the AUC and a 0.170 (39.6%) improvement in the S90 compared to the baseline model.

An important advantage of our proposed classifier is its computational efficiency. The classifiers were developed in the Bayesian paradigm, which typically requires MCMC methods to obtain an approximation to the posterior distribution. We are able to avoid MCMC by estimating nuisance parameters using an EB approach that allows us to derive closed-form solutions for the voxel-wise cancer predictions. This results in a substantial reduction in the computational burden of our classifier compared to a fully Bayesian analysis, but with minimal loss of classification accuracy. For our data, it took approximately 4.76 seconds on average for Msmooth, our final model, to conduct voxel-wise cancer classification for a new prostate slice with 3467 voxels using R parallel computing with 15 cores. For comparison, a fully Bayesian version implemented in JAGS with 5000 draws was 138 times slower, with no improvement in the classification accuracy.

An additional advantage of our classifier is that cancer classifications for voxels with missing mpMRI parameters are available without the need for an imputation step. Applying the classifier to voxels with missing mpMRI resulted in only a slight decrease in AUC and S90 relative to the complete case analysis. Missing parameters were observed for 8.553% of voxels in our data, which highlights the practical importance of obtaining predictions for voxels with missing parameters. However, direct application of our classifier to voxels with missing parameters relies on an assumption of MCAR, which is unlikely to be true but resulted in reasonable classification accuracy in our data. Modeling the missing data mechanism can potentially result in higher classification accuracy by allowing us to leverage information found in the missing data process, therefore will be one of our future considerations.

Our final model, which includes ADC, K^{trans}, k_{ep} and AUGC90, was selected by comparing the AUC for models with different combinations of the mpMRI parameters and identifying the combination that maximized the average AUC by LOOCV. As a result, the mpMRI parameter T2, which is commonly used for prostate cancer detection, was not included in the model. However, several combinations that include T2 demonstrated similar performance and can not be ruled-out as significantly worse than our final model. Further work is needed to identify the optimal approach to model selection in our current framework. Alternately, Bayesian model averaging could be used in the place of model

selection, with the model weights a function of 2-D location, as the optimal combination of parameters could potentially depend on the location within the prostate.

Our proposed classifier utilizes a hierarchical modeling framework that allows us to incorporate further structural complexity in terms of the mpMRI parameters and voxel-wise cancer risk. We exploited this flexibility to account for the anatomic structure of the prostate in the current models, but this framework could be extended to account for other features of the data, as well. For example, we observed strong, subject-specific shifts from the overall mean for the mpMRI parameters. In principle, this could be accommodated in our modeling framework by including random effects into the likelihood but additional work is needed to understand the practical implications of this approach. In this paper, we focused our model development on 2-D slices due to technical limitations that currently prevent 3-D co-registration of the MRI and pathology data sets. However, our method could be easily extended to 3-D by replacing the 2-D splines with3D splines,[41, 42,43] and the computation time should scale linearly by the number of voxels in the image.

Currently, our classifier utilizes an ad hoc Spatial Gaussian Kernal Smoothing step to account for the spatial dependence, which was shown in Section 5.2 to result in substantial improvement in the classification accuracy. Accounting for spatial dependence in the data can also be achieved by assigning spatially correlated priors on $p(c^*|a^*, b^*)$ in the Bayesian hierarchical modeling framework, which induces spatial correlation through a pre-specified covariance function. However, this results in two problems. First, the spatial model fitting will involve inversion of 34 covariance matrices (one for each slice), each of which is of high dimensionality ranging from 2098 by 2098 to 5756 by 5756, thus making the model computationally infeasible for our data. Second, specifying spatial structure on the voxel-wise cancer probabilities leads to intractable posterior distributions, which would require MCMC sampling to approximate the posterior distributions and result in a loss of computational efficiency, which is a strength of our proposed method. However, we do think that more sophisticated modeling of the complex spatial structure via non-stationary spatial processes has the potential to improve our classifier but we leave this task for future research.

7 | SOFTWARE

The R code for implementing our simulations is available for download at https://github.com/Jin93/R-code-for-voxel-wise-classifiers-of-prostate-cancer-with-mpMRI-2018.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NCI R01 CA155268, NCI P30 CA077598, NIBIB P41 EB015894 and the Assistant Secretary of Defense for Health affairs, through the Prostate Cancer Research Program under Award No. W81XWH-15-1-0478. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense.

References

1. John KurhanewiczDaniel VigneronPeter CarrollFergus Coakley. Multiparametric magnetic resonance imaging in prostate cancer: present and future. Current opinion in urology. 2008; 18(1): 71. [PubMed: 18090494]
2. Louise DickinsonAhmed Hashim U, Clare Allen, et al. Magnetic resonance imaging for the detection, localisation, and characterisation of prostate cancer: recommendations from a European consensus meeting. European urology. 2011; 59(4):477–494. [PubMed: 21195536]
3. Kirema Garcia-ReyesPassoni Niccolò M, Palmeri Mark L. , et al. Detection of prostate cancer with multiparametric MRI (mpMRI): effect of dedicated reader education on accuracy and confidence of index and anterior cancer diagnosis. Abdominal imaging. 2015; 40(1):134–142. [PubMed: 25034558]
4. Rosenkrantz Andrew B, Sooah Kim, Lim Ruth P, et al. Prostate cancer localization using multiparametric MR imaging: comparison of Prostate Imaging Reporting and Data System (PI-RADS) and Likert scales. Radiology. 2013; 269(2):482–492. [PubMed: 23788719]
5. Barentsz Jelle O, Jonathan Richenberg, Richard Clements, et al. ESUR prostate MR guidelines 2012. European radiology. 2012; 22(4):746–757. [PubMed: 22322308]
6. Michael Kelm B, Menze Bjoern H, Zechmann Christian M, Baudendistel Klaus T, Hamprecht Fred A. Automated estimation of tumor probability in prostate magnetic resonance spectroscopic imaging: pattern recognition vs quantification. Magnetic resonance in medicine. 2007; 57(1):150–159. [PubMed: 17191229]
7. Yahui PengYulei JiangTatjana AnticGiger Maryellen L, Scott EggenerAytekin Oto. A study of T2-weighted MR image texture features and diffusion-weighted MR image features for computer-aided diagnosis of prostate cancer. 2013:86701H.
8. Viswanath Satish E, Bloch Nicholas B, Chappelow Jonathan C, et al. Central gland and peripheral zone prostate tumors have significantly different quantitative imaging signatures on 3 tesla endorectal, in vivo T2-weighted MR imagery. Journal of Magnetic Resonance Imaging. 2012; 36(1):213–224. [PubMed: 22337003]
9. Simone MazzettiDe Luca MassimoChristian Bracco, et al. A CAD system based on multi-parametric analysis for cancer prostate detection on DCE-MRI. 2011:79633Q–79633Q.
10. Lukasz MatulewiczJansen Jacobus FA, Louisa Bokacheva, et al. Anatomic segmentation improves prostate cancer detection with artificial neural networks analysis of 1H magnetic resonance spectroscopic imaging. Journal of Magnetic Resonance Imaging. 2014; 40(6):1414–1421. [PubMed: 24243554]
11. Geert LitjensOscar DebatsJelle BarentszNico KarssemeijerHenkjan Huisman. Computer-aided detection of prostate cancer in MRI. IEEE transactions on medical imaging. 2014; 33(5):1083–1092. [PubMed: 24770913]
12. Pallavi TiwariAnant MadabhushiMark Rosen. A hierarchical unsupervised spectral clustering scheme for detection of prostate cancer from magnetic resonance spectroscopy (MRS). Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007. 2007:278–286.
13. Guillaume LemaîtreRobert MartíJordi FreixenetVilanova Joan C, Walker Paul M, Fabrice Meriaudeau. Computer-Aided detection and diagnosis for prostate cancer based on mono and multi-parametric MRI: a review. Computers in biology and medicine. 2015; 60:8–31. [PubMed: 25747341]
14. Andrew CameronFarzad KhalvatiHaider Masoom A, Alexander Wong. MAPS: a quantitative radiomics approach for prostate cancer detection. IEEE Transactions on Biomedical Engineering. 2016; 63(6):1145–1156. [PubMed: 26441442]
15. Farzad KhalvatiAlexander WongHaider Masoom A. Automated prostate cancer detection via comprehensive multi-parametric magnetic resonance imaging texture feature models. BMC medical imaging. 2015; 15(1):27. [PubMed: 26242589]
16. Duc FehrHarini VeeraraghavanAndreas Wibmer, et al. Automatic classification of prostate cancer Gleason Scores from multiparametric magnetic resonance images. Proceedings of the National Academy of Sciences. 2015; 112(46):E6265–E6273.

17. Wasserman Neil F, Benjamin Spilseth, Jafar Golzarian, Metzger Gregory J. Use of MRI for lobar classification of benign prostatic hyperplasia: potential phenotypic biomarkers for research on treatment strategies. American Journal of Roentgenology. 2015; 205(3):564–571. [PubMed: 26295642]
18. Myers Robert P. Structure of the adult prostate from a clinician's standpoint. Clinical anatomy. 2000; 13(3):214–215. [PubMed: 10797630]
19. Metzger Gregory J, Chaitanya Kalavagunta, Benjamin Spilseth, et al. Detection of Prostate Cancer: Quantitative Multiparametric MR Imaging Models Developed Using Registered Correlative Histopathology. Radiology. 2016; 279(3):805–816. [PubMed: 26761720]
20. Litjens GJS, Barentsz JO, Karssemeijer N, Huisman HJ. Automated computer-aided detection of prostate cancer in MR images: from a whole-organ to a zone-based approach. 2012:83150G.
21. Computer-aided detection of prostate cancer in MRI, Litjens GeertDebats OscarBarentsz JelleKarssemeijer NicoHuisman Henkjan. IEEE transactions on medical imaging. 2014; 33(5): 1083–1092. [PubMed: 24770913]
22. Xin LiuSamil Yetik Imam. Automated prostate cancer localization without the need for peripheral zone extraction using multiparametric MRI. Medical physics. 2011; 38(6):2986–2994. [PubMed: 21815372]
23. Robert TothJustin RibaultJohn GentileDan SperlingAnant Madabhushi. Simultaneous segmentation of prostatic zones using active appearance models with multiple coupled levelsets. Computer Vision and Image Understanding. 2013; 117(9):1051–1060. [PubMed: 23997571]
24. Chaitanya KalavaguntaXiangmin ZhouSchmechel Stephen C, Metzger Gregory J. Registration of in vivo prostate MRI and pseudo-whole mount histology using Local Affine Transformations guided by Internal Structures (LATIS). Journal of Magnetic Resonance Imaging. 2015; 41(4): 1104–1114. [PubMed: 24700476]
25. Yuan WangHobbs Brian P, Jianhua HuNg Chaan S, Kim-Anh Do. Predictive classification of correlated targets with application to detection of metastatic cancer using functional CT imaging. Biometrics. 2015; 71(3):792–802. [PubMed: 25851056]
26. Geoffrey McLachlan. Discriminant analysis and statistical pattern recognition. John Wiley & Sons; 2004.
27. Chris FraleyRaftery Adrian E. Model-based clustering, discriminant analysis, and density estimation. Journal of the American statistical Association. 2002; 97(458):611–631.
28. Hamsici Onur C, Martinez Aleix M. Bayes optimality in linear discriminant analysis. IEEE transactions on pattern analysis and machine intelligence. 2008; 30(4):647–657. [PubMed: 18276970]
29. Epileptic seizure detection using lacunarity and Bayesian linear discriminant analysis in intracranial EEG, Zhou WeidongLiu YinxiaYuan QiLi Xueli. IEEE Transactions on Biomedical Engineering. 60(12):3375–3381.2013; IEEE [PubMed: 23629837]
30. Bayes optimal template matching for spike sorting–combining fisher discriminant analysis with optimal filtering, Franke FelixQuiroga Rodrigo QuianHierlemann AndreasObermayer Klaus. Journal of computational neuroscience. 2015; 38(3):439–459. [PubMed: 25652689]
31. Carlin Bradley P, Louis Thomas A. Bayesian methods for data analysis. CRC Press; 2008.
32. De Boor CarlEtats-Unis Mathématicien. A practical guide to splines. Springer-Verlag; New York: 1978.
33. Charles KooperbergSmarajit BoseStone Charles J. Polychotomous regression. Journal of the American Statistical Association. 1997; 92(437):117–127.
34. Little Roderick JA, Rubin Donald B. Statistical Analysis with Missing Data. 2. 2002. 200–220.
35. HÅÅVARD RUE, Hååkon Tjelmeland. Fitting Gaussian Markov random fields to Gaussian fields. Scandinavian journal of Statistics. 2002; 29(1):31–49.
36. Stein Michael L, Zhiyi Chi, Welty Leah J. Approximating likelihoods for large spatial data sets. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2004; 66(2):275–296.
37. Nadaraya Elizbar A. On estimating regression. Theory of Probability & Its Applications. 1964; 9(1):141–142.
38. Watson Geoffrey S. Smooth regression analysis. Sankhyā: The Indian Journal of Statistics, Series A. 1964:359–372.

39. Adrian BaddeleyRolf Turner, et al. Spatstat: an R package for analyzing spatial point patterns. Journal of statistical software. 2005; 12(6):1–42.

40. Smarajit BoseHastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference and Prediction. 2003

41. Ruppert David W, Matt P, Carroll Raymond J. Semiparametric regression during 2003–2007. Electronic journal of statistics. 2009; 3:1193. [PubMed: 20305800]

42. Luciano SpinelloOliver Arras KaiRudolph TriebelRoland Siegwart. A Layered Approach to People Detection in 3D Range Data. 2010:1–1.

43. Junghoon LeeJonghye WooFangxu XingMurano Emi Z, Maureen StonePrince Jerry L. Semi-automatic segmentation for 3D motion analysis of the tongue with dynamic MRI. Computerized Medical Imaging and Graphics. 2014; 38(8):714–724. [PubMed: 25155697]

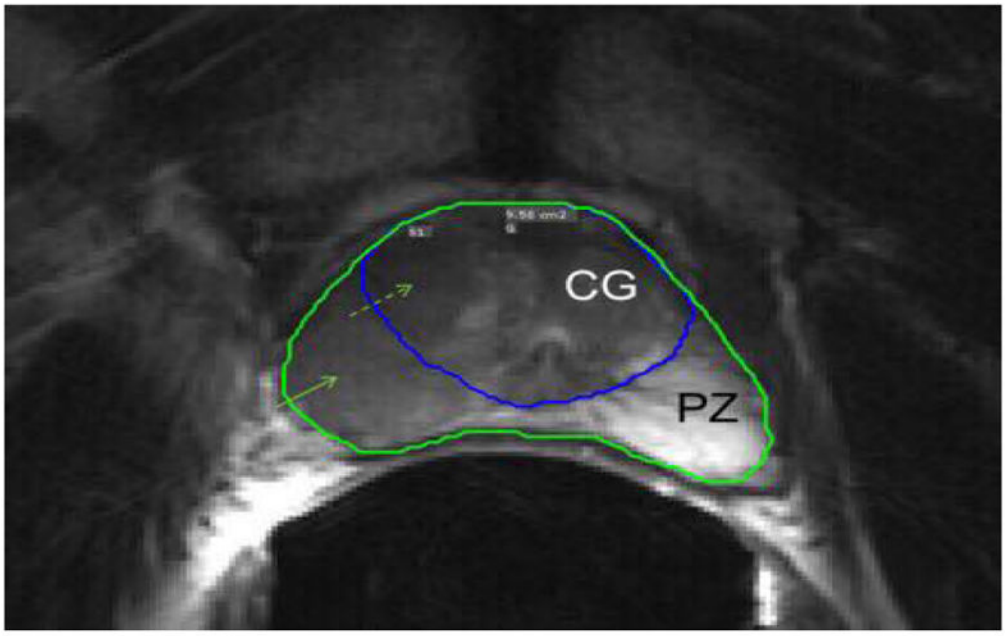

FIGURE 1.

Manually guided segmentation of the T2-weighted anatomic images showing the prostate capsule (green curve) with the blue curve demarcating the division between the peripheral zone (PZ) and central gland (CG). Histopathologically identified cancer is indicated in the PZ and CG by the solid and broken green arrows, respectively.

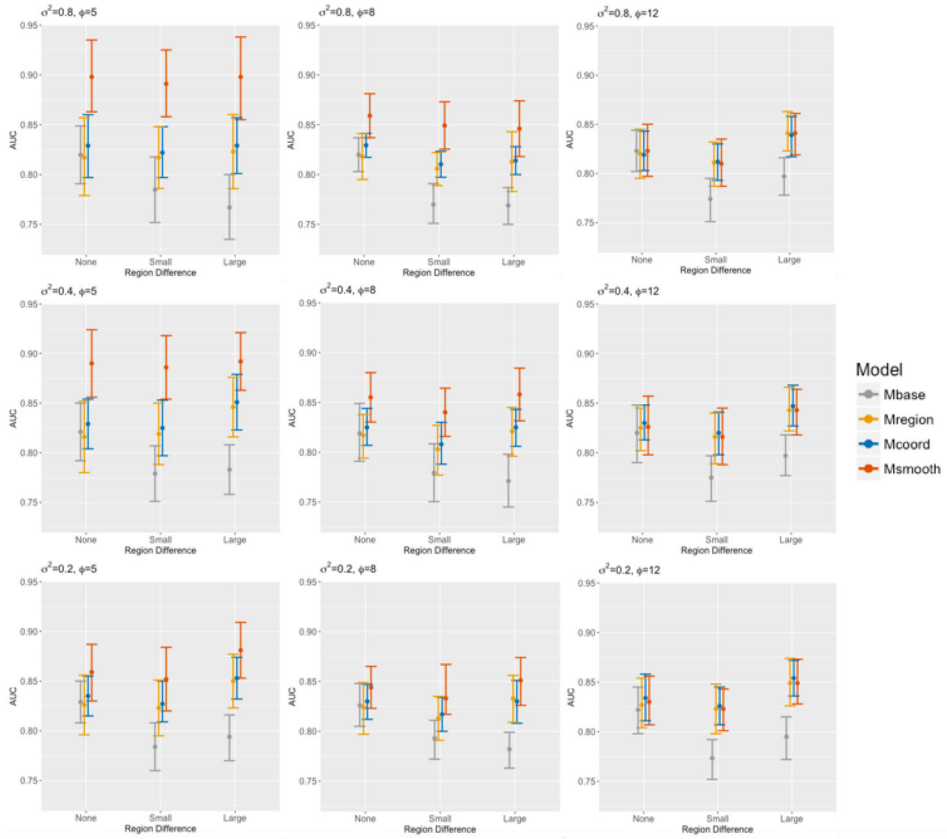

FIGURE 2.

Comparison of AUC between Mbase, Mregion, Mcoord and Msmooth versus different magnitudes of regional heterogeneity under different settings of σ^2 (spatial variance) and ϕ (range of spatial correlation).

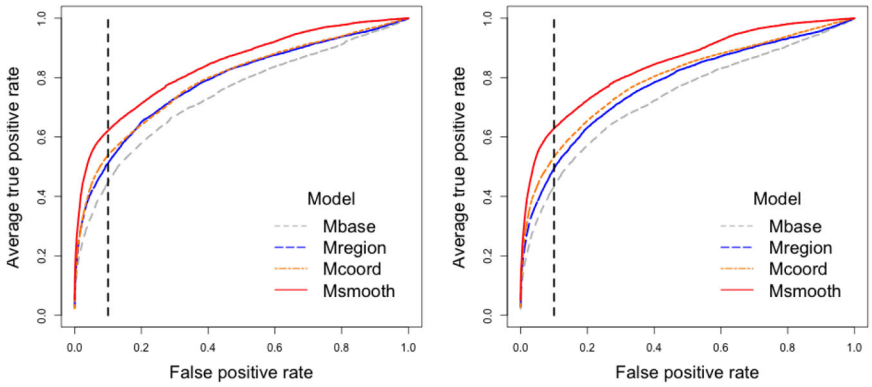

FIGURE 3.

Average ROC curves by LOOCV evaluating the classification accuracy of the proposed models using complete voxels (“complete case” data set, on the left) and all voxels (“all case” data set, on the right). The dashed lines overlaying the ROC curves indicate the S90 values.

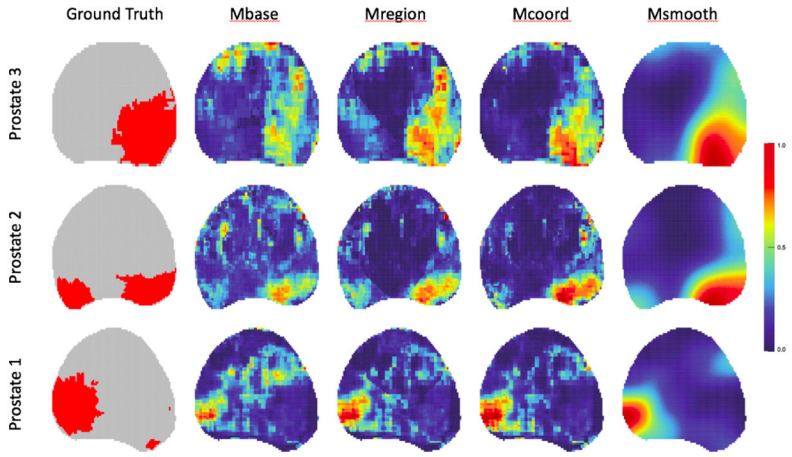

FIGURE 4.

Maps of three representative prostate slices: ground truth (column 1, red regions represent the registered areas of cancer), heatmaps of predicted cancer probabilities for Mbase, Mregion, Mcoord and Msmooth, respectively (column 2–5, where warmer color represents higher predicted cancer probability, and the color was scaled by the range of the predicted cancer probabilities, separately for each prostate slice).

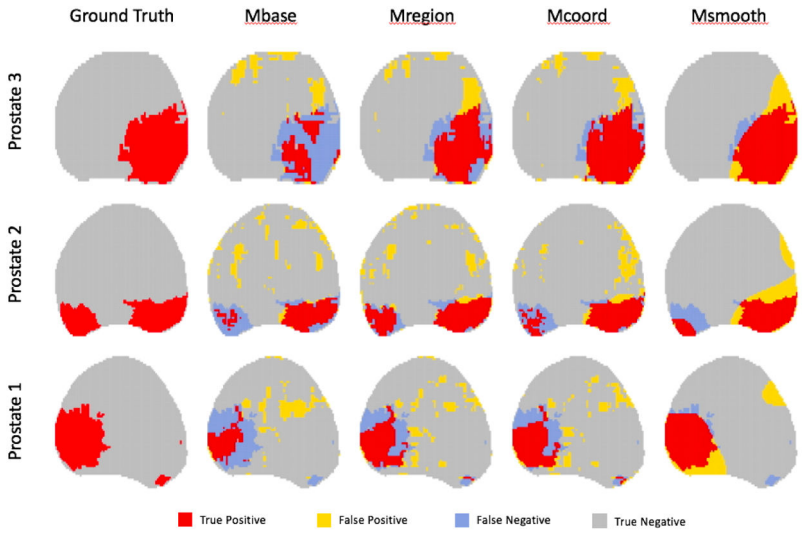

FIGURE 5.

Maps of three representative prostate slices: ground truth (column 1), maps of binary cancer classification using the S90 threshold for Mbase, Mregion, Mcoord and Msmooth, respectively (column 2–5).

TABLE 1

Comparison of the data components that are required for model training (column 2–5) and for prospective classification of cancer status for voxels in a new prostate slice (column 6–8) between Mbase, Mregion and Mcoord.

	Model Training				Prospective Classification of Cancer		
	Y	C	R	(A,B)	Y^*	R^*	(A^*,B^*)
Mbase	✓	✓			✓		
Mregion	✓	✓	✓	✓	✓		✓
Mcoord	✓	✓		✓	✓		✓

* is the indicator of a new prostate slice.

TABLE 2

Classification accuracy of the proposed classifiers. Each point estimate is the average value obtained by LOOCV. Data in parentheses are the 95% Bootstrap confidence intervals generated from 1000 Bootstrap samples. For each model, the Bootstrap p-values for AUC and S90 comparisons with other models are listed in the third and fifth column, respectively. The superscript is used to indicate the two models being compared. 1–4 in the superscripts represent Mbase, Mregion, Mcoord and Msmooth, respectively.

Model	AUC	P-value		S90	P-value	
Complete Cases						
Mbase	.729(.679,.781)			429(.337,.515)		
Mregion	.767(.715,.816)	$.013^{1,2}$		.475(.375,.577)	$.015^{1,2}$	
Mcoord	.772(.716,.826)	$.011^{1,3}$	$.413^{2,3}$	.494(.391,.597)	$.012^{1,3}$	$.320^{2,3}$
Msmooth	.822(.756,.880)	$<.001^{1,4}$	$.001^{3,4}$	.599(.485,.712)	$.001^{1,4}$	$<.001^{3,4}$
All Cases						
Mbase	.721(.669,.774)			.419(.332,.506)		
Mregion	.760(.709,.808)	$.036^{1,2}$		.468(.366,.568)	$.039^{1,2}$	
Mcoord	.766(.709,.821)	$.012^{1,3}$	$.293^{2,3}$	.486(.387,.584)	$.013^{1,3}$	$.209^{2,3}$
Msmooth	.818(.752,.880)	$.002^{1,4}$	$<.001^{3,4}$	.596(.482,.709)	$.003^{1,4}$	$<.001^{3,4}$