Extended Excess Hazard Models for Spatially Dependent Survival Data

Journal Title
XX(X):3–44
c The Author(s) 2016

Reprints and permission: sagepub.co.uk/journalsPermissions.nav DOI: 10.1177/ToBeAssigned

www.sagepub.com/



ndré Victor Ribeiro maral ¹ and Francisco Javier Rubio ² and Manuela Quaresma ³ and Francisco J. Rodríguez-Cortés ⁴ and Paula Moraga ¹

bstract

Relative survival represents the preferred framework for the analysis of population cancer survival data. The aim is to model the survival probability associated to cancer in the absence of information about the cause of death. Recent data linkage developments have allowed for incorporating the place of residence into the population cancer data bases; however, modeling this spatial information has received little attention in the relative survival setting. We propose a flexible parametric class of spatial excess hazard models (along with inference tools), named "Relative Survival Spatial General Hazard" (RS-SGH), that allows for the inclusion of fixed and spatial effects in both timelevel and hazard-level components. We illustrate the performance of the proposed model using an extensive simulation study, and provide guidelines about the interplay of sample size, censoring, and model misspecification. We present a case study using real data from colon cancer patients in England. This case study illustrates how a spatial model can be used to identify geographical areas with low cancer survival, as well as how to summarize such a model through marginal survival quantities and spatial effects.

Keywords

Censored data; Excess Hazard; Net Survival; Relative Survival; Spatial frailty models.

1 Introduction

Survival analysis represents one of the main branches in Statistics, which concerns the study of times-to-event, potentially subject to censoring. The main quantity of interest in survival analysis is the probability of survival beyond a specific time point, associated with either the entire population under study or for subgroups of such a population. The most relevant approaches for analyzing survival data are: (i) the overall survival framework, which aims at analyzing all-cause mortality; (ii) the cause-specific survival framework, which incorporates information about the cause of death; and (iii) the relative survival framework, which aims at quantifying the survival associated to a cause of death of interest (such as cancer) in the absence of information about the cause of death. In the context of cancer epidemiology, national and international health agencies are interested in monitoring the survival probability of cancer patients at the population level (lemani et al. 2015). The preferred approach for population-based cancer survival analysis is the "relative survival" framework (Estève et al. 1990; Perme et al. 2012).

The relative survival approach aims at estimating the survival (or hazard) function associated to cancer, in the absence of reliable information about the cause of death for the whole population (since information about the cause of death is typically unreliable at the population level). The main idea is to assume an additive decomposition of the hazard function $h(\cdot)$ into two components, namely the hazard

Corresponding author:

André Victor Ribeiro Amaral. CEMSE Division, King Abdullah University of Science and Technology. Thuwal, Saudi Arabia.

Email: andre.ribeiroamaral@kaust.edu.sa

¹CEMSE Division, King Abdullah University of Science and Technology. Thuwal, Saudi Arabia.

²Department of Statistical Science, University College London, UK.

³Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine. London, UK.

⁴ Escuela de Estadística, Universidad Nacional de Colombia. Medellín, Colombia.

 associated to other causes of death $h_{\rm O}(\cdot)$, and the hazard associated to cancer $h_{\rm E}(\cdot)$. The latter is typically referred to as the "excess hazard." That is,

$$h(t; \mathbf{x}) = h_{\mathcal{O}}(\text{age} + t) + h_{\mathcal{E}}(t; \mathbf{x}), \ t \ge 0, \tag{1}$$

where t is the time measured from the date of diagnosis, "age" is age at diagnosis of cancer, and $\mathbf{x} \in \mathbb{R}^p$ is the vector of available covariates. Since $h_0(age + t)$ is unknown in practice, it is usually estimated using the population hazard $h_{\rm P}(\text{age} + t; \mathbf{z})$, which is obtained from the life tables based on the characteristics $\mathbf{z} \in \mathbb{R}^q \subset \mathbf{x}$. Depending on the country, the available life tables may be stratified by age, calendar year, sex, deprivation level, et cetera. Several excess hazard models have been proposed using both parametric and nonparametric estimation approaches (see Eletti et al. (2022) for a recent review). The main quantity of interest in the relative survival framework is the "net survival," which is the survival associated to the excess hazard $S_{\rm N}(t;\mathbf{x})=\exp\Big\{-\int_0^t h_{\rm E}(r;\mathbf{x})dr\Big\}$. The net survival only depends on the excess hazard function. Thus, it is a useful quantity for comparing the performance of cancer management between different populations since it is not affected by differences in population mortality hazards. For that reason, comparisons between different countries, regions, or periods of time are based on the marginal net survival

$$S_{\mathbf{N}}(t) = \frac{1}{m} \sum_{i=1}^{m} S_{\mathbf{N}}(t; \mathbf{x}_i),$$

where $\{\mathbf{x}_i\}_{i=1}^m$ represents the covariates associated to the (sub-)population of interest, such that m denotes the population size.

The utilization of spatial information regarding the residence of cancerdiagnosed patients may enable the identification of regional inequalities

45

46 47 48

49

50 51

in cancer survival (Public Health England 2020; Exarchakou et al. 2018; Quaresma et al. 2022). Furthermore, such information may facilitate the identification of areas with low cancer survival, which can be be used to inform policymaking aiming at improving cancer survival. Indeed, cancer registry data may present a spatially dependent structure, as individuals from adjacent neighborhoods are likely to share environmental and socioeconomical factors (Li and Ryan 2002). In that case, the individuals' locations would act like a proxy for non-observed regional characteristics (Zhou and Hanson 2018).

Spatial survival modeling has received a great deal of attention in the overall survival framework. The main idea is usually to incorporate a spatial term into a survival regression model (see Klein et al. (2014) for a book-length review on classical survival models). For instance, Li and Ryan (2002) propose adding a spatial frailty, modeled as a zero-mean Gaussian process (GP), into a semiparametric Proportional Hazard (PH) model. Banerjee et al. (2003b) fit a PH model with spatially dependent random effects for geostatistical and lattice data. Carlin and Banerjee (2003) propose a Bayesian semiparametric survival model for spatio-temporal correlated data based on including generalized multivariate conditionally autoregressive (MC R) region-specific frailties into a hazard regression model. Li et al. (2015) propose a normal transformation model of the General Hazard (GH) model (Chen and Jewell 2001, also known as Extended Hazard (EH) model). The spatial variability is modeled through the covariance matrix of the normal transformation. Zhou and Hanson (2018) propose a framework for fitting PH, Proportional Odds (PO), and accelerated failure time (FT) models, accounting for different types of censoring, including random effects with intrinsic conditionally autoregressive (IC R) priors. Basak et al. (2022) propose a semiprametric model for clustered interval-censored survival data. In that case, the hazard function is written as a product of the baseline

40

41

42

43 44

45

46 47 48

49

50 51

hazard component and a non-parametric component modeled as a soft Bayesian additive regression tree (sB RT) that is used to incorporate the possible clustering effects. Rubio and Drikvandi (2022) consider the GH structure with random effects at the cluster level; however, they do not account for the spatial structure and limit their proposal to modeling the clustering components.

In contrast, spatial survival models in the relative survival framework have received less attention. For instance, Charvat et al. (2016) propose a parametric frailty model for the excess hazard function using different types of splines or parametric baseline hazards. However, the frailties are assumed to be independent, thus only accounting for clustering but not for the spatial dependence. Cramb et al. (2016) propose a Bayesian spatial frailty approach based on modeling the cumulative excess hazard using splines, thus requiring a different interpretation for the estimated effects. Their proposal does not include time-scale effects, and the frailties are modeled using an IC R normal distribution. This method was later applied in Cramb et al. (2017). Finally, Eletti et al. (2022) propose a link-based additive excess hazard model that allows for the inclusion of non-linear effects, temporal-dependent effects, and spatial effects via Markov random fields. In a slightly different vein, Yu and Tiwari (2012) studied cure mixture models in the relative survival framework. They adopted a mixture of three accelerated failure time models for the excess hazard, with spatial frailties modeled using multivariate conditionally autoregressive (MC R) distributions.

In this paper, we introduce a general class of parametric spatial frailty models for survival data under the relative survival framework. The basic idea consists of adding spatial effects, at two levels (time and hazard), into the General Hazard (GH) model (Etezadi- moli and Ciampi 1987; Chen and Jewell 2001), which is a hazard structure that generalizes the PH model, the accelerated failure time (FT) model (Buckley and James

 1979), the accelerated hazard (H) model (Chen and Wang 2000), and others, as we will describe in Section 2.

To do so, we extend the existing approaches by modeling the dependence structure through spatial smoothing methods, namely utoregressive (IC R) and Besag-York-Mollié Intrinsic Conditional (BYM2) model priors. It also allows for incorporating fixed and spatial effects at the time-scale and at the hazard scale without requiring numerical integration. By taking such an approach, we can easily compute credible intervals as a measure of uncertainty, and further investigate other quantities of interest. For instance, similar to the discussion in Moraga (2019), we can compute the *relative* exceedance probabilities, which are useful for assessing unusual elevation in any function of the linear predictor terms, such as the excess hazard, net survival, among others. The term "relative" is important, since we are extending the concept of exceedance probabilities to the relative survival framework. This quantity also helps detecting high-risk areas based on the analysis of the spatial random effects, as the possibly non-observed spatial heterogeneity is captured by such components. The R (R Development Core Team 2022) and ST N (Carpenter et al. 2017) scripts containing the implementation of the examples presented here, as well as additional examples using real data are available at https://github.com/avramaral/relative survival.

The remainder of this paper is organized as follows. Section 2 introduces notation and presents the proposed modeling approach. We also discuss some particular sub-models of interest. We introduce two spatial smoothing methods that account for non-observed spatial characteristics and list all implemented models. In Section 3, we detail the inference procedure and present a brief discussion on the prior distributions specification for our class of models. In Section 4, we provide a simulation study that illustrate the performance of our model under

 different scenarios, and present guidelines about the interplay of sample size, censoring, and model misspecification. Section 5 presents a case study that analyzes the variation of colon cancer survival for different geographic regions in England. Finally, in Section 6, we present a general discussion, and comment on the limitations and possible extensions of our work.

2 Spatial models

In this section, we introduce the proposed general model structure, and discuss the particular models that can be derived from it. lso, we discuss different spatial smoothing methods that can be used with our approach and list all possible modeling scenarios.

2.1 Excess hazard model

Let us first introduce some notation. Let $o_{ij} \in \mathbb{R}_+$ be a sample of timesto-event, where $i=1,\ldots,r$ indicates the region and $j=1,\ldots,n_i$ denotes the individuals. Iso, let $c_{ij} \in \mathbb{R}_+$ be the right-censoring times, and $t_{ij} = \min\{o_{ij},c_{ij}\}$ be the observed survival times. Let $\delta_{ij}=(o_{ij}< c_{ij})$ be the vital status indicators (that is, $\delta_{ij}=1$, if dead, and $\delta_{ij}=0$, if right-censored or alive), and $n=\sum_{i=1}^r n_i$ be the total sample size across the r regions. Let $\mathbf{x}_{ij} \in \mathbb{R}^p$ be the vector of available covariates. Similar to the mixed effects survival regression model (for overall survival) proposed in Rubio and Drikvandi (2022), we consider the excess hazard model

$$h_{E}(t; \mathbf{x}_{ij} \mid \boldsymbol{\theta}, , \boldsymbol{\beta}, \boldsymbol{\gamma}, \tilde{u}_{i}, u_{i}) = h_{0}(t \exp{\{\tilde{\mathbf{x}}_{ij}^{\top} + \tilde{u}_{i}\}} \mid \boldsymbol{\theta}) \exp{\{\mathbf{s}_{ij}^{\top} \boldsymbol{\gamma} + \mathbf{x}_{ij}^{\top} \boldsymbol{\beta} + u_{i}\}},$$
(2)

where $h_0(\cdot \mid \boldsymbol{\theta})$ is the baseline excess hazard function, defined through a flexible parametric distribution, $\boldsymbol{\theta}$ represents the corresponding distribution parameters, \mathbf{x}_{ij} play the role of hazard-level effects, $\tilde{\mathbf{x}}_{ij} \subseteq \mathbf{x}_{ij}$ represent the time-level effects, where $\tilde{\mathbf{x}}_{ij} \in \mathbb{R}^{\tilde{p}}$, and $= \begin{pmatrix} 1, \dots, \tilde{p} \end{pmatrix}^{\top}$

and $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^{\top}$ are the regression coefficients associated to $\tilde{\mathbf{x}}_{ij}$ and \mathbf{x}_{ij} , respectively. dditionally, $\mathbf{s}_{ij} = (\mathbf{s}_{ij1}^{\top}, \dots, \mathbf{s}_{ijk}^{\top})^{\top} \in \mathbb{R}^q$ and $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_q)^{\top}$, where $q = \sum_{l=1}^k q_l$, such that q_l is the dimension of s_{ijl} , and s_{ijl} is the spline expansion of a (continuous) covariate x_{ijl} . Lastly, we assume that $\tilde{\mathbf{u}}$ and \mathbf{u} are independent, with $\tilde{\mathbf{u}} = (\tilde{u}_1, \dots, \tilde{u}_r)^\top \sim \tilde{G}$ and $\mathbf{u} = (u_1, \dots, u_r)^{\top} \sim G$, such that \tilde{G} and G are multivariate distributions that account the spatial dependence among regions. The spatial models used to define \tilde{G} and G will be introduced in Section 2.2. Thus, our proposal can be seen as an extension of the MEGH model proposed in Rubio and Drikvandi (2022) to the relative survival framework, but also with the incorporation of spatial effects.

We will denote Model (2) as the RS-SGH (Relative Survival Spatial General Hazard) model, and we will also consider eight particular submodels that might be useful for researchers and practitioners when fitting this class of models. These alternative modeling approaches are described in Table 3 (ppendix).

Let $\boldsymbol{\xi} = (\boldsymbol{\theta}^{\top}, \quad {}^{\top}, \boldsymbol{\beta}^{\top}, \boldsymbol{\gamma}^{\top})^{\top}$, then the cumulative hazard function $H(\cdot \mid$ $\mathbf{x}_{ij}, \boldsymbol{\xi}, \tilde{u}_i, u_i$) associated with Model (1) can be written in the following manner

$$H(t; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_i, u_i) = \int_0^t h(\zeta; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_i, u_i) d\zeta$$

$$= H_{\mathsf{P}}(\mathsf{age}_{ij} + t; \mathbf{z}_{ij}) \quad H_{\mathsf{P}}(\mathsf{age}_{ij}; \mathbf{z}_{ij}) + H_{\mathsf{E}}(t; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_i, u_i),$$

where $H_{\mathrm{P}}(\cdot;\mathbf{z}_{ij})$ is the cumulative population hazard, and $H_{\mathrm{E}}(\cdot;\mathbf{x}_{ij}\mid$ $\boldsymbol{\xi}, \tilde{u}_i, u_i)$ is the cumulative excess hazard function. Moreover, the cumulative excess hazard function can be written in closed-form as

$$H_{E}(t; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_{i}, u_{i}) = \int_{0}^{t} h_{E}(\zeta; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_{i}, u_{i}) d\zeta$$

$$= H_{0}(t \exp{\{\tilde{\mathbf{x}}_{ij}^{\top} + \tilde{u}_{i}\} \mid \boldsymbol{\theta}\}) \exp{\{\mathbf{x}_{ij}^{\top} \boldsymbol{\beta} \quad \tilde{\mathbf{x}}_{ij}^{\top} + u_{i} \quad \tilde{u}_{i}\}},$$

where $H_0(\cdot \mid \boldsymbol{\theta})$ is the cumulative baseline excess hazard.

We can now adapt the concept of individual net survival based on the proposed spatial excess hazard model. The net survival, for a specific covariate and conditional on model parameters, and random effects, can be defined as

$$S_{N}(t; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_{i}, u_{i}) = \exp\{ H_{E}(t; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_{i}, u_{i}) \}.$$
(3)

Consequently, the region-specific net survival associated to the *i*-th region is defined as follows

$$S_{\mathbf{N},i}(t \mid \boldsymbol{\xi}) = \frac{1}{n_i} \sum_{j=1}^{n_i} \int_{\mathbb{R}^2} S_{\mathbf{N}}(t; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_i, u_i) d\tilde{G}(u_i) dG(u_i). \tag{4}$$

Let us now discuss some specific choices for modeling the parametric baseline hazard function $h_0(\cdot \mid \boldsymbol{\theta})$. Since the Weibull baseline hazard is the only choice that leads to a non-identifiable model (Chen and Jewell 2001), we will adopt distributions that do not belong to the Weibull family. We will focus on 2-parameter and 3-parameter distribution that can account for a variety of shapes. These include the Log-normal (LN), Loglogistic (LL), Power Generalized Weibull (PGW), Gamma (G M), and Generalized Gamma (GG) distributions. In Web ppendix 1 (Supporting Information), we specify all possible distributions for such a baseline component.

Lastly, notice that, by setting $h_P(age_{ij} + t; \mathbf{z}_{ij}) = 0$, for all individuals in all regions, we shift to the overall survival framework. Therefore, the RS-SGH model generalizes several well-known modeling approaches in different directions and under different frameworks.

Spatial effects 2.2

We aim at incorporating spatial effects in the excess hazard Model (2) by incorporating the neighborhood structure into the distribution of the

 random effects $\tilde{\mathbf{u}} = (\tilde{u}_1, \dots, \tilde{u}_r)^{\top}$ and $\mathbf{u} = (u_1, \dots, u_r)^{\top}$. To this end, we will define them based on two approaches: the Intrinsic Conditional utoregressive (IC R) and Besag-York-Mollié (BYM2) models. To formulate these models, we need to introduce the concept of adjacency matrix. Briefly, given two regions k and l, we will say that k and l are neighbors (written $k \sim l$, with $k \neq l$) if those regions share any boundary. Notice that if $k \sim l$, then $l \sim k$. However, a region will not be its own neighbor. Based on this "neighbor operator" (\sim), we can define an adjacency matrix A, such that $a_{kl} = 1$, if $k \sim l$, and $a_{kl} = 0$, otherwise. The diagonal of A is defined as zero, that is $\operatorname{diag}(\mathbf{A}) = -\mathbf{s}$ s a remark, Freni-Sterrantino et al. (2018) present guidelines on how to adapt these models if the corresponding spatial graph is disconnected. dditionally, Morris et al. (2019) (Section 3.5) comment on how to implement these extensions using ST N.

2.2.1 Intrinsic Conditional Autoregressive ICAR) model For the Intrinsic Conditional utoregressive (IC R) model, the conditional distribution of u_k given all other random effects u_l , such that $l \neq k$ (written \mathbf{u}_k), is

$$\pi(u_k \mid \mathbf{u}_k) = \text{Normal}\left(\frac{\sum\limits_{s \in k} u_s}{\lambda_k}, \frac{1}{\lambda_k \tau_u}\right),$$

where $_k$ and λ_k correspond to the neighbors and the number of neighbors of region k, respectively, and τ_u is the precision term. Besag (1974) proved that the corresponding joint specification of \mathbf{u} follows a multivariate normal distribution with mean—and precision matrix $\mathbf{Q} = \tau_u(\mathbf{D} - \mathbf{A})$, where \mathbf{D} is a $(r \times r)$ diagonal matrix with d_{kk} containing the number of neighbors of k, and $d_{kl} = 0$, $\forall k \neq l$. Moreover, as shown in Besag et al. (1991, 1995), the joint distribution of \mathbf{u} as specified above can be further

 simplified to the following pairwise difference

$$\pi(\mathbf{u}) \propto \tau_u^{\frac{r}{2}} \exp \left\{ -\frac{\tau_u}{2} \sum_{k \sim l} (u_k - u_l)^2 \right\}.$$
 (5)

However, from Equation (5), one can notice that the joint distribution of \mathbf{u} is non-identifiable (adding any constant to all elements of \mathbf{u} does not change the joint distribution). To overcome this issue, it suffices to impose the constraint $\sum_{k=1}^r u_k = 0$. From a practical point of view, and under the Bayesian framework, the approximate condition $\sum_{k=1}^r u_k \approx 0$ is implemented instead, using a "soft sum-to-zero constraint". That is, when implementing the model, we assign a zero-mean prior distribution to the mean of \mathbf{u} with very small variance. Such an approach is recommended by Morris et al. (2019), as the ST N's Hamiltonian Monte Carlo sampler runs faster under this setting. Finally, the same modeling procedure will be adopted for $\tilde{\mathbf{u}}$.

2.2.2 Besag-York-Mollié BYM2) model Iternatively, unstructured (or non-spatial) random effects could be added, along with the structured IC R components, to the excess hazard Model (2). This approach is known as a Besag-York-Mollié (BYM)-type model (Besag et al. 1991). However, as commented in Mahmood et al. (2022), such a parameterization might present some shortcomings. For instance, a model expressed based on such a convolution of the structured and unstructured random effects may fail to fit, as one of the two components can account for almost all observed variance (Morris et al. 2019). Iso, it might be difficult to set reasonable priors for the corresponding scale parameters (Banerjee et al. 2003a). iming to avoid these issues, instead, the BYM2 model is often used (Riebler et al. 2016).

To formulate the BYM2 model, the unstructured and structured random effects $(\mathbf{v} = (v_1, \dots, v_r)^{\top})$ and $\mathbf{s} = (s_1, \dots, s_r)^{\top}$, respectively), can be

written as

$$\mathbf{u} = \mathbf{v} + \mathbf{s} = \sigma(\sqrt{1 - \rho} \mathbf{v}^* + \sqrt{\rho} \mathbf{s}^*),$$

where σ is the overall standard deviation, $\rho \in [0,1]$ determines the proportion of the variance that comes from the structured random effects, $\mathbf{v}^{\star} \sim \text{Normal}(\ ,\ _r)$, such that $\ _r$ is an $(r \times r)$ identity matrix, and \mathbf{s}^{\star} is the scaled IC R model (Sørbye and Rue 2017), such that $\text{Var}(s_i) \approx 1, \forall i$. s before, similar reasoning is applied to define $\tilde{\mathbf{u}}$ in terms $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{s}}$.

2.2.3 IID model One last alternative would be defining $\tilde{\mathbf{u}}$ and \mathbf{u} purely based on an "independent and identically distributed" (i.i.d.) model; that is, $\mathbf{u} \sim \text{Normal}(\ , \sigma_u^2 \ _r)$ and $\tilde{\mathbf{u}} \sim \text{Normal}(\ , \sigma_{\tilde{u}}^2 \ _r)$. This would be the same including a clustering effect per region. Under the overall survival framework, this idea has been explored by Rubio and Drikvandi (2022) using likelihood inference, and we will also implement such a model in a Bayesian setting. Il implemented models for the possible baseline hazard distributions, spatial random effects, and overall model structure are detailed in Table 4 in ppendix B.

2.2.4 Point data model Ithough our focus is on employing areal data to model spatial dependence, one might also be interested in using latitude-longitude coordinates (if available) to determine a patient's location. In such cases, the spatial structure may be accounted for by a model for point data. For instance, "penalized spline regression" (Fahrmeir et al. 2013) is a popular method for spatial smoothing. Iternatively, Diggle and Ribeiro (2007) proposed a geostatistical framework to model the spatial correlation structure in the point data while enabling rigorous statistical inference. For the latter, additional assumptions about the sampling scheme can be made, e.g., "preferential sampling" (Diggle et al. 2010)—with a non-stationary extension proposed

 by maral et al. (2023). Throughout this paper, we will employ the methods described in Sections 2.2.1–2.2.3.

3 Inference

In this section, we introduce the inference procedure used for fitting Model (1) with excess hazard given by Model (2). lso, we present some guidelines for setting the prior distributions, and define a model selection measure.

3.1 Likelihood function

Let $\mathcal{D} = \{(t_{ij}, \delta_{ij}, \mathbf{x}_{ij}, \mathbf{z}_{ij}); i = 1, ..., r, \text{ and } j = 1, ..., n_i\}$ be the observed data. In that case, the likelihood function for the vector of unknown parameters can be written as proportional to

$$\prod_{i=1}^{r} \prod_{j=1}^{n_i} h_{\mathsf{P}}(\mathsf{age}_{ij} + t_{ij}; \mathbf{z}_{ij}) + h_{\mathsf{E}}(t_{ij}; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_i, u_i) \right\}^{\delta_{ij}} \times$$

$$\exp \left\{ H_{\mathsf{E}}(t_{ij}; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_i, u_i) \right\},$$
(6)

where $h_P(age_{ij} + t_{ij}; \mathbf{z}_{ij})$ is obtained from the life tables. From Equation (6), notice that the only term in the likelihood function that distinguishes an overall survival model from a relative survival model is $h_P(age_{ij} + t_{ij}; \mathbf{z}_{ij})$, therefore, by setting it to zero, we could also retrieve the overall survival framework.

Nevertheless, as proved by Chen and Jewell (2001), the General Hazard model (and thus, the RS-SHG model, as it extends the GH approach) is non-identifiable if the baseline hazard in $h_{\rm E}(t_{ij}; {\bf x}_{ij} \mid {\boldsymbol \xi}, \tilde{u}_i, u_i)$ is Weibull. However, this scenario is not of concern since, if the true model is Weibull, it means that a simpler model would fit the data well—see Rubio et al. (2019) for further details. Furthermore, our capability to simultaneously recover the two spatial structures in Equation (6) is noteworthy.

 demonstrated in Sections 4.3 and 5, we can estimate $\sigma_u = 1/\sqrt{\tau_u}$ (and $\sigma_{\tilde{u}} = 1/\sqrt{\tau_{\tilde{u}}}$) for all proposed models. To do so, in practice, we must have a certain number of *uncensored* observations per cluster (in addition to avoiding the Weibull distribution when defining the baseline hazard). In this case, there is an interplay between the number of individuals in each region and the censoring rate in these areas.

The next section presents our prior elicitation strategy. Inference is performed by sampling from the corresponding posterior distributions based on the RStan's implementation (ST N Development Team 2021) of the Hamiltonian Monte Carlo algorithm (Betancourt and Girolami 2015).

3.2 Prior distributions

Ithough we acknowledge that other choices can be made, in this section, we recommend some weakly informative priors for the model parameters. For the linear fixed effects, we set $_{\tilde{p}} \sim \text{Normal}(0, \sigma_{\tilde{p}}^2)$, $\forall \tilde{p}$, and $\beta_p \sim \text{Normal}(0, \sigma_{\beta_p}^2)$, such that $\sigma_{\tilde{p}}^2$ and $\sigma_{\beta_p}^2$ are large enough to reflect the vague prior information. On the other hand, for the nonlinear fixed effects, we adopted a novel choice of g-priors (Zellner 1986) that account for censoring; that is, letting \mathbf{S}_k be the design matrix associated with the spline basis expansion of the k-th covariate \mathbf{x}_k , and defining $\mathbf{M}_k = g \ (\mathbf{S}_k^{\top} \mathbf{S}_k)^{-1}$, we set $\boldsymbol{\gamma}_k \sim \text{MVN}(\ , \sigma^2 \ \mathbf{M}_k)$, where $g = (n \ (0.5 \times (n \ n_{\text{obs}})))/q$, n_{obs} corresponds to the number of uncensored observations, and $\sigma^2 \sim \text{Half-Cauchy}(0,\tau_\sigma)$, such that $\tau_\sigma > 0$. In that case, notice that g accounts, to some extent, for the effective number of observations—as the rescaled number of censored patients is subtracted from the total number of collected data points.

In our setting, the g-priors can be seen as a type of shrinkage prior, where the induced shrinkage is mild as we only include a few variables in the models. Iternative prior specifications could have been employed

 to induce higher levels of shrinkage in place of the selected g-priors. However, in the context of our problem, we do not aim to induce higher levels of shrinkage since there are only a few variables available at the population level, and all of these variables are typically relevant for cancer survival. s an alternative, in the Bayesian smoothing literature, it is also common to assign priors to the spline coefficients that enforce smoothness between adjacent spline coefficients (similarly to what the IC R model does in the spatial setting). These priors typically take the form of random walks or intrinsic Gaussian Markov random fields (Fahrmeir et al. 2013; Rue and Held 2005).

Regarding the spatial smoothing distributions, for the IC R model, we set $\tau_u \sim \text{Gamma}(\theta_{\tau_u}, \theta_{\tau_u})$ (same for $\tau_{\tilde{u}}$), such that $\theta_{\tau_u} > 0$ is a small lthough the Gamma distribution with such scale and shape parameters is commonly found in the literature—mainly due to The BUGS (Bayesian inference Using Gibbs Sampling) project implementation (Lunn et al. 2009), we again emphasize that it is possible to use other types of priors. For instance, Gelman (2006) suggests the usage of a distribution from the half-t family for the variance parameter in Iternatively, the penalized complexity (PC) priors hierarchical models. (Simpson et al. 2017) could also be explored in our setting. In Section 6, we briefly discuss possible extensions of our work concerning prior elicitation. For the BYM2 model, we set $\sigma \sim \text{Half-Normal}(0,1)$ and $\rho \sim$ Beta(0.5, 0.5) (both when defining u and $\tilde{\mathbf{u}}$), such that the latter is based on the recommendations given by Morris et al. (2019). Finally, for each of the baseline hazard distributions listed in Section 2.1, and based on the parameterization given in Web ppendix 1 (Supporting Information), we set the priors as follows

1. Log-normal: $\mu \sim \text{Normal}(0, \sigma_{\mu}^2)$, where σ_{μ}^2 is a large number, and $\sigma \sim \text{Half-Cauchy}(0, \tau_{\sigma})$, with $\sigma_{\mu}^2, \tau_{\sigma} > 0$ (Rubio and Steel 2018).

- 2. Log-logistic: as for the LN model, $\mu \sim \text{Normal}(0, \sigma_{\mu}^2)$, where σ_{μ}^2 is a large number, and $\sigma \sim \text{Half-Cauchy}(0, \tau_{\sigma})$, with $\sigma_{\mu}^2, \tau_{\sigma} > 0$.
- 3. Power Generalized Weibull: $\eta \sim \text{Half-Cauchy}(0, \tau_{\eta}), \quad \nu \sim \text{Half-Cauchy}(0, \tau_{\nu}), \quad \text{and} \quad \kappa \sim \text{Gamma}(0.65, \quad 1.83), \quad \text{with scale parameters } \tau_{\eta}, \tau_{\nu} > 0.$ The prior specification for κ has been proven to be weakly informative (Dette et al. 2018; lvares and Rubio 2021).
- 4. Gamma: $\eta \sim \text{Half-Cauchy}(0, \tau_{\eta})$ and $\nu \sim \text{Half-Cauchy}(0, \tau_{\nu})$, with scale parameters $\tau_{\eta}, \tau_{\nu} > 0$.
- 5. Generalized Gamma: as for the PGW model, $\eta \sim$ Half-Cauchy $(0, \tau_{\eta}), \nu \sim$ Half-Cauchy $(0, \tau_{\nu})$, and $\kappa \sim$ Gamma(0.65, 1.83), with scale parameters $\tau_{\eta}, \tau_{\nu} > 0$.

3.3 Model selection

To compare the fitted models, we will use a leave-one-out cross validation (LOO CV) procedure; that is, we will use the likelihood evaluated at the parameters' posterior samples as a goodness-of-fit measure. In particular, we will use the Pareto-smoothed importance sampling (PSIS) implementation (Vehtari et al. 2017) and compute the corresponding quantities using the 100 package (Vehtari et al. 2022). Under the Bayesian framework, the LOO estimate of out-of-sample predictive fit will be computed as

$$elpd_{LOO} = \sum_{i=1}^{r} \sum_{j=1}^{n_i} \log \left[\pi(t_{ij} \mid \mathbf{t}_{ij}) \right],$$

where $\pi(t_{ij} \mid \mathbf{t}_{ij})$ is the LOO predictive density given \mathbf{t}_{ij} , such that \mathbf{t}_{ij} corresponds to the vector of all observed time points, except t_{ij} . However, instead of re-fitting the model $n = \sum_{i=1}^{r} n_i$ times, $\pi(t_{ij} \mid \mathbf{t}_{ij})$ will be approximately computed, $\forall i, j$, using the PSIS technique. For the details, the reader can refer to Vehtari et al. (2015) and Vehtari et al. (2017).

 Throughout this paper, we will denote such an estimate as $elpd_{PSIS-LOO}$. s a final remark, assuming well-specified and -fitted models, when comparing different approaches, the larger $elpd_{PSIS-LOO}$, the better—as such a quantity sums over the posterior predictive model evaluated at a new observation t_{ij} , for each i and j.

4 Simulation study

In this simulation section, we will assess the performance of our RS-SGH model in three directions. First, we will evaluate our fitted models with respect to their ability to recover the true net survival function, as in Equation (3). For that case, we will analyze our models performance based on different sample sizes, different censoring rates, and misspecified distributions for the baseline hazard. Second, fixing all components but the random effect structures, we will compare and select models based on the elpd_{PSIS-LOO} criterion, as in Section 3.3. Third, we will use our fitted models to identify riskier areas based on the analysis of the spatial effects. To do so, we will simulate data from the RS-SGH model as described in Web ppendix 3 (Supporting Information).

4.1 Marginal quantities

Our first analysis concerns the estimation of marginal quantities, such as the net survival in Equation (3). From that equation, notice that we are integrating out the effects of the spatial components; therefore, provided the model is well fitted and for a sufficiently large sample size in all regions, all random effect structures are expected to produce similar results—as the random effects are assumed to be zero-mean for all models in Section 2.2. In that case, we will benchmark our fitted spatial model with respect to the true corresponding curves.

We will focus on analyzing the effect of (i) different sample sizes; (ii) different censoring rates; and (iii) misspecified distributions for the

baseline hazard. In this regard, we will set the sample size to 200, 500, 1000, and 2000 patients, the censoring rate to 25% and 50%, and we will simulate and fit Model (2) with the baseline hazard component defined by the Log-normal (LN) and Power Generalized Weibull (PGW) distributions. The simulation details are given in ppendix C. From that section, note that the considered spatial structure is defined based on the map of England, split into 9 regions (see Figure 7). This choice is not arbitrary, as it is based on genuine epidemiological questions about cancer survival and how the England territory is administrated. Furthermore, for our simulation study, the number of regions was set to 9 to ensure that we can repeatedly fit the model based on Equation (6) within "reasonable" computational time—the average fitting times are presented in Table ST2, Supporting Information. For comparison, the fitting times for Section 5, where a larger number of regions is considered (with also more patients), are presented in Table ST10. Lastly, recall that there must exist an interplay between the number of regions, censoring rate, and the sample size. s shown by Rubio and Drikvandi (2022), we must have a sufficiently large number of uncensored observations in each cluster to correctly estimate the model parameters. s a final remark, all simulation and fitting scenarios are listed in Table 1.

Table 1. All simulated scenarios for Section 4.1. "Data Generating model" refers to the model assumed for the data generating procedure, and "Fitted model" is defined as per Table 4.

#	Data Generating model	Censoring rate	Sample size	Fitted model	#	Data Generating model	Censoring rate	Sample size	Fitted model
01	RS-SGH LN IC R	25%	200	RS-SGH LN IC R	13	RS-SGH PGW IC R	25%	200	RS-SGH LN IC R
02	RS-SGH LN IC R	25%	500	RS-SGH LN IC R	14	RS-SGH PGW IC R	25%	500	RS-SGH LN IC R
03	RS-SGH LN IC R	25%	1000	RS-SGH LN IC R	15	RS-SGH PGW IC R	25%	1000	RS-SGH LN IC R
04	RS-SGH LN IC R	25%	2000	RS-SGH LN IC R	16	RS-SGH PGW IC R	25%	2000	RS-SGH LN IC R
05	RS-SGH LN IC R	25%	200	RS-SGH PGW IC R	17	RS-SGH PGW IC R	25%	200	RS-SGH PGW IC R
06	RS-SGH LN IC R	25%	500	RS-SGH PGW IC R	18	RS-SGH PGW IC R	25%	500	RS-SGH PGW IC R
07	RS-SGH LN IC R	25%	1000	RS-SGH PGW IC R	19	RS-SGH PGW IC R	25%	1000	RS-SGH PGW IC R
08	RS-SGH LN IC R	25%	2000	RS-SGH PGW IC R	20	RS-SGH PGW IC R	25%	2000	RS-SGH PGW IC R
09	RS-SGH LN IC R	50%	200	RS-SGH LN IC R	21	RS-SGH PGW IC R	50%	200	RS-SGH LN IC R
10	RS-SGH LN IC R	50%	500	RS-SGH LN IC R	22	RS-SGH PGW IC R	50%	500	RS-SGH LN IC R
11	RS-SGH LN IC R	50%	1000	RS-SGH LN IC R	23	RS-SGH PGW IC R	50%	1000	RS-SGH LN IC R
12	RS-SGH LN IC R	50%	2000	RS-SGH LN IC R	24	RS-SGH PGW IC R	50%	500	RS-SGH LN IC R

From Table 1, notice that we are not fitting the PGW distribution model with 50% censoring rate. We did it in this way since we identified that, for 3-parameter distributions (e.g., Power Generalized Weibull), it

 might be difficult to obtain well-mixed posterior chains for models fitted based on highly censored data sets. When generating the data, we will simulate 100 data sets for each combination of sample size, censoring rate, and baseline hazard distribution. Next, for the fitting step, we will write Model (2) using the same covariates as the selected ones for the data generating scheme. Iso, for the MCMC-based (Markov chain Monte Carlo) code (implemented using RStan in the background) from Web ppendix 2 (Supporting Information), we set the number of chains, the number of iterations and the burn-in size as 4, 4000, and 2000, respectively (after fitting the models, the chains for the posterior sampled values were observed to be well mixed in all cases). Then, to assess the fitted models, we will plot the estimated net survival curves (averaged over all regions) along with an error measure defined as

Error =
$$\int_{\mathcal{T}} |f(t)| \hat{f}(t) dt$$
, (7)

where f is the true function, \hat{f} is the corresponding estimated function, and $\mathcal{T} = [t_1, t_2]$ is the analyzed time interval.

First, Figures 1 and 2 show the net survival curves and the corresponding errors, as per Equation (7), in estimating the true functions for $\mathcal{T}=[0,4]$, respectively, for data generated from the RS-SGH LN IC R model with 25% and 50% censoring rates for all sample sizes. In that case, we fit the same model as the generating scheme; thus, here, we aim to assess the impact of the censoring rate and the sample size when employing such an approach. From these figures, we can see that our models recover well the true net survival functions for a 25% censoring rate, with decreasing uncertainty as the sample size increases. In particular, we observe reasonable results for a sample size larger than 500–1000 patients. In a similar manner, for scenarios with 50% censoring rate, the estimates get better as we increase the number of patients; however, if the sample

 size is too small (e.g., 200 patients), the observed bias (and the variability represented in Figure 2) when estimating the net survival curves is larger than before—although such a high censoring effect vanishes as the sample size gets larger.

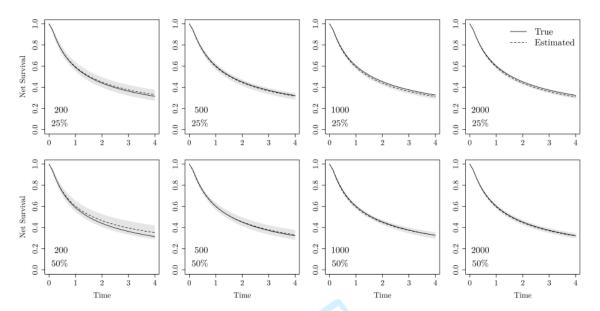
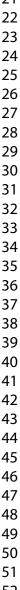


Figure 1. True and estimated (along with a 95 equal-tailed credible interval) net survival curves based on the fitted RS-SGH LN ICAR model. The data were generated from the same model with 25 and 50 censoring rates and sample size set to $200,\,500,\,1000,\,$ and 2000 patients. Such estimates were obtained by averaging over the 100 simulated data sets and all regions for each scenario (the corresponding uncertainty was computed based on the percentiles for the curves that average the regions' net survival).

Second, Figures 3 and 4 show similar plots to before, however, for data generated from the RS-SGH LN IC R and RS-SGH PGW IC R models with 25% censoring rate for all sample sizes. In these two cases, the fitted models were RS-SGH PGW IC R and RS-SGH LN IC R, respectively; that is, we are fitting misspecified models for the baseline hazard component. From such figures, we can notice that misspecified baseline hazard distributions do not seem to be an important issue if one is solely interested in computing marginal quantities, as the net survival curves. However, as mentioned before, depending on the censoring rate, 3-parameter distributions (e.g., PGW) might require larger samples to fit.



0.4 Error (Net Survival) 0.30.1 200 500 1000 2000 200 500 1000 2000 25%25%25%25%50% 50%50%50%

Figure 2. Error in estimating the true net survival function based on the fitted RS-SGH LN ICAR model. The data were generated from the same model with 25 and 50 censoring rates and sample size set to 200, 500, 1000, and 2000 patients. The computed errors aggregate the 100 simulated data sets and all regions for each scenario. The crosses (\times) correspond to the boxplot values mean.

Lastly, Figures SF1, SF2, SF3, and SF4 (Supporting Information) show the corresponding results for the remaining scenarios from Table 1.

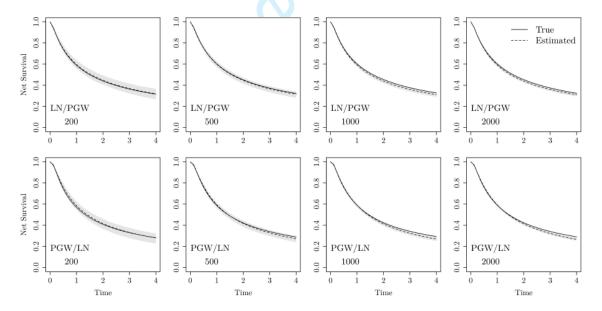


Figure 3. True and estimated (along with a 95 equal-tailed credible interval) net survival curves based on the fitted RS-SGH PGW ICAR (first row) and RS-SGH LN (second row) models. In these two cases, the data were generated from models RS-SGH LN ICAR (first row) and RS-SGH PGW ICAR (second row), respectively, with 25 censoring rate and sample size set to 200, 500, 1000, and 2000 patients. Such estimates were obtained by averaging over the 100 simulated data sets and all regions for each scenario (the corresponding uncertainty was computed based on the percentiles for the curves that average the regions' net survival).



49

50 51

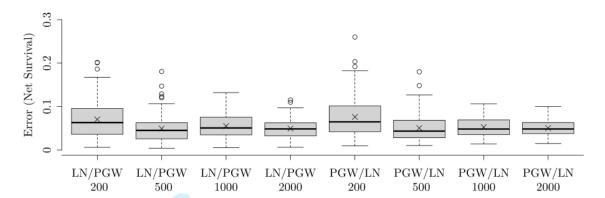


Figure 4. Error in estimating the true net survival function based on the fitted RS-SGH PGW ICAR and RS-SGH LN models. In these two cases, the data were generated from models RS-SGH LN ICAR (four first boxes) and RS-SGH PGW ICAR (four last boxes), respectively, with 25 censoring rate and sample size set to $200,\,500,\,1000,\,$ and 2000 patients. The computed errors aggregate the 100 simulated data sets and all regions for each scenario. The crosses (\times) correspond to the boxplot values mean.

4.2 Random effects selection

In this section, we will analyze the role of the spatial effects in model selection. That is, fixing all components but the random effects, we will select the most appropriate model according the the estimated elpd_{PSIS-LOO}. To do so, we will, for the same data generating scenarios from Table 1 with 25% censoring rate, fit models with no random effects (RS-SGH), IID random effects (RS-SGH IID), IC R random effects (RS-SGH IC R), and BYM2 random effects (RS-SGH BYM2) with different distributions for the baseline hazard component. Table ST3 (Supporting Information) lists all considered combinations for data generation and model fitting. Then, similarly to Section 4.1, we will rank the models for the 100 different simulated data sets in all scenarios.

fter fitting all models (the ones that were not fitted yet in Section 4.1), we compute $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$ (using the 100 package) and compare such estimated quantities for all equivalent scenarios. s a remark, when fitting the models, all posterior chains were well mixed. Table 2 reports the "best-model proportions" (i.e., the number of times, out of the 100 data sets, that a model was selected based on the estimated $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$) for all

 scenarios with sample size set to 2000 patients. Tables ST4, ST5, and ST6 (Supporting Information) report similar results, but based on the data sets containing 200, 500, and 1000 patients, respectively.

Table 2. "Best-model proportions" for model selection based on the estimated $elpd_{PSIS-LOO}$. In all scenarios, we assumed a 25 censoring rate and set the sample size to 2000 patients.

#	Data Generating model	Fitted model	Best-model proportions	#	Data Generating model	Fitted model	Best-model proportions
01	RS-SGH LN IC R	RS-SGH LN —	2%	09	RS-SGH PGW IC R	RS-SGH LN —	1%
02	RS-SGH LN IC R	RS-SGH LN IID	6%	10	RS-SGH PGW IC R	RS-SGH LN IID	79%
03	RS-SGH LN IC R	RS-SGH LN IC R	51%	11	RS-SGH PGW IC R	RS-SGH LN IC R	9%
04	RS-SGH LN IC R	RS-SGH LN BYM2	41%	12	RS-SGH PGW IC R	RS-SGH LN BYM2	11%
05	RS-SGH LN IC R	RS-SGH PGW —	15%	13	RS-SGH PGW IC R	RS-SGH PGW —	7%
06	RS-SGH LN IC R	RS-SGH PGW IID	32%	14	RS-SGH PGW IC R	RS-SGH PGW IID	26%
07	RS-SGH LN IC R	RS-SGH PGW IC R	28%	15	RS-SGH PGW IC R	RS-SGH PGW IC R	41%
08	RS-SGH LN IC R	RS-SGH PGW BYM2	25%	16	RS-SGH PGW IC R	RS-SGH PGW BYM2	26%

From Table 2 (and Tables ST4–ST6, Supporting Information), we can see that models that account for some random effects structure were selected more often in all scenarios (except for 200 patients, as in Table ST4). lso, as the sample size increases (500 patients or more), not only the models with spatial effects were selected with higher proportions, but also the correct model (RS-SGH IC R) was the most frequently selected approach for some of the specified settings with generating model based on the Log-normal distribution for the baseline hazard. On the contrary, for the misspecified scenarios with generating scheme based on the Power Generalized Weibull distribution, the model with clustering effects seemed to perform better than the competing approaches—as a reason for this to happen, recall that the Log-normal model might fail to recover the PGW hazard shape; in that case, it is possible that the spatial structure for the random effects gets suppressed by the error from the poorly fitted fixed components and the IID model performs better. In that way, under the assumption that the baseline hazard distribution can capture the corresponding hazard shape from the data, the employed model selection approach seems to work well when selecting an appropriate random structure, provided that we have a minimum of 500–1000 data points (as we also identified in Section 4.1).

4.3 Spatial effects analysis

 For the following analysis, we showcase the insights we can obtain from the estimated spatial structure. For that, we will generate data from a model with manually set spatial effects. In particular, we will simulate data from the RS-SGH model with a Log-normal distribution for the baseline hazard, such that $\mu=0.65$ and $\sigma=1.15$, as in Section 4.1. Iso, we will choose the covariates and set the corresponding coefficients as in ppendix C. Lastly, we will set $\tilde{\bf u}={\bf u}=(2.0,\,1.5,\,1.0,\,0.5,\,0,\,0.5,\,1.0,\,1.5,\,2.0)^{\rm T}$ for the 1–9 regions in England (as per Figure 7), respectively. Then, for a data set with 10000 individuals and censoring rate of 25%, we will fit the RS-SGH LN model, with no random effects and with IID, IC R, and BYM2 random structures, once, such that the MCMC setting parameters will be defined as in Section 4.1, and analyze the estimated spatial effects (if any) based on the corresponding posterior distributions.

Similar to previous simulations, when fitting the models, all posterior chains were well mixed. However, when comparing such approaches according to the elpd_{PSIS-LOO} criterion, the results pointed out to the IC R random structure as the most appropriate model—although the pairwise differences between the IC R model and the IID and BYM2 models seem to be non-significant, as shown in Table ST7 (Supporting Information). The model with no spatial effects was ranked in the lowest position. This means that, although we do need to account for non-observed spatial heterogeneity, for such a large data set, all random structures captured well the spatial effects. Figure 5 shows the true and estimated spatial effects (at both time- and hazard-levels) for the IC R random structure (Figures SF5 and SF6, in Web ppendix 6 (Supporting Information), show the corresponding maps for the IID and BYM2 structures, respectively), such that the plotted estimates were computed based on the mean of the sample obtained from the corresponding random effect posterior

distributions. The first thing we can observe from these maps is that we were able to recover the spatial effects reasonably well. Table ST8 (Supporting Information) shows the estimates for all RS-SGH LN IC R model parameters along with a 95% equal-tail credible interval for the same model (Tables ST8 also presents similar results for models RS-SGH LN IID and RS-SGH LN BYM2). Second, based on these estimates only, we can study the geographical inequalities for different population groups. That is, fixing all terms but $\tilde{\mathbf{u}}$ and \mathbf{u} , the risk of dying is larger for patients who live in regions with positive estimates for the hazard-level spatial effects. The time-level spatial effects have a similar interpretation, but they have to be analyzed along with the baseline hazard shape; i.e., if $h_0(t; \boldsymbol{\theta})$ increases with t, then positive \tilde{u}_i 's imply in riskier areas; contrarily, if the baseline hazard is a decreasing function, a positive timelevel random effect decreases the risk of dying in i, for patients with the same characteristics, in comparison to other regions with smaller effects of the same kind.

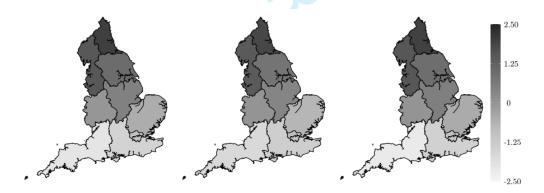


Figure 5. Spatial effects for the RS-SGH LN ICAR model. Left panel: True spatial effects $\tilde{} = 2\ 0, 1\ 5, 1\ 0, 0\ 5, 0, 0\ 5, 1\ 0, 1\ 5, 2\ 0)^{\top}$. Middle panel: Estimated time-level spatial effects $\tilde{} = 1\ 86, 1\ 45, 0\ 81, 0\ 41, 0\ 04, 0\ 69, 0\ 90, 1\ 33, 1\ 65)^{\top}$. Right panel: Estimated hazard-level spatial effects $= 1\ 96, 1\ 54, 0\ 97, 0\ 53, 0\ 04, 0\ 45, 1\ 01, 1\ 43, 2\ 15)^{\top}$.

However, as suggested by Taylor (2017), analyzing the spatial effects as in Figure 5 ignores the precision of the estimates—recall that we would be more precise in estimating $\tilde{\mathbf{u}}$ and \mathbf{u} in regions with more patients.

 Iternatively, we could compute and analyze the relative exceedance probability $(u_i > c)$, for all i and some threshold c (the same applies to \tilde{u}_i , $\forall i$). Under the Bayesian setting, such a probability can be estimated based on the posterior sample for the spatial effects. This measure quantifies the variability of the random effect estimate around c, but it is also useful to assess unusual elevations in such a model component. For instance, we might be interested in computing $(u_i > 0)$, for all regions i (the same for \tilde{u}_i , $\forall i$). This quantity can be used as a proxy for the risk level to which a group of patients (defined by their geographical location) is subjected, compared to the general population. Figures SF7, SF8, and SF9 (Supporting Information) show the computed relative exceedance probabilities for both time-level and hazard-level random effects, such that the threshold c = 0, for models RS-SGH LN IC R, RS-SGH LN IID, and RS-SGH LN BYM2, respectively.

Lastly, we can also analyze the estimates for $\sigma_u = 1/\sqrt{\tau_u}$ (and $\sigma_{\tilde{u}} = 1/\sqrt{\tau_{\tilde{u}}}$) in all cases. Figures SF10, SF11, and SF12 (Supporting Information) show the estimated posterior densities for such parameter when fitting the RS-SGH LN model with IC R, IID, and BYM2 random effects, respectively. From these figures (and Table ST8), we can see that the two spatial structures were simultaneously and successfully estimated. s discussed in Section 3.1, if (i) we have enough uncensored patients in each region and (ii) the baseline hazard does not belong to the Weibull family, then we may estimate well the random effects.

5 Case study

In this section, we will analyze a data set that contains survival information about male and female patients diagnosed with colon cancer between 2015 and 2016 in England. ppendix D presents a complete description of the data set. More specifically, we analyze the survival of colon cancer patients in England with spatial structure defined in two

10 11 12

13

14

15 16

17

18 19

> 28 29

> 30

> 40 41

> 42

43 44

50 51

different manners: (i) based on the administrative boundaries given by the Government Office Regions (as per Figure 7) and (ii) based on the health boundaries determined by the Cancer Iliances (Office for National Statistics 2022). The main goal is assessing the impact of different geographies when accounting for the possible spatial correlation in the data.

For all scenarios, we know subject-specific prognostic factors, which include age at diagnosis, sex, deprivation level, and cancer stage. The population hazard term $h_P(age_{ij} + t; \mathbf{z}_{ij})$ was determined based on the life tables for England defined for the corresponding calendar year, and stratified by age, sex, deprivation level (according to the computed quintiles of such a score), and region of residence (Rachet et al. 2015; Inequalities in Cancer Outcomes Network 2022). lso, for all models separately fitted for male and female individuals, we always set the time-level linear predictor to age_{ij} + \tilde{u}_i , and the hazard-level linear predictor to $age_{ij}\beta_1 + \sum_{k=2}^{K} stage_{ij}(k)\beta_k + deprivation_{ij}\beta_{(K+1)} +$ u_i , where $_{\text{stage}_{ij}(k)}$, for $2 \leq k \leq K$, is an indicator function for individuals who belong to the k-th cancer tumour stage, and, as in Section 4, the \tilde{u}_i and u_i components (if any) are defined as one of the (spatially dependent) random structures introduced in Section 2.2. Finally, the variables "age" and "deprivation" were standardized for numerical stability.

Given the setting we just described, we will fit Model (2) for 10,936 males and 9,586 females with a diagnosis of colon cancer in 2016 in England. The linear predictor terms will be defined as mentioned above, such that we have K = 4 levels for the cancer stage ("1" being early stage and "4" late stage). lso, regarding the baseline hazard distributions and the random effects, for each scenario we fit the following models: RS-SGH LL IC R, RS-SGH LL BYM2, RS-SGH LN IC R, RS-SGH LN BYM2, RS-SGH PGW IC R, and RS-SGH PGW BYM2, as per the notation introduced in Table 4. We selected these models for two reasons.

 First, due to the computational cost associated with fitting them to such a large data set (Table ST10, Supporting Information, shows the fitting times for all cases), we decided to limit our investigation to models with clinical motivation (in particular, models that present some underlying spatial structure). Second, such models are flexible enough to cover many different hazard shapes and possible spatially dependent random effects. For the MCMC-based code, we set the number of chains, the number of iterations and the burn-in size as 4, 10000, and 8000, respectively (the posterior chains were well mixed in all cases, except for the RS-SGH LL BYM2 model with male patients spatially distributed over the Government Office Regions—see Table ST9, Supporting Information). Next, the best model is selected according to the elpd_{PSIS-LOO} criterion, as in Section 3.3. Lastly, the spatial structure is defined according to two geographies: (i) the 9 Government Office Regions (GOR), as in Figure 7, and (ii) the 19 Cancer lliances Regions delimited during the calendar year of 2016.

Considering these fitted models, Table ST9 (Supporting Information) shows the selected model (according to the $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$ criterion) for each scenario. Thus, the following results are based on the highest-ranked modeling alternatives. Then, we compute the net survival for t=1 and 3 years (along with the estimated 95% equal-tailed credible interval) for all regions. Here, it is important mentioning how the uncertainty for this quantity is being estimated; for each sampled vector of parameters $\boldsymbol{\xi}^s$, where s is the index for the posterior sample, we determine the 2.5^{th} and 97.5^{th} percentiles of $S_{\text{N},i}(t\mid\boldsymbol{\xi}^s)$, for all t and i, as per Equation (4). Figure 6 shows the estimated net survival for the male and female groups, the two different geographies, and t=3 years; also, Figures SF13 and SF14 (Supporting Information) report the associated uncertainty. Similarly, SF15, SF16, and SF17 (Supporting Information) present the corresponding maps for t=1 year.

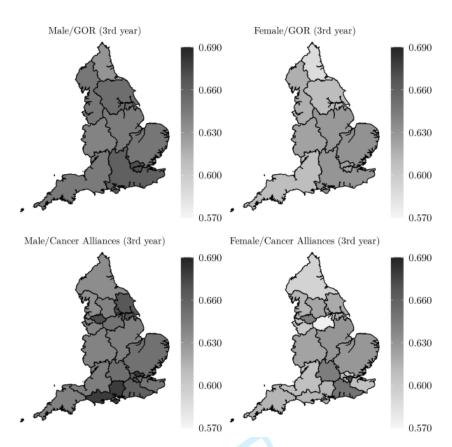


Figure 6. Net survival point estimate for t=3 based on the (i: top-left panel) "Government Office Regions" spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right panel) "Government Office Regions" spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for female patients.

From these figures, we can analyze (i) the rate of change for the net survival estimates, (ii) the difference between the net survival for male and female patients, and (iii) the impact of the chosen administrative boundaries when estimating the quantities of interest. Firstly, the net survival seems to decrease faster during the first years after diagnosis, such that the corresponding estimates after 1 and 3 years of diagnosis are, approximately, 0.75 and 0.63, respectively. Secondly, regarding the differences between male and female patients, the female individuals are shown to be slightly more likely of dying than men—this difference can be seen in the maps for all follow-up time windows. Thirdly, analyzing

 the England territory based on a finer resolution (e.g., Cancer Iliances Regions) brings us more information if compared to the GOR-based results. In particular, "South Yorkshire, Bassetlaw, North Derbyshire and Hardwick" (as per Figure SF18 Supporting Information) shows a lower (for the female group) estimated net survival than the other regions—notice that, just by inspecting the GOR-based estimates, it would be difficult to identify these locations, and practical implications (e.g., allocation of resources for medically underserved areas) could not happen.

Furthermore, given that the random effects play a major role in the description of the results, Table ST11 (Supporting Information) shows the estimated values (with reported uncertainty) of $\sigma_u = 1/\sqrt{\tau_u}$ (same for $\sigma_{\tilde{u}}$) and ρ (same for $\tilde{\rho}$) for all highest-ranked BYM2 models. dditionally, Figures SF19–SF22 show the estimated posterior densities of σ_u and $\sigma_{\tilde{u}}$ for the same models. From that table and figures, note that, first, the uncertainty associated with the estimation of σ_u and $\sigma_{\tilde{u}}$ is not large. This indicates that we did not encounter any identifiability issues when recovering the underlying spatial structures. Second, the point estimates for ρ and $\tilde{\rho}$ are around 0.55 (with standard error around 0.35) for all scenarios, which indicates that the proportion of the variance that comes from the structured random effects is similar to the contribution from the unstructured random effects. That is, the flexibility that comes with the BYM2 structure seems to be important to correctly characterize the variability that cannot be explained by the fixed effects, both in time- and hazard-level.

We can also compute the net survival stratified by the categorical variables; in this section, we will consider the "deprivation level" and the "cancer stage" as the population strata. However, recall that the deprivation level is a continuously defined score, thus, if we want to aggregate the patients based on such information, we can compute its quintiles and classify the individuals according to the obtained intervals. In

 that way, the deprivation score will have 5 levels ("1" being *least deprived* and "5" *most deprived*). s before, Figures SF23 and SF26 (Supporting Information) show the estimated net survival maps, for t=3 years, stratified by the "deprivation level" and the "cancer stage," respectively, such that we plotted and compared the deprivation levels "1" (least deprived level) and "5" (most deprived level), and the cancer stages "1, 2, and 3" (early stages) and "4" (late stage). Figures SF24, SF25, SF27, and SF28 (Supporting Information) report the associated uncertainty. Lastly, Figures SF29–SF34 Supporting Information) present the corresponding maps for t=1 year.

Firstly, based on the figures for the estimated net survival curves stratified by deprivation level, we can notice that, for all time points, not only the estimates vary over space (with less homogeneous spatial distribution when we consider the finer spatial resolution), but also the net survival for the different population strata decrease as the deprivation level gets larger; in particular, patients with a deprivation score of 5 have higher chances of dying than the least deprived group regardless of the time span, gender, and the Government Office Region (or Cancer lliance)—which is likely to be associated with sub-optimal treatment strategies offered to this group. Secondly, by analyzing the figures that show the net survival estimates stratified by the cancer stage, we have similar conclusions, that is, the net survival decrease as the patients are diagnosed with later stages for the colon cancer. However, for this stratified analysis, we can notice a much larger difference in the chances of surviving between the groups with cancer stages "1, 2, and 3" (early stages) and "4" (late stage), regardless of the other factors. In fact, when plotting the corresponding net survival maps, we had to present the results in different scales for each level of severity (as stressed in all figures captions); otherwise, the spatial variability in each of these two groups would not be captured in the maps.

 This occurs since patients with *stage 4*-cancer are less likely to be cured and, instead, only receive palliative care.

6 Discussion

In this work, we introduced the Relative Survival Spatial General Hazard (RS-SGH) class of models that generalizes, under the relative survival framework, other survival models. The proposed RS-SGH models account for spatial random effects both in the time-level and hazard-level components, such that these random structures can be modeled, among other approaches, according to the IC R or BYM2 smoothing priors. The proposed class of models was implemented using R (R Development Core Team 2022) and ST N (Carpenter et al. 2017) and made available in a public repository, which allows for reproducibility of our research. ppendix 2 (Supporting Information) provides an example on how to use the scripts; in particular, Table ST1 lists all models that are currently possible to implement. lso, regarding model selection, we computed and used the elpd_{PSIS-LOO} estimates (as per Section 3.3), and tested its performance in Section 4.2. This work also contains other minor contributions, such as (i) the prior distribution recommendations (as per Section 3.2) for the model parameters and hyperparameters, (ii) some guidelines about the sample size, baseline hazard distribution misspecification, and censoring rate when fitting models of this kind (as per Section 4.1), and (iii) a simple extension of the "exceedance probability" idea to the computed (and interpreted) "relative exceedance probabilities" (as per Section 4.3).

iming to validate the proposed model and inference tools, we conducted a simulation study that analyzed the effects of the sample size, censoring rate, and the baseline hazard distribution when estimating the model parameters and recovering the net survival curves. In this regard, the sample size and the censoring rate were shown to be the

most important factors to control; for instance, in most cases, a minimum sample size of 500–1000 patients provided estimates with less variability for the net survival curves. lso, higher censoring rates (e.g., 50%) with not large sample sizes (e.g., 200–500 patients) produced biased estimates for this same quantity. In fact, for 3-parameter distributions (e.g., Power Generalized Weibull), it might be difficult even to obtain well-mixed posterior chains when fitting the model. However, the misspecification of the baseline hazard distribution, provided that we have enough noncensored observations and a model that can capture the true hazard shape, s part of had little impact in the estimation of marginal quantities. the simulation study, we also assessed the model selection performance and the ability to recover the true spatial effects. lso, based on these estimated random structures, we could compute the relative exceedance probabilities, which are functions of the spatial effects that can be used, depending on the set threshold, to compare specific locations to the general population with respect to their net survival. s a note, our simulation study was conducted based on a spatial graph defined by 9 regions. Hence, it may be of interest to explore scenarios with a larger number of areas.

We have also presented a case study aiming at answering genuine questions of interest in cancer epidemiology. In particular, we found that a finer spatial resolution brings us more important information about areas that present lower net survival than the overall country. Identifying these locations is crucial as, based on such knowledge, decision-makers can focus their resources on improving the lives of the vulnerable groups of the population. Moreover, we have illustrated how to produce summaries for subgroups of the population of interest, such as those defined by "deprivation level" and "cancer stage." For the former, we have found that most-deprived patients (deprivation level 5) exhibit lower chances of survival compared to the least deprived groups. For the latter, patients with late-stage cancer (stage 4) experience a significant reduction in their

 survival prospects; in fact, as we have briefly mentioned in Section 5, in most cases, these patients only receive palliative care.

The proposed methodology and results presented in this work can be extended in different directions. Firstly, we could also include a timedependent component in Equation (2) that explains the non-observed temporal variability associated to the year of diagnosis. Such an extension could be mainly useful for studies that take individuals diagnosed over a very large time window, as the treatment (thus, the chances of surviving) is likely to improve in the long term. Secondly, in Section 5, when modeling survival, it may be useful to simultaneously include spatial information not only about the patients' place of residence, but also about their local of treatment. s pointed out by Quaresma et al. (2022), cancer incidence depends on where you live (as this is related to deprivation, and deprivation has a strong relationship with geographies), while survival also depends on where you are treated (as it depends on the quality of healthcare). Thus, future work might extend our model into this direction. Thirdly, missing data is a prevalent problem in population studies. Thus, a possible extension of our work consists of developing multiple imputation strategies to account for missing data, while also accounting for spatial variability. Fourthly, less common smoothing priors for describing the possible spatial autocorrelations among regions could have been used; for instance, the directed acyclic graph auto-regressive (D G R) model (Datta et al. 2019) is an alternative to the IC R model that can also be used for modeling other data structures (e.g., images and networks). However, while still using the IC R formulation, the PC priors can be employed when specifying the precision parameters in the spatial random effects. s discussed by Simpson et al. (2017) (and references therein), a Gamma prior may not be the most suitable choice for this problem. Similarly, PC priors may also be used in the context of the BYM2 model (Riebler et al. 2016). Therefore, the implementation of such penalized complexity

 priors is a consideration for future work. Finally, the idea of incorporating spatial (or spatio-temporal) random structures into the hazard model can also be implemented in other survival modeling frameworks, such as the competing risks models, cure models, and overall survival models.

Funding

Manuela Quaresma is funded by the Cancer Research UK Population Research Committee Funding Scheme: Cancer Research UK Population Research Committee-Programme ward C7923/ 29018. The research work of Francisco J. Rodríguez-Cortés has been partially supported by Universidad Nacional de Colombia, HERMES projects, Grant/ ward Number: 56470.

Declaration of conflicting interests

The authors declare that they have no conflict of interest.

Data vailability Statement

Data may be obtained from a third party and are not publicly available. This study uses English cancer registration data managed by the National Cancer Registration and nalysis Service (NCR S). The authors do not own these data and hence are not permitted to share them in the original form. The data are available from the Data ccess Request Service (D RS) at NHS Digital (merged with NHS England from February 2023).

References

llemani, C., Weir, H. K., Carreira, H., Harewood, R., Spika, D., Wang, X. S., Bannon, F., hn, J. V., Johnson, C. J., Bonaventure, and the CONCORD Working Group (2015). Global surveillance of cancer survival 1995–2009: analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2). The Lancet *385*, 977–1010.

lvares, D. and Rubio, F. J. (2021). tractable Bayesian joint model for longitudinal and survival data. Statistics in Medicine *40*, 4213–4229.

- maral, V. R., Krainski, E. T., Zhong, R. and Moraga, P. (2023). Model-Based Geostatistics Under Spatially Varying Preferential Sampling. Journal of gricultural, Biological and Environmental Statistics, 1–27.
- Banerjee, S., Carlin, B. P. and Gelfand, . E. (2003a). Hierarchical modeling and analysis for spatial data. Chapman and Hall/CRC.
- Banerjee, S., Wall, M. M. and Carlin, B. P. (2003b). Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota. Biostatistics *4*, 123–142.
- Basak, P., Linero, ., Sinha, D. and Lipsitz, S. (2022). Semiparametric analysis of clustered interval-censored survival data using soft Bayesian additive regression trees (SB RT). Biometrics 78, 880–893.
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society: Series B (Methodological) *36*, 192–225.
- Besag, J., Green, P., Higdon, D. and Mengersen, K. (1995). Bayesian computation and stochastic systems. Statistical science *10*, 3–41.
- Besag, J., York, J. and Mollie, . (1991). Bayesian image restoration, with two applications in spatial statistics. nnals of the Institute of Statistical Mathematics *43*, 1–20.
- Betancourt, M. and Girolami, M. (2015). Hamiltonian Monte Carlo for hierarchical models. Current trends in Bayesian methodology with applications 79, 2–4.
- Buckley, J. and James, I. (1979). Linear regression with censored data. Biometrika 66, 429–436.
- Carlin, B. P. and Banerjee, S. (2003). Models for Spatio-Temporally Correlated Survival Data. In Bayesian Statistics 7: Proceedings of the Seventh Valencia International Meeting p. 45, Oxford University Press.
- Carpenter, B., Gelman, ., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P. and Riddell, . (2017). Stan: Probabilistic Programming Language. Journal of Statistical Software 76, 132.
- Charvat, H., Remontet, L., Bossard, N., Roche, L., Dejardin, O., Rachet, B., Launoy, G., Belot, and Group, C. W. S. (2016). multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates. Statistics in Medicine 35, 3066–3084.
- Chen, Y. Q. and Jewell, N. P. (2001). On a general class of semiparametric hazards regression models. Biometrika 88, 687–702.
- Chen, Y. Q. and Wang, M. C. (2000). nalysis of accelerated hazards models. Journal of the merican Statistical ssociation *95*, 608–618.

- Cramb, S., Moraga, P., Mengersen, K. and Baade, P. (2017). Spatial variation in cancer incidence and survival over time across Queensland, ustralia. Spatial and Spatio-temporal Epidemiology *23*, 59–67.
- Cramb, S. M., Mengersen, K. L., Lambert, P. C., Ryan, L. M. and Baade, P. D. (2016). flexible parametric approach to examining spatial variation in relative survival. Statistics in Medicine *35*, 5448–5463.
- Datta, ., Banerjee, S., Hodges, J. S. and Gao, L. (2019). Spatial disease mapping using directed acyclic graph auto-regressive (D G R) models. Bayesian analysis *14*, 1221.
- Dette, H., Ley, C. and Rubio, F. (2018). Natural (Non-) Informative Priors for Skew-symmetric Distributions. Scandinavian Journal of Statistics *45*, 405–420.
- Diggle, P. J., Menezes, R. and Su, T. (2010). Geostatistical inference under preferential sampling. Journal of the Royal Statistical Society: Series C (pplied Statistics) *59*, 191–232.
- Diggle, P. J. and Ribeiro, P. J. (2007). Model-based Geostatistics. Springer.
- Eletti, ., Marra, G., Quaresma, M., Radice, R. and Rubio, F. J. (2022). Unifying Framework for Flexible Excess Hazard Modeling with pplications in Cancer Epidemiology. Journal of the Royal Statistical Society: Series C *71*, 1044–1062.
- Estève, J., Benhamou, E., Croasdale, M. and Raymond, L. (1990). Relative survival and the estimation of net survival: elements for further discussion. Statistics in Medicine *9*, 529–538.
- Etezadi- moli, J. and Ciampi, . (1987). Extended hazard regression for censored survival data with covariates: a spline approximation for the baseline hazard function. Biometrics *43*, 181–192.
- Exarchakou, ., Rachet, B., Belot, ., Maringe, C. and Coleman, M. (2018). Impact of national cancer policies on cancer survival trends and socioeconomic inequalities in England, 1996-2013: population based study. BMJ *360*.
- Fahrmeir, L., Kneib, T., Lang, S. and Marx, B. (2013). Regression: models, methods and applications. Springer.
- Freni-Sterrantino, ., Ventrucci, M. and Rue, H. (2018). note on intrinsic conditional autoregressive models for disconnected graphs. Spatial and spatio-temporal epidemiology 26, 25–34.
- Gelman, . (2006). Prior distributions for variance parameters in hierarchical models. Bayesian nalysis *1*, 515–533.
- Henderson, R., Shimakura, S. and Gorst, D. (2002). Modeling spatial variation in leukemia survival data. Journal of the merican Statistical ssociation *97*, 965–972.

- Inequalities in Cancer Outcomes Network (2022). Life tables for England and Wales by sex, calendar period, region and deprivation. https://icon.lshtm.ac.uk/tools-analysis/uk-life-tables/.
- Klein, J. P., Van-Houwelingen, H. C., Ibrahim, J. G. and Scheike, T. H. (2014). Handbook of Survival nalysis. CRC Press, Boca Raton, FL.
- Li, L., Hanson, T. and Zhang, J. (2015). Spatial extended hazard model with application to prostate cancer survival. Biometrics *71*, 313–322.
- Li, Y. and Ryan, L. (2002). Modeling spatial survival data using semiparametric frailty models. Biometrics *58*, 287–297.
- Lunn, D., Spiegelhalter, D., Thomas, . and Best, N. (2009). The BUGS project: Evolution, critique and future directions. Statistics in medicine 28, 3049–3067.
- Mahmood, M., maral, V. R., Mateu, J. and Moraga, P. (2022). Modeling infectious disease dynamics: Integrating contact tracing-based stochastic compartment and spatio-temporal risk models. Spatial Statistics *51*, 100691.
- Moraga, P. (2019). Geospatial health data: Modeling and visualization with R-INL and Shiny. CRC Press.
- Morris, M., Wheeler-Martin, K., Simpson, D., Mooney, S., Gelman, and DiMaggio, C. (2019). Bayesian hierarchical spatial models: Implementing the Besag York Mollie model in Stan. Spatial and Spatio-Temporal Epidemiology *31*, 100301.
- Office for National Statistics (2022). C L and NCV (Feb 2017) Full Clipped Boundaries in England. https://geoportal.statistics.gov.uk/datasets/ons:: cal-and-ncv-feb-2017-full-clipped-boundaries-in-england/.
- Perme, M. P., Stare, J. and Estève, J. (2012). On estimation in relative survival. Biometrics *68*, 113–120.
- Public Health England (2020).Geographical patterns of cancer survival England: adults diagnosed 2013 2017 in to and followed up to 2018. https://www.gov.uk/government/statistics/ geographic-patterns-of-cancer-survival-in-england-adults-diagnosed-201
- Quaresma, M., Carpenter, J. R., Turculet, . and Rachet, B. (2022). Variation in colon cancer survival for patients living and receiving care in London, 2006–2013: does where you live matter? Journal of Epidemiology and Community Health *76*, 196–205.
- Rachet, B., Maringe, C., Woods, L. M., Ellis, L., Spika, D. and llemani, C. (2015). Multivariable flexible modelling for estimating complete, smoothed life tables for sub-national populations. BMC Public Health *15*, 1–9.

- R Development Core Team (2022). R: Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna, ustria. ISBN 3-900051-07-0.
- Riebler, ., Sørbye, S. H., Simpson, D. and Rue, H. (2016). n intuitive Bayesian spatial model for disease mapping that accounts for scaling. Statistical Methods in Medical research 25, 1145–1165.
- Rubio, F., Remontet, L., Jewell, N. and Belot, (2019). On a general structure for hazard-based regression models: an application to population-based cancer research. Statistical Methods in Medical Research 28, 2404–2417.
- Rubio, F. J. (2022). SimLT. R package version 0.1.0.
- Rubio, F. J. and Drikvandi, R. (2022). MEGH: parametric class of general hazard models for clustered survival data. Statistical Methods in Medical Research *31*, 1603–1616.
- Rubio, F. J. and Steel, M. F. J. (2018). Flexible linear mixed models with improper priors for longitudinal and survival data. Electronic Journal of Statistics *12*, 572–598.
- Rue, H. and Held, L. (2005). Gaussian Markov random fields: theory and applications. CRC press.
- Simpson, R., Rue, H., Riebler, ., Martins, T. G. and Sørbye, S. H. (2017). Penalising model component complexity: principled, practical approach to constructing priors. Statistical Science *32*, 1–28.
- Sørbye, S. H. and Rue, H. (2017). Penalised complexity priors for stationary autoregressive processes. Journal of Time Series nalysis *38*, 923–935.
- ST N Development Team (2021). RStan: the R interface to ST N. R package version 2.21.3.
- Taylor, B. M. (2017). Spatial Survival nalysis pp. 2101–2109. Cham: Springer International Publishing.
- Vehtari, ., Gabry, J., Magnusson, M., Yao, Y., Burkner, P., Paananen, T. and Gelman, . (2022). 100: Efficient leave-one-out cross-validation and W IC for Bayesian models. R package version 2.5.1.
- Vehtari, ., Gelman, . and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and W IC. Statistics and computing , 27, 1413–1432.
- Vehtari, ., Simpson, D., Gelman, ., Yao, Y. and Gabry, J. (2015). Pareto smoothed importance sampling.
- Yu, B. and Tiwari, R. (2012). Bayesian approach to mixture cure models with spatial frailties for population-based cancer relative survival data. Canadian Journal of Statistics, 40, 40–54.
- Zellner, . (1986). On assessing prior distributions and Bayesian regression analysis with g-prior distributions.

 Zhou, H. and Hanson, T. (2018). unified framework for fitting Bayesian semiparametric models to arbitrarily censored survival data, including spatially referenced data. Journal of the merican Statistical ssociation, 113, 571–581.

Sub-models based on the RS-SGH approach

Table 3 shows eight possible sub-models that can be derived from the Relative Survival Spatial General Hazard (RS-SGH) model and that we believe are useful for researchers and practitioners working with survival data.

Table 3. Eight simpler models based on the Relative Survival Spatial General Hazard (RS-SGH) modeling approach. The "Description" column refers to the corresponding terms in Equation (2).

Name	Description	Name	Description
RS-SGH-I	$\tilde{\mathbf{u}} =$	RS-GH	$\tilde{\mathbf{u}} = \mathbf{u} =$
RS-SGH-II	$\tilde{\mathbf{u}} = \mathbf{u}$	RS-PH	$\tilde{\mathbf{u}} = \mathbf{u} = , =$
RS-SPH	$\tilde{\mathbf{u}} = , =$	RS- FT	$ \tilde{\mathbf{u}} = \mathbf{u} = , = \boldsymbol{\beta}$
RS-S FT	$ \tilde{\mathbf{u}} = \mathbf{u}, = \boldsymbol{\beta} $	RS- H	$\tilde{\mathbf{u}} = \mathbf{u} = \ , oldsymbol{eta} =$

B Implemented models

Referring back to Model (2) (and all variations from Table 3) and assuming a parametric form for the baseline hazard function given by the models listed in Section 2.1, and considering all possible structures for the random effects defined in Section 2.2, we can enumerate at least 95 models to choose from. Table 4 lists all possible models.

The code, available on https://github.com/avramaral/relative_survival, implements all such models, and the fitting procedure, using RStan (ST N Development Team 2021) in the background, can be performed as in Web ppendix 2. In that section, we provide the code snippet that can be used for fitting Model (2) for

Table 4. All implemented models. The "Dist." column refers to the possible distributions for the baseline hazard function, "Model" refers to the implemented excess hazard models—as per Table (3), and "R.E." refers to the spatially structured (and unstructured) random effects described in Section 2.2.

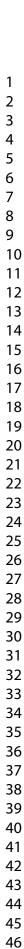
#	Dist.	Model	R.E.																
01	LN	RS-SGH	BYM2	20	LL	RS-SGH	BYM2	39	PGW	RS-SGH	BYM2	58	G M	RS-SGH	BYM2	77	GG	RS-SGH	BYM2
02	LN	RS-SGH	IC R	21	LL	RS-SGH	IC R	40	PGW	RS-SGH	IC R	59	G M	RS-SGH	IC R	78	GG	RS-SGH	IC R
03	LN	RS-SGH	IID	22	LL	RS-SGH	IID	41	PGW	RS-SGH	IID	60	G M	RS-SGH	IID	79	GG	RS-SGH	IID
04	LN	RS-SGH-I	BYM2	23	LL	RS-SGH-I	BYM2	42	PGW	RS-SGH-I	BYM2	61	G M	RS-SGH-I	BYM2	80	GG	RS-SGH-I	BYM2
05	LN	RS-SGH-I	IC R	24	LL	RS-SGH-I	IC R	43	PGW	RS-SGH-I	IC R	62	G M	RS-SGH-I	IC R	81	GG	RS-SGH-I	IC R
06	LN	RS-SGH-I	IID	25	LL	RS-SGH-I	IID	44	PGW	RS-SGH-I	IID	63	G M	RS-SGH-I	IID	82	GG	RS-SGH-I	IID
07	LN	RS-SGH-II	BYM2	26	LL	RS-SGH-II	BYM2	45	PGW	RS-SGH-II	BYM2	64	G M	RS-SGH-II	BYM2	83	GG	RS-SGH-II	BYM2
08	LN	RS-SGH-II	IC R	27	LL	RS-SGH-II	IC R	46	PGW	RS-SGH-II	IC R	65	G M	RS-SGH-II	IC R	84	GG	RS-SGH-II	IC R
09	LN	RS-SGH-II	IID	28	LL	RS-SGH-II	IID	47	PGW	RS-SGH-II	IID	66	G M	RS-SGH-II	IID	85	GG	RS-SGH-II	IID
10	LN	RS-SPH	BYM2	29	LL	RS-SPH	BYM2	48	PGW	RS-SPH	BYM2	67	G M	RS-SPH	BYM2	86	GG	RS-SPH	BYM2
11	LN	RS-SPH	IC R	30	LL	RS-SPH	IC R	49	PGW	RS-SPH	IC R	68	G M	RS-SPH	IC R	87	GG	RS-SPH	IC R
12	LN	RS-SPG	IID	31	LL	RS-SPH	IID	50	PGW	RS-SPH	IID	69	G M	RS-SPG	IID	88	GG	RS-SPG	IID
13	LN	RS-S FT	BYM2	32	LL	RS-S FT	BYM2	51	PGW	RS-S FT	BYM2	70	G M	RS-S FT	BYM2	89	GG	RS-S FT	BYM2
14	LN	RS-S FT	IC R	33	LL	RS-S FT	IC R	52	PGW	RS-S FT	IC R	71	G M	RS-S FT	IC R	90	GG	RS-S FT	IC R
15	LN	RS-S FT	IID	34	LL	RS-S FT	IID	53	PGW	RS-S FT	IID	72	G M	RS-S FT	IID	91	GG	RS-S FT	IID
16	LN	RS-GH	-	35	LL	RS-GH	_	54	PGW	RS-GH	_	73	G M	RS-GH	_	92	GG	RS-GH	-
17	LN	RS-PH	_	36	LL	RS-PH	_	55	PGW	RS-PH	_	74	G M	RS-PH	_	93	GG	RS-PH	
18	LN	RS- FT	_	37	LL	RS- FT	_	56	PGW	RS- FT	_	75	G M	RS- FT	_	94	GG	RS- FT	-
19	LN	RS- H	—	38	LL	RS- H	_	57	PGW	RS- H	_	76	G M	RS- H	_	95	GG	RS- H	-

an example based on observed leukemia-diagnosed patients (Henderson et al. 2002).

C Simulation details

The covariates for the simulated data will be based on the lung cancer estimates in London, obtained (and implemented in the SimLT package) by Rubio (2022). In particular, we will generate synthetic data for n patients—0.5n male and 0.5n female patients, such that we will have information about the "date of diagnosis," "deprivation level" (1 to 5, where 1 is "least deprived" and 5 is "most deprived"), "region" (9 regions of England, as per Figure 7), and "age." Based on it, and given the life tables for England (for the corresponding period), we can simulate the survival times t_{ij}^P associated to the population hazard.

Next, we can simulate the survival times $t_{ij}^{\rm E}$ associated to the excess hazard with parameters that we detail now. The excess hazard model was



49

50 51



Figure 7. Map of England divided into the 1–9 Government Office Regions, namely, North East, North West, Yorkshire and The Humber, East Midlands, West Midlands, East of England, London, South East, and South West, respectively.

defined as follows

$$h_{E}(t; \mathbf{x}_{ij} \mid \boldsymbol{\theta}, , \boldsymbol{\beta}, \tilde{u}_{i}, u_{i}) = h_{0}(t \exp\{age_{ij} + \tilde{u}_{i}\} \mid \boldsymbol{\theta}) \times \exp\left\{age_{ij}\beta_{1} + \sum_{k=2}^{5} dep_{ij}(k)\beta_{k} + sex_{ij}\beta_{6} + u_{i}\right\},$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)^{\top}$ and $\mathbf{x}_{ij} = (\mathrm{age}_{ij}, \mathrm{dep}_{ij}(2), \mathrm{dep}_{ij}(3), \mathrm{dep}_{ij}(4), \mathrm{dep}_{ij}(5), \mathrm{sex}_{ij})^{\top}$, such that $\mathrm{dep}_{ij}(k)$, for $2 \leq k \leq 5$, is an indicator function for individuals who belong to the k-th deprivation level group (notice that "deprivation level 1" is our reference class). For LN baseline hazard distribution, we set the parameters, according to the parameterization in Web ppendix 1 (Supporting Information), as $\mu = 0.65$ and $\sigma = 1.15$; and for the PGW, we set them as $\eta = 0.5$, $\nu = 3.75$, and $\kappa = 8$. The true coefficients were = 1.0 and $\boldsymbol{\beta} = (1.0, 1.0, 1.0, 1.0, 1.0, 2.0)^{\top}$ in all cases. For the spatial effects, we set both $\tilde{\mathbf{u}}$ and \mathbf{u} as following the IC R model with $\tau_{\tilde{u}} = \tau_u = 10$. Finally, for the excess hazard simulated survival times, we apply two sources of censoring: 1) 1.5-year (50% censoring rate) and 4-year (25% censoring rate) administrative censoring for all individuals

 (that corresponds to the end of the study), and 2) a random censoring given by an Exponential(rate = 0.01) model (that represents the individuals who, for any reason, dropped the study). The final survival times were set as $t_{ij} = \min(t_{ij}^P, t_{ij}^E)$, $\forall i, j$, with the corresponding censoring indicators.

D Data description for pplications

For Section 5, we obtained information on all adult (aged 15–99 years) colon cancer patients (International Classification of Diseases for Oncology, third edition, ICD-O-3 codes 18.0–18.9) diagnosed in England between 2015 and 2016, such that we extracted the data from the National Cancer Registration and nalysis Service (NCR S) data base linked to Hospital Episode Statistics (HES), including basic information on patient, tumour characteristics, and area of residence. ll patients were followed up to update their vital status until 31 December 2018. The data variables available for analysis were sex, age at diagnosis, follow-up time (measured in years from diagnosis), vital status indicator (dead or censored as alive at the end of follow-up), Government Office Region (GOR) of residence at diagnosis, Cancer lliance of residence at diagnosis, deprivation score (based on the Income Domain scores of the 2011 Indices of Multiple Deprivation, IMD), deprivation category (defined according to the quintiles of the IMD Income Domain scores distribution, such that "1" is the *least deprived* group and "2, 3, 4, 5" are the *most* deprived groups), and colon cancer stage at diagnosis ("1" being localised cancer stage, and "2, 3, 4" corresponding to the metastatic cancer stage).

Supporting Information for "Extended Excess Hazard Model for Spatially Dependent Survival Data"

André Victor Ribeiro Amaral †, Francisco Javier Rubio ‡, Manuela Quaresma §, Francisco J. Rodr´guez-Cortés ¶ and Paula Moraga †

- † CEMSE Division, King Abdullah University of Science and Technology. Thuwal, Saudi Arabia.
 - ‡ Department of Statistical Science, University College London. London, UK.
- § Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine. London, UK.
 - ¶ Escuela de Estadística, Universidad Nacional de Colombia. Medellín, Colombia.
 - * Corresponding author. E-mail: andre.ribeiroamaral@kaust.edu.sa



Web Appendix 1 Baseline distributions

Below, we detail the chosen distributions for the baseline function $h_0(\cdot)$. In particular, we define the Log-normal (LN), Log-logistic (LL), Power Generalized Weibull (PGW), Gamma (GAM), and Generalized Gamma (GG) distributions. For each distribution, the probability density function $f(\cdot)$, hazard function $h(\cdot)$, cumulative hazard function $H(\cdot)$, and survival function $S(\cdot)$ are specified as in the following sections.

Log-normal distribution

$$f(t \mid \boldsymbol{\theta}) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left\{ -\frac{(\log(t) - \mu)^2}{2\sigma^2} \right\}, \text{ for } t > 0,$$

$$h(t \mid \boldsymbol{\theta}) = \frac{\frac{1}{t\sigma} \phi\left(\frac{\log t}{\sigma}\right)}{\Phi\left(-\frac{\log x}{\sigma}\right)},$$

$$H(t \mid \boldsymbol{\theta}) = -\log\left(1 - \Phi\left(\frac{\log(t)}{\sigma}\right)\right), \text{ and}$$

$$S(t \mid \boldsymbol{\theta}) = 1 - \Phi\left(\frac{\log(t)}{\sigma}\right),$$

where $\boldsymbol{\theta} = (\mu, \sigma^2)$, such that $\mu \in \mathbb{R}$ and $\sigma^2 > 0$, $\phi(\cdot)$ is the probability density function of the standard Normal distribution, and $\Phi(\cdot)$ is the cumulative distribution function of the standard Normal distribution.

Log-logistic distribution

$$f(t \mid \boldsymbol{\theta}) = \frac{g\left(\frac{\log t\right) - \mu}{\sigma}\right)}{t\sigma}, \text{ for } t > 0,$$

$$h(t \mid \boldsymbol{\theta}) = \frac{g\left(\frac{\log t\right) - \mu}{\sigma}\right)}{t\sigma\left(1 - G\left(\frac{\log t\right) - \mu}{\sigma}\right)},$$

$$H(t \mid \boldsymbol{\theta}) = -\log\left(1 - G\left(\frac{\log(t) - \mu}{\sigma}\right)\right), \text{ and }$$

$$S(t \mid \boldsymbol{\theta}) = 1 - G\left(\frac{\log(t) - \mu}{\sigma}\right),$$

where $\theta = (\mu, \sigma)$, such that $\mu \in \mathbb{R}$ and $\sigma > 0$, $g(t) = \exp\{-t\}(1 + \exp\{-t\})^{-2}$, and $G(t) = (1 + \exp\{-t\})^{-1}$.

Power Generalized Weibull distribution

$$f(t \mid \boldsymbol{\theta}) = \frac{\nu}{\kappa \eta^{\nu}} t^{\nu-1} \left(1 + \left(\frac{t}{\eta} \right)^{\nu} \right)^{\left(-1 \right)} \exp \left\{ 1 - \left(1 + \left(\frac{t}{\eta} \right)^{\nu} \right)^{-} \right\}, \text{ for } t > 0,$$

$$h(t \mid \boldsymbol{\theta}) = \frac{\nu}{\kappa \eta^{\nu}} t^{\nu-1} \left(1 + \left(\frac{t}{\eta} \right)^{\nu} \right)^{\left(-1 \right)},$$

$$H(t \mid \boldsymbol{\theta}) = -1 + \left(1 + \left(\frac{t}{\eta} \right)^{\nu} \right)^{-}, \text{ and}$$

$$S(t \mid \boldsymbol{\theta}) = \exp \left\{ 1 - \left(1 + \left(\frac{t}{\eta} \right)^{\nu} \right)^{-} \right\},$$

where $\theta = (\eta, \nu, \kappa)$, such that $\eta > 0$ is a scale parameter and $\nu, \kappa > 0$ are shape parameters.

Gamma distribution

$$\begin{split} f(t \mid \boldsymbol{\theta}) &= \frac{1}{(\nu)\eta^{\nu}} t^{\nu-1} \exp\left\{-\frac{t}{\eta}\right\}, \text{ for } t > 0, \\ h(t \mid \boldsymbol{\theta}) &= \frac{t^{\nu-1} \exp\left\{-\frac{t}{\eta}\right\}}{\eta^{\nu} \left[-(\nu) - \gamma\left(\nu, \frac{t}{\tau}\right)\right]} \\ H(t \mid \boldsymbol{\theta}) &= -\log\left(1 - \frac{\gamma\left(\nu, \frac{t}{\tau}\right)}{(\nu)}\right), \text{ and} \\ S(t \mid \boldsymbol{\theta}) &= 1 - \frac{\gamma\left(\nu, \frac{t}{\tau}\right)}{(\nu)}, \end{split}$$

where $\theta = (\eta, \nu)$, such that $\eta > 0$ is a scale parameter and $\nu > 0$ is a shape parameter, (\cdot) is gamma function, and $\gamma(\cdot, \cdot)$ is the lower incomplete gamma function.

Generalized Gamma distribution

$$f(t \mid \boldsymbol{\theta}) = \frac{\kappa}{\frac{\nu}{\kappa} \eta^{\nu}} t^{\nu-1} \exp\left\{-\left(\frac{t}{\eta}\right)^{\kappa}\right\}, \text{ for } t > 0,$$

$$h(t \mid \boldsymbol{\theta}) = \frac{t^{\nu-1} \exp\left\{-\left(\frac{t}{\eta}\right)^{\kappa}\right\}}{\eta^{\nu} \left[-\frac{\nu}{\kappa}\right) - \gamma \left(\frac{\nu}{\kappa}, \left(\frac{t}{\eta}\right)^{\kappa}\right)\right]},$$

$$H(t \mid \boldsymbol{\theta}) = -\log\left(1 - \frac{\gamma \left(\frac{\nu}{\kappa}, \left(\frac{t}{\eta}\right)^{\kappa}\right)}{\frac{\nu}{\kappa}\right)}\right), \text{ and}$$

$$S(t \mid \boldsymbol{\theta}) = 1 - \frac{\gamma \left(\frac{\nu}{\kappa}, \left(\frac{t}{\eta}\right)^{\kappa}\right)}{\frac{\nu}{\kappa}},$$

where $\boldsymbol{\theta}=(\eta,\nu,\kappa)$, such that $\eta>0$ is a scale parameter and $\nu,\kappa>0$ are shape parameters, (\cdot) is gamma function, and $\gamma(\cdot,\cdot)$ is the lower incomplete gamma function.

From a practical point of view, it might be useful to use that $F(t \mid \eta, \nu, \kappa) = G(t^{\kappa} \mid \eta^{\kappa}, (\nu/\kappa)),$ such that $F(\cdot)$ is the cumulative distribution function (CDF) of the Generalized Gamma distribution (as defined above), and $G(t \mid \eta, \nu)$ is the CDF of the Gamma distribution with scale and shape parameters given by η and ν , respectively.



Web Appendix 2 Code snippet

As an example, we will fit the RS-SGH model for leukemia-diagnosed patients (Henderson et al., 2002), such that the excess hazard component will be given by the following expression

```
h_{\mathrm{E}}(t; \mathbf{x}_{ij} \mid \boldsymbol{\theta}, , \boldsymbol{\beta}, \tilde{u}_i, u_i) = h_0(t \exp\{\mathrm{age}_{ij} + \tilde{u}_i\} \mid \boldsymbol{\theta}) \exp\{\mathrm{age}_{ij}\beta_1 + \mathrm{wbc}_{ij}\beta_2 + \mathrm{sex}_{ij}\beta_3 + \mathrm{dep}_{ij}\beta_4 + u_i\},
```

where $\boldsymbol{\theta}$ collects the corresponding distribution parameters, $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)^{\top}$ and $\mathbf{x}_{ij} = (\text{age}_{ij}, \text{wbc}_{ij}, \text{sex}_{ij}, \text{dep}_{ij})^{\top}$, such that "wbc" stands for "white blood count," and "dep" corresponds to the Townsend Score (a index of social deprivation, such that higher values indicates less affluent areas). The baseline hazard $h_0(\cdot|\boldsymbol{\theta})$ will be specified according to a log-normal distribution, and the random effects $\tilde{\mathbf{u}}$ and \mathbf{u} will follow the ICAR model.

Based on the code from this repository (https://github.com/avramaral/relative_survival), we can fit such a model as follows:

```
source( header.R ) # load libraries and needed functions
data <- readRDS(file = DATA/leuk.rds ) # load the leukemia data
# Optional
data age <- scale(data age)</pre>
data wbc <- scale(data wbc)</pre>
data dep <- scale(data dep)</pre>
map <- readRDS(file = DATA/nwengland_map.rds ) # load the England map</pre>
adj_info <- adj_list(map = map) # create an object with information about the neighborhood</pre>
    structure
model <- LN_ABST
dist <- gsub(pattern = _ , replacement = _ , x = substring(text = model, first = c(1, 4), last
    = c(3, 7))[1]) # extract the distribution code from model
d \leftarrow data_stan(data = data, model = model, cov.tilde = c(age), cov = c(age, wbc, sex, age)
     dep ), nonlinear = c(), adj_info = adj_info) # create the data object
m <- compile_model(model = model) # compile the Stan model</pre>
r <- fit_stan(mod = m, data = d) # fit the model
```

From the above code, notice that the variable model specifies the fitted model. In particular, we can set it as XXXYYZZ, where XXX specifies the model for the baseline term (the options are LN_, LL_, PGW, GAM, and GG_, such that they correspond to the LN, LL, PGW, Gamma, and GG distributions, respectively), YY specifies the structure for the fixed coefficients (these two characters refer to the vector of coefficients in the time-level and hazard-level components, respectively, such that the allowed combinations are listed in the models.R file), and ZZ specifies the possibly different random effects structures for $\tilde{\mathbf{u}}$ and \mathbf{u} ,

respectively (the letters C and D refer to the IID model, S and T refer to the ICAR model, and Y and Z refer to the BYM2 model. If one does not want to include the random effects structure in either time- or hazard-level components, they can set the corresponding character as X). Table ST1 lists all implemented models (along with the model codes).

Web Table ST1: All implemented models with the corresponding model code. For each of the RS-SGH, RS-SGH-II, RS-SGH-II, RS-SPH, RS-SAFT, RS-GH, RS-PH, RS-AFT, and RS-AH models, we specify the baseline hazard distribution, and the random effects structure.

-#	Model	Code	#	Model	Code	#	Model	Code	#	Model	Code	#	Model	Code
01	RS-SGH LN BYM2	LN_ABYZ	20	RS-SGH LL BYM2	LL_ABYZ	39	RS-SGH PGW BYM2	PGWABYZ	58	RS-SGH Gamma BYM2	GAMABYZ	77	RS-SGH GG BYM2	GG_ABYZ
02	RS-SGH LN ICAR	LN_ABST	21	RS-SGH LL ICAR	LL_ABST	40	RS-SGH PGW ICAR	PGWABST	59	RS-SGH Gamma ICAR	GAMABST	78	RS-SGH GG ICAR	GG_ABST
03	RS-SGH LN IID	LN_ABCD	22	RS-SGH LL IID	LL_ABCD	41	RS-SGH PGW IID	PGWABCD	60	RS-SGH Gamma IID	GAMABCD	79	RS-SGH GG IID	GG_ABCD
04	RS-SGH-I LN BYM2	LN_ABXZ	23	RS-SGH-I LL BYM2	LL_ABXZ	42	RS-SGH-I PGW BYM2	PGWABXZ	61	RS-SGH-I Gamma BYM2	GAMABXZ	80	RS-SGH-I GG BYM2	GG_ABXZ
05	RS-SGH-I LN ICAR	LN_ABXT	24	RS-SGH-I LL ICAR	LL_ABXT	43	RS-SGH-I PGW ICAR	PGWABXT	62	RS-SGH-I Gamma ICAR	GAMABXT	81	RS-SGH-I GG ICAR	GG_ABXT
06 07	RS-SGH-I LN IID RS-SGH-II LN BYM2	LN_ABXD LN_ABYY	25 26	RS-SGH-I LL IID RS-SGH-II LL BYM2	LL_ABXD LL_ABYY	44 45	RS-SGH-II PGW IID RS-SGH-II PGW BYM2	PGWABXD PGWABYY	63	RS-SGH-I Gamma IID RS-SGH-II Gamma BYM2	GAMABXD GAMABYY	82	RS-SGH-I GG IID RS-SGH-II GG BYM2	GG_ABXD
08	RS-SGH-II LN ICAR	LN_ABY1	27	RS-SGH-II LL ICAR	LL_ABSS	46	RS-SGH-II PGW ICAR	PGWABYY	64 65	RS-SGH-II Gamma ICAR	GAMABSS	83 84	RS-SGH-II GG ICAR	GG_ABYY GG_ABSS
09	RS-SGH-II LN IID	LN_ABCC	28	RS-SGH-II LL IID	LL_ABCC	47	RS-SGH-II PGW IID	PGWABCC	66	RS-SGH-II Gamma IID	GAMABCC	85	RS-SGH-II GG ICAR	GG_ABCC
10	RS-SPH LN BYM2	LN_XBXZ	29	RS-SPH LL BYM2	LL_XBXZ	48	RS-SPH PGW BYM2	PGWXBXZ	67	RS-SPH Gamma BYM2	GAMXBXZ	86	RS-SPH GG BYM2	GG_XBXZ
11	RS-SPH LN ICAR	LN_XBXT	30	RS-SPH LL ICAR	LL_XBXT	49	RS-SPH PGW ICAR	PGWXBXT	68	RS-SPH Gamma ICAR	GAMXBXT	87	RS-SPH GG ICAR	GG_XBXT
12	RS-SPH LN IID	LN_XBXD	31	RS-SPH LL IID	LL_XBXD	50	RS-SPH PGW IID	PGWXBXD	69	RS-SPH Gamma IID	GAMXBXD	88	RS-SPH GG IID	GG_XBXD
13 14	RS-SAFT LN BYM2	LN_AAYY LN_AASS	32 33	RS-SAFT LL BYM2 RS-SAFT LL ICAR	LL_AAYY LL_AASS	51 52	RS-SAFT PGW BYM2 RS-SAFT PGW ICAR	PGWAAYY PGWAASS	70 71	RS-SAFT Gamma BYM2 RS-SAFT Gamma ICAR	GAMAAYY GAMAASS	89 90	RS-SAFT GG BYM2 RS-SAFT GG ICAR	GG_AAYY GG_AASS
15	RS-SAFT LN ICAR RS-SAFT LN IID	LN_AACC	34	RS-SAFT LL HD	LL_AACC	53	RS-SAFT PGW ICAR	PGWAACC	72	RS-SAFT Gamma IID	GAMAACC	91	RS-SAFT GG ICAR	GG_AACC
16	RS-GH LN —	LN_ABXX	35	RS-GH LL —	LL_ABXX	54	RS-GH PGW —	PGWABXX	73	RS-GH Gamma —	GAMABXX	92	RS-GH GG —	GG_ABXX
17	RS-PH LN —	LN_XBXX	36	RS-PH LL —	LL_XBXX	55	RS-PH PGW —	PGWXBXX	74	RS-PH Gamma —	GAMXBXX	93	RS-PH GG —	GG_XBXX
18	RS-AFT LN —	LN_AAXX	37	RS-AFT LL —	LL_AAXX	56	RS-AFT PGW —	PGWAAXX	75	RS-AFT Gamma —	GAMAAXX	94	RS-AFT GG —	GG_AAXX
19	RS-AH LN —	LN_AXXX	38	RS-AH LL —	LL_AXXX	57	RS-AH PGW —	PGWAXXX	76	RS-AH Gamma —	GAMAXXX	95	RS-AH GG —	GG_AXXX

Web Appendix 3 RS-SGH data simulation

In this section, we will see how to simulate the survival data from the Relative Survival Spatial General Hazard (RS-SGH) model. To do so, first, we have to simulate the survival times $t_{ij}^{\rm P}$ associated to the population hazard, and second, we have to simulate the survival times $t_{ij}^{\rm E}$ that corresponds to the excess hazard model. Then, at the end, all we have to do is setting $t_{ij} = \min\{t_{ij}^{\rm P}, t_{ij}^{\rm E}\}$. Also, throughout this section, we will assume known covariates for all individuals and life table for the corresponding region and time-window. To simulate (colon and lung) cancer-related covariates in England, one can refer, for example, to Rubio (2022).

To simulate $t_{ij}^{\rm P}$, notice that the problem boils down to simulating from a piecewise constant hazard model, as this is often the information we have available from the life tables. Finally, recall that this is equivalent to simulate from a piecewise exponential (piecewise constant rate) distribution. The rsim.pwexp) function from the SimLT package (Rubio, 2022) implements such a procedure.

Secondly, we will simulate the survival times associated to the excess hazard model. To do so, we will rely on the Probability Integral Transform (PIT). Here, the idea is simulating $z \sim \text{Uniform}(0,1)$, and apply the PIT for the corresponding survival model. This means solving Equation (1) for t_{ij}^{E} :

$$S_{\rm N}(t_{ij}^{\rm E}; \mathbf{x}_{ij} \mid \boldsymbol{\xi} \tilde{u}_i, u_i) = \exp\{ H_{\rm E}(t_{ij}; \mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_i, u_i) \} = (1 \quad z),$$
 (1)

where $\boldsymbol{\xi} = (\boldsymbol{\theta}^{\top}, \quad {}^{\top}, \boldsymbol{\beta}^{\top}, \boldsymbol{\gamma}^{\top})^{\top}$, and $\boldsymbol{\theta}$ collects the corresponding distribution parameters. To do so, we can proceed as follows,

$$\exp \left\{ \begin{array}{cc} H_{\mathrm{E}}(t_{ij}^{\mathrm{E}};\mathbf{x}_{ij} \mid \boldsymbol{\xi}, \tilde{u}_i, u_i) \right\} = (1 \quad z) \\ \Longrightarrow \exp \left\{ \begin{array}{cc} H_0(t_{ij}^{\mathrm{E}} \exp\{\tilde{\mathbf{x}}_{ij}^\top + \tilde{u}_i\} \mid \boldsymbol{\theta}) \exp\{\mathbf{x}_{ij}^\top \boldsymbol{\beta} \quad \tilde{\mathbf{x}}_{ij}^\top + u_i \quad \tilde{u}_i\} \right\} = (1 \quad z) \\ \Longrightarrow S_0(t_{ij}^{\mathrm{E}} \exp\{\tilde{\mathbf{x}}_{ij}^\top + \tilde{u}_i\} \mid \boldsymbol{\theta})^{\exp\{\begin{array}{cc} i_j \boldsymbol{\beta} & \tilde{i}_{ij} + u_i & \tilde{u}_i \} \\ \end{array} \right\} = (1 \quad z). \end{array}$$

This implies that

$$\Rightarrow 1 \quad F_0(t_{ij}^{\mathrm{E}} \exp\{\tilde{\mathbf{x}}_{ij}^{\top} + \tilde{u}_i\} \mid \boldsymbol{\theta}) = \exp\left\{\frac{\log(1 \quad z)}{\exp\{\mathbf{x}_{ij}^{\top}\boldsymbol{\beta} \quad \tilde{\mathbf{x}}_{ij}^{\top} + u_i \quad \tilde{u}_i\}}\right\}$$

$$\Rightarrow 1 \quad \exp\left\{\frac{\log(1 \quad z)}{\exp\{\mathbf{x}_{ij}^{\top}\boldsymbol{\beta} \quad \tilde{\mathbf{x}}_{ij}^{\top} + u_i \quad \tilde{u}_i\}}\right\} = F_0(t_{ij}^{\mathrm{E}} \exp\{\tilde{\mathbf{x}}_{ij}^{\top} + \tilde{u}_i\} \mid \boldsymbol{\theta})$$

$$\Rightarrow F_0^{-1} \left[1 \quad \exp\left\{\frac{\log(1 \quad z)}{\exp\{\mathbf{x}_{ij}^{\top}\boldsymbol{\beta} \quad \tilde{\mathbf{x}}_{ij}^{\top} + u_i \quad \tilde{u}_i\}}\right\} \mid \boldsymbol{\theta}\right] = t_{ij}^{\mathrm{E}} \exp\{\tilde{\mathbf{x}}_{ij}^{\top} + \tilde{u}_i\},$$

such that $F_0^{-1}(\cdot;\boldsymbol{\theta})$ is the quantile function for the baseline hazard model. Lastly,

$$t_{ij}^{\mathrm{E}} = \frac{F_0^{-1} \begin{bmatrix} 1 & \exp\left\{\log(1-z)\exp\{\tilde{\mathbf{x}}_{ij}^\top & \mathbf{x}_{ij}^\top\boldsymbol{\beta} + \tilde{u}_i & u_i\}\right\} \mid \boldsymbol{\theta} \end{bmatrix}}{t_{ij}\exp\{\tilde{\mathbf{x}}_{ij}^\top + \tilde{u}_i\}}.$$

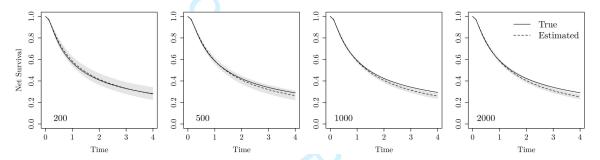
As mentioned before, once we have simulated t_{ij}^{P} and t_{ij}^{E} , we simply set $t_{ij} = \min(t_{ij}^{\mathrm{P}}, t_{ij}^{\mathrm{E}}), \forall i, j.$

Analysis of marginal quantities (Simulation) Web Appendix 4

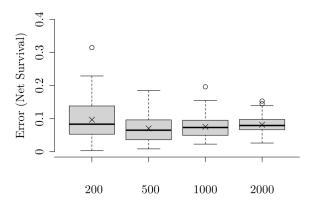
This section presents complementary results for Section 4.1. Table ST2 shows the computational cost for fitting the models. Figures SF1 and SF3 show the estimated and true net survival curves for different scenarios, and SF2 and SF4 plot the corresponding errors based on these simulated data sets.

Web Table ST2: Fitting time for all simulated scenarios for Section 4.1. "Fitting time (sec.)" presents the average time (in seconds) to fit the corresponding model based on the 100 simulated data sets for each scenario. The models were fitted on a Intel-Xeon Gold 6230R CPU at 2.10 Ghz.

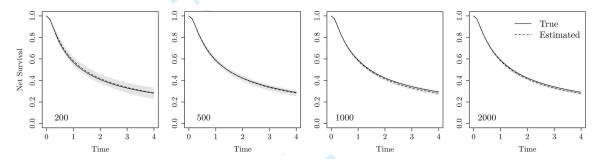
#	Data Generating model	Cens. rate	Sample size	Fitted model	Fitting time (sec.)	#	Data Generating model	Cens. rate	Sample size	Fitted model	Fitting time (sec.)
01	RS-SGH LN ICAR	25%	200	RS-SGH LN ICAR	816.46	13	RS-SGH PGW ICAR	25%	200	RS-SGH LN ICAR	639.69
02	RS-SGH LN ICAR	25%	500	RS-SGH LN ICAR	1269.51	14	RS-SGH PGW ICAR	25%	500	RS-SGH LN ICAR	988.41
03	RS-SGH LN ICAR	25%	1000	RS-SGH LN ICAR	1873.57	15	RS-SGH PGW ICAR	25%	1000	RS-SGH LN ICAR	1600.08
04	RS-SGH LN ICAR	25%	2000	RS-SGH LN ICAR	2612.08	16	RS-SGH PGW ICAR	25%	2000	RS-SGH LN ICAR	2198.06
05	RS-SGH LN ICAR	25%	200	RS-SGH PGW ICAR	1139.34	17	RS-SGH PGW ICAR	25%	200	RS-SGH PGW ICAR	848.25
06	RS-SGH LN ICAR	25%	500	RS-SGH PGW ICAR	1849.68	18	RS-SGH PGW ICAR	25%	500	RS-SGH PGW ICAR	1271.65
07	RS-SGH LN ICAR	25%	1000	RS-SGH PGW ICAR	2618.80	19	RS-SGH PGW ICAR	25%	1000	RS-SGH PGW ICAR	1989.32
08	RS-SGH LN ICAR	25%	2000	RS-SGH PGW ICAR	3535.60	20	RS-SGH PGW ICAR	25%	2000	RS-SGH PGW ICAR	2374.49
09	RS-SGH LN ICAR	50%	200	RS-SGH LN ICAR	908.27	21	RS-SGH PGW ICAR	50%	200	RS-SGH LN ICAR	674.08
10	RS-SGH LN ICAR	50%	500	RS-SGH LN ICAR	1391.55	22	RS-SGH PGW ICAR	50%	500	RS-SGH LN ICAR	1039.07
11	RS-SGH LN ICAR	50%	1000	RS-SGH LN ICAR	2016.23	23	RS-SGH PGW ICAR	50%	1000	RS-SGH LN ICAR	1654.65
10	DC CCILIN ICAD	E007	2000	DC CCILIN ICAD	0720.00	0.4	DC CCII DCW ICAD	E007	500	DC CCILINICAD	0214 54



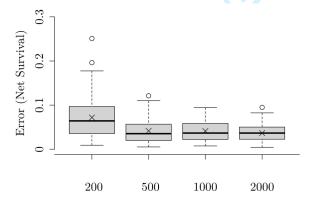
Web Figure SF1: True and estimated (along with a 95% equal-tailed credible interval) net survival curves based on the fitted RS-SGH LN ICAR model. The data were generated from the RS-SGH PGW ICAR model with 50% censoring rate and sample size set to 200, 500, 1000, and 2000 patients. Such estimates were obtained by averaging over the 100 simulated data sets and all regions for each scenario (the corresponding uncertainty was computed based on the percentiles for the curves that average the regions' net survival).



Web Figure SF2: Error in estimating the true net survival function based on the fitted RS-SGH LN ICAR model. The data were generated from the RS-SGH PGW ICAR model with 50% censoring rate and sample size set to 200, 500, 1000, and 2000 patients. The computed errors aggregate the 100 simulated data sets and all regions for each scenario. The crosses (\times) correspond to the boxplot values mean.



Web Figure SF3: True and estimated (along with a 95% equal-tailed credible interval) net survival curves based on the fitted RS-SGH PGW ICAR model. The data were generated from the same model with 25% censoring rate and sample size set to 200, 500, 1000, and 2000 patients. Such estimates were obtained by averaging over the 100 simulated data sets and all regions for each scenario (the corresponding uncertainty was computed based on the percentiles for the curves that average the regions' net survival).



Web Figure SF4: Error in estimating the true net survival function based on the fitted RS-SGH PGW ICAR model. The data were generated from the same model with 25% censoring rate and sample size set to 200, 500, 1000, and 2000 patients. The computed errors aggregate the 100 simulated data sets and all regions for each scenario. The crosses (×) correspond to the boxplot values mean.

Web Appendix 5 Model selection (Simulation)

This section presents complementary material for Section 4.2. Table ST3 lists all scenarios for which we will simulate data and fit the corresponding model. Then, based on such results, we will compare the equivalent scenarios with respect to the estimated $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$. Also, tables ST4, ST5, and ST6 report the "best-model proportions" for all scenarios with sample size set to 200, 500, and 1000 patients, respectively.

Web Table ST3: All simulated scenarios for Section 4.2. "Data Generating model" refers to the model assumed for the data generating procedure, and "Fitted model" defines the baseline hazard distribution and the random effects structure. For all scenarios, the censoring rate is assumed to be 25%.

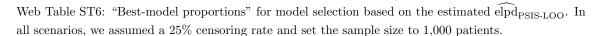
#	Data Generating model	Sample size	Fitted model	#	Data Generating model	Sample size	Fitted model
01	RS-SGH LN ICAR	200	RS-SGH LN —	33	RS-SGH PGW ICAR	200	RS-SGH LN —
02	RS-SGH LN ICAR	500	RS-SGH LN —	34	RS-SGH PGW ICAR	500	RS-SGH LN —
03	RS-SGH LN ICAR	1000	RS-SGH LN —	35	RS-SGH PGW ICAR	1000	RS-SGH LN —
04	RS-SGH LN ICAR	2000	RS-SGH LN —	36	RS-SGH PGW ICAR	2000	RS-SGH LN —
05	RS-SGH LN ICAR	200	RS-SGH LN IID	37	RS-SGH PGW ICAR	200	RS-SGH LN IID
06	RS-SGH LN ICAR	500	RS-SGH LN IID	38	RS-SGH PGW ICAR	500	RS-SGH LN IID
07	RS-SGH LN ICAR	1000	RS-SGH LN IID	39	RS-SGH PGW ICAR	1000	RS-SGH LN IID
08	RS-SGH LN ICAR	2000	RS-SGH LN IID	40	RS-SGH PGW ICAR	2000	RS-SGH LN IID
09	RS-SGH LN ICAR	200	RS-SGH LN ICAR	41	RS-SGH PGW ICAR	200	RS-SGH LN ICAR
10	RS-SGH LN ICAR	500	RS-SGH LN ICAR	42	RS-SGH PGW ICAR	500	RS-SGH LN ICAR
11	RS-SGH LN ICAR	1000	RS-SGH LN ICAR	43	RS-SGH PGW ICAR	1000	RS-SGH LN ICAR
12	RS-SGH LN ICAR	2000	RS-SGH LN ICAR	44	RS-SGH PGW ICAR	500	RS-SGH LN ICAR
13	RS-SGH LN ICAR	200	RS-SGH LN BYM2	45	RS-SGH PGW ICAR	200	RS-SGH LN BYM2
14	RS-SGH LN ICAR	500	RS-SGH LN BYM2	46	RS-SGH PGW ICAR	500	RS-SGH LN BYM2
15	RS-SGH LN ICAR	1000	RS-SGH LN BYM2	47	RS-SGH PGW ICAR	1000	RS-SGH LN BYM2
16	RS-SGH LN ICAR	2000	RS-SGH LN BYM2	48	RS-SGH PGW ICAR	2000	RS-SGH LN BYM2
17	RS-SGH LN ICAR	200	RS-SGH PGW —	49	RS-SGH PGW ICAR	200	RS-SGH PGW —
18	RS-SGH LN ICAR	500	RS-SGH PGW —	50	RS-SGH PGW ICAR	500	RS-SGH PGW —
19	RS-SGH LN ICAR	1000	RS-SGH PGW —	51	RS-SGH PGW ICAR	1000	RS-SGH PGW —
20	RS-SGH LN ICAR	2000	RS-SGH PGW —	52	RS-SGH PGW ICAR	2000	RS-SGH PGW —
21	RS-SGH LN ICAR	200	RS-SGH PGW IID	53	RS-SGH PGW ICAR	200	RS-SGH PGW IID
22	RS-SGH LN ICAR	500	RS-SGH PGW IID	54	RS-SGH PGW ICAR	500	RS-SGH PGW IID
23	RS-SGH LN ICAR	1000	RS-SGH PGW IID	55	RS-SGH PGW ICAR	1000	RS-SGH PGW IID
24	RS-SGH LN ICAR	2000	RS-SGH PGW IID	56	RS-SGH PGW ICAR	500	RS-SGH PGW IID
25	RS-SGH LN ICAR	200	RS-SGH PGW ICAR	57	RS-SGH PGW ICAR	200	RS-SGH PGW ICAR
26	RS-SGH LN ICAR	500	RS-SGH PGW ICAR	58	RS-SGH PGW ICAR	500	RS-SGH PGW ICAR
27	RS-SGH LN ICAR	1000	RS-SGH PGW ICAR	59	RS-SGH PGW ICAR	1000	RS-SGH PGW ICAR
28	RS-SGH LN ICAR	2000	RS-SGH PGW ICAR	60	RS-SGH PGW ICAR	2000	RS-SGH PGW ICAR
29	RS-SGH LN ICAR	200	RS-SGH PGW BYM2	61	RS-SGH PGW ICAR	200	RS-SGH PGW BYM2
30	RS-SGH LN ICAR	500	RS-SGH PGW BYM2	62	RS-SGH PGW ICAR	500	RS-SGH PGW BYM2
31	RS-SGH LN ICAR	1000	RS-SGH PGW BYM2	63	RS-SGH PGW ICAR	1000	RS-SGH PGW BYM2
32	RS-SGH LN ICAR	2000	RS-SGH PGW BYM2	64	RS-SGH PGW ICAR	2000	RS-SGH PGW BYM2

Web Table ST4: "Best-model proportions" for model selection based on the estimated $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$. In all scenarios, we assumed a 25% censoring rate and set the sample size to 200 patients.

#	Data Generating model	Fitted model	Best-model proportions	#	Data Generating model	Fitted model	Best-model proportions
01	RS-SGH LN ICAR	RS-SGH LN —	30%	09	RS-SGH PGW ICAR	RS-SGH LN —	26%
02	RS-SGH LN ICAR	RS-SGH LN IID	7%	10	RS-SGH PGW ICAR	RS-SGH LN IID	8%
03	RS-SGH LN ICAR	RS-SGH LN ICAR	22%	11	RS-SGH PGW ICAR	RS-SGH LN ICAR	24%
04	RS-SGH LN ICAR	RS-SGH LN BYM2	41%	12	RS-SGH PGW ICAR	RS-SGH LN BYM2	42%
05	RS-SGH LN ICAR	RS-SGH PGW —	32%	13	RS-SGH PGW ICAR	RS-SGH PGW —	30%
06	RS-SGH LN ICAR	RS-SGH PGW IID	22%	14	RS-SGH PGW ICAR	RS-SGH PGW IID	21%
07	RS-SGH LN ICAR	RS-SGH PGW ICAR	24%	15	RS-SGH PGW ICAR	RS-SGH PGW ICAR	25%
08	RS-SGH LN ICAR	RS-SGH PGW BYM2	22%	16	RS-SGH PGW ICAR	RS-SGH PGW BYM2	24%

Web Table ST5: "Best-model proportions" for model selection based on the estimated $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$. In all scenarios, we assumed a 25% censoring rate and set the sample size to 500 patients.

#	Data Generating model	Fitted model	Best-model proportions	#	Data Generating model	Fitted model	Best-model proportions
01	RS-SGH LN ICAR	RS-SGH LN —	18%	09	RS-SGH PGW ICAR	RS-SGH LN —	10%
02	RS-SGH LN ICAR	RS-SGH LN IID	3%	10	RS-SGH PGW ICAR	RS-SGH LN IID	20%
03	RS-SGH LN ICAR	RS-SGH LN ICAR	32%	11	RS-SGH PGW ICAR	RS-SGH LN ICAR	23%
04	RS-SGH LN ICAR	RS-SGH LN BYM2	47%	12	RS-SGH PGW ICAR	RS-SGH LN BYM2	47%
05	RS-SGH LN ICAR	RS-SGH PGW —	24%	13	RS-SGH PGW ICAR	RS-SGH PGW —	16%
06	RS-SGH LN ICAR	RS-SGH PGW IID	23%	14	RS-SGH PGW ICAR	RS-SGH PGW IID	22%
07	RS-SGH LN ICAR	RS-SGH PGW ICAR	29%	15	RS-SGH PGW ICAR	RS-SGH PGW ICAR	30%
08	RS-SGH LN ICAR	RS-SGH PGW BYM2	24%	16	RS-SGH PGW ICAR	RS-SGH PGW BYM2	32%



#	Data Generating model	Fitted model	Best-model proportions	#	Data Generating model	Fitted model	Best-model proportions
01	RS-SGH LN ICAR	RS-SGH LN —	4%	09	RS-SGH PGW ICAR	RS-SGH LN —	1%
02	RS-SGH LN ICAR	RS-SGH LN IID	8%	10	RS-SGH PGW ICAR	RS-SGH LN IID	35%
03	RS-SGH LN ICAR	RS-SGH LN ICAR	47%	11	RS-SGH PGW ICAR	RS-SGH LN ICAR	18%
04	RS-SGH LN ICAR	RS-SGH LN BYM2	41%	12	RS-SGH PGW ICAR	RS-SGH LN BYM2	46%
05	RS-SGH LN ICAR	RS-SGH PGW —	14%	13	RS-SGH PGW ICAR	RS-SGH PGW —	3%
06	RS-SGH LN ICAR	RS-SGH PGW IID	5%	14	RS-SGH PGW ICAR	RS-SGH PGW IID	18%
07	RS-SGH LN ICAR	RS-SGH PGW ICAR	36%	15	RS-SGH PGW ICAR	RS-SGH PGW ICAR	49%
08	RS-SGH LN ICAR	RS-SGH PGW BYM2	45%	16	RS-SGH PGW ICAR	RS-SGH PGW BYM2	30%

Web Appendix 6 Analysis of the spatial e ects (Simulation)

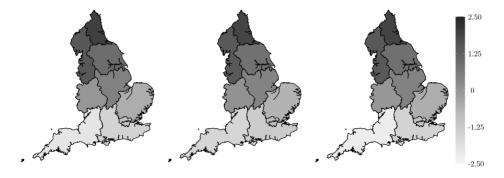
This section presents complementary results for Section 4.3. In Figures SF5 and SF6, we plot the true and estimated spatial effects for models RS-SGH LN IID and RS-SGH LN BYM2, respectively. Similarly, Table ST8 show all estimated parameters (with a 95% equal-tail credible interval) for models RS-SGH LN IID, RS-SGH LN ICAR, and RS-SGH LN BYM2. Moreover, Figures SF7, SF8, and SF9 present the estimated relative exceedance probabilities for the same models as before. Lastly, Figures SF10, SF11, and SF12, show the estimated posterior of $\sigma_u = 1/\sqrt{\tau_u}$ (same for $\sigma_{\tilde{u}} = 1/\sqrt{\tau_{\tilde{u}}}$) for the ICAR, IID, and BYM2 random effects, respectively, when fitting the RS-SGH LN model.

Web Table ST7: Results for the competing models in Section 4.3 according to the $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$. The $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$ difference (with standard error) represents the pairwise difference between the others models and the reference model (RS-SGH LN ICAR).

Model	RS-SGH LN ICAR	RS-SGH LN IID	RS-SGH LN BYM2	RS-SGH LN —
elpd _{PSIS-LOO} difference (SE)	0.0 (0.0)	-0.1 (1.2)	-0.3 (0.5)	-3187.7 (67.6)



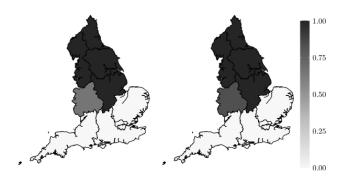
Web Figure SF5: Spatial effects for the RS-SGH LN IID model presented in Section 4.3. Left panel: True spatial effects $\tilde{\mathbf{u}} = \mathbf{u} = (2.0, 1.5, 1.0, 0.5, 0, 0.5, 1.0, 1.5, 2.0)^{\top}$. Middle panel: Estimated time-level spatial effects $\tilde{\mathbf{u}} = (1.94, 1.54, 0.90, 0.52, 0.15, 0.56, 0.78, 1.23, 1.51)^{\top}$. Right panel: Estimated hazard-level spatial effects $\mathbf{u} = (1.86, 1.44, 0.86, 0.43, 0.07, 0.56, 1.12, 1.54, 2.27)^{\top}$.



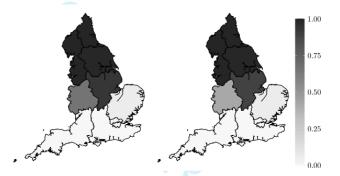
Web Figure SF6: Spatial effects for the RS-SGH LN BYM2 model presented in Section 4.3. Left panel: True spatial effects $\tilde{\mathbf{u}} = \mathbf{u} = (2.0, 1.5, 1.0, 0.5, 0, 0.5, 1.0, 1.5, 2.0)^{\top}$. Middle panel: Estimated time-level spatial effects $\tilde{\mathbf{u}} = (1.90, 1.49, 0.85, 0.46, 0.09, 0.63, 0.85, 1.29, 1.59)^{\top}$. Right panel: Estimated hazard-level spatial effects $\mathbf{u} = (1.91, 1.49, 0.91, 0.48, 0.01, 0.51, 1.07, 1.49, 2.21)^{\top}$.

Web Table ST8: Estimated parameters (with standard deviation and a 95% equal-tail credible interval) for the RS-SGH LN model with IID, ICAR, and BYM2 random effects, presented in the simulation study (Section 4.3).

Random Effect	Parameter	Mean	SD	95% equal-tail CI	Parameter	Mean	SD	95% equal-tail CI
IID	1	0.83	0.03	(0.76; 0.90)	β_5	0.96	0.04	(1.04; 0.87)
IID	β_1	0.88	0.02	(0.85; 0.91)	β_6	1.98	0.03	(1.92; 2.05)
IID	β_2	0.90	0.05	(0.93; 0.81)	μ	0.55	0.36	(0.14; 1.27)
IID	β_3	0.97	0.05	(1.06; 0.88)	σ	1.15	0.05	(1.06; 1.25)
IID	β_4	0.97	0.04	(1.06; 0.88)		_		
IID	\tilde{u}_1	1.94	0.36	(1.26; 2.68)	u_1	1.86	0.36	(1.17; 2.59)
IID	\tilde{u}_2	1.54	0.35	(0.86; 2.26)	u_2	1.44	0.36	(0.75; 2.17)
IID	\tilde{u}_3	0.90	0.36	(0.22; 1.63)	u_3	0.86	0.36	(0.17; 1.59)
IID	\tilde{u}_4	0.52	0.36	(0.17; 1.26)	u_4	0.43	0.36	(0.26; 1.16)
IID	\tilde{u}_5	0.15	0.36	(0.55; 0.89)	u_5	0.07	0.36	(0.75; 0.67)
IID	\tilde{u}_6	0.56	0.37	(1.27; 0.19)	u_6	0.56	0.36	(1.26; 0.16)
IID	\tilde{u}_7	0.78	0.37	(1.50; 0.02)	u_7	1.12	0.36	(1.81; 0.39)
IID	\tilde{u}_8	1.23	0.37	(1.93; 0.49)	u_8	1.54	0.36	(2.24; 0.81)
IID	$ ilde{u}_9$	1.51	0.40	(2.31; 0.71)	u_9	2.27	0.37	(2.97; 1.52)
ICAR	1	0.84	0.03	(0.77; 0.91)	β_5	0.97	0.04	(1.05; 0.88)
ICAR	β_1	0.88	0.02	(0.85; 0.91)	β_6	1.99	0.03	(1.92; 2.05)
ICAR	β_2	0.91	0.05	(1.00; 0.82)	μ	0.65	0.04	(0.57; 0.74)
ICAR	β_3	0.98	0.04	(1.07; 0.89)	σ	1.21	0.02	(1.18; 1.25)
ICAR	β_4	0.98	0.04	(1.07; 0.89)		_		_
ICAR	\tilde{u}_1	1.86	0.09	(1.69; 2.04)	u_1	1.96	0.04	(1.89; 2.04)
ICAR	$ ilde{u}_2$	1.45	0.07	(1.32; 1.58)	u_2	1.54	0.03	(1.48; 1.60)
ICAR	\tilde{u}_3	0.81	0.08	(0.66; 0.97)	u_3	0.97	0.03	(0.90; 1.03)
ICAR	$ ilde{u}_4$	0.41	0.09	(0.23; 0.60)	u_4	0.53	0.04	(0.46; 0.61)
ICAR	\tilde{u}_5	0.04	0.10	(0.15; 0.23)	u_5	0.04	0.04	(0.04; 0.11)
ICAR	\tilde{u}_6	0.69	0.10	(0.88; 0.49)	u_6	0.45	0.04	(0.53; 0.37)
ICAR	\tilde{u}_7	0.90	0.12	(1.14; 0.65)	u_7	1.01	0.05	(1.10; 0.91)
ICAR	$ ilde{u}_8$	1.33	0.11	(1.55; 1.11)	u_8	1.43	0.05	(1.52; 1.34)
ICAR	$ ilde{u}_9$	1.65	0.19	(2.01; 1.28)	u_9	2.15	0.07	(2.29; 2.01)
BYM2	1	0.83	0.03	(0.77; 0.90)	β_5	0.96	0.04	(1.04; 0.88)
BYM2	β_1	0.88	0.02	(0.85; 0.91)	β_6	1.99	0.03	(1.92; 2.05)
BYM2	β_2	0.91	0.05	(1.00; 0.82)	μ	0.60	0.12	(0.32; 0.81)
BYM2	β_3	0.97	0.05	(1.06; 0.88)	σ	1.18	0.04	(1.10; 1.26)
BYM2	β_4	0.98	0.04	(1.06; 0.89)		_		
BYM2	$\tilde{ ho}$	0.81	0.22	(0.20; 1.00)	ρ	0.82	0.21	(0.22; 1.00)
BYM2	\tilde{u}_1	1.90	0.14	(1.66; 2.20)	u_1	1.91	0.12	(1.63; 2.12)
BYM2	\tilde{u}_2	1.49	0.13	(1.27; 1.78)	u_2	1.49	0.12	(1.21; 1.70)
BYM2	\tilde{u}_3	0.85	0.13	(0.61; 1.14)	u_3	0.91	0.12	(0.63; 1.12)
BYM2	\tilde{u}_4	0.46	0.14	(0.21; 0.78)	u_4	0.48	0.12	(0.19; 0.69)
BYM2	\tilde{u}_5	0.09	0.15	(0.18; 0.41)	u_5	0.01	0.12	(0.30; 0.20)
BYM2	\tilde{u}_6	0.63	0.16	(0.91; 0.29)	u_6	0.51	0.13	(0.80; 0.29)
BYM2	\tilde{u}_7	0.85	0.17	(1.16; 0.48)	u_7	1.07	0.13	(1.37; 0.85)
BYM2	\tilde{u}_8	1.29	0.16	(1.57; 0.95)	u_8	1.49	0.13	(1.79; 1.27)
BYM2	\tilde{u}_9	1.59	0.22	(2.01; 1.14)	u_9	2.21	0.14	(2.54; 1.97)
			-	•	•			•



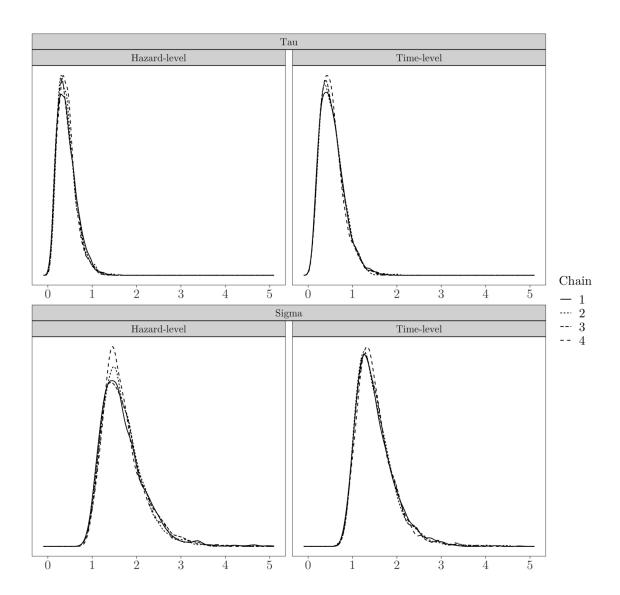
Web Figure SF7: Relative exceedance probability for the RS-SGH LN ICAR model presented in Section 4.3. Left panel: Estimated time-level relative exceedance probabilities $f(\tilde{\mathbf{u}}) = (1, 1, 1, 1, 0.654750, 0, 0, 0, 0)^{\top}$, such that $f(\tilde{u}_i) = (\tilde{u}_i > 0)$, $\forall i$. Right panel: Estimated hazard-level relative exceedance probabilities $f(\mathbf{u}) = (1, 1, 1, 1, 0.841625, 0, 0, 0, 0)^{\top}$, such that $f(u_i) = (u_i > 0)$, $\forall i$.



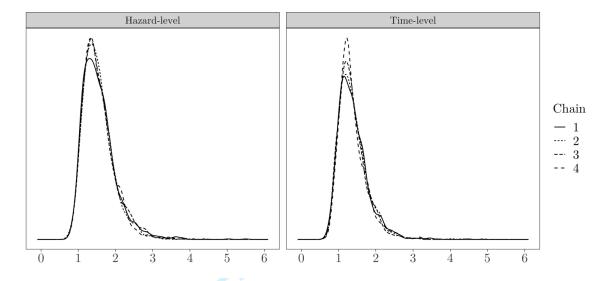
Web Figure SF8: Relative exceedance probability for the RS-SGH LN IID model presented in Section 4.3. Left panel: Estimated time-level relative exceedance probabilities $f(\tilde{\mathbf{u}}) = (1, 1, 0.992375, 0.932875, 0.663250, 0.061500, 0.023125, 0.001375, 0.000375)^{\top}$, such that $f(\tilde{u}_i) = (\tilde{u}_i > 0)$, $\forall i$. Right panel: Estimated hazard-level relative exceedance probabilities $f(\mathbf{u}) = (1, 1, 0.990875, 0.889750, 0.422500, 0.060500, 0.001625, 0, 0)^{\top}$, such that $f(u_i) = (u_i > 0)$, $\forall i$.



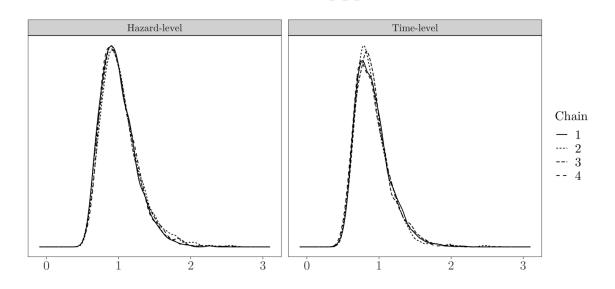
Web Figure SF9: Relative exceedance probability for the RS-SGH LN BYM2 model presented in Section 4.3. Left panel: Estimated time-level relative exceedance probabilities $f(\tilde{\mathbf{u}}) = (1, 1, 1, 0.998875, 0.715125, 0.001125, 0.000125, 0, 0)^{\top}$, such that $f(\tilde{u}_i) = (\tilde{u}_i > 0)$, $\forall i$. Right panel: Estimated hazard-level relative exceedance probabilities $f(\mathbf{u}) = (1, 1, 1, 0.997375, 0.523250, 0.000125, 0, 0, 0)^{\top}$, such that $f(u_i) = (u_i > 0)$, $\forall i$.



Web Figure SF10: Estimated posterior densities for τ_u (as it appears in the **hazard level** component) and $\tau_{\tilde{u}}$ (as it appears in the **time level** component) for the ICAR random effects when fitting the RS-SGH LN model. However, aiming to make these results comparable with the estimates showed in Figures SF11 and SF12 (IID and BYM2 random effects, respectively), we also display the estimated posterior density for $\sigma_u = 1/\sqrt{\tau_u}$ and $\sigma_{\tilde{u}} = 1/\sqrt{\tau_{\tilde{u}}}$. The curves are plotted separately for each posterior chain.



Web Figure SF11: Estimated posterior densities for σ_u (as it appears in the hazard level component) and $\sigma_{\tilde{u}}$ (as it appears in the **time level** component) for the IID random effects when fitting the RS-SGH LN model. The curves are plotted separately for each posterior chain.



Web Figure SF12: Estimated posterior densities for σ_u (as it appears in the hazard level component) and $\sigma_{\tilde{u}}$ (as it appears in the **time level** component) for the BYM2 random effects when fitting the RS-SGH LN model. The curves are plotted separately for each posterior chain.

Web Appendix 7 Case study

This section presents complementary results for Section 5. Table ST9 shows the ranked models (according to the $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$ criterion), Table ST10 displays the fitting time for all models, and Table ST11 presents the estimated hyperparameters of the random effects for the highest-ranked BYM2 models in the first case study. Figures SF13 and SF14 report the uncertainty for the estimated net survival, such that t=3 years. Figures SF15, SF16, and SF17 present similar results (including the net survival point estimate), however, for t=1 year. Figures SF23, SF24, SF25, SF29, SF30 and SF31 report the net survival estimates with data stratified by "deprivation level," and Figures SF26, SF27, SF28, SF32, SF33, and SF34 report equivalent results with data stratified by "cancer stage." Lastly, Figure SF18 shows the England map according to the Cancer Alliance Regions.

Web Table ST9: Results for the competing models in Section 5 according to the $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$. The $\widehat{\text{elpd}}_{\text{PSIS-LOO}}$ difference (with standard error) represents the pairwise difference between the others models and the reference model. The "Geography 01" refers to the Government Office Regions and the "Geography 02" refers to the Cancer Alliances Regions in 2016. The strikethrough scenario corresponds to the fitted model for which we did not observe well mixed posterior chains.

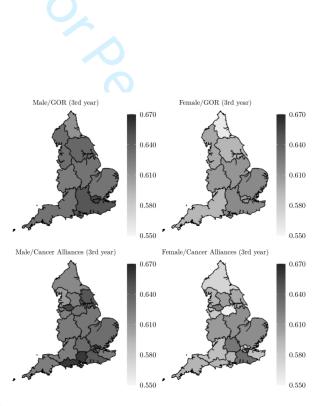
Sex	Geography	Model	$\widehat{\mathrm{elpd}}_{\mathrm{PSIS-LOO}}$ diff. (SE)	Sex	Geography	Model	elpd _{PSIS-LOO} diff. (SE)
Male	01	RS-SGH LN BYM2	0.0 (0.0)	Female	01	RS-SGH LN BYM2	0.0 (0.0)
Male	01	RS-SGH LN ICAR	-1.4 (1.4)	Female	01	RS-SGH LN ICAR	-0.3 (1.4)
Male	01	RS-SGH PGW ICAR	-14.6 (3.6)	Female	01	RS-SGH LL BYM2	-9.8 (4.0)
Male	01	RS-SGH PGW BYM2	-19.7 (5.4)	Female	01	RS-SGH PGW BYM2	-9.8 (2.4)
Male	01	RS-SGH LL ICAR	-24.6 (6.5)	Female	01	RS-SGH PGW ICAR	-20.6 (3.5)
Male	01	RS-SGH-LL-BYM2	-21284.1 (576.0)	Female	01	RS-SGH LL ICAR	-42.0 (5.8)
Male	02	RS-SGH LN BYM2	0.0 (0.0)	Female	02	RS-SGH LN BYM2	0.0 (0.0)
Male	02	RS-SGH LN ICAR	-2.1 (1.8)	Female	02	RS-SGH LN ICAR	-2.5 (1.8)
Male	02	RS-SGH PGW ICAR	-15.4 (3.8)	Female	02	RS-SGH PGW BYM2	-9.9 (2.3)
Male	02	RS-SGH PGW BYM2	-16.0 (4.6)	Female	02	RS-SGH PGW ICAR	-12.0 (3.0)
Male	02	RS-SGH LL BYM2	-23.3 (6.3)	Female	02	RS-SGH LL BYM2	-41.6 (6.0)
Male	02	RS-SGH LN ICAR	-25.2 (6.6)	Female	02	RS-SGH LL ICAR	-44.2 (6.2)

Web Table ST10: Fitting time for all models in Section 5. "Fitting time (sec.)" presents the time (in seconds) to fit the corresponding model. The models are ordered as in Table ST9. The models were fitted on a AMD EPYC 7402 CPU at 2.8 Ghz.

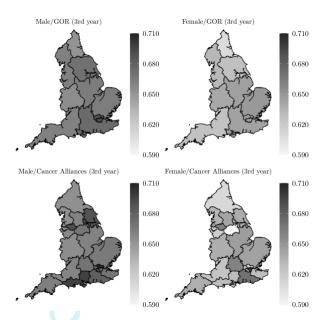
Sex	Patients' location	Model	Fitting time (sec.)	Sex	Patients' location	Model	Fitting time (sec.)
Male	01	RS-SGH LN BYM2	70950.50	Female	01	RS-SGH LN BYM2	63060.69
Male	01	RS-SGH LN ICAR	24974.23	Female	01	RS-SGH LN ICAR	27092.05
Male	01	RS-SGH PGW ICAR	35546.03	Female	01	RS-SGH LL BYM2	143318.70
Male	01	RS-SGH PGW BYM2	134614.80	Female	01	RS-SGH PGW BYM2	83139.05
Male	01	RS-SGH LL ICAR	77084.76	Female	01	RS-SGH PGW ICAR	44532.04
Male	01	RS-SGH LL BYM2	154491.00	Female	01	RS-SGH LL ICAR	158862.70
Male	02	RS-SGH LN BYM2	52695.24	Female	02	RS-SGH LN BYM2	45791.73
Male	02	RS-SGH LN ICAR	15219.16	Female	02	RS-SGH LN ICAR	14379.65
Male	02	RS-SGH PGW ICAR	21843.38	Female	02	RS-SGH PGW BYM2	60938.90
Male	02	RS-SGH PGW BYM2	68693.83	Female	02	RS-SGH PGW ICAR	17720.56
Male	02	RS-SGH LL BYM2	113003.30	Female	02	RS-SGH LL BYM2	99978.17
Male	02	RS-SGH LN ICAR	44572.72	Female	02	RS-SGH LL ICAR	38042.61

Web Table ST11: Estimated parameters (with standard deviation and a 95% equal-tail credible interval) for the highest-ranked BYM2 models in Section 5. We reported $\sigma_u = 1/\sqrt{\tau_u}$ (and $\sigma_{\tilde{u}} = 1/\sqrt{\tau_{\tilde{u}}}$) and ρ (and $\tilde{\rho}$). The "Geography 01" refers to the Government Office Regions and the "Geography 02" refers to the Cancer Alliances Regions in 2016.

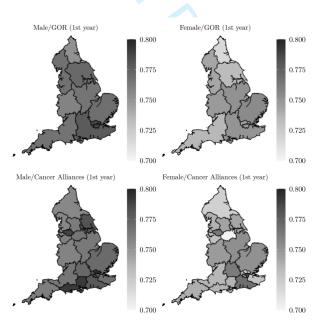
Sex	Geography	Model	Parameter	Mean	SD	95% equal-tail CI	Parameter	Mean	SD	95% equal-tail CI
Male	01	RS-SGH LN BYM2	$\sigma_{\tilde{u}}$	0.09	0.07	(0.00; 0.26)	σ_u	0.03	0.03	(0.00; 0.09)
Male	01	RS-SGH LN BYM2	$\tilde{ ho}$	0.53	0.35	(0.00; 1.00)	ρ	0.52	0.36	(0.00; 1.00)
Male	02	RS-SGH LN BYM2	$\sigma_{ ilde{u}}$	0.08	0.06	(0.00; 0.23)	σ_u	0.04	0.03	(0.00; 0.11)
Male	02	RS-SGH LN BYM2	$\tilde{ ho}$	0.49	0.35	(0.00; 1.00)	ρ	0.46	0.35	(0.00; 1.00)
Female	01	RS-SGH LN BYM2	$\sigma_{ ilde{u}}$	0.13	0.11	(0.01; 0.41)	σ_u	0.04	0.03	(0.00; 0.13)
Female	01	RS-SGH LN BYM2	$\tilde{ ho}$	0.52	0.35	(0.00; 1.00)	ρ	0.52	0.35	(0.00; 1.00)
Female	02	RS-SGH LN BYM2	$\sigma_{\tilde{u}}$	0.09	0.07	(0.00; 0.26)	σ_u	0.03	0.03	(0.00; 0.09)
Female	02	RS-SGH LN BYM2	õ	0.53	0.35	(0.00; 1.00)	ρ	0.52	0.36	(0.00: 1.00)



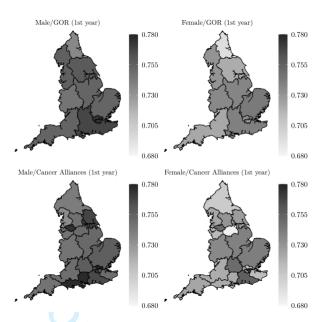
Web Figure SF13: $2.5^{\rm th}$ net survival percentile for t=3 based on the (i: top-left panel) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right panel) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for female patients.



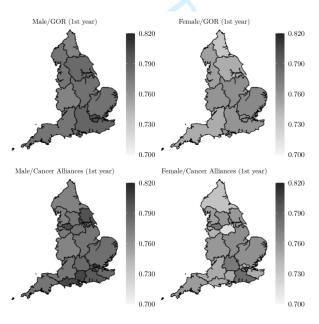
Web Figure SF14: $97.5^{\rm th}$ net survival percentile for t=3 based on the (i: top-left panel) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right panel) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for female patients.



Web Figure SF15: Net survival point estimate for t=1 based on the (i: top-left panel) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right panel) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for female patients.



Web Figure SF16: $2.5^{\rm th}$ net survival percentile for t=1 based on the (i: top-left panel) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right panel) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for female patients.

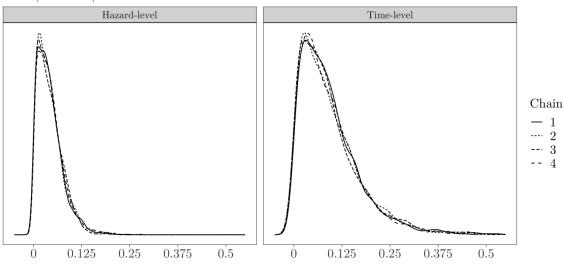


Web Figure SF17: $97.5^{\rm th}$ net survival percentile for t=1 based on the (i: top-left panel) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right panel) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right panel) "Cancer Alliances Regions" spatial structure with fitted model RS-SGH LN BYM2 for female patients.



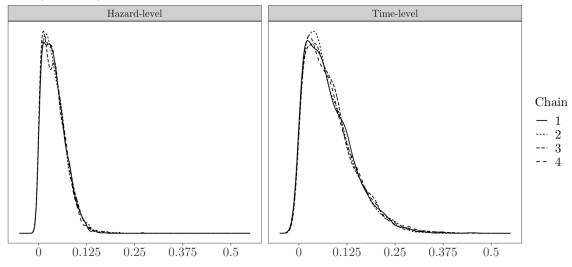
Web Figure SF18: Map of England divided into the 1–19 Cancer Alliances Regions, namely, West Yorkshire, "Humber, Coast and Vale," "Cheshire and Merseyside," "South Yorkshire, Bassetlaw, North Derbyshire and Hardwick," West Midlands, East Midlands, East of England, South East London, "Kent and Medway," "Surrey and Sussex," Thames Valley, Peninsula, "Somerset, Wiltshire, Avon and Gloucestershire," Wessex, "North East and Cumbria," "Lancashire and South Cumbria," "National Cancer Vanguard: Greater Manchester," "National Cancer Vanguard: North Central and North East London," and "National Cancer Vanguard: North West and South West London," respectively.

Male, Geo 01, RS-SGH LN BYM2



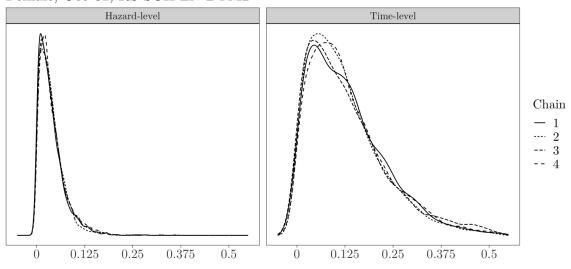
Web Figure SF19: Estimated posterior densities for σ_u (as it appears in the **hazard level** component) and $\sigma_{\tilde{u}}$ (as it appears in the **time level** component) for the BYM2 random effects when fitting the RS-SGH LN model for m le patients based on the "Government Office Regions" ("Geo 01"). The curves are plotted separately for each posterior chain.

Male, Geo 02, RS-SGH LN BYM2



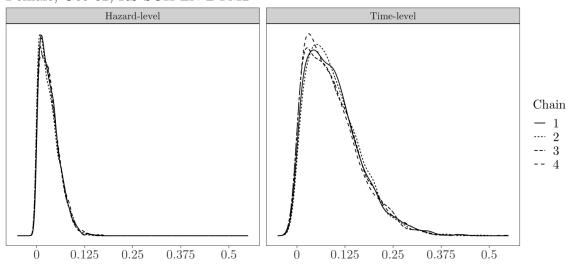
Web Figure SF20: Estimated posterior densities for σ_u (as it appears in the hazard level component) and $\sigma_{\tilde{u}}$ (as it appears in the **time level** component) for the BYM2 random effects when fitting the RS-SGH LN model for m le patients based on the "Cancer Alliances Regions in 2016" ("Geo 02"). The curves are plotted separately for each posterior chain.

Female, Geo 01, RS-SGH LN BYM2

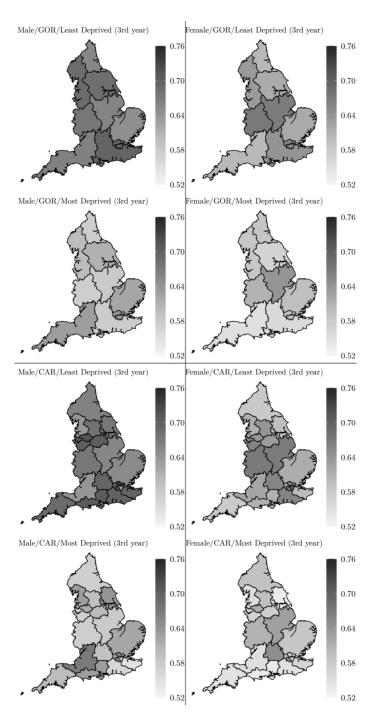


Web Figure SF21: Estimated posterior densities for σ_u (as it appears in the hazard level component) and $\sigma_{\tilde{u}}$ (as it appears in the **time level** component) for the BYM2 random effects when fitting the RS-SGH LN model for fem le patients based on the "Government Office Regions" ("Geo 01"). The curves are plotted separately for each posterior chain.

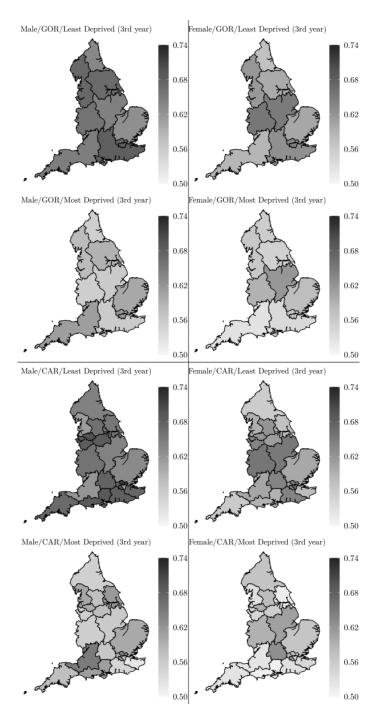
Female, Geo 02, RS-SGH LN BYM2



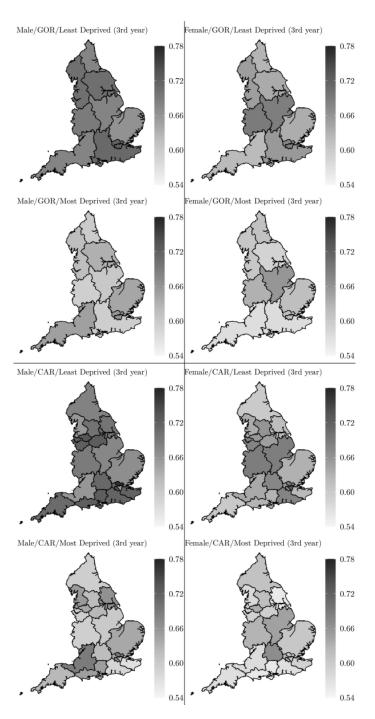
Web Figure SF22: Estimated posterior densities for σ_u (as it appears in the hazard level component) and $\sigma_{\tilde{u}}$ (as it appears in the **time level** component) for the BYM2 random effects when fitting the RS-SGH LN model for fem le patients based on the "Cancer Alliances Regions in 2016" ("Geo 02"). The curves are plotted separately for each posterior chain.



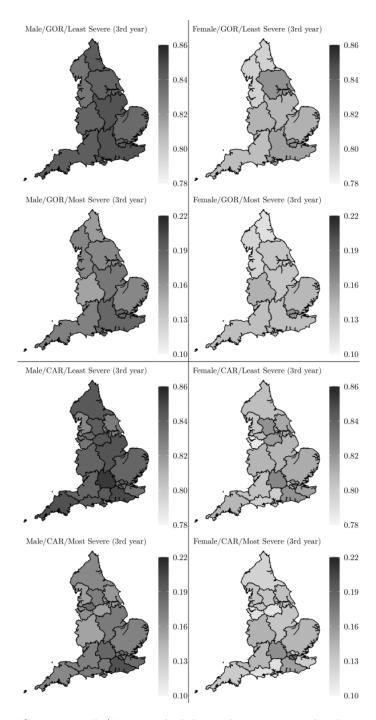
Web Figure SF23: "Deprivation level" ("1" being le st deprived and "5" most deprived) stratified net survival point estimate for t=3 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the le st deprived level, and the lower map represents the most deprived level.



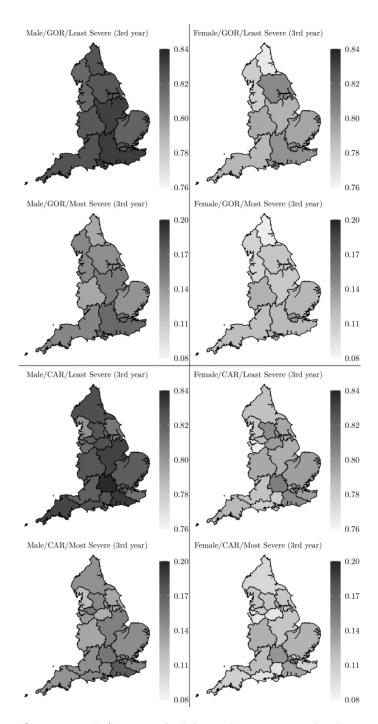
Web Figure SF24: "Deprivation level" ("1" being $le\ st\ deprived\$ and "5" $most\ deprived\$) stratified 2.5th net survival percentile for t=3 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the $le\ st\ deprived$ level, and the lower map represents the $most\ deprived$ level.



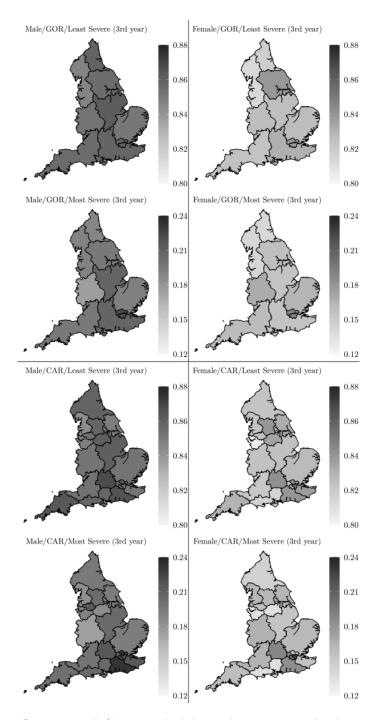
Web Figure SF25: "Deprivation level" ("1" being le st deprived and "5" most deprived) stratified 97.5th net survival percentile for t=3 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the le st deprived level, and the lower map represents the most deprived level.



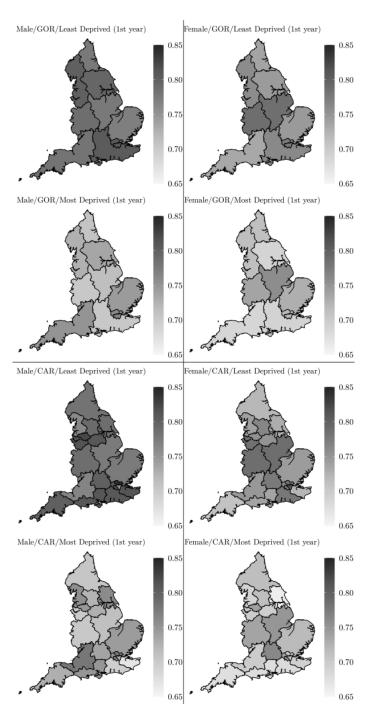
Web Figure SF26: "Cancer stage" ("1, 2, and 3" being le st severe and "4" most severe) stratified net survival point estimate for t=3 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the le st severe level, and the lower map represents the most severe level. Important: different from all other analyses, notice that the maps for different cancer stages are plotted in different scales.



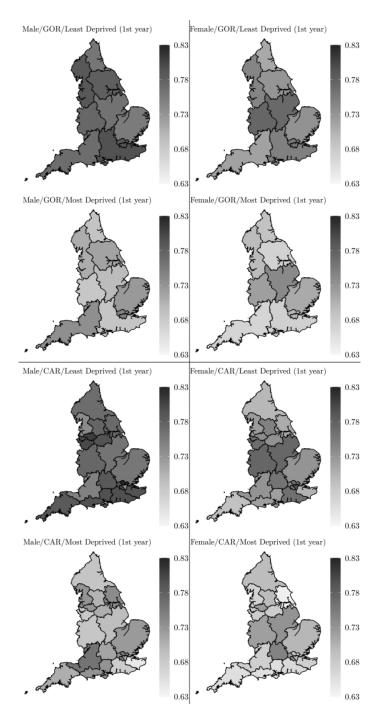
Web Figure SF27: "Cancer stage" ("1, 2, and 3" being le st severe and "4" most severe) stratified 2.5th net survival percentile for t=3 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the le st severe level, and the lower map represents the most severe level. Important: different from all other analyses, notice that the maps for different cancer stages are plotted in different scales.



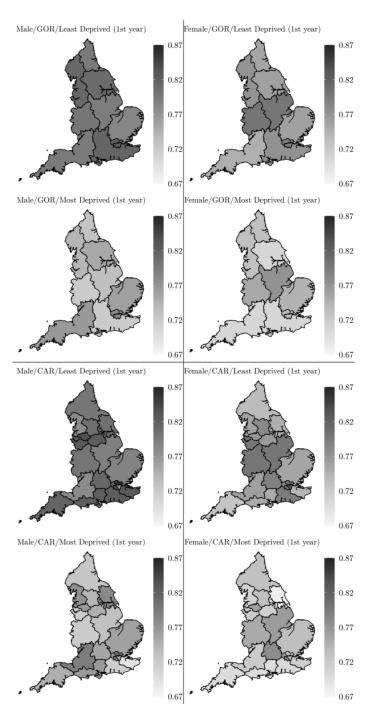
Web Figure SF28: "Cancer stage" ("1, 2, and 3" being le st severe and "4" most severe) stratified 97.5th net survival percentile for t=3 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the le st severe level, and the lower map represents the most severe level. Important: different from all other analyses, notice that the maps for different cancer stages are plotted in different scales.



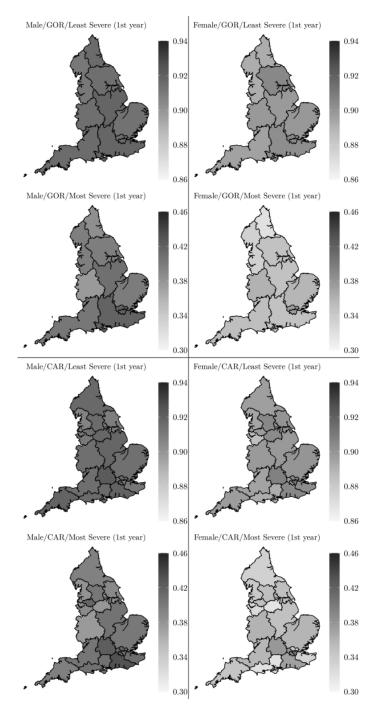
Web Figure SF29: "Deprivation level" ("1" being le st deprived and "5" most deprived) stratified net survival point estimate for t=1 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the le st deprived level, and the lower map represents the most deprived level.



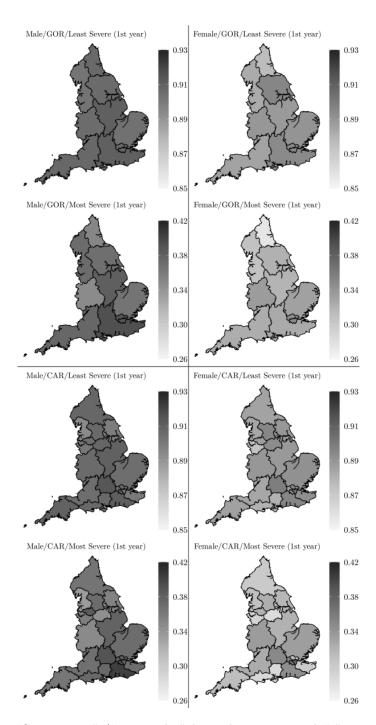
Web Figure SF30: "Deprivation level" ("1" being $le\ st\ deprived\$ and "5" $most\ deprived\$) stratified 2.5th net survival percentile for t=1 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the $le\ st\ deprived$ level, and the lower map represents the $most\ deprived$ level.



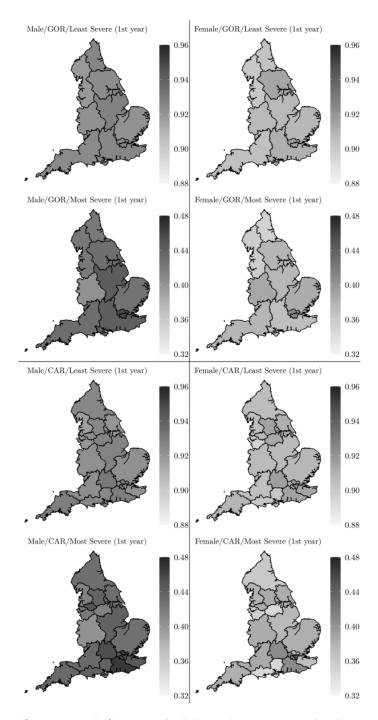
Web Figure SF31: "Deprivation level" ("1" being le st deprived and "5" most deprived) stratified 97.5th net survival percentile for t=1 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the le st deprived level, and the lower map represents the most deprived level.



Web Figure SF32: "Cancer stage" ("1, 2, and 3" being $le\ st\ severe$ and "4" $most\ severe$) stratified net survival point estimate for t=1 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the $le\ st\ severe$ level, and the lower map represents the $most\ severe$ level. Important: different from all other analyses, notice that the maps for different cancer stages are plotted in different scales.



Web Figure SF33: "Cancer stage" ("1, 2, and 3" being le st severe and "4" most severe) stratified 2.5th net survival percentile for t=1 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the le st severe level, and the lower map represents the most severe level. Important: different from all other analyses, notice that the maps for different cancer stages are plotted in different scales.



Web Figure SF34: "Cancer stage" ("1, 2, and 3" being le st severe and "4" most severe) stratified 97.5th net survival percentile for t=1 based on the (i: top-left maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, (ii: top-right maps) "Government Office Regions" (GOR) spatial structure with fitted model RS-SGH LN BYM2 for female patients, (iii: bottom-left maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for male patients, and (iv: bottom-right maps) "Cancer Alliances Regions" (CAR) spatial structure with fitted model RS-SGH LN BYM2 for female patients. For each panel, the upper map represents the le st severe level, and the lower map represents the most severe level. Important: different from all other analyses, notice that the maps for different cancer stages are plotted in different scales.

References

Henderson, R., Shimakura, S. and Gorst, D. (2002). Modeling spatial variation in leukemia survival data. Journal of the American Statistical Association 97, 965–972.

Rubio, F. J. (2022). SimLT. R package version 0.1.0.

