

UCLA

UCLA Previously Published Works

Title

Hierarchical multivariate directed acyclic graph autoregressive models for spatial diseases mapping.

Permalink

<https://escholarship.org/uc/item/0vf7x67k>

Journal

Statistics in Medicine, 41(16)

Authors

Gao, Leiwen

Datta, Abhirup

Banerjee, Sudipto

Publication Date

2022-07-20

DOI

10.1002/sim.9404

Peer reviewed



Published in final edited form as:

Stat Med. 2022 July 20; 41(16): 3057–3075. doi:10.1002/sim.9404.

Hierarchical multivariate directed acyclic graph autoregressive models for spatial diseases mapping

Leiwen Gao¹, Abhirup Datta², Sudipto Banerjee¹

¹Department of Biostatistics, University of California, Los Angeles, Los Angeles, California, USA

²Department of Biostatistics, Johns Hopkins University, Baltimore, Maryland, USA

Abstract

Disease mapping is an important statistical tool used by epidemiologists to assess geographic variation in disease rates and identify lurking environmental risk factors from spatial patterns. Such maps rely upon spatial models for regionally aggregated data, where neighboring regions tend to exhibit similar outcomes than those farther apart. We contribute to the literature on multivariate disease mapping, which deals with measurements on multiple (two or more) diseases in each region. We aim to disentangle associations among the multiple diseases from spatial autocorrelation in each disease. We develop multivariate directed acyclic graphical autoregression models to accommodate spatial and inter-disease dependence. The hierarchical construction imparts flexibility and richness, interpretability of spatial autocorrelation and inter-disease relationships, and computational ease, but depends upon the order in which the cancers are modeled. To obviate this, we demonstrate how Bayesian model selection and averaging across orders are easily achieved using bridge sampling. We compare our method with a competitor using simulation studies and present an application to multiple cancer mapping using data from the Surveillance, Epidemiology, and End Results program.

Keywords

areal data analysis; Bayesian hierarchical models; directed acyclic graphical autoregression; multiple disease mapping; multivariate areal data models

1 | INTRODUCTION

Spatially referenced data comprising regional aggregates of health outcomes over delineated administrative units such as counties or zip codes are widely used by epidemiologists to map mortality or morbidity rates and better understand their geographic variation. Disease mapping, as this exercise is customarily called, employs statistical models to present smoothed maps of rates or counts of a disease. Such maps can assist investigators in

Correspondence Sudipto Banerjee, Department of Biostatistics, University of California, Los Angeles, Los Angeles, CA, USA. sudipto@ucla.edu.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

identifying lurking risk factors¹ and in detecting “hot-spots” or spatial clusters emerging from common environmental and socio-demographic effects shared by neighboring regions. By interpolating estimates of health outcome from areal data onto a continuous surface, disease mapping also generates smoothed maps for the small-area scale, adjusting for the sparsity of data or low population size.^{2,3}

For a single disease, there has been a long tradition of employing Markov random fields (MRFs)⁴ to introduce conditional dependence for the outcome in a region given its neighbors. Two conspicuous examples are the conditional autoregression (CAR)^{5,6} and simultaneous autoregression (SAR) models⁷ that build dependence using undirected graphs to model geographic maps. More recently, a class of directed acyclic graphical autoregressive (DAGAR) models was proposed as a preferred alternative to CAR or SAR models in allowing better identifiability and interpretation of spatial autocorrelation parameters.⁸

Multivariate disease mapping is concerned with the analysis of multiple diseases that are associated among themselves and across space. It is not uncommon to find substantial associations among different diseases sharing genetic and environmental risk factors. Quantification of genetic correlations among multiple cancers has revealed associations among several cancers including lung, breast, colorectal, ovarian, and pancreatic cancers.⁹ Disease mapping exercises with lung and esophageal cancers have also evinced associations among them.¹⁰ When the diseases are inherently related so that the prevalence of one encourages (or inhibits) occurrence of the other, there can be substantial inferential benefits in jointly modeling the diseases rather than fitting independent univariate models for each disease.^{10–21}

The existence of multivariate MRFs can be demonstrated using a multivariate extension of the so called “Brook’s lemma,” which attempts to derive a joint distribution from specified full conditionals.^{13,22,23} McNab, in a series of papers, has delivered substantial insights into the construction, computation, and properties of different classes of multivariate CAR models.^{21,24–26} Rather than work with full conditionals, an alternate approach builds joint distributions using linear transformations of a set of univariate CAR models.^{14,16,19,27,28} A different class merges from hierarchical constructions^{10,29} where each disease enters the model in a given sequence (or order) of conditional probability models. This produces simple yet flexible and interpretable association structures, but every ordering produces a different model resulting in an explosion of models even for a modest number of cancers (say, more than 2 or 3 diseases). While multivariate MRF models constructed from undirected graphs are invariant to ordering, hence obviate the issue of order dependence, they impose restrictions to ensure positive-definiteness of covariance matrices, can be computationally onerous and render covariance structures that are challenging to interpret.

We introduce a class of multivariate DAGAR (MDAGAR) models for multiple diseases mapping by building the joint distribution hierarchically using univariate DAGAR models. This approach is analogous to generalized MCAR (GMCAR) models.¹⁰ The objective here is to retain the interpretation of spatial autocorrelation offered by the DAGAR, which is challenging for the CAR³⁰ and order-free MCAR models. Our methodological

innovation is devising a hierarchical MDAGAR model in conjunction with a bridge sampling algorithm^{31,32} for choosing among differently ordered hierarchical models and, more importantly, offering Bayesian model averaged (BMA) inference to neutralize the effect of order dependent inference. The idea is to begin with a fixed ordered set of cancers, posited to be associated with each other and a cross space, and build a hierarchical model. The DAGAR specification produces a comprehensible association structure, while bridge sampling allows us to rank differently ordered models using their marginal posterior probabilities. Since each model corresponds to an assumed conditional dependence, the marginal posterior probabilities will indicate the tenability of such assumptions given the data. Epidemiologists, then, will be able to use this information to establish relationships among the diseases and spatial autocorrelation for each disease.

The article proceeds as follows. Section 2 develops the hierarchical MDAGAR model and introduces a bridge sampling method to select the MDAGAR with the best hierarchical order. Section 3 presents simulation studies comparing MDAGAR with GMCAR and order-free MCAR models and also illustrates model averaged inference from the bridge sampling. Section 4 applies our MDAGAR to age-adjusted incidence rates of four cancers from the Surveillance, Epidemiology, and End Results (SEER) database and discusses different cases with respect to predictors. Finally, in Section 5, we summarize with some concluding remarks and pointers for future research.

2 | METHODS

2.1 | Overview of univariate DAGAR modeling

Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ be a graph corresponding to a geographic map, where the vertices $\mathcal{V} = \{1, 2, \dots, k\}$ represent clearly delineated regions on the map and $\mathcal{E} = \{(i, j): i \sim j\}$ is the collection of edges between the vertices representing neighboring pairs of regions. We denote two neighboring regions i and j by $i \sim j$. We assume that the vertices in \mathcal{V} are ordered in a fixed sequence according to their number labels. The DAGAR model builds a spatial autocorrelation model for a single outcome on \mathcal{G} using the ordered set of vertices in \mathcal{V} ⁸. Let $N(1)$ be the empty set and let $j \in \mathcal{V} \setminus \{1\}$ be the index for any region except 1. We define $N(j)$ to be the set of labels of geographic neighbors of j that precede j in \mathcal{V} , that is, $N(j) = \{l \in \mathcal{V}: l < j; l \sim j\}$.⁸ Let $\{w_i: i \in \mathcal{V}\}$ be a collection of k random variables defined over the map. DAGAR specifies the following autoregression,

$$w_1 = \epsilon_1; \quad w_j = \sum_{l \in N(j)} b_{jl} w_l + \epsilon_j, \quad j = 2, \dots, k, \quad (1)$$

where $\epsilon_j \stackrel{\text{ind}}{\sim} N(0, \lambda_j)$ with the precision λ_j , and $b_{jl} = 0$ if $l \notin N(j)$. This implies that $\mathbf{w} \sim N(\mathbf{0}, \tau \mathbf{Q}(\rho))$, where $\mathbf{Q}(\rho)$ is a spatial precision matrix that depends only upon a spatial autocorrelation parameter ρ and τ is a positive scale parameter. The precision matrix $\mathbf{Q}(\rho) = (\mathbf{I} - \mathbf{B})^\top \mathbf{F} (\mathbf{I} - \mathbf{B})$, \mathbf{B} is a $k \times k$ strictly lower-triangular matrix and \mathbf{F} is a $k \times k$ diagonal matrix. The elements of \mathbf{B} and \mathbf{F} are denoted by b_{jl} and λ_j , respectively, where

$$b_{jl} = \begin{cases} 0, & \text{if } l \notin N(j) \\ \frac{\rho}{1 + (n_{<j} - 1)\rho^2}, & \text{if } j = 2, 3, \dots, k, l \in N(j) \end{cases} \quad \text{and} \quad \lambda_j = \frac{1 + (n_{<j} - 1)\rho^2}{1 - \rho^2}, j = 1, 2, \dots, k. \quad (2)$$

$n_{<j}$ is the number of members in $N(j)$ and $n_{<1} = 0$. The above definition of b_{jl} is consistent with the lower-triangular structure of $c\mathbf{B}$ because $l \notin N(j)$ for any $l \geq j$. The derivation of \mathbf{B} and \mathbf{F} as functions of a spatial correlation parameter ρ is based upon forming local autoregressive models on embedded spanning trees of subgraphs of \mathcal{G} .⁸

DAGAR and CAR are both examples of MRFs.⁴ They are similar in that both models use a graph to model geographic neighbors, but they are different in how they model spatial dependencies. DAGAR, as the name suggests, builds dependencies using a directed acyclic graph (DAG). This produces a joint likelihood using sequential construction of the partial conditional distributions $w_i | w_{<i}$. CAR builds a joint model by specifying Gaussian full conditional distributions $w_i | w_{-i}$ by treating the underlying map as an undirected graph, where absence of an edge between two regions denotes conditional independence of their spatial effects given other geographic neighbors. These two approaches yield different structures for the precision matrix $\mathbf{Q}(\rho)$ with different interpretations for the parameter ρ . DAGAR retains the interpretation of ρ as an autocorrelation parameter,⁸ while the interpreting spatial autocorrelation in CAR is challenging.³⁰

2.2 | Motivating multivariate disease mapping

There is a substantial literature on joint modeling of multiple spatially oriented outcomes, some of which have been cited in Section 1. While it is possible to model each disease separately using a univariate DAGAR, hence independent of each other, the resulting inference will ignore the association among the diseases. This will be manifested in model assessment because the less dependence among diseases that a model accommodates, the farther away it will be from the joint model in the sense of Kullback-Leibler divergence.

More formally, suppose we have two mutually exclusive sets A and B that contain labels for diseases. Let \mathbf{y}_A and \mathbf{y}_B be the vectors of spatial outcomes over all regions corresponding to the diseases in set A and set B , respectively. A full joint model $p(\mathbf{y})$, where $\mathbf{y} = (\mathbf{y}_A^\top, \mathbf{y}_B^\top)^\top$, can be written as $p(\mathbf{y}) = p(\mathbf{y}_A) \times p(\mathbf{y}_B | \mathbf{y}_A)$. Let C_1 and C_2 be two nested subsets of diseases in A such that $C_2 \subset C_1 \subset A$. Consider two competing models, $p_1(\mathbf{y}) = p(\mathbf{y}_A) \times p(\mathbf{y}_B | \mathbf{y}_{C_1})$ and $p_2(\mathbf{y}) = p(\mathbf{y}_A) \times p(\mathbf{y}_B | \mathbf{y}_{C_2})$, where $p_1(\cdot)$ and $p_2(\cdot)$ are probability densities constructed from the joint probability measure $p(\cdot)$ by imposing conditional independence such that $p(\mathbf{y}_B | \mathbf{y}_A) = p(\mathbf{y}_B | \mathbf{y}_{C_1})$ and $p(\mathbf{y}_B | \mathbf{y}_A) = p(\mathbf{y}_B | \mathbf{y}_{C_2})$, respectively. Both $p_1(\cdot)$ and $p_2(\cdot)$ suppress dependence by shrinking the conditional set A , but $p_2(\cdot)$ suppresses more than $p_1(\cdot)$. We show below that $p_2(\cdot)$ is farther away from $p(\cdot)$ than $p_1(\cdot)$.

A straightforward application of Jensen's inequality yields $\mathbb{E}_{B|C_1} \left[\log \frac{p(\mathbf{y}_B | \mathbf{y}_{C_1})}{p(\mathbf{y}_B | \mathbf{y}_{C_2})} \right] \geq 0$, where $\mathbb{E}_{B|C_1}[\cdot]$ denotes the conditional expectation with respect to $p(\mathbf{y}_B | \mathbf{y}_{C_1})$. Therefore,

$$\begin{aligned} \text{KL}(p||p_2) - \text{KL}(p||p_2) &= \mathbb{E}_{A,B} \left[\log \left(\frac{p(\mathbf{y})}{p_2(\mathbf{y})} \right) - \log \left(\frac{p(\mathbf{y})}{p_2(\mathbf{y})} \right) \right] = \mathbb{E}_{A,B} \left[\log \frac{p(\mathbf{y})}{p_2(\mathbf{y})} \right] \\ &= \mathbb{E}_{A,B} \left[\log \frac{p(\mathbf{y}_B | \mathbf{y}_{C_1})}{p(\mathbf{y}_B | \mathbf{y}_{C_2})} \right] \\ &= \mathbb{E}_{B,C_1} \left[\log \frac{p(\mathbf{y}_B | \mathbf{y}_{C_1})}{p(\mathbf{y}_B | \mathbf{y}_{C_2})} \right] = \mathbb{E}_{C_1} \left\{ \mathbb{E}_{B|C_1} \left[\log \frac{p(\mathbf{y}_B | \mathbf{y}_{C_1})}{p(\mathbf{y}_B | \mathbf{y}_{C_2})} \right] \right\} \geq 0. \end{aligned} \quad (3)$$

The equality $\mathbb{E}_{A,B}[\cdot] = \mathbb{E}_{B,C_1}[\cdot]$ in the last row follows from the fact that the argument is a function of diseases in B, C_1 and C_2 and, hence, in B and C_1 because $C_2 \subset C_1$. The argument given in (3) is free of distributional assumptions and is linked to the submodularity of entropy and the ‘‘information never hurts’’ principle.^{33,34} Equation (3) shows that models built upon hierarchical dependence structures depend upon the order in which the diseases enter the model. While this is a disadvantage, hierarchical dependencies are easier to interpret, easier to compute using currently available Bayesian modeling software such as BUGS or JAGS and have been shown to be very competitive in inferential performance.³⁵ Hence, we develop and implement Bayesian model averaging over different ordered models in a computationally efficient manner.

2.3 | Multivariate DAGAR model

Modeling multiple diseases will introduce associations among the diseases and spatial dependence for each disease. Let y_{ij} be a disease outcome of interest for disease i in region j . For sake of clarity, we assume that y_{ij} is a continuous variable (eg, incidence rates) related to a set of explanatory variables through the regression model,

$$y_{ij} = \mathbf{x}_{ij}^\top \boldsymbol{\beta}_i + \mathbf{w}_{ij} + e_{ij}, \quad (4)$$

where \mathbf{x}_{ij} is a $p_i \times 1$ vector of explanatory variables specific to disease i within region j , $\boldsymbol{\beta}_i$ are the slopes corresponding to disease i , \mathbf{w}_{ij} is a random effect for disease i in region j , and $e_{ij} \stackrel{\text{ind}}{\sim} N(0, (\sigma_i^2)^{-1})$ is the random noise arising from uncontrolled imperfections in the data.

Part of the residual from the explanatory variables is captured by the spatial-temporal effect w_{ij} . Let $\mathbf{w}_i = (w_{i1}, w_{i2}, \dots, w_{ik})^\top$ for $i = 1, 2, \dots, q$. We adopt a hierarchical approach,¹⁰ where we specify the joint distribution of $\mathbf{w} = (\mathbf{w}_1^\top, \mathbf{w}_2^\top, \dots, \mathbf{w}_q^\top)^\top$ as $p(\mathbf{w}) = p(\mathbf{w}_1) \prod_{i=2}^q p(\mathbf{w}_i | \mathbf{w}_{<i})$. We model $p(\mathbf{w}_1)$ and each of the conditional densities $p(\mathbf{w}_i | \mathbf{w}_{<i})$ with $\mathbf{w}_{<i} = (\mathbf{w}_1^\top, \dots, \mathbf{w}_{i-1}^\top)^\top$ for $i \geq 2$ as univariate spatial models. The merits of this approach include simplicity and computational efficiency while ensuring that richness in structure is accommodated through the $p(\mathbf{w}_i | \mathbf{w}_{<i})$'s.

We point out two important distinctions from the GMCAR model¹⁰: (i) instead of using CAR for the spatial dependence, we use DAGAR; and (ii) we apply a computationally

efficient bridge sampling algorithms³² to compute the marginal posterior probabilities for each ordered model. The first distinction allows better interpretation of spatial autocorrelation than the CAR models. The second distinction is of immense practical value and makes this approach feasible for a much larger number of outcomes. Without this distinction, analysts would be dealing with $q!$ models for q diseases and choose among them based upon a model-selection metric. That would be overly burdensome for more than 2 or 3 diseases.

2.4 | A conditional multivariate DAGAR model

The multivariate DAGAR (or MDAGAR) model is constructed as

$$\mathbf{w}_1 = \epsilon_1; \quad \mathbf{w}_i = A_{i1}\mathbf{w}_1 + A_{i2}\mathbf{w}_2 + \dots + A_{i,i-1}\mathbf{w}_{i-1} + \epsilon_i \text{ for } i = 2, 3, \dots, q, \tag{5}$$

where $\epsilon_i \sim N(\mathbf{0}, \tau_i \mathbf{Q}(\rho_i))$ and $\tau_i \mathbf{Q}(\rho_i)$ are univariate DAGAR precision matrices with \mathbf{B} and \mathbf{F} as in (2). In (5), we model \mathbf{w}_1 as a univariate DAGAR and, progressively, the conditional density of each \mathbf{w}_i given $\mathbf{w}_1, \dots, \mathbf{w}_{i-1}$ is also as a DAGAR for $i = 2, 3, \dots, q$.

Each disease has its own distribution with its own spatial autocorrelation parameter. There are q spatial autocorrelation parameters, $\{\rho_1, \rho_2, \dots, \rho_q\}$, corresponding to the q diseases. Given the differences in the geographic variation of different diseases, this flexibility is desirable. Each matrix $A_{i'}$ in (5) with $i' = 1, \dots, i - 1$ models the association between diseases i and i' . We specify $A_{i'} = \eta_{0i'} \mathbf{I}_k + \eta_{ii'} \mathbf{M}$, where \mathbf{M} is the binary adjacency matrix for the map, that is, $m_{j'j} = 1$ if $j' \sim j$ and 0 otherwise. Coefficients $\eta_{0i'}$ and $\eta_{ii'}$ associate w_{ij} with $w_{i'j}$ and $w_{i'j'}$. In other words, $\eta_{0i'}$ is the diagonal element in $A_{i'i'}$, while $\eta_{ii'}$ is the element in the j th row and j' th column if $j' \sim j$. Therefore, for the joint distribution of \mathbf{w} , if \mathbf{A} is the $kq \times kq$ strictly block-lower triangular matrix with $(i'j')$ th block being $A_{i'i'}$ whenever $i' \geq i$, and $\epsilon = (\epsilon_1^\top, \dots, \epsilon_q^\top)^\top$, then (5) renders $\mathbf{w} = \mathbf{A}\mathbf{w} + \epsilon$.

Since $\mathbf{I} - \mathbf{A}$ is still lower triangular with 1s on the diagonal, it is nonsingular with $\det(\mathbf{I} - \mathbf{A}) = 1$. Writing $\mathbf{w} = (\mathbf{I} - \mathbf{A})^{-1} \epsilon$, where $\epsilon \sim N(\mathbf{0}, \Lambda)$ and the block diagonal matrix Λ has $\tau_i \mathbf{Q}(\rho_i), \dots, \tau_q \mathbf{Q}(\rho_q)$ on the diagonal, we obtain $\mathbf{w} \sim N(\mathbf{0}, \mathbf{Q}_w)$ for $\rho = (\rho_1, \dots, \rho_q)^\top$ with

$$\mathbf{Q}_w = (\mathbf{I} - \mathbf{A})^\top \Lambda (\mathbf{I} - \mathbf{A}). \tag{6}$$

We say that \mathbf{w} follows MDAGAR if $\mathbf{w} \sim N(\mathbf{0}, \mathbf{Q}_w)$.

Interpretation of ρ_1, \dots, ρ_q is clear: ρ_1 measures the spatial association for the first disease, while $\rho_i, i \geq 2$, is the residual spatial correlation in the disease i after accounting for the first $i - 1$ diseases. Similarly, τ_1 is the spatial precision for the first disease, while $\tau_i, i \geq 2$, is the residual spatial precision for disease i after accounting for the first $i - 1$ diseases.

2.4.1 | Model implementation—We extend (4) to the following Bayesian hierarchical framework with the posterior distribution

$$\begin{aligned}
p(\boldsymbol{\beta}, \mathbf{w}, \boldsymbol{\eta}, \boldsymbol{\rho}, \boldsymbol{\tau}, \boldsymbol{\sigma} \mid \mathbf{y}) &\propto p(\boldsymbol{\rho}) \times p(\boldsymbol{\eta}) \times \prod_{i=1}^q \\
&\{IG(1/\tau_i \mid a_\tau, b_\tau) \times IG(\sigma_i^2 \mid a_\sigma, b_\sigma) \times N(\boldsymbol{\beta}_i \mid \boldsymbol{\mu}_\beta, \mathbf{V}_\beta^{-1})\} \\
&\times N(\mathbf{w} \mid \mathbf{0}, \mathbf{Q}_w) \times \prod_{i=1}^q \prod_{j=1}^k N(y_{ij} \mid \mathbf{x}_{ij}^\top \boldsymbol{\beta}_i + \mathbf{w}_{ij}, 1/\sigma_i^2),
\end{aligned} \tag{7}$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top, \dots, \boldsymbol{\beta}_q^\top)^\top$, $\boldsymbol{\tau} = \{\tau_1, \tau_2, \dots, \tau_q\}$, $\boldsymbol{\sigma} = \{\sigma_1^2, \sigma_2^2, \dots, \sigma_q^2\}$ and $\boldsymbol{\eta} = \{\boldsymbol{\eta}_2, \boldsymbol{\eta}_3, \dots, \boldsymbol{\eta}_q\}$ with $\boldsymbol{\eta}_i = (\boldsymbol{\eta}_{i1}^\top, \boldsymbol{\eta}_{i2}^\top, \dots, \boldsymbol{\eta}_{i,i-1}^\top)^\top$ and $\boldsymbol{\eta}_{i,i'} = (\boldsymbol{\eta}_{0i,i'}^\top, \boldsymbol{\eta}_{1i,i'}^\top)^\top$ for $i = 2, \dots, q$ and $i' = 1, \dots, i-1$. For variance parameters $1/\tau_i$ and σ_i^2 , $IG(\cdot \mid a, b)$ is the inverse-gamma distribution with shape and rate parameters a and b , respectively. For each element in $\boldsymbol{\eta}_i$ we choose a normal prior $N(\boldsymbol{\mu}_{ij}, \sigma_{\eta_{ij}}^2)$, while the prior $N(\mathbf{w} \mid \mathbf{0}, \mathbf{Q}_w)$ can also be written as

$$\begin{aligned}
p(\mathbf{w} \mid \boldsymbol{\tau}, \boldsymbol{\eta}_2, \dots, \boldsymbol{\eta}_q, \boldsymbol{\rho}) &\propto \tau_1^k |\mathbf{Q}(\rho_1)|^{\frac{1}{2}} \exp\left\{-\frac{\tau_1}{2} \mathbf{w}_1^\top \mathbf{Q}(\rho_1) \mathbf{w}_1\right\} \\
&\times \prod_{i=2}^q \tau_i^k |\mathbf{Q}(\rho_i)|^{\frac{1}{2}} \exp\left\{-\frac{\tau_i}{2} \left(\mathbf{w}_i - \sum_{i'=1}^{i-1} \mathbf{A}_{i'} \mathbf{w}_{i'}\right)^\top \mathbf{Q}(\rho_i) \left(\mathbf{w}_i - \sum_{i'=1}^{i-1} \mathbf{A}_{i'} \mathbf{w}_{i'}\right)\right\},
\end{aligned} \tag{8}$$

where $\det(\mathbf{Q}(\rho_i)) = \prod_{j=1}^k \lambda_{ij}$, and $\mathbf{w}_i^\top \mathbf{Q}(\rho_i) \mathbf{w}_i = \lambda_{i1} w_{i1}^2 + \sum_{j=2}^k \lambda_{ij} (w_{ij} - \sum_{j' \in N(j)} b_{ijj'} w_{ij'})^2$.

We sample the parameters from the posterior distribution in (7) using Markov chain Monte Carlo (MCMC) with Gibbs sampling and random walk metropolis³⁶ as implemented in the `rjags` package within the `R` statistical computing environment. Web Appendix B S.2.1 presents details on the MCMC updating scheme.

2.5 | Model selection via bridge sampling

It is clear from (5) that each ordering of diseases in MDAGAR will produce a different model. For the bivariate situation, it is convenient to compare only two models (orders) by the significance of parameter estimates as well as model performance. However, when there are more than two diseases involved in the model, at least six models (for three diseases) will be fitted and comparing all models become cumbersome or even impracticable.

Instead, we pursue model averaging of MDAGAR models. Given a set of $T = q!$ candidate models, say M_1, \dots, M_T , Bayesian model selection and model averaging calculates

$$p(M = M_t \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid M = M_t) p(M = M_t)}{\sum_{j=1}^T p(\mathbf{y} \mid M = M_j) p(M = M_j)}, \tag{9}$$

for $t = 1, \dots, T$.³⁷ Computing the marginal likelihood $p(\mathbf{y} \mid M_i)$ in (9) is challenging. Methods such as importance sampling³⁸ and generalized harmonic mean³⁹ have been proposed as stable estimators with finite variance, but finding the required importance density with strong constraints on the tail behavior relative to the posterior distribution is often challenging. Bridge sampling estimates the marginal likelihood (ie, thenormalizing constant)

by combining samples from two distributions: a bridge function $h(\cdot)$ and a proposal distribution $g(\cdot)$.⁴⁰ Let $\theta_i = \{\beta_i, \sigma_i, \rho_i, \tau_i, \eta_{2,i}, \dots, \eta_{q,i}\}$ be the set of parameters in model M_i with prior $p(\theta_i | M_i)$ as defined in the first row of (7). Based on the identity,

$$1 = \frac{\int p(\mathbf{y} | \theta_i, M_i) p(\theta_i | M_i) h(\theta_i | M_i) g(\theta_i | M_i) d\theta_i}{\int p(\mathbf{y} | \theta_i, M_i) p(\theta_i | M_i) h(\theta_i | M_i) g(\theta_i | M_i) d\theta_i},$$

a current version of the bridge sampling estimator is

$$\begin{aligned} p(\mathbf{y} | M = M_i) &= \frac{E_{g(\theta_i | M_i)}[p(\mathbf{y} | \theta_i, M_i) p(\theta_i | M_i) h(\theta_i | M_i)]}{E_{p(\theta_i | \mathbf{y}, M_i)}[h(\theta_i | M_i) g(\theta_i | M_i)]} \\ &\approx \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} p(\mathbf{y} | \tilde{\theta}_{i,i}, M_i) p(\tilde{\theta}_{i,i} | M_i) h(\tilde{\theta}_{i,i} | M_i)}{\frac{1}{N_1} \sum_{j=1}^{N_1} h(\theta_{i,j}^* | M_i) g(\theta_{i,j}^* | M_i)}, \end{aligned} \tag{10}$$

where $\theta_{i,j}^* \sim p(\theta_i | \mathbf{y}, M_i)$, $j = 1, \dots, N_1$, are N_1 posterior samples and $\tilde{\theta}_{i,i} \sim g(\theta_i | M_i)$, $i = 1, \dots, N_2$, are N_2 samples drawn from the proposal distribution.³² The likelihood $p(\mathbf{y} | \theta_i, M = M_i)$ is obtained by integrating out \mathbf{w} from (7) as

$$N(\mathbf{y} | \mathbf{X}\beta, [\mathbf{Q}_w^{-1}(\rho_i, \tau_i, \eta_{2,i}, \dots, \eta_{q,i}) + \text{diag}(\sigma_i) \otimes \mathbf{I}_k]^{-1}), \tag{11}$$

given that $\mathbf{y} = (\mathbf{y}_1^\top, \dots, \mathbf{y}_q^\top)^\top$ with $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{ik})^\top$, $\text{diag}(\sigma)$ is a diagonal matrix with σ_i^2 , $i = 1, \dots, q$, on the diagonal, and \mathbf{X} is the design matrix with \mathbf{X}_i as block diagonal where $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{ik})^\top$. The bridge function $h(\theta_i | M_i)$ is specified by the optimal choice³¹,

$$h(\theta_i | M_i) = C \frac{1}{s_1 p(\mathbf{y} | \theta_i, M_i) p(\theta_i | M_i) + s_2 p(\mathbf{y} | M_i) g(\theta_i | M_i)}, \tag{12}$$

where C is a constant. Inserting (12) in (10) yields the estimate of $p(\mathbf{y} | M = M_i)$ after convergence of an iterative scheme³¹ as

$$\hat{p}(\mathbf{y} | M_i)^{(t+1)} = \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} \frac{l_{2,i}}{s_1 l_{2,i} + s_2 \hat{p}(\mathbf{y} | M_i)^{(t)}}}{\frac{1}{N_1} \sum_{j=1}^{N_1} \frac{1}{s_1 l_{1,j} + s_2 \hat{p}(\mathbf{y} | M_i)^{(t)}}}, \tag{13}$$

where $l_{1,j} = \frac{p(\mathbf{y} | \theta_{i,j}^*, M_i) p(\theta_{i,j}^* | M_i)}{g(\theta_{i,j}^* | M_i)}$, $l_{2,i} = \frac{p(\mathbf{y} | \tilde{\theta}_{i,i}, M_i) p(\tilde{\theta}_{i,i} | M_i)}{g(\tilde{\theta}_{i,i} | M_i)}$, $s_1 = \frac{N_1}{N_1 + N_2}$, and $s_2 = \frac{N_2}{N_1 + N_2}$.

Given the log marginal likelihood estimates from `bridgesampling`, the posterior model probability for each model is calculated from (9) by setting prior probability of each model $p(M = M_i)$. For Bayesian model averaging (BMA), the model averaged posterior distribution of a quantity of interest Δ is obtained as $p(\Delta | \mathbf{y}) = \sum_{i=1}^T p(\Delta | M = M_i, \mathbf{y}) p(M = M_i | \mathbf{y})$,³⁷ and the posterior mean is

$$E(\Delta | \mathbf{y}) = \sum_{i=1}^T E(\Delta | M = M_i, \mathbf{y}) p(M = M_i | \mathbf{y}). \quad (14)$$

Setting $\Delta = \{\boldsymbol{\beta}, \mathbf{w}\}$ fetches us the model averaged posterior estimates for spatial random effects as well as calculating the posterior mean incidence rates as discussed in Section 4.

3 | SIMULATION

We simulate three different experiments. The first is designed to evaluate MDAGAR's inferential performance against GMCAR. The second compares MDAGAR, GMCAR and order-free MCAR for data generated from the latter. The third experiment illustrates the effectiveness of bridge sampling (Section 2.5) in preferring models with a correct “ordering” of the diseases.

3.1 | Data generation

We compare MDAGAR's inferential performance with GMCAR¹⁰ (Section 3.2) and order-free MCAR¹⁶ (Section 3.3). We choose the 48 states of the contiguous United States as our underlying map, where two states are treated as neighbors if they share a common geographic boundary. We generated our outcomes y_{ij} using the model in (4) with $q = 2$, that is, two outcomes, and two covariates, x_{1j} and x_{2j} , with $p_1 = 2$ and $p_2 = 3$. We fixed the values of the covariates after generating them from $N(\mathbf{0}, \mathbf{I}_{p_i}), i = 1, 2$, independent across regions.

The regression slopes were set to $\boldsymbol{\beta}_1 = (1, 5)^\top$ and $\boldsymbol{\beta}_2 = (2, 4, 5)^\top$.

Turning to the spatial random effects, we generated values of $\mathbf{w} = (\mathbf{w}_1^\top, \mathbf{w}_2^\top)^\top$ from a $N(\mathbf{0}, \mathbf{Q}_w)$ distribution, where the precision matrix is

$$\mathbf{Q}_w = \begin{bmatrix} \tau_1 \mathbf{Q}(\rho_1) + \tau_2 \mathbf{A}_{21}^\top \mathbf{Q}(\rho_2) \mathbf{A}_{21} & \tau_2 \mathbf{A}_{21}^\top \mathbf{Q}(\rho_2) \\ \tau_2 \mathbf{Q}(\rho_2) \mathbf{A}_{21} & \tau_2 \mathbf{Q}(\rho_2) \end{bmatrix}. \quad (15)$$

We set $\tau_1 = \tau_2 = 0.25$, $\rho_1 = 0.2$, and $\rho_2 = 0.8$ in (15) and take $\mathbf{Q}(\rho_i) = \mathbf{D}(\rho_i)^{-1}$, where $\mathbf{D}(\rho_i) = \exp(-\phi_i d(j, j'))$, $\phi_i = -\log(\rho_i)$ is the spatial decay for disease i and $d(j, j')$ refers to the distance between the embedding of the j th and j' th vertex. The vertices are embedded on the Euclidean plane and the centroid of each state is used to create the distance matrix. Using this exponential covariance matrix to generate the data offers a “neutral” ground to compare the performance of MDAGAR with GMCAR. We specified \mathbf{A}_{12} using fixed values of $\boldsymbol{\eta} = \{\eta_{021}, \eta_{121}\}$. Here, we considered three sets of values for $\boldsymbol{\eta}$ to correspond to low, medium and high correlation among diseases. We fixed $\boldsymbol{\eta} = \{0.05, 0.1\}$ to ensure an average correlation of 0.15 (range 0.072–0.31); $\boldsymbol{\eta} = \{0.5, 0.3\}$ with an average correlation of 0.55 (range 0.45–0.74); and $\boldsymbol{\eta} = \{2.5, 0.5\}$ with a mean correlation of 0.89 (range 0.84–0.94). We generated w_{ij} 's for each of the above specifications for $\boldsymbol{\eta}$ and, with the values of w_{ij} generated as above, we generated the outcome $y_{ij} \sim N(\mathbf{x}_{ij}^\top \boldsymbol{\beta}_i + w_{ij}, 1/\sigma_i^2)$, where $\sigma_1^2 = \sigma_2^2 = 0.4$. We repeat the above procedure to replicate 85 datasets for each of the three specifications of $\boldsymbol{\eta}$.

For our third experiment (Sections 3.4 and 3.5), we generate a dataset with $q = 3$ cancers. We extend the above setup to include one more disease. We generate y_{ij} 's from (4) with the value of x_{3j} fixed after being generated from $N(\mathbf{0}, \mathbf{I}_3)$, $\beta_3 = (5, 3, 6)^\top$, and $\sigma_3^2 = 0.4$. Let $[i, j, k]$ denote the model $p(\mathbf{w}_i) \times p(\mathbf{w}_j | \mathbf{w}_i) \times p(\mathbf{w}_k | \mathbf{w}_j, \mathbf{w}_i)$. For three diseases the six resulting models are denoted as $M_1 = [1, 2, 3]$, $M_2 = [1, 3, 2]$, $M_3 = [2, 1, 3]$, $M_4 = [2, 3, 1]$, $M_5 = [3, 1, 2]$, and $M_6 = [3, 2, 1]$.

Each of the six models imply a corresponding joint distribution $\mathbf{w} \sim N(\mathbf{0}, \mathbf{Q}_w)$ which is used to generate the \mathbf{w}_{ij} 's. Let the parenthesized suffix (i) denote the disease in the i th order. For example, in $M_2 = [1, 3, 2]$, we write \mathbf{w} in the form of (5) as

$$\mathbf{w}_1 \sim \epsilon_{(1)}; \quad \mathbf{w}_3 = \mathbf{A}_{(21)}\mathbf{w}_1 + \epsilon_{(2)}; \quad \mathbf{w}_2 = \mathbf{A}_{(31)}\mathbf{w}_1 + \mathbf{A}_{(32)}\mathbf{w}_3 + \epsilon_{(3)},$$

where $\epsilon_{(i)} \sim N(\mathbf{0}, \tau_{(i)}\mathbf{Q}(\rho_{(i)}))$ with $\mathbf{Q}(\rho_{(i)}) = D(\rho_{(i)})^{-1}$ as in the first experiment, and $\mathbf{A}_{(i'j')} = \eta_{0(i'j')}\mathbf{I} + \eta_{1(i'j')}\mathbf{M}$ is the coefficient matrix associating random effects for diseases in the i th and j' th order. We set

$$\tau_{(1)} = \tau_{(2)} = \tau_{(3)} = 0.25, \rho_{(1)} = 0.2, \rho_{(2)} = 0.8, \rho_{(3)} = 0.5, \eta_{0(21)} = 0.5, \eta_{1(21)} = 0.3, \eta_{0(31)} = 1, \eta_{1(31)} = 0.6, \eta_{0(32)} = 1.5,$$

and $\eta_{1(32)} = 0.9$ to completely specify \mathbf{Q}_w for each of the 6 models. For each M_i , we generate 50 datasets by first generating $\mathbf{w} \sim N(\mathbf{0}, \mathbf{Q}_w)$ and then generating y_{ij} 's from (4) using the above specifications. Details on the algorithms and the computing environments for each model are provided in Section S.2.1.

3.2 | Comparisons between MDAGAR and GMCAR

In our first experiment, we analyzed the 85 replicated datasets using (7) with

$$p(\rho) \times p(\boldsymbol{\eta}) \propto \prod_{i=1}^{q=2} \{\text{Unif}(\rho_i | 0, 1)\} \times N(\boldsymbol{\eta}_{21} | \mathbf{0}, 0.01\mathbf{I}_2), \quad (16)$$

where $\boldsymbol{\eta}_{21} = (\eta_{021}, \eta_{121})^\top$ and Unif is the uniform density. Prior specifications are completed by setting $a_r = 2, b_r = 0.4, a_\sigma = 2, b_\sigma = 0.4, \boldsymbol{\mu}_\beta = \mathbf{0}, V_\beta = 1000\mathbf{I}$ in (7). The same set of priors was used for both MDAGAR and GMCAR as they have the same number of parameters with similar interpretations. Both models are fast to compute; MDAGAR reported an average running time of 3.87 minutes for each dataset in the bivariate disease analysis, while that for GMCAR was 6.25 minutes.

We compare models using the widely applicable information criterion (WAIC)^{41,42} and a model comparison score D based on a balanced loss function for replicated data.⁴³ Both WAIC and D reward goodness of fit and penalize model complexity. Details on how these metrics are computed are provided in Web Appendix B S.2.2. In addition, we also computed the average mean squared error (AMSE) of the spatial random effects estimated from each of the 85 datasets. We found the mean (standard deviation) of the AMSEs to be 1.69 (0.034) from the 85 low-correlation datasets, 1.47 (0.030) from the 85 medium-correlation datasets, and 2.35 (0.059) from the 85 high-correlation datasets. The corresponding numbers for

GMCAR were 1.83 (0.033), 1.59 (0.031), and 2.14 (0.050), respectively. The MDAGAR tends to have smaller AMSE for low and medium correlations, while GMCAR's AMSE tends to be pronouncedly lower than MDAGAR's when the correlations are high. We also compute the mean values of WAICs and D scores for each simulated dataset. Figure 1 plots the values of WAICs (A-C) and D scores (D-F) for the 85 datasets corresponding to each of the three correlation settings. Here, MDAGAR outperforms GMCAR in all three correlation settings with respect to both WAICs and D scores. While MDAGAR outperforms GMCAR in overall model fitting scores for most correlation settings, GMCAR can yield better estimates of spatial effects in high correlation settings.

Figure 2 presents scatter plots for the true values (x -axis) of spatial random effects against their posterior estimates (y -axis). To be precise, each panel plots $85 \times 48 \times 2 = 8160$ true values of the elements of the 96×1 vector \mathbf{w} for 85 datasets against their corresponding posterior estimates. We see strong agreements between the true values and their estimates for both MDAGAR and GMCAR. The agreement is more pronounced for the datasets corresponding to medium and high correlations. For the low-correlation datasets, MDAGAR still exhibits strong agreement which is better than GMCAR.

We compute $D_{\text{KL}}(N(\mathbf{0}, \mathbf{Q}_{\text{true}}) \| N(\mathbf{0}, \mathbf{Q}_w)) = \frac{1}{2} \left[\log \left(\frac{\det(\mathbf{Q}_{\text{true}})}{\det(\mathbf{Q}_w)} \right) + \text{tr}(\mathbf{Q}_w \mathbf{Q}_{\text{true}}^{-1}) - qk \right]$, which is the Kullback-Leibler divergence between the model for \mathbf{w} with the true generative precision matrix (\mathbf{Q}_{true}) and those with MDAGAR and GMCAR precisions (\mathbf{Q}_w). Using the posterior samples in the precision matrix, we evaluate the posterior probability that $D_{\text{KL}}(N(\mathbf{0}, \mathbf{Q}_{\text{true}}) \| N(\mathbf{0}, \mathbf{Q}_{\text{MDAGAR}}))$ is smaller than $D_{\text{KL}}(N(\mathbf{0}, \mathbf{Q}_{\text{true}}) \| N(\mathbf{0}, \mathbf{Q}_w))$. Figure 3 depicts a density plot of these probabilities over the 85 datasets. In low and medium correlations, the MDAGAR has a mean probability of around 69% to be closer to the true model than the GMCAR, while for high correlations GMCAR excels with an average probability of 72% to be closer to the true model. These findings are consistent with the AMSEs, where GMCAR tended to perform better when correlations were high. Additional comparative diagnostics from MDAGAR and GMCAR, such as coverage probabilities for parameters and correlations between random effects for two diseases in the same state, are presented in Web Appendix B S.2.2.2.

3.3 | Comparisons between MDAGAR and order-free MCAR

We also generated data using an order-free MCAR model¹⁶ to evaluate MDAGAR and GMCAR when the underlying structure is different from the proposed conditional scheme. For the MCAR model, we specified the joint covariance matrix of \mathbf{w} as

$$\mathbf{Q}_w^{-1} = (\mathbf{A} \otimes \mathbf{I}_{k \times k}) \mathbf{\Gamma}^{-1} (\mathbf{A} \otimes \mathbf{I}_{k \times k})^{\top}, \quad (18)$$

where $\mathbf{\Sigma} = \mathbf{A} \mathbf{A}^{\top}$ is a $q \times q$ matrix corresponding to disease dependence, \mathbf{A} is the upper triangular Cholesky decomposition of $\mathbf{\Sigma}$ and $\mathbf{\Gamma}$ is a $kq \times kq$ block diagonal matrix with $\mathbf{\Gamma}_{ii}^{-1} = \tau_i^2 (\mathbf{D} - \rho_i \mathbf{W})$ ($k \times k$ precision matrix for a proper CAR) for each $i = 1, \dots, q$.

This corresponds to the MCAR generated from $\mathbf{w} = (\mathbf{A} \otimes \mathbf{I}) \mathbf{v}$, where $\mathbf{v} = (\mathbf{v}_1, \dots, \mathbf{v}_q)^{\top}$ and $\mathbf{v}_i \stackrel{\text{ind}}{\sim} N(\mathbf{0}, \mathbf{D} - \rho_i \mathbf{W})$ for $i = 1, \dots, q$, \mathbf{D} is the diagonal matrix with number of neighbors along the diagonal and \mathbf{W} is a binary adjacency matrix. Therefore, \mathbf{w} is generated from

independent but not identically distributed latent proper CAR distributions (see Reference 16, Section 3.2).

Keeping other model specifications same as in Section 3.1 (so $q = 2$ and ρ_j^2 's are as in Section 3.1), we fixed $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0.7 & 1 \end{bmatrix}$. Computing (17) with these specifications yields a mean correlation of 0.52 among the entries of the matrix (range: 0.48–0.54). The above procedure is replicated for 50 datasets for each model. We estimated the MDAGAR and GMCAR models in two opposite orders, denoted MDAGAR1, MDAGAR2, GMCAR1, and GMCAR2, and compared with the order-free MCAR. We estimate (7) with the respective specifications for \mathbf{Q}_w for each model. For the MDAGAR and GMCAR models, we used the priors specified in the previous section using (16). For the MCAR, we assigned $\log a_{ii}$, $i = 1, 2$, and a_{21} with normal priors with variances 0.0625 and 100, respectively. The order-free MCAR is also fast to compute and reported an average running time of 5.89 minutes for each dataset in this experiment.

Figure 4 plots values of (A) WAICs, (B) D scores, and (C) the posterior mean of the Kullback-Leibler divergence between a given model and the true density, $D_{\text{KL}}(p(\mathbf{y}_{\text{true}})||p(\mathbf{y}))$, for each of the 50 datasets (indexed in the x -axis) computed for each of the five models. For model fitting, GMCAR1 exhibits better performance with smaller values for WAIC and D , while GMCAR2, MDAGAR1, and MDAGAR2 are all comparable with MCAR. GMCAR1 and MDAGAR1 exhibit slightly better performance in D scores compared with GMCAR2 and MDAGAR2, respectively, but the two orders produce similar WAICs. In terms of the posterior means of $D_{\text{KL}}(p(\mathbf{y}_{\text{true}})||p(\mathbf{y}))$, MCAR is expectedly closer to the true model (having the same data generating structure), but MDAGAR is still very competitive performance in spite of being a misspecified model. The variability in the posterior means of $D_{\text{KL}}(p(\mathbf{y}_{\text{true}})||p(\mathbf{y}))$ for the different models reveal substantial overlap so conditional models have the ability to compete with order-free MCAR even when data are generated from the latter.

3.4 | Analyses using different orderings for spatial units

The MDAGAR model in Section 3.2 is analyzed using an ordering of spatial units (counties) from the southwest to the northeast. Here, we repeat the analysis for the MDAGAR model using three other orderings that start in the southeast, northwest, and northeast, respectively. We present results from these differently ordered DAGAR models using the 85 low-correlation simulated datasets. For the random effects, the mean (standard deviation) of the AMSEs for three different orderings (southeast, northwest, and northeast) are 1.61 (0.029), 1.28 (0.026), and 1.43 (0.027), respectively, without significantly differing from the original ordering in Section 3.2.

Figure 5 plots the densities of mean WAICs, D scores, and $D_{\text{KL}}(p(\mathbf{y}_{\text{true}})||p(\mathbf{y}))$ over the 85 datasets for the MDAGAR model using three different orderings and the original ordering in Section 3.2. In computing $D_{\text{KL}}(p(\mathbf{y}_{\text{true}})||p(\mathbf{y}))$, we specify $p(\mathbf{y}_{\text{true}}) = N(\mathbf{X}\boldsymbol{\beta}_{\text{true}} + \mathbf{w}_{\text{true}}, \text{diag}(\boldsymbol{\sigma}_{\text{true}}) \otimes \mathbf{I}_k)$, which is the density of the true \mathbf{y} and $p(\mathbf{y}) = N(\mathbf{X}\boldsymbol{\beta} + \mathbf{w}, \text{diag}(\boldsymbol{\sigma}) \otimes \mathbf{I}_k)$ is the density for \mathbf{y} from MDAGAR. While the ordering of the diseases does not appear to have a significant impact

on model fitting as the density plots for the four orderings almost overlap with each other, (3) suggests that some order dependence may be expected.

3.5 | Model selection for different disease orders

We now evaluate the effectiveness of the method in Section 2.5 at selecting the model with the correct ordering of diseases. We used the `bridgesampling` package in R to compute $p(M_i | \mathbf{y}^{(n)}) = \max_{t=1, \dots, 6} p(M_t | \mathbf{y}^{(n)})$ for each of $n = 50 \times 6$ datasets generated as described in Section 3.1. Table 1 presents the probability of each model being selected for different true model scenarios. The probability of selecting the true model is shown in bold along the diagonal. Our experiment reveals that bridge sampling is extremely effective at choosing the correct order. It was able to identify the correct order between 78% and 90%, which is substantially larger than any of the probability of choosing any of the misspecified models.

4 | MULTIPLE CANCER ANALYSIS FROM SEER

We now turn to analyzing an areal dataset using the MDAGAR model for four different cancers: lung, esophagus, larynx, and colorectal. The incidence of adenocarcinoma of lung and esophageal cancer have been found to share common risk factors⁴⁴ and metabolic mechanisms.⁴⁵ Lung cancer appears to be among the most common second primary cancers in patients with colon cancer.⁴⁶ Meanwhile, patients with laryngeal cancer have also been reported to possess high risks of developing second primary lung cancer.⁴⁷ The dataset is extracted from the SEER*Stat database using the SEER*Stat statistical software.⁴⁸ The dataset consists of the four cancers: lung, esophagus, larynx, and colorectal, where the outcome is the 5-year average age-adjusted incidence rates (age-adjusted to the 2000 U.S. Standard Population) per 100000 population in the years from 2012 to 2016 across 58 counties in California, USA, as mapped in Figure 6. The maps exhibit preliminary evidence of correlation across space and among cancers. Cutoffs for the different levels of incidence rates are quantiles for each cancer. For all four cancers, incidence rates are relatively higher in counties concentrated in the middle northern areas including Shasta, Tehama, Glenn, Butte, and Yuba than those other areas. In general, northern areas have higher incidence rates than in the south. This is especially pronounced for lung cancer and esophagus cancer. For larynx cancer, while the highest incidence rates are in the northwest (Del Norte and Siskiyou counties), the incidence rates in the south are also at somewhat higher levels. For colorectal cancer, the edge areas at the bottom also exhibit high incidence rates.

As an exploratory tool to assess associations among the cancers, we calculate Pearson's correlation for each pair of cancers by regarding incidence rates in different counties as independent samples and find Pearson's correlation coefficient between the incidence of lung cancer and those of esophageal, larynx, and colorectum cancers to be 0.55, 0.46, and 0.46, respectively. Meanwhile, the correlation between esophageal and larynx cancer is 0.27. Next, to explore the spatial association for each disease, we calculate Moran's I based upon r th order neighbors for each cancer and plot the areal correlogram.⁴⁹ Defining distance intervals, $(0, d_1], (d_1, d_2], (d_2, d_3], \dots$, the r th order neighbors refer to units with distance in $(d_{r-1}, d_r]$, that is, within distance d_r but separated by more than d_{r-1} . The distance is the Euclidean distance from an Albers map projection of California. As shown in Figure 7,

lung, esophageal, and colorectum cancers all present spatial patterns that initially diminish with increasing r and eventually flatten close to 0. Overall, counties with similar levels of incidence rates tend to depict some spatial clustering.

We turn to model based inference using (7). We return to the MDAGAR, GMCAR, and MCAR, where neighbors are defined using shared borders. We analyze this dataset and separate the spatial correlation for each cancer from association among cancers with the following prior specifications,

$$\begin{aligned}
 p(\boldsymbol{\eta}, \boldsymbol{\rho}, \boldsymbol{\tau}, \boldsymbol{\sigma}, \boldsymbol{w}) &= \prod_{i=1}^q \text{Unif}(\rho_i | 0, 1) \times \prod_{i=2}^q \prod_{j=1}^{i-1} N(\boldsymbol{\eta}_{ij} | 0, 0.01\mathbf{I}_2) \\
 &\times \prod_{i=1}^q N(\boldsymbol{\beta}_i | 0, 0.001\mathbf{I}) \times \prod_{i=1}^q \text{IG}(1/\tau_i | 2, 0.1) \times \prod_{i=1}^q \text{IG}(\sigma_i^2 | 2, 1) \\
 &\times N(\boldsymbol{w} | \mathbf{0}, \mathbf{Q}_w).
 \end{aligned} \tag{19}$$

We also discuss a case excluding the risk factor (see Web Appendix B Section S.2.2.3).

For covariates, we include county attributes that possibly affect the incidence rates, including percentages of residents younger than 18 years old (young_{ij}), older than 65 years old (old_{ij}), with education level below high school (edu_{ij}), percentages of unemployed residents (unemp_{ij}), black residents (black_{ij}), male residents (male_{ij}), uninsured residents (uninsure_{ij}), and percentages of families below the poverty threshold (poverty_{ij}). All covariates are common for different cancers and extracted from the SEER*Stat database⁴⁸ for the same period, 2012 to 2016.

Since cigarette smoking is a common risk factor for cancers, adult smoking rates (smoke_{ij}) for 2014 to 2016 were obtained from the California Tobacco Facts and Figures 2018 database.⁵⁰ Spatial patterns in the map of adult cigarette smoking rates, shown in Figure 8, are similar to the incidence of cancers, especially lung and esophageal cancers, the highest smoking rates are concentrated in the north. While some central California counties (eg, Stanislaus, Tuolumne, Merced, Mariposa, Fresno, and Tulare) also exhibit high rates, although there is clearly less spatial clustering of the high rates than in the north.

Since the order of cancers in the DAG specify the model, we fit all $4! = 24$ models using (7) and compute the marginal likelihoods using bridge sampling (Section 2.5). By setting the prior model probabilities as $p(M = M_t) = \frac{1}{24}$ for $t = 1, 2, \dots, 24$, we compute the posterior model probabilities using (9). These are presented in Table 2. We obtain BMA estimates using (14) with the weights in Table 2. Among all models, model M_{10} is selected as the best model with the largest posterior probability 0.577 and the corresponding conditional structure is [esophageal] \times [larynx | esophageal] \times [colorectal | esophageal, larynx] \times [lung | esophageal, larynx, colorectal].

Table 3 is a summary of the parameter estimates including regression coefficients, spatial auto correlation (ρ_i), spatial precision (τ_i), and noise variance (σ_i^2) for each cancer. From M_{10} and BMA, we find the regression slopes for the percentage of smokers and uninsured

residents are significantly positive and negative, respectively, for esophageal cancer. The negative association between percentage of uninsured and esophageal cancer may seem surprising, but is likely a consequence of counties exhibiting low incidence rates for esophageal cancer having a relatively large number of uninsured residents (see top right in Figure 6 and the right most figure in Figure 8). Since esophageal cancer has low incidence rates, this association could well be spurious due to spatial confounding. Percentage of smokers is, unsurprisingly, found to be a significant risk factor for lung cancer, while the percentage of blacks seems to be significantly associated with elevated incidence of larynx cancer. In addition, we tend to see that percentage of population below the poverty level has a pronounced association with higher rates of lung and esophageal cancer.

Recall from Section 2.4 that ρ_1 is the residual spatial autocorrelation for esophageal cancer after accounting for the explanatory variables, while ρ_i for $i = 2, 3, 4$ are residual spatial autocorrelations after accounting for the explanatory variables and the preceding cancers in the model M_{10} . From Table 3, we see that esophageal cancer exhibits relatively weaker spatial autocorrelation, while the residual spatial autocorrelations for larynx and colorectal cancers after accounting for preceding cancers are both at moderate levels of around 0.5. Similarly for the spatial precision τ , larynx appears to have the smallest conditional variability while that for colorectal and lung are slightly larger.

For the posterior mean incidence rates and spatial random effects w_{ij} , we present estimates from model M_{10} and BMA. Figure 9A,B is maps of posterior mean spatial random effects and model fitted incidence rates for four cancers obtained from BMA, while Figure 10A,B shows maps of those from model M_{10} . The posterior mean incidence rates from BMA and M_{10} are in accord with each other, and both present DAGAR-smoothed versions of the original patterns in Figure 6. For posterior means of spatial random effects, in general, the estimates from M_{10} are similar to model averaged estimates, especially for lung and colorectal cancers, exhibiting relatively large positive values in the northern counties, where the incidence rates are high. However, for esophageal and larynx cancers we see slight discrepancies between M_{10} and BMA in the north. The BMA estimates produce larger positive random effects, ranging between 0.1 and 0.5, in most counties, while M_{10} produces estimates between 0 and 0.1 for esophageal cancer. More counties with random effects larger than 0.1 are estimated from M_{10} for larynx cancer. We believe this is attributable, at least in part, to another competitive model, $M_{15} = [\text{larynx}] \times [\text{esophagus} \mid \text{larynx}] \times [\text{lung} \mid \text{larynx, esophagus}] \times [\text{colorectal} \mid \text{larynx, esophagus, lung}]$ (posterior probability 0.342), which contributes to the BMA. On the other hand, the effects of some important county-level covariates play an essential role in the discrepancy between the estimates of random effects and model fitted incidence rates for each cancer.

Recall from Section 2 that $\eta_{0ii'}$ and $\eta_{1ii'}$ reflect the associations among cancers that can be attributed to spatial structure. Specifically, larger values of $\eta_{0ii'}$ will indicate inherent associations unrelated to spatial structure, while the magnitude of $\eta_{1ii'}$ reflects associations due to spatial structure. Figure S.2 presents posterior distributions of η for all pairs of cancers. We see from the distribution of η_{043} that there is a pronounced nonspatial

component in the association between lung and colorectal cancers. Similar, albeit somewhat less pronounced, nonspatial associations are seen between larynx and esophageal cancers and between lung and larynx cancers. Analogously, the posterior distributions for η_{143} and η_{132} tend to have substantial positive support suggesting substantial spatial cross-correlations between lung and colorectal cancers and between colorectal and larynx cancers. Interestingly, we find negative support in the posterior distributions for η_{121} and η_{142} . The negative mass implies that the covariance among cancers within a region is suppressed by strong dependence with neighboring regions. This seems to be the case for associations between lung and esophageal cancers and between lung and larynx cancers.

Web Appendix B also presents supplementary analysis that excludes adult smoking rates from the covariates, which we refer to as “Case 2.” Figure S.3 shows estimated correlations between pairwise cancers in each of the 58 counties. The top row presents the correlations including smoking rates (“Case 1”) as has been analyzed here. The bottom row presents the corresponding maps for “Case 2.” Interestingly, accounting for smoking rates substantially diminishes the associations among esophageal, colorectal and lung cancers. These are significantly associated in “Case 2” but only lung and colorectal retain their significance after accounting for smoking rates.

We also implemented the order-free MCAR model (as described in Section 3.3) and presented the estimates of posterior mean incidence rates and spatial random effects in Figure 11. Compared with MDAGAR, the MCAR exhibits better fitting for colorectal cancer since the posterior incidence rates in Figure 11B is closer to those in the raw map (Figure 6), while MDAGAR seems to outperform MCAR for larynx cancer. Overall, the model fitting is comparable between MDAGAR and MCAR.

5 | DISCUSSION

We have developed a multivariate “MDAGAR” model in conjunction with a bridge sampling method to estimate spatial correlations for multiple correlated diseases. The MDAGAR is constructed hierarchically over areal units based on univariate DAGAR models. We demonstrate that MDAGAR tends to outperform GMCAR when association between spatial random effects for different diseases is weak or moderate. MDAGAR retains the interpretability of spatial autocorrelations, as in univariate DAGAR, separating the spatial correlation for each disease from any inherent or endemic association among diseases. While MDAGAR, like all DAG based models, is specified according to a fixed order of the diseases, we show that bridge sampling can effectively choose among the different orders and also provide BMA inference in a computationally efficient manner.

Our data analysis elicits how correlations between incidence rates for different cancers are impacted by risk factors. For example, eliminating adult cigarette smoking rates produces similar spatial patterns for the incidence rates of esophageal, lung and colorectal cancer. In addition, the significant correlation between lung and esophageal cancer, even after accounting for smoking rates, implies other inherent or endemic association such as latent risk factors and metabolic mechanisms. We also see that the MDAGAR based posterior estimates of the latent spatial effects in Figures 9A and 10A resemble those from MDAGAR

without covariates (Figure 12), while the maps for the estimated incidence rates in Figures 9B and 10B account for the spatial variability of the covariates.

Future research will look at different constructions of graphical models for areal data. Examples can include defining r th order neighbors using distance metrics, as in Figure 7, and deriving alternate precision matrices. We also intend to address scalability with very large number of diseases. Here, common spatial factor models for areal data⁵¹ can be adapted to model the factors as DAGAR, thereby yielding classes of DAGAR based factor models. A very different approach will be to build scalable graphical models using two different graphs: one for areal units (CAR or DAGAR) and another undirected graph representing conditional independence among cancers. Multidimensional MRFs as well as developments analogous to recently introduced graphical Gaussian processes⁵² can be pursued for high-dimensional disease mapping. Finally, spatial confounding in multivariate disease mapping^{53–55} will be explored in the context of MDAGAR.

DATA AVAILABILITY STATEMENT

All computer programs implementing the examples in this paper can be found in the public domain and downloaded from https://github.com/LeiwenG/Multivariate_DAGAR.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The work of the first and third authors has been supported in part by the Division of Mathematical Sciences (DMS) of the National Science Foundation (NSF) under grant 1916349 and by the National Institute of Environmental Health Sciences (NIEHS) under grants R01ES030210 and 5R01ES027027. The work of the second author was supported by the Division of Mathematical Sciences (DMS) of the National Science Foundation (NSF) under grant 1915803.

Funding information

Division of Mathematical Sciences, Grant/Award Numbers: 1915803, 1916349; National Institute of Environmental Health Sciences, Grant/Award Numbers: 5R01ES027027, R01ES030210

REFERENCES

1. Koch T. Cartographies of Disease: Maps, Mapping, and Medicine. Redlands, CA: Esri Press; 2005.
2. Berke O. Exploratory disease mapping: kriging the spatial risk function from regional count data. *Int J Health Geogr.* 2004;3(1):18. [PubMed: 15333131]
3. Richardson S, Thomson A, Best N, Elliott P. Interpreting posterior relative risk estimates in disease-mapping studies. *Environ Health Perspect.* 2004;112(9):1016–1025. [PubMed: 15198922]
4. Rue H, Held L. Gaussian Markov Random Fields : Theory and Applications. Monographs on Statistics and Applied Probability. Boca Raton, FL: Chapman and Hall/CRC Press; 2005.
5. Besag J. Spatial interaction and the statistical analysis of lattice systems. *J Royal Stat Soc Ser B (Methodol).* 1974;36(2):192–225.
6. Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Stat Math.* 1991;43(1):1–20.
7. Kissling WD, Carl G. Spatial autocorrelation and the selection of simultaneous autoregressive models. *Glob Ecol Biogeogr.* 2008;17(1): 59–71.

8. Datta A, Banerjee S, Hodges JS, Gao L. Spatial disease mapping using directed acyclic graph auto-regressive (DAGAR) models. *Bayesian Anal.* 2019;14(4):1221–1244. doi:10.1214/19-BA1177 [PubMed: 33859772]
9. Lindström S, Finucane H, Bulik-Sullivan B, et al. Quantifying the genetic correlation between multiple cancer types. *Cancer Epidemiol Prev Biomark.* 2017;26(9):1427–1435.
10. Jin X, Carlin BP, Banerjee S. Generalized hierarchical multivariate CAR models for areal data. *Biometrics.* 2005;61(4):950–961. [PubMed: 16401268]
11. Knorr-Held L, Best NG. A shared component model for detecting joint and selective clustering of two diseases. *J R Stat Soc A Stat Soc.* 2001;164(1):73–85.
12. Kim H, Sun D, Tsutakawa RK. A bivariate Bayes method for improving the estimates of mortality rates with a twofold conditional autoregressive model. *J Am Stat Assoc.* 2001;96(456):1506–1521.
13. Gelfand AE, Vounatsou P. Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics.* 2003;4(1):11–15. [PubMed: 12925327]
14. Carlin BP, Banerjee S. Hierarchical multivariate CAR models for spatio-temporally correlated survival data. *Bayesian Stat.* 2003;7(7):45–63.
15. Held L, Natário I, Fenton SE, Rue H, Becker N. Towards joint disease mapping. *Stat Methods Med Res.* 2005;14(1):61–82. [PubMed: 15691000]
16. Jin X, Banerjee S, Carlin BP. Order-free co-regionalized areal data models with application to multiple-disease mapping. *J Royal Stat Soc Ser B (Stat Methodol).* 2007;69(5):817–838.
17. Zhang Y, Hodges JS, Banerjee S. Smoothed ANOVA with spatial effects as a competitor to MCAR in multivariate spatial smoothing. *Ann Appl Stat.* 2009;3(4):1805. [PubMed: 20596299]
18. Diva U, Dey DK, Banerjee S. Parametric models for spatially correlated survival data for individuals with multiple cancers. *Stat Med.* 2008;27(12):2127–2144. [PubMed: 18167633]
19. Martinez-Beneito MA. A general modelling framework for multivariate disease mapping. *Biometrika.* 2013;100(3):539–553.
20. Mari-Dell’Olmo M, Martinez-Beneito MA, Gotsens M, Palència L. A smoothed ANOVA model for multivariate ecological regression. *Stoch Env Res Risk A.* 2014;28(3):695–706.
21. Lawson B, Banerjee S, Haining R, Ugarte D. *Handbook of Spatial Epidemiology.* Boca Raton, FL: CRC press; 2016.
22. Mardia K. Multi-dimensional multivariate Gaussian Markov random fields with application to image processing. *J Multivar Anal.* 1988;24(2):265–284.
23. Sain SR, Furrer R, Cressie N. A spatial analysis of multivariate output from regional climate models. *Ann Appl Stat.* 2011;5(1):150–175. doi:10.1214/10-AOAS369
24. MacNab YC. Linear models of coregionalization for multivariate lattice data: a general framework for coregionalized multivariate CAR models. *Stat Med.* 2016;35(21):3827–3850. doi:10.1002/sim.6955 [PubMed: 27091685]
25. MacNab YC. Some recent work on multivariate Gaussian Markov random fields (with discussion). *TEST Offic J Spanish Soc Stat Oper Res.* 2018;27(3):497–541. doi:10.1007/s11749-018-0605-3
26. MacNab YC. Bayesian estimation of multivariate Gaussian Markov random fields with constraint. *Stat Med.* 2020;39(30):4767–4788. doi:10.1002/sim.8752 [PubMed: 32935375]
27. Zhu J, Eickhoff JC, Yan P. Generalized linear latent variable models for repeated measures of spatially correlated multivariate data. *Biometrics.* 2005;61(3):674–683. doi:10.1111/j.1541-0420.2005.00343.x [PubMed: 16135018]
28. Bradley JR, Holan SH, Wikle CK. Multivariate spatio-temporal models for high-dimensional areal data with application to longitudinal employer-household dynamics. *Ann Appl Stat.* 2015;9(4):1761–1791. doi:10.1214/15-AOAS862
29. Daniels MJ, Zhou Z, Zou H. Conditionally specified space-time models for multivariate processes. *J Comput Graph Stat.* 2006;15(1):157–177. doi:10.1198/106186006X100434
30. Wall MM. A close look at the spatial structure implied by the CAR and SAR models. *J Stat Plan Infer.* 2004;121(2):311–324.
31. Meng XL, Wong WH. Simulating ratios of normalizing constants via a simple identity: a theoretical exploration. *Stat Sin.* 1996;6:831–860.

32. Gronau QF, Sarafoglou A, Matzke D, et al. A tutorial on bridge sampling. *J Math Psychol.* 2017;81:80–97. [PubMed: 29200501]
33. Cover TM, Thomas JA. *Elements of Information Theory.* Wiley Series in Telecommunications and Signal Processing. Hoboken, NJ: Wiley Interscience; 1991.
34. Banerjee S. Modeling massive spatial datasets using a conjugate Bayesian linear modeling framework. *Spat Stat.* 2020;37:100417. [PubMed: 35265456]
35. Cressie N, Zammit-Mangion A. Multivariate spatial covariance models: a conditional approach. *Biometrika.* 2016;103(4):915–935. doi:10.1093/biomet/asw045
36. Gamerman D, Lopes HF. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference.* Boca Raton, FL: Chapman & Hall/CRC Press; 2006.
37. Hoeting JA, Madigan D, Raftery AE, Volinsky CT. Bayesian model averaging: a tutorial. *Stat Sci.* 1999;14:382–401.
38. Perrakis K, Ntzoufras I, Tsionas EG. On the use of marginal posteriors in marginal likelihood estimation via importance sampling. *Comput Stat Data Anal.* 2014;77:54–69.
39. Gelfand AE, Dey DK. Bayesian model choice: asymptotics and exact calculations. *J Royal Stat Soc Ser B (Methodol).* 1994;56(3):501–514.
40. Gronau QF, Singmann H, Wagenmakers EJ. Bridgesampling: an R package for estimating normalizing constants; 2017. arXiv preprint arXiv:1710.08162.
41. Watanabe S. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J Mach Learn Res.* 2010;11(Dec):3571–3594.
42. Gelman A, Hwang J, Vehtari A. Understanding predictive information criteria for Bayesian models. *Stat Comput.* 2014;24(6):997–1016.
43. Gelfand AE, Ghosh SK. Model choice: a minimum posterior predictive loss approach. *Biometrika.* 1998;85(1):1–11.
44. Agrawal K, Markert RJ, Agrawal S. Risk factors for adenocarcinoma and squamous cell carcinoma of the esophagus and lung. *Hypertension.* 2018;61(46):0–09.
45. Shi WX, Chen SQ. Frequencies of poor metabolizers of cytochrome P450 2C19 in esophagus cancer, stomach cancer, lung cancer and bladder cancer in Chinese population. *World J Gastroenterol.* 2004;10(13):1961. [PubMed: 15222046]
46. Kurishima K, Miyazaki K, Watanabe H, et al. Lung cancer patients with synchronous colon cancer. *Mol Clin Oncol.* 2018;8(1):137–140. [PubMed: 29387406]
47. Akhtar J, Bhargava R, Shameem M, et al. Second primary lung cancer with glottic laryngeal cancer as index tumor—A case report. *Case Rep Oncol.* 2010;3(1):35–39. [PubMed: 20740156]
48. Surveillance Research Program. National Cancer Institute. SEER*Stat software. Surveillance Res Program. 2019. <https://seer.cancer.gov/seerstat/>
49. Banerjee S, Carlin BP, Gelfand AE. *Hierarchical Modeling and Analysis for Spatial Data.* Boca Raton, FL: CRC Press; 2014.
50. California Department of Public Health. California Tobacco control program California tobacco facts and figures; 2018.
51. Wang F, Wall MM. Generalized common spatial factor model. *Biostatistics.* 2003;4(4):569–582. doi:10.1093/biostatistics/4.4.569 [PubMed: 14557112]
52. Dey D, Datta A, Banerjee S. Graphical Gaussian processes for highly multivariate spatial data. *Biometrika.* 2020. doi:10.1093/biomet/asab061
53. Azevedo D, Prates M, Bandyopadhyay D. MSPOCK: alleviating spatial confounding in multivariate disease mapping models. *J Agric Biol Environ Stat.* 2021;26:464–491.
54. Khan K, Calder CA. Restricted spatial regression methods: implications for inference. *J Am Stat Assoc.* 2020;1–13. doi:10.1080/01621459.2020.1788949
55. Zimmerman DL, Hoef JMV. On deconfounding spatial confounding in linear models. *Am Stat.* 2021;1–9. doi:10.1080/00031305.2021.1946149

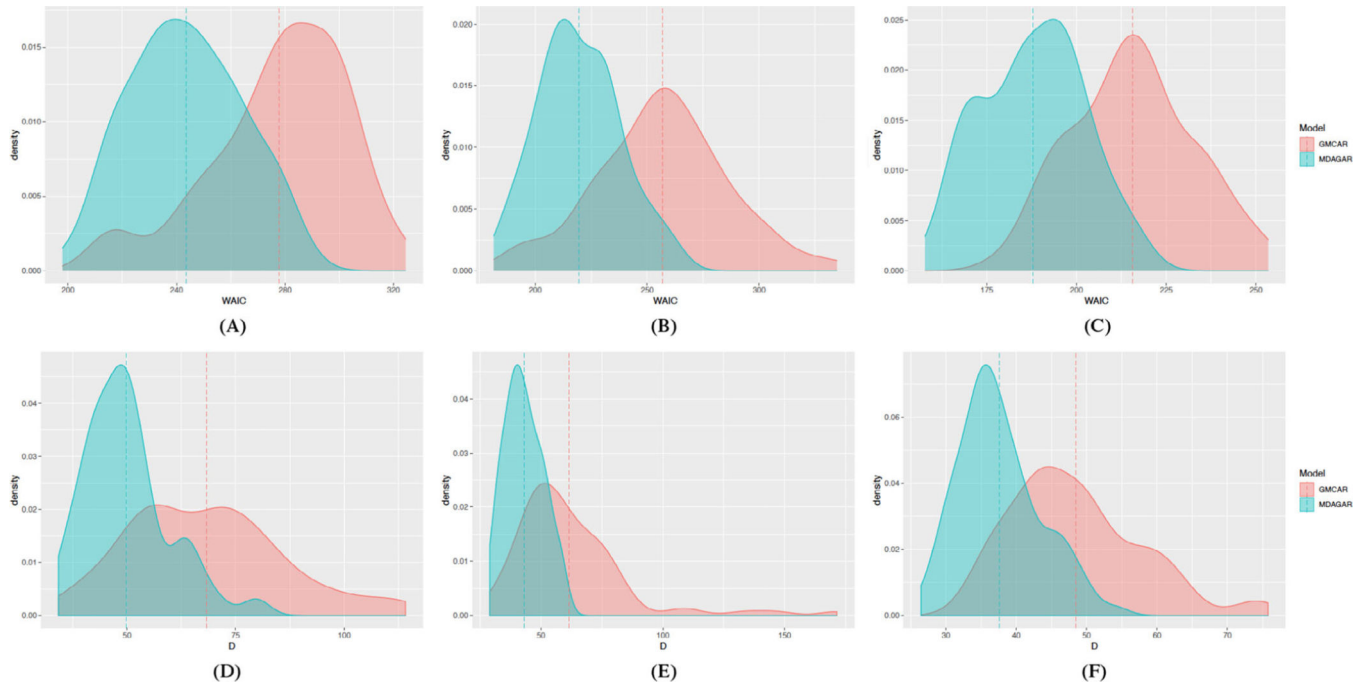


FIGURE 1. Density plots for WAICs and D scores over 85 datasets. (A-C) Density plots of WAIC for MDAGAR (blue) and GMCAR (red) models with low, medium, and high correlation, respectively, (D-F) the corresponding density plots for D scores. The dotted vertical line shows the mean for WAIC and D in each plot

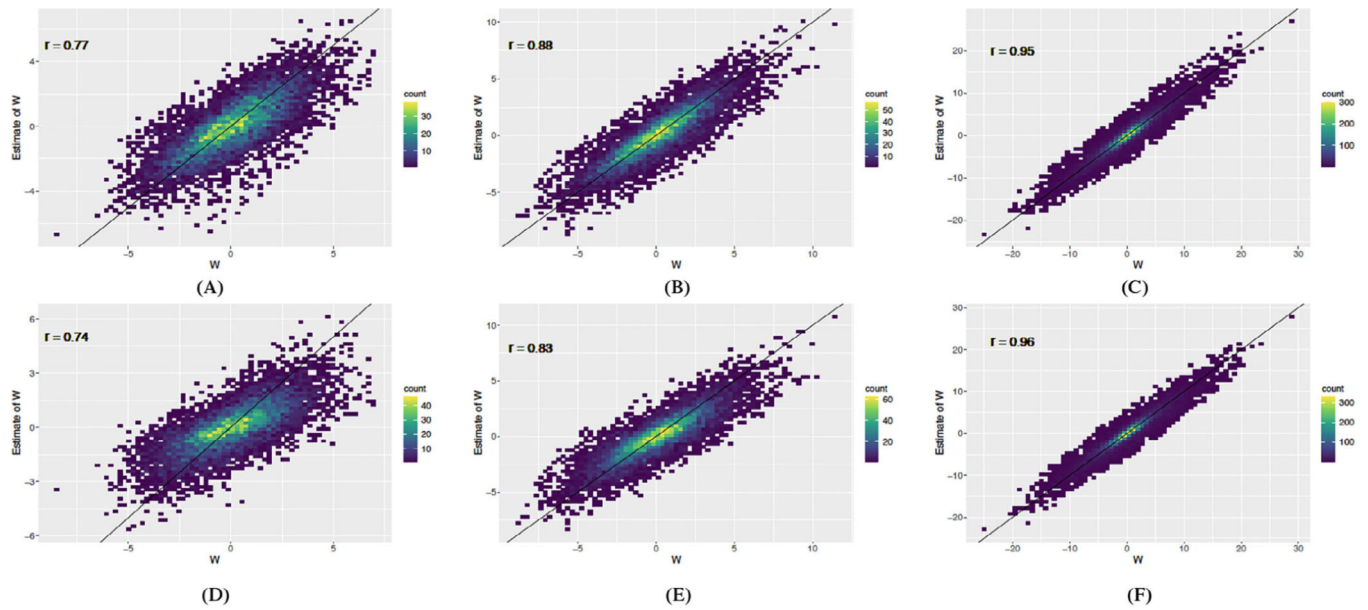


FIGURE 2.

Scatter plots for estimates of spatial random effects (y -axis) against the true values (x -axis) with 45° lines over 85 datasets: (A-C) Estimates from MDAGAR model with low, medium, and high correlation, (D-F) the corresponding estimates from GMCAR. Pearson's correlation coefficient for each plot is indicated as " r "

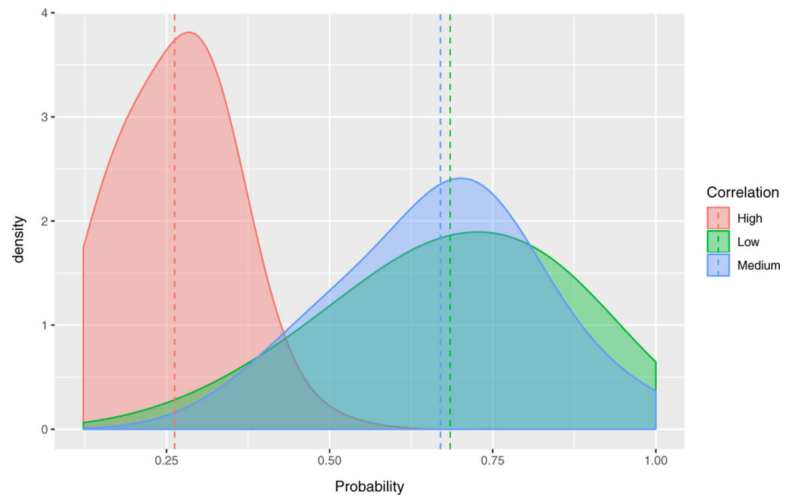
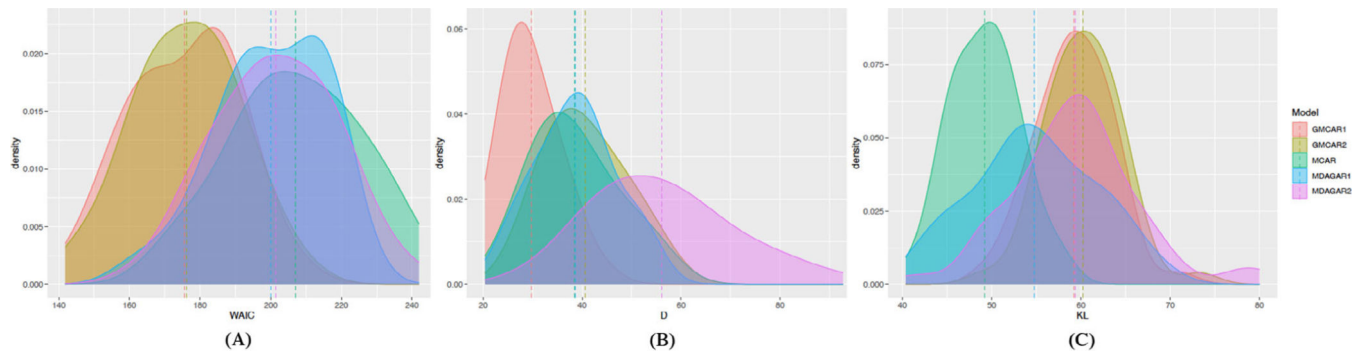


FIGURE 3. Density plots for probability that the KL-divergence between the MDAGAR and the true model is smaller than that between GMCAR and the true model with three levels of correlation for two diseases: Low (purple), medium (green), and high (red)

**FIGURE 4.**

Density plots for (A) WAICs, (B) D scores, and (C) the posterior mean of $D_{KL}(p(y_{true})||p(y))$ over 50 datasets, respectively, using MDAGAR1, MDAGAR2, GMCAR1, GMCAR2, and MCAR. The dot vertical line shows the mean for each plot

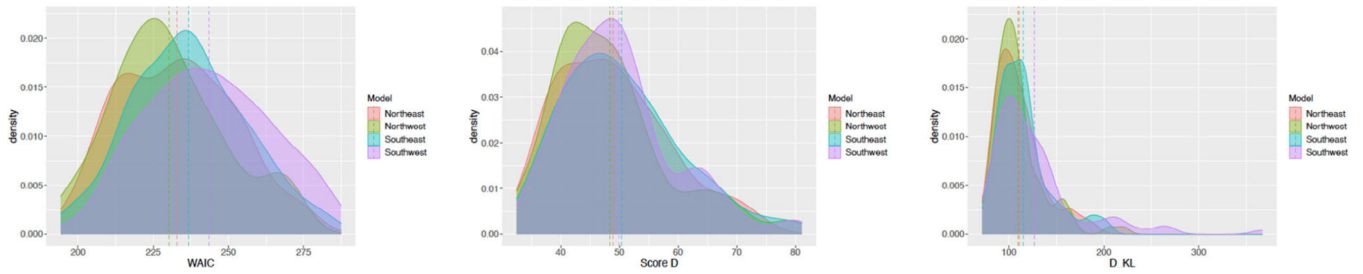


FIGURE 5.

Density plots for WAICs, D scores, and $D_{\text{KL}}(p(\mathbf{y}_{\text{true}})||p(\mathbf{y}))$ over 85 datasets for the MDAGAR model using four different orderings: Northeast (red), northwest (green), southeast (blue), and southwest (purple). The dotted vertical line shows the mean for each plot

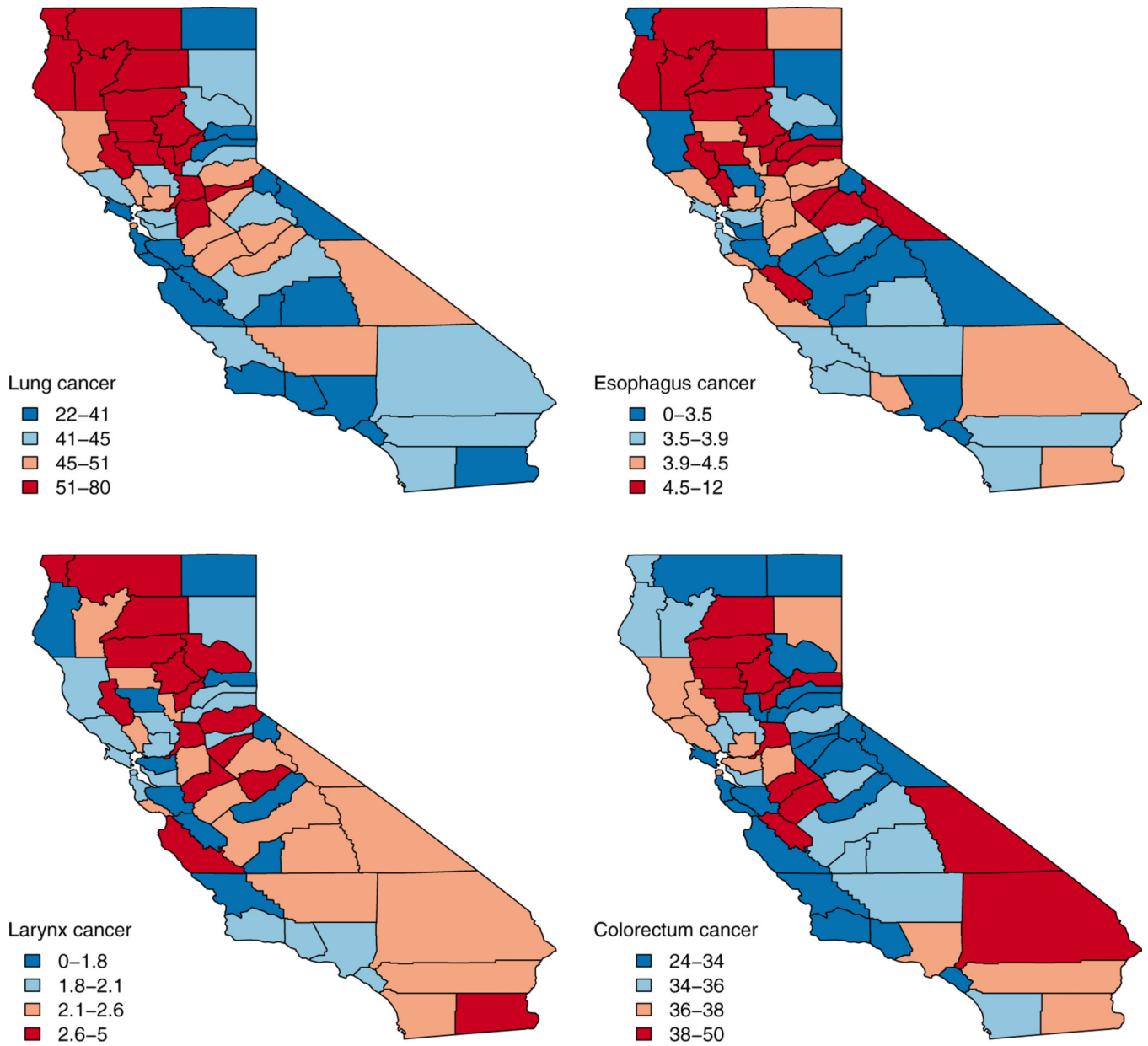


FIGURE 6. Maps of 5-year average age-adjusted incidence rates per 100 000 population for lung, esophagus, larynx, and colorectal cancer in California, 2012 to 2016

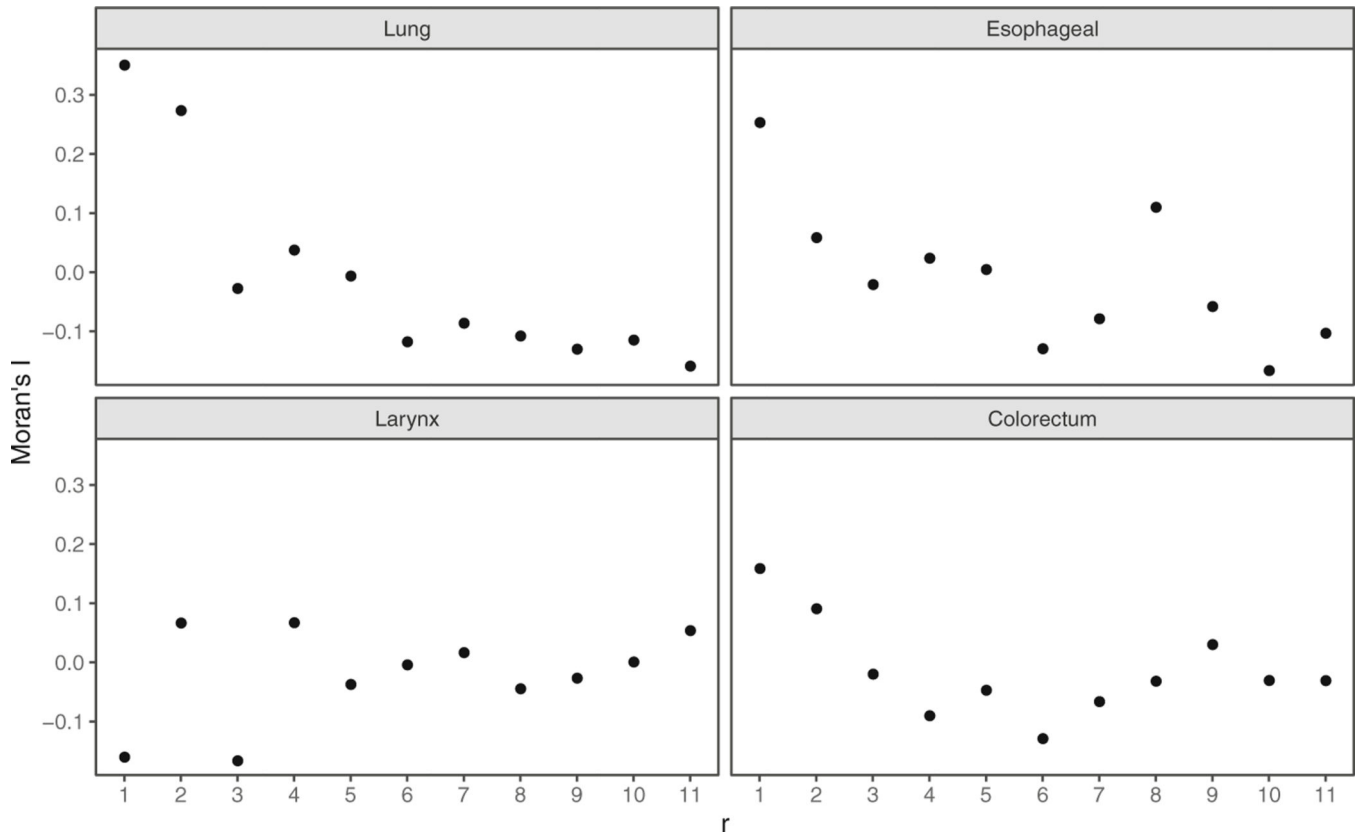


FIGURE 7. Moran's I of r th order neighbors for lung, esophageal, larynx, and colorectum cancer

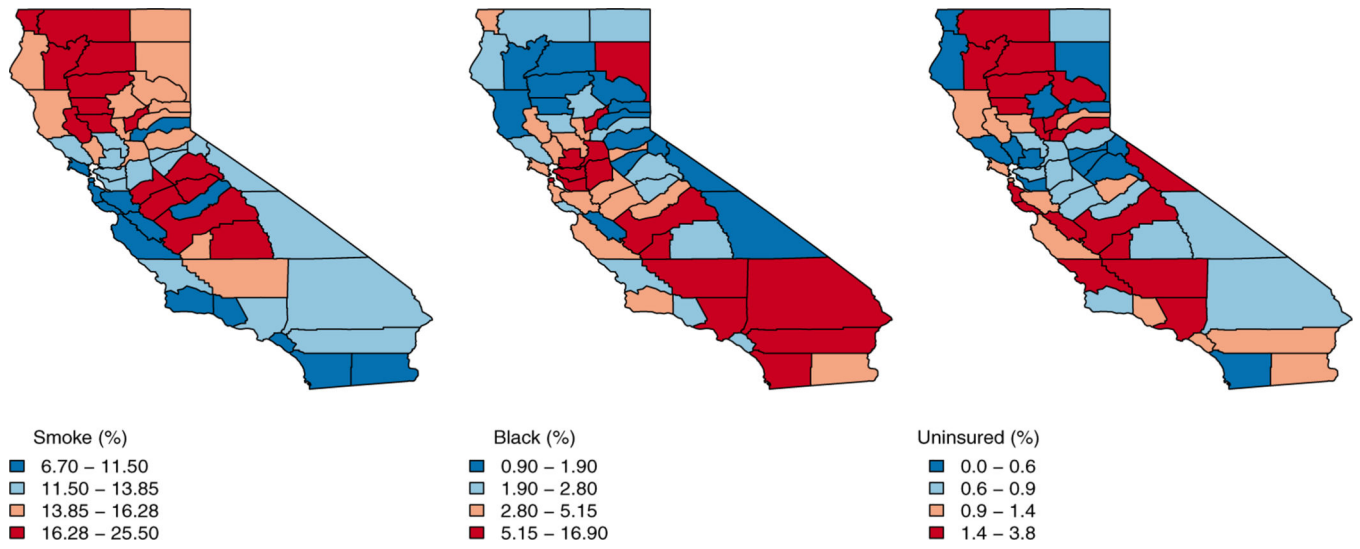


FIGURE 8. Important county-level covariates with significant effects: Adult cigarette smoking rates (left), percentage of black residents (middle), and uninsured residents (right)

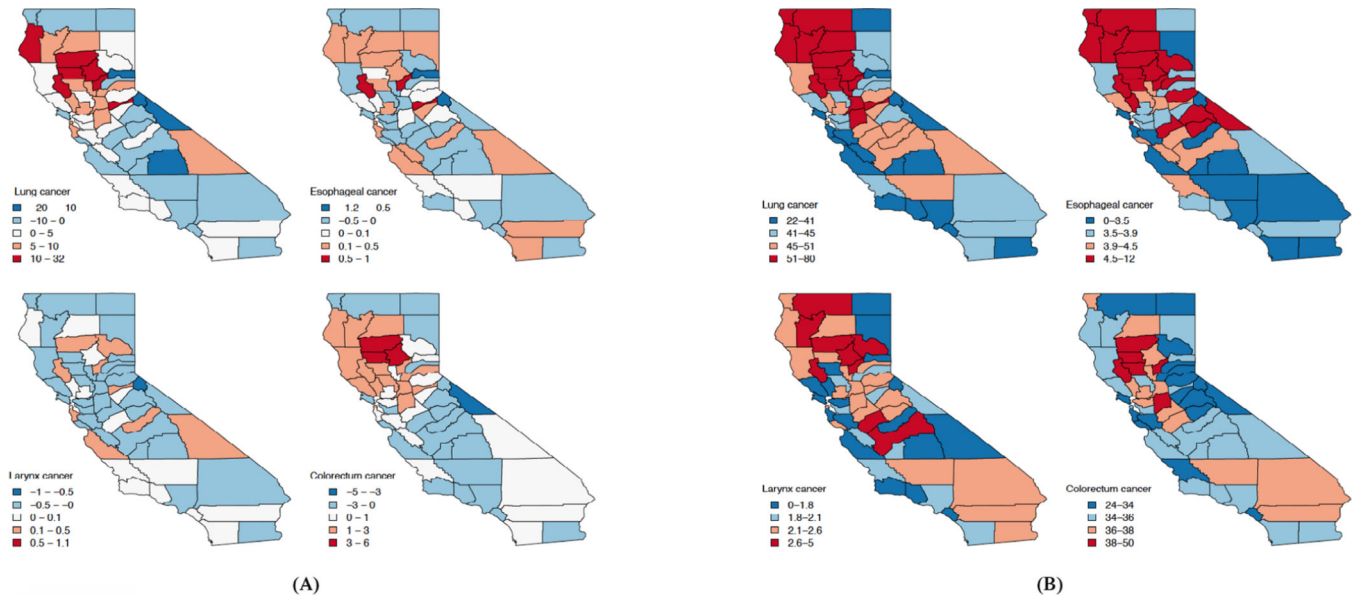


FIGURE 9. Maps of posterior results using BMA for lung, esophagus, larynx, and colorectal cancer in California including (A) posterior mean spatial random effects and (B) posterior mean incidence rates

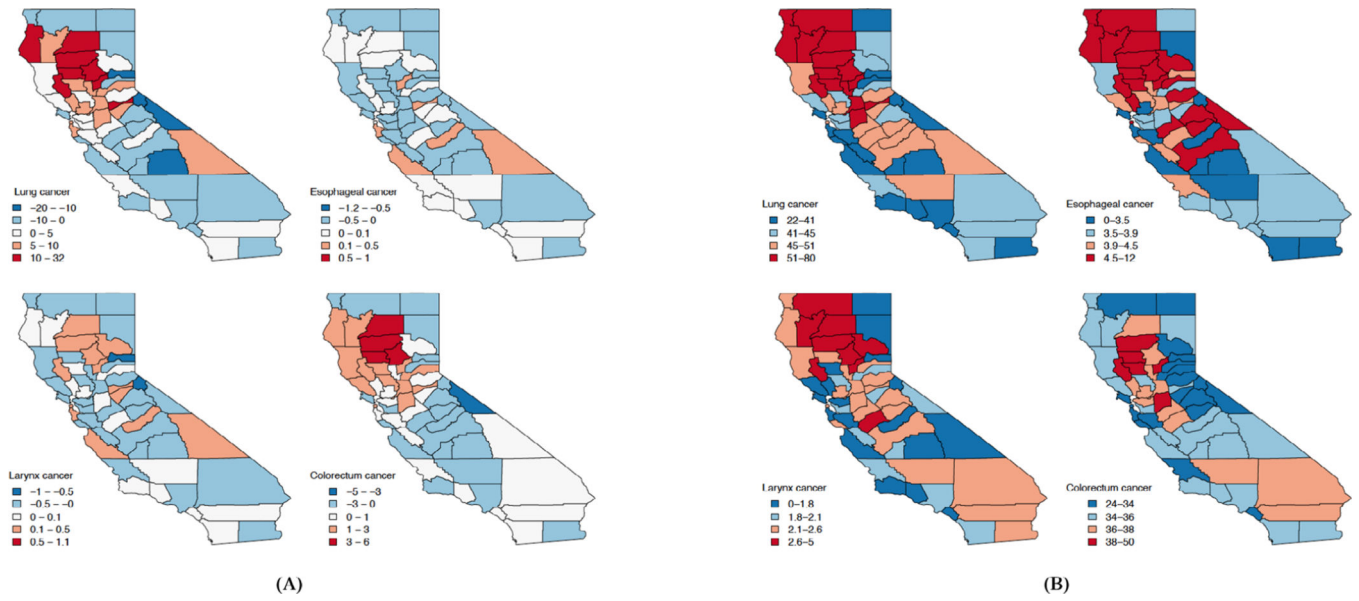


FIGURE 10. Maps of posterior results using the highest probability model M_{10} for lung, esophagus, larynx, and colorectal cancer in California including (A) posterior mean spatial random effects and (B) posterior mean incidence rates

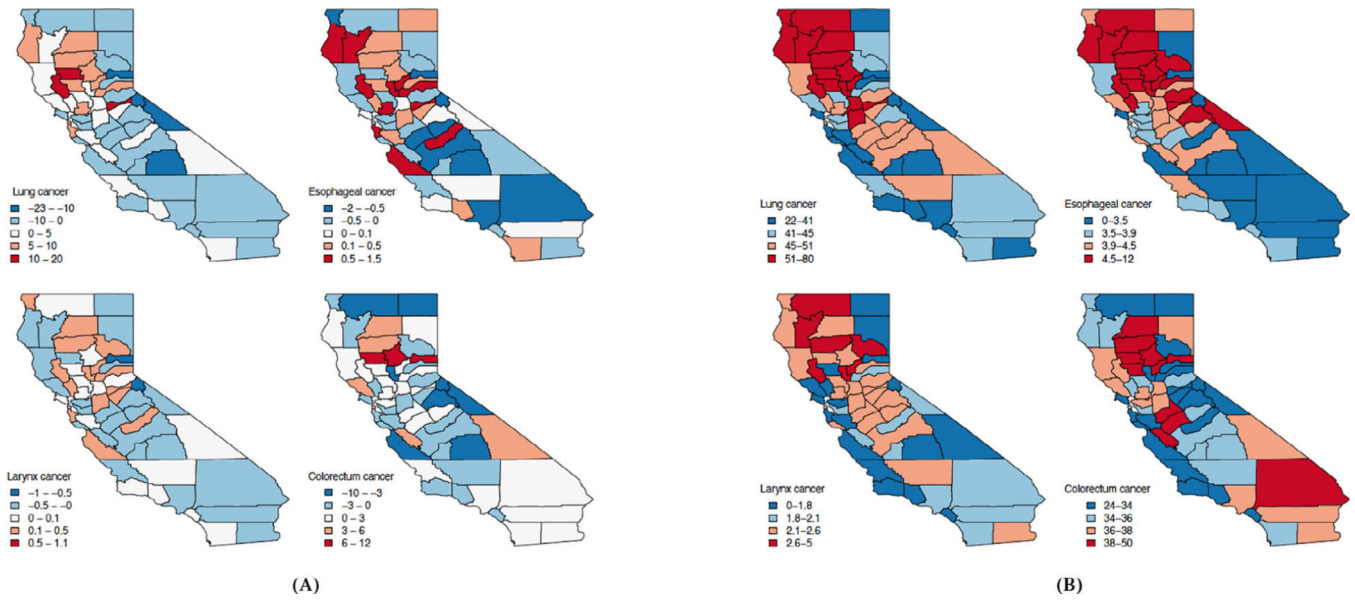


FIGURE 11. Maps of posterior results (Case 1) using MCAR for lung, esophagus, larynx, and colorectal cancer in California including (A) posterior mean spatial random effects and (B) posterior mean incidence rates

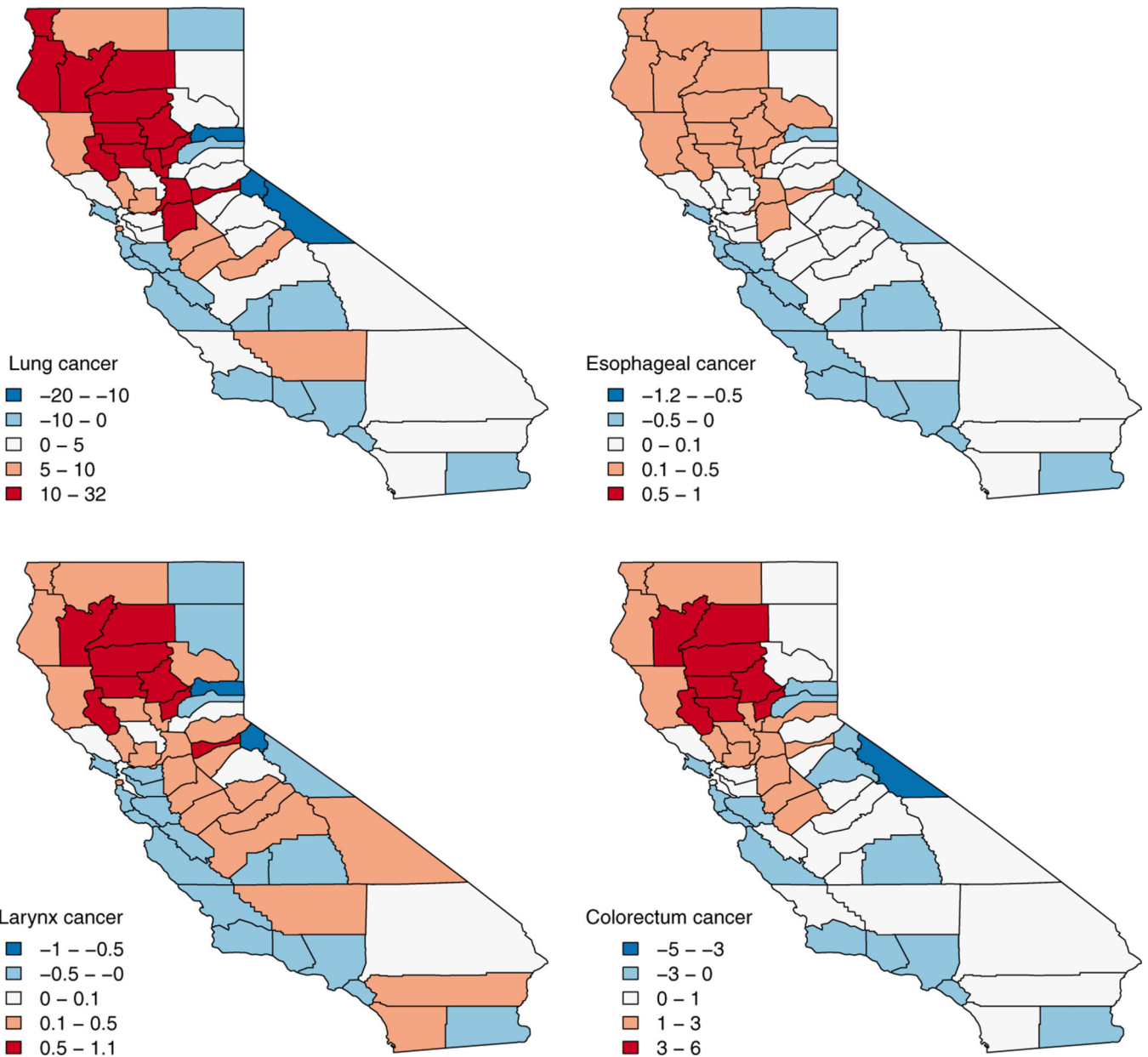


FIGURE 12. Maps of posterior mean spatial random effects (with no covariates) using the same order as M_{10}

TABLE 1

Proportion of times ($\pi(M_i)$) bridge sampling chose the model with the correct order out of the 50 datasets with that order

True model	$\pi(M_1)$	$\pi(M_2)$	$\pi(M_3)$	$\pi(M_4)$	$\pi(M_5)$	$\pi(M_6)$
M_1	0.90	0.00	0.10	0.00	0.00	0.00
M_2	0.00	0.86	0.00	0.00	0.14	0.00
M_3	0.14	0.00	0.86	0.00	0.00	0.00
M_4	0.00	0.00	0.00	0.90	0.00	0.10
M_5	0.00	0.22	0.00	0.00	0.78	0.00
M_6	0.00	0.00	0.00	0.16	0.00	0.84

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 2

The posterior model probabilities for 24 models

$p(\mathbf{M}_1 \mathbf{y})$	$p(\mathbf{M}_2 \mathbf{y})$	$p(\mathbf{M}_3 \mathbf{y})$	$p(\mathbf{M}_4 \mathbf{y})$	$p(\mathbf{M}_5 \mathbf{y})$	$p(\mathbf{M}_6 \mathbf{y})$	$p(\mathbf{M}_7 \mathbf{y})$	$p(\mathbf{M}_8 \mathbf{y})$
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$p(\mathbf{M}_9 \mathbf{y})$	$p(\mathbf{M}_{10} \mathbf{y})$	$p(\mathbf{M}_{11} \mathbf{y})$	$p(\mathbf{M}_{12} \mathbf{y})$	$p(\mathbf{M}_{13} \mathbf{y})$	$p(\mathbf{M}_{14} \mathbf{y})$	$p(\mathbf{M}_{15} \mathbf{y})$	$p(\mathbf{M}_{16} \mathbf{y})$
0.000	0.577	0.000	0.000	0.000	0.000	0.342	0.079
$p(\mathbf{M}_{17} \mathbf{y})$	$p(\mathbf{M}_{18} \mathbf{y})$	$p(\mathbf{M}_{19} \mathbf{y})$	$p(\mathbf{M}_{20} \mathbf{y})$	$p(\mathbf{M}_{21} \mathbf{y})$	$p(\mathbf{M}_{22} \mathbf{y})$	$p(\mathbf{M}_{23} \mathbf{y})$	$p(\mathbf{M}_{24} \mathbf{y})$
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002

Note: Bold values signify that the 95% credible intervals exclude 0.

TABLE 3

Posterior means (95% credible intervals) for parameters estimated from M_{10} and BMA estimates for regression coefficients only for the SEER four cancer dataset

Parameters	Model	Esophageal	Larynx	Colorectal	Lung
Intercept	M_{10}	16.76 (4.06, 29.56)	6.37 (-1.16, 13.89)	19.16 (-11.94, 49.72)	28.68 (-18.3, 74.93)
	BMA	15.87 (2.92, 28.63)	6.85 (-0.71, 14.38)	18.21 (-14.03, 49.07)	28.25 (-18.12, 74.52)
Smokers(%)	M_{10}	0.25 (0.12, 0.37)	0.04 (-0.03, 0.12)	0.23 (-0.12, 0.57)	0.81 (0.08, 1.62)
	BMA	0.23 (0.10, 0.36)	0.05 (-0.03, 0.12)	0.22 (-0.13, 0.58)	0.80 (0.08, 1.59)
Young(%)	M_{10}	-0.12 (-0.31, 0.07)	-0.07 (-0.18, 0.04)	0.27 (-0.2, 0.76)	-0.08 (-0.90, 0.74)
	BMA	-0.11 (-0.3, 0.08)	-0.08 (-0.19, 0.03)	0.29 (-0.18, 0.78)	-0.01 (-0.86, 0.82)
Old (%)	M_{10}	-0.11 (-0.25, 0.04)	-0.05 (-0.14, 0.03)	0.10 (-0.28, 0.48)	-0.09 (-0.81, 0.67)
	BMA	-0.10 (-0.25, 0.05)	-0.05 (-0.14, 0.03)	0.10 (-0.29, 0.49)	-0.08 (-0.79, 0.66)
Edu (%)	M_{10}	0.02 (-0.08, 0.12)	-0.02 (-0.08, 0.04)	0.16 (-0.12, 0.43)	-0.20 (-0.75, 0.31)
	BMA	0.02 (-0.09, 0.12)	-0.02 (-0.07, 0.04)	0.15 (-0.14, 0.42)	-0.24 (-0.79, 0.27)
Unemp (%)	M_{10}	-0.13 (-0.29, 0.03)	0.01 (-0.08, 0.10)	-0.09 (-0.54, 0.37)	0.60 (-0.47, 1.55)
	BMA	-0.12 (-0.28, 0.05)	0.01 (-0.08, 0.1)	-0.08 (-0.54, 0.38)	0.61 (-0.43, 1.56)
Black (%)	M_{10}	0.14 (-0.06, 0.34)	0.14 (0.03, 0.26)	-0.16 (-0.73, 0.39)	0.15 (-1.06, 1.29)
	BMA	0.13 (-0.07, 0.33)	0.15 (0.03, 0.27)	-0.18 (-0.75, 0.39)	0.14 (-1.02, 1.25)
Male (%)	M_{10}	-0.04 (-0.17, 0.09)	0.00 (-0.07, 0.08)	0.24 (-0.12, 0.60)	0.14 (-0.57, 0.79)
	BMA	-0.04 (-0.17, 0.09)	0 (-0.07, 0.08)	0.24 (-0.12, 0.62)	0.14 (-0.55, 0.82)
Uninsured (%)	M_{10}	-0.24 (-0.44, -0.04)	-0.08 (-0.20, 0.04)	0.07 (-0.44, 0.58)	0.01 (-0.82, 0.86)
	BMA	-0.23 (-0.42, -0.02)	-0.08 (-0.2, 0.04)	0.09 (-0.42, 0.61)	0 (-0.81, 0.82)
Poverty (%)	M_{10}	0.30 (-0.24, 0.84)	0.20 (-0.12, 0.51)	-0.06 (-1.51, 1.45)	0.85 (-2.15, 3.85)
	BMA	0.32 (-0.23, 0.87)	0.2 (-0.12, 0.51)	-0.08 (-1.54, 1.42)	0.8 (-2.14, 3.75)
ρ_{cancer}	M_{10}	0.25 (0.01, 1.00)	0.33 (0.01, 0.96)	0.50 (0.03, 0.97)	0.52 (0.03, 0.99)
τ_{cancer}	M_{10}	25.27 (5.08, 61.57)	27.60 (8.05, 60.42)	19.97 (3.06, 55.61)	20.31 (1.77, 55.92)
σ_{cancer}^2	M_{10}	1.67 (1.11, 2.47)	0.49 (0.28, 0.75)	8.22 (1.09, 14.23)	1.19 (0.18, 5.21)

Note: Bold values signify that the 95% credible intervals exclude 0.