Modeling Spatial Variation in Leukemia Survival Data

Robin HENDERSON, Silvia SHIMAKURA, and David GORST

In this article we combine ideas from spatial statistics with lifetime data analysis techniques to investigate possible spatial variation in survival of adult acute myeloid leukemia patients in northwest England. Exploratory analysis suggests both clinically and statistically significant variation in survival rates across the region. A multivariate gamma frailty model incorporating spatial dependence is proposed and applied, with results confirming the dependence of hazard on location.

KEY WORDS: Cancer; Frailty; Geostatistics; Hierarchic model; Latent process; Semiparametric model.

1. INTRODUCTION

Although leukemia survival rates continue to improve as more effective therapies are introduced, considerable between-patient heterogeneity remains conditional on treatment and known prognostic factors (see, e.g., Cassileth et al. 1992; Estey, Shen, and Thall 2001; Schoch et al. 2001). In this article we investigate whether at least part of this heterogeneity might be linked to spatial effects, using data maintained by the North West Leukemia Register in the United Kingdom. This is a high-quality database that holds records of incidence and subsequent survival status of all leukemia cases in northwest England. In a previous informal study, Gorst (1995) suggested that there could be district-to-district variation in survival rates above and beyond what might be expected to occur by chance alone. Such a finding, if substantiated, would be of considerable interest. It could be due to patient management differences between treatment centers, which could have an important influence on future clinical practice, or due to background variation in population or environmental characteristics, necessitating further epidemiologic study.

We investigate whether the survival distribution for acute myeloid leukemia (AML) in adults is homogeneous across the region after allowing for known risk factors. We use register data on the 1,043 cases recorded between 1982 and 1998. AML represents the biggest single category of adult leukemia in the register. Figure 1 shows residential locations of the AML cases in the study period, together with the 24 administrative districts that make up the region. The boxed area is 100 km $\times$ 120 km, and the numbers are district identifiers, used for later reference. Apparent clustering is of course due in large part to the population distribution. In this work we do not discuss the detection and modeling of spatial variation in disease *incidence*, for which there are now well-established methods (Elliott, Wakefield, Best, and Briggs 2000). Instead, we concentrate on subsequent *survival* by extending standard survival models to the spatial setting. The simple cloropeth map in Figure 2 of estimated relative risks between districts, explained more fully later, suggests substantial variability between districts and also some apparent clustering of districts with similar risks. There seems to be a region of high risk running from northeast to southwest, with a low-risk region to the west. To investigate, we adopt a multivariate frailty approach that incorporates the effects of known covariates, individual heterogeneity, and spatial traits. Our ultimate goal is to model possible residual spatial variation in survival after accounting for known subject-specific prognostic factors and unobserved individual heterogeneity.

The article is organized as follows. In Section 2 we summarize an initial survival analysis of the AML data using standard univariate methods based on a Cox model with and without frailty. In Section 3 we investigate possible variation in survival across the region after allowing for covariate effects, using a lattice structure based on the 24 districts. We use a Bayesian hierarchic multivariate gamma model and Markov Chain Monte Carlo (MCMC) methodology for estimation, and the deviance information criterion (DIC) (Spiegelhalter et al. 2002) to compare competing models. In Section 4 we take an alternative approach, using the exact locations of the subjects' residences rather than knowledge only of their district, using an additive gamma frailty model that allows a proportion of the total frailty to be explained by a spatially varying component. We provide closing remarks and conclusions in Section 5.

2. INITIAL SURVIVAL ANALYSIS

To set the scene, we begin with a standard survival analysis ignoring any spatial variation across the region. Data consist of observation times $\boldsymbol{t}$, death/censoring indicators $\boldsymbol{\delta}$, and covariates $\boldsymbol{x}$ for 1,043 patients. Median survival time was just over 6 months, though some patients survived for more than 13 years. Some 16% of cases were censored. Complete information is available for four covariates: age; sex (0 = F, 1 = M); white blood cell count (WBC) at diagnosis, truncated at 500 units with 1 unit = 50×10^9/L; and a measure of deprivation for the enumeration district of residence. For this, we use the Townsend score, which is a quantitative measure with a range of -7 to 10 in the AML data, higher values indicating less affluent areas (Townsend, Phillimore, and Beattie 1988). The Townsend score is available for each of the 8,131 enumeration

Robin Henderson is Reader, Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, U.K. (E-mail: *Robin.Henderson@lancaster.ac.uk*). Silvia Shimakura is Lecturer, Departamento de Estatística, Universidade Federal do Paraná, Caixa Postal 19081, 81531-990, Curitiba, PR, Brazil (E-mail: *Silvia.Shimakura@est.ufpr.br*). David Gorst is Consultant Haematologist, Department of Haematology, Royal Lancaster Infirmary, Lancaster LA1 4RP, U.K. This work is supported in part by the CAPES Foundation, Brazil. The authors thank the editor and three reviewers for constructive comment on earlier versions of this article. Silvia Shimakura was sponsored by CAPES grant BEX1139/96-7.

Journal of the American Statistical Association
December 2002, Vol. 97, No. 460, Application and Case Studies
DOI 10.1198/016214502388618753

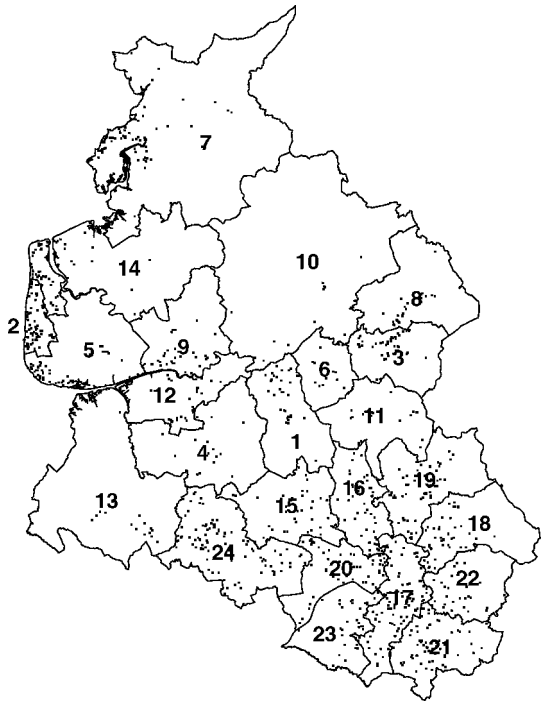

Figure 1. District Boundaries and Location of AML Cases in Northwest England, 1981–1998. Numbers are district identifiers.

districts that make up the study region. The size of each enumeration district is negligible in comparison with the entire region, and thus the Townsend score can be considered an individual-level covariate.

Results of fitting a standard Cox proportional hazards model,

$$\alpha(t|\boldsymbol{x}) = \alpha_0(t)\exp(\boldsymbol{\beta x}),$$

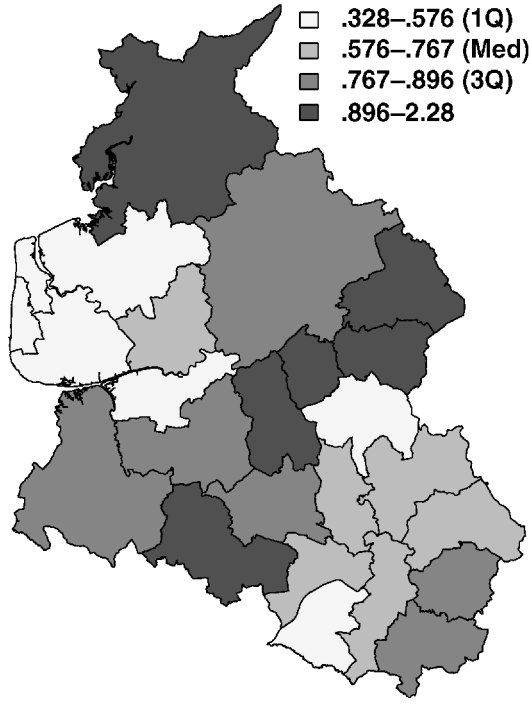

Figure 2. Estimated District-Relative Risks Under the Independent Fixed-Effects Model.

Table 1. Cox and Cox/Gamma Survival Results

	Cox			Cox/Gamma		
	β	Standard error	Ratio	β	Standard error	Ratio
Age	.0296	.0021	14.04	.0470	.0045	10.44
Sex	.0261	.0339	0.77	.0563	.0505	1.11
WBC	.0031	.0004	6.91	.0057	.0008	7.12
Townsend	.0292	.0090	3.23	.0547	.0187	2.92
	Frailty variance $\xi = .772$, standard error $= .168$					
Log-likelihood	−5,273.02			−5,255.44		

are summarized in the left part of Table 1. All covariates except sex have statistically significant effects, with higher risk for older patients, those with high WBC counts, and those living in more deprived areas.

Table 1 also includes results obtained when the Cox model is augmented by assuming a subject-specific random frailty term Z acting multiplicatively on the hazard, that is,

$$\alpha(t|z, \boldsymbol{x}) = z\alpha_0(t)\exp(\boldsymbol{\beta x}).$$

We take the common assumption (Hougaard 2000) of a gamma distribution of mean 1 and variance ξ for Z, written as $Z \sim \Gamma(1/\xi, 1/\xi)$, with the understanding that the parameterization $\Gamma(\alpha, \beta)$ implies mean α/β. As expected (Henderson and Oman 1999), estimates of the covariate effects $\boldsymbol{\beta}$ are all increased when frailty is included, although because the standard errors are also higher, the significance levels are not consistently greater. The estimated frailty variance is $\xi = .772$ (standard error, .168); from this and the log-likelihoods with and without frailty, it is clear that there is highly significant evidence of unexplained heterogeneity in the data. Standard diagnostic assessment procedures indicate an excellent fit of this frailty model to the marginal data. Details of this approach are omitted for space reasons.

3. DISTRICT-LEVEL ANALYSES

Our approach to investigating spatial variation follows standard geostatistical practice in postulating spatial association through underlying random effects, in our case the individual frailties Z_i. In many spatial applications, a standard assumption is that spatial effects can be described by a Gaussian random field or some function thereof (e.g., Diggle, Tawn, and Moyeed 1998). It is straightforward to adapt these methods to spatial survival problems by writing $Z_i = \exp(W_i)$ and assuming a multivariate Gaussian structure for $W = (W_1, W_2, \ldots, W_n)'$ with correlations depending on locations. In the absence of spatial variation, all W_i's are independent, and the model corresponds to a log-normal mixture of proportional hazards models. However, log-normal frailty is rarely used in survival, probably because there is no closed form for the marginal survival function. Moreover, we found that a gamma frailty model provides an excellent fit to the AML data when location is not considered, as mentioned in Section 2. Hence we prefer to develop a multivariate model that allows spatial dependence but reduces to the standard gamma frailty model if there is no spatial variation. For this, we begin with a Bayesian hierarchic approach with a lattice structure for any

spatial component, based on the partition of the region into 24 districts.

Letting μ_j be the mean frailty effect in district j, we assume conditional independence of the individual frailties Z_i given μ_j with

$$Z_i|\mu_j \sim \Gamma(1/\xi, 1/(\xi\mu_j)).$$

The district effects $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_{24})'$ are assumed to have multivariate Gaussian distribution with correlation matrix R and $N(1, \nu)$ marginals conditional on $\mu_j > 0$. We allow the correlation r_{jk} between μ_j and μ_k to depend on the distance d_{jk} between the locations of the district centroids s_j and s_k and investigate the following forms:

1. Between-district independence:

$$r_{jk} = 0.$$

2. Powered exponential (e.g., Diggle, Tawn, and Moyeed 1998):

$$r_{jk} = \exp\{-(d_{jk}/\phi)^\kappa\}, \qquad \phi > 0, \qquad 0 < \kappa \leq 2.$$

3. Matérn (Matérn 1986, p. 18):

$$r_{jk} = \frac{(d_{jk}/\phi)^\kappa K_\kappa(d_{jk}/\phi)}{2^{\kappa-1}\Gamma(\kappa)}, \qquad \phi > 0, \qquad \kappa > 0,$$

where $K_\kappa(\cdot)$ is a modified Bessel function of the third kind.

4. A nonisotropic version of the Matérn structure, replacing d_{jk}/ϕ with the square root of

$$(s_j - s_k)' \begin{pmatrix} 1/\phi_1^2 & 0 \\ 0 & 1/\phi_2^2 \end{pmatrix} (s_j - s_k),$$

as suggested by Williams in a discussion of the work of Diggle et al. (1998).

The parameter ν measures the amount of spatial variation between districts, and the model reduces to the previous independent gamma model with frailty mean 1 and variance ξ when $\nu = 0$. Otherwise, frailty effects within district j have gamma distribution with a district-specific mean μ_j and a common coefficient of variation $\sqrt{\xi}$. This enables subject-specific residual heterogeneity after allowing for spatially correlated district effects. Note that modeling $\log\boldsymbol{\mu}$ rather than $\boldsymbol{\mu}$ as multivariate Gaussian would guarantee positivity, but because the parameter ν is expected to be small, the conditioning on $\mu_j > 0$ will have negligible effect, and we prefer the slightly more direct approach.

Assigning priors to all parameters in the model completes the Bayesian specification of the model. For our application, we used the results obtained from the marginal analysis in Section 2 (Table 1) to guide the choice of vague but proper priors for $\boldsymbol{\beta}$, ν, and ξ. Independent Gaussian priors for $\boldsymbol{\beta}$ centered at the marginal estimates were made vague by setting their variances at 100. For ν and ξ, we used independent vague inverse gamma $IG(2, .77)$ priors, with mean .77 (as estimated in Sec. 2) but infinite variance. We used fixed values for the powered exponential and Matérn shape parameters κ but estimated the correlation scale parameter ϕ, using a vague $IG(2, .2)$ prior, with mean .2 and infinite variance.

For estimation, we used MCMC methods (Gilks, Richardson, and Spiegelhalter 1996), using a Metropolis–Hastings random walk algorithm (Hastings 1970) to update all parameters except the baseline hazards estimates. For this we took a profile approach, updating at each iteration of the chain using the Breslow estimator (Breslow 1974) conditional on parameters and frailties. We used a block-acceptance strategy for regression parameters $\boldsymbol{\beta}$ and (separately) the frailty parameters $\boldsymbol{\theta}$, say, and individual conditional updating for frailties Z_i and district means μ_j.

In Table 2 we use the deviance information criterion and the effective number of parameters p_D (Spiegelhalter et al. 2002) for an informal comparison between various correlation structures for the district effects $\boldsymbol{\mu}$. With just 24 districts, the chain runs very quickly, and so values are based on 10 separate runs of 500,000 iterations after a burn-in of 50,000 iterations. Even so, Monte Carlo standard error for the DIC measures is around .1. For reference, we also include in Table 2 the estimated DIC under an alternative conditional autoregressive (CAR) approach. Here a conditional Gaussian distribution is assumed for the mean response μ_j in district j given the mean $\bar{\mu}_{(j)}$ of its m_j neighbors $\mu_{(j)}$,

$$\mu_j|\mu_{(j)} \sim N(\bar{\mu}_{(j)}, \sigma^2/m_j),$$

with a vague $IG(2, .77)$ prior for σ_2 in this application.

From Table 2, we see that the spatial frailty models have smaller p_D and DIC values than the independence model, suggesting spatial association. The p_D values for all models are much smaller than the actual number of parameters (24 district effects μ_j, 1,043 individual frailty effects Z_i, and other parameters), suggesting shrinkage of the frailties towards their grand mean. This effect is more pronounced for increasing values of κ. Among the spatial frailty models, whereas the Matérn correlation functions seem to be the preferred choice under the two criteria, it appears that there is difficulty in selecting the exact choice of correlation structure, because the p_D and DIC values for these models are similar. This is confirmed by Figure 3, which compares fitted variograms (at posterior mean parameter estimates) under the alternative models with the median empirical variogram (Cressie 1993) obtained from samples from the posterior distribution of $\boldsymbol{\mu}$ under the independence model, together with 2.5% and 97.5% percentiles. Indeed, all fitted variograms (including the powered exponential models) fall within the reference range.

Table 2. DIC and Effective Number of Parameters (p_D) for District-Level Analyses

Model		p_D	*DIC*
Between-district independence		19.3	10,521.2
CAR		16.2	10,518.8
Powered exponential	$\kappa = 0.5$	16.4	10,518.4
	$\kappa = 1.0$	15.3	10,517.3
	$\kappa = 1.5$	14.1	10,516.0
Matérn	$\kappa = 1.0$	13.2	10,515.5
	$\kappa = 2.0$	12.1	10,515.0
	$\kappa = 5.0$	11.0	10,515.2
Asymmetric Matérn	$\kappa = 1.0$	14.0	10,516.3

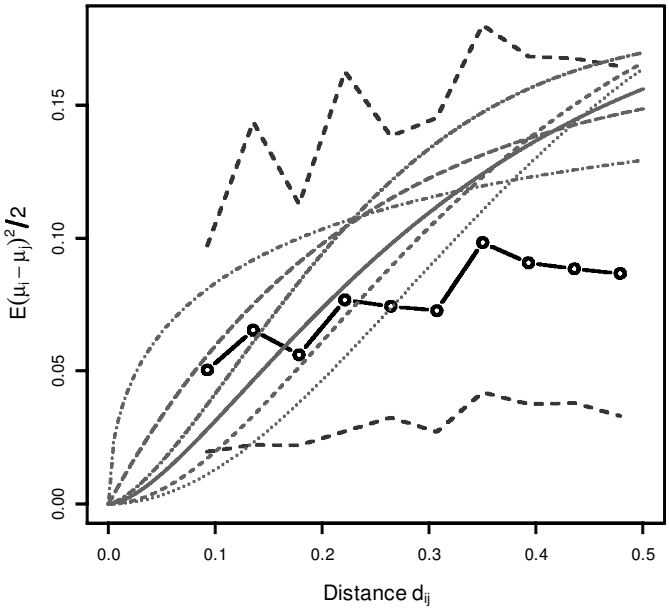

Figure 3. District-Level Semivariograms. Median, 2.5%, and 97.5% binned posterior values under independence assumption (bold lines), together with fitted curves under various Matérn (—, M1; - - -, M2; · · · ·, M5) and powered exponential (·-·-·, PE .5; – –; PE 1.0; ∘—∘—∘—, PE 1.5) models.

Combining the results in Table 2 and the fact that Matérn functions are often preferred to powered exponential functions because, unlike the latter, they have the advantage of being differentiable for $\kappa \geq 1$, for the remainder of this section we concentrate on results obtained using the Matérn model with $\kappa = 1$. Summaries of posterior distributions are given in Table 3. Results for the covariate coefficients and frailty variance ξ are comparable with those obtained from the nonspatial analysis given in the right part of Table 1. The posterior median of ν, the between-district relative risk variance, appears small at .18, although converting to standard deviation $\sqrt{.18} = .42$, we see evidence of substantial between-district variation. Specimen values of the between-district correlations at the posterior median of ϕ are given in Table 4.

Posterior means for the district effects μ_j were shrunk toward 1, in comparison with their estimated variance ν, ranging from about .75 to 1.25. The posterior distributions were all almost symmetric with medians close to the means and standard deviations between .10 and .18. A cloropeth map of the posterior means as estimated district-level relative risks is shown in Figure 4. In comparison with the within-subject heterogeneity ξ, the variation between districts is small—although, given that these relative risks can apply to large numbers of patients, the estimated ±25% spread between districts is clinically important.

Table 3. Summary of Matérn, $\kappa = 1$, District-Level Analysis

		2.5%	50%	Mean	97.5%
Age	β_1	.0396	.0460	.0461	.0532
Sex	β_2	−.0346	.0548	.0547	.1407
WBC	β_3	.0037	.0051	.0051	.0065
Townsend	β_4	.0277	.0519	.0525	.0797
Frailty variance	ξ	.4555	.7207	.7291	1.0580
Spatial variance	ν	.1000	.1818	.1972	.4000
Correlation scale	ϕ	.0662	.2030	.2226	.4600

Table 4. Specimen Values for the Correlation Function: From Matérn, $\kappa = 1$, District-Level Analysis

	Distance (km)			
	10	20	30	40
r	.832	.608	.423	.286

Figure 5 compares the posterior means of the district parameters μ_j with corresponding relative risks between districts obtained by simply including 23 indicator variables to represent districts as covariates in a Cox/gamma frailty analysis as a simple extension of the results of Section 2. This is the method used to obtain Figure 2. The largest district, 24, with 102 cases, is taken as the baseline. In general there is strong association between the two sets of estimates, although the spatial analysis leads to shrinkage toward the overall mean, and districts 6 and 11 stand out as anomalous. These are contiguous (Fig. 1) with small numbers of subjects (12 and 17, respectively). District 6 has subjects with relatively short survival times, and district 11 has relatively long term survivors (Fig. 2). The spatial analysis clearly borrows strength from neighboring districts to smooth out the pattern.

4. POINT LOCATION ANALYSIS

One disadvantage of the district-based approach is that the boundaries are political and essentially arbitrary in terms of potential risk factors. Treatment centers are not linked to these districts, and so any between-center differences would not necessarily show up in a district-level analysis. (For reference, we

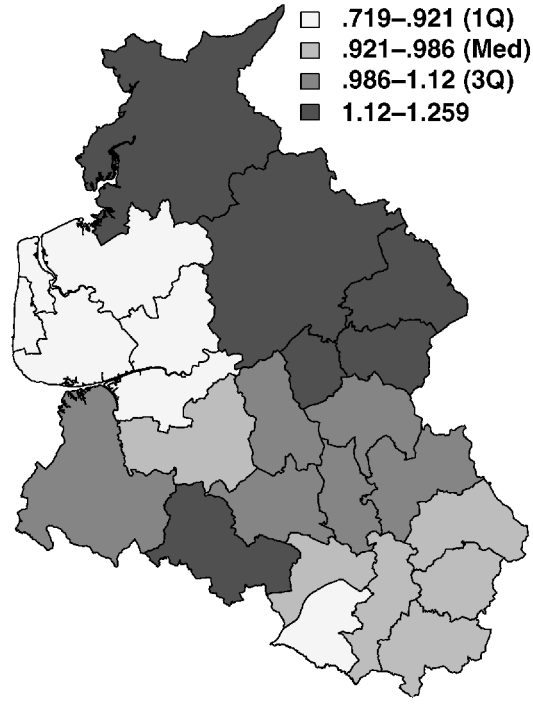

Figure 4. Estimated District-Relative Risks Under the Matérn Model With $\kappa = 1$.

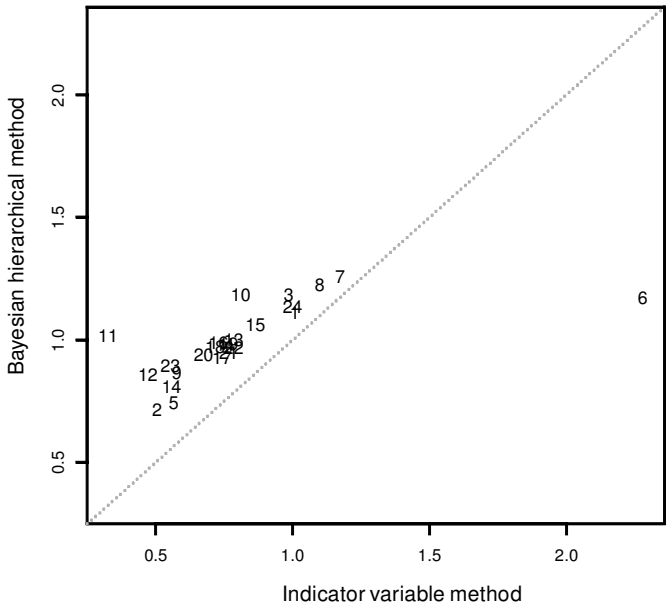

Figure 5. District Effects by Indicator Variable and Bayesian Hierarchic Methods. Numbers are district identifiers.

have no information on treatment center, which in any case can change over the course of treatment.) A second disadvantage is that aggregation into districts can lead to loss of fine-scale spatial structure (Diggle 1983). Finally, a third disadvantage is that the district-level approach can lead to potential ecological fallacies (Cressie 1993) induced by averaging spatial effects over subregions. In this section, therefore, we report an analysis that overcomes the problems just described by making use of the point locations of the AML cases. To do so, we consider an alternative to the hierarchic model of the previous section, under which we found severe identifiability problems when the mean parameters μ_j were allowed to be person-specific rather than common within districts. Instead, we adopt an additive-type frailty model that retains the gamma marginal structure with or without spatial variation.

In the Appendix we show how under a certain condition, a distribution for the n-vector Z of frailty terms can be constructed with the following properties:

- P1: gamma marginals with mean 1 and variance ξ, that is, $Z_i \sim \Gamma(1/\xi, 1/\xi)$, as assumed in Section 2.
- P2: covariance matrix Σ with elements $(\Sigma)_{ii} = \xi$ and $(\Sigma)_{ij} = \tau\xi r_{ij}$, $i \neq j$, for $0 < \tau < 1$ and with r_{ij} depending on distance d_{ij} as in Section 3.

Because $\tau < 1$, we have a so-called "nugget effect," appropriate when individuals at the same location do not share the same frailty. The interpretation is that a proportion τ of the heterogeneity variance is explained by spatial effects. Informally, the condition required for Z to have a proper multivariate distribution for all $\xi > 0$ and nonnegative correlation matrices $R = (r_{ij})$ is that this nugget effect is not small. More formally, there should exist a number ξ_0 with $\tau\xi < \xi_0 < \xi/\tau$ such that $2/\xi_0$ is a positive integer. This is appropriate when, as in our application, there is a reasonable degree of between-subject heterogeneity after allowing for spatial effects. Further explanation is provided in the Appendix.

Our model is additive in that Z can be decomposed into a spatially varying component Z_S and a mutually independent component Z_I, with $Z = (\xi/\xi_0)(Z_S + Z_I)$. Frailty models with an additive structure for clustered survival data have been proposed by a number of authors (Pickles and Crouchley 1994; Yashin, Vaupel, and Iachine 1995; Petersen, Andersen, and Gill 1996; Zahl 1997; Korsgaard and Andersen 1998; Petersen 1998). In previous applications, however, there has been an assumption of a relatively large number of relatively small clusters, such as found in twin studies, with within-cluster association but between-cluster independence. In our application we have a single cluster of possibly mutually associated lifetimes, with the association structure depending on the spatial distribution. Note that the additive structure is not necessary for interpretation; properties P1 and P2 are adequate in this respect.

For the AML application, we took $\xi_0 = 2$, forcing τ to be less than $\xi/2$ and preventing ξ from exceeding 2. Given the results of the previous section this seemed reasonable. Again, an isotropic Matérn correlation function with $\kappa = 1$ was assumed for the spatial frailty component. Here, interindividual distances are used rather than interdistrict ones.

Independent proper priors were adopted for the parameters. These were made vague by setting independent Gaussian priors for $\boldsymbol{\beta}$ centered at the marginal estimates, each with variance 100, a $U(0, 1)$ prior for τ, an $IG(2, .2)$ prior for ϕ (with mean .2 and infinite variance), and an $IG(3, 2)$ prior for ξ constrained to $\xi \leq 2$. Metropolis–Hastings random walk steps with Gaussian proposals were then used for updating $\boldsymbol{\beta}$ and the log-transformation of ξ, ϕ, and τ. The log-transformation makes the parameters in the updating more orthogonal, which speeds up mixing of the chain.

The analysis based on point locations requires the inversion of a $1{,}043 \times 1{,}043$ matrix at each iteration of the chain. Hence we ran 25,000 iterations only, with an additional burn-in of 5,000, by which time convergence was judged to have occurred. Thinning by retaining 1 in every 10 iterations yielded a final sample of size 2,500 for posterior analysis. Sample autocorrelations within the chain, cross-correlations between parameters, and plots of sample traces were used in monitoring the chain. Starting values were taken from the district-level analysis.

Table 5 provides posterior summaries of the parameters. The estimates of covariate effects $\boldsymbol{\beta}$ are typically more dispersed than those under the district-level approach. Inferences are mostly unaltered, however, with the exception of the deprivation index, which is now at the border of significance thanks to an increase in variance. The posterior distribution for ξ is also more dispersed than that under the district-level model,

Table 5. Summary of Matérn, $\kappa = 1$, Point-Level Analysis

		2.5%	50%	Mean	97.5%
Age	β_1	.0158	.0427	.0433	.0730
Sex	β_2	−.1733	.0342	.0358	.2463
WBC	β_3	.0016	.0048	.0048	.0082
Townsend	β_4	−.0197	.0428	.0435	.1102
Frailty variance	ξ	.3399	.7608	.8432	1.7452
Relative variance	τ	.0351	.2391	.2666	.6685
Correlation scale	ϕ	.0443	.1078	.1270	.3257

but there is still a significant frailty effect. A fair amount of this frailty is explained by spatial effects, as indicated by the posterior summaries for τ and ϕ.

Besides parameter estimates, it is also desirable to map the predicted mean frailty surface. This requires sampling from the posterior distribution of $(\boldsymbol{Z}^*|\boldsymbol{Z}, \boldsymbol{\theta})$, where $\boldsymbol{Z}^*$ is the set of values of $\boldsymbol{Z}$ at the locations for which prediction is required. This is straightforward when $\xi_0 = 2$, because $\boldsymbol{Z}^*$ can be derived as a simple transformation of an underlying Gaussian variable (see the Appendix). Figure 6 shows the predicted mean frailty surface at 2,815 sites within the region. For reference, the variance of prediction (not shown here) has a very similar spatial pattern, with values ranging between .45 and .75 and areas with high risk also having relatively low precision. Overall, Figure 6 is consistent with the map obtained under the district-level approach (see Fig. 4). It shows regions of high risk toward the northeast and the southwest and regions of low risk toward the west and southeast. An advantage of the point-level approach is that it allows assessment of local spatial variation in frailty effects. Two clusters with lower frailties (one toward the west and the other in the southeast), and four with higher frailties (one toward the north, another toward the southwest, and two to the east) can be identified from this map. In particular, notice that one of the areas with somewhat high risk is within the boundaries of district 6, a result consistent with previous analyses. Also, it is interesting to point out that the largest city in the region, Manchester, covers the eight districts to the southeast, and the suggestion is of low risk in the central parts of the city with higher risk around the suburbs. This would not be identified from a district-only analysis.

The predicted mean frailty surface (Fig. 6) can be compared with a district-aggregated map of deprivation index (Fig. 7) to investigate whether the predicted surface mirrors to some extent the spatial variation of deprivation index. In particular, it is interesting to assess whether areas of high risk correspond to areas of high social deprivation, even after including deprivation as a covariate. If this is so, it could explain the change in significance of the deprivation index. The comparison does not reveal a clear correspondence between these two maps, however.

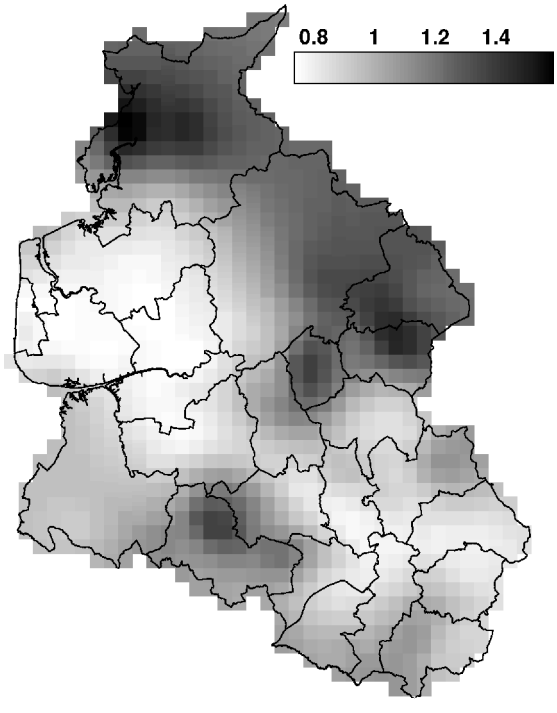

Figure 6. Predicted Frailty Effect Under the Matérn Model With $\kappa = 1$.

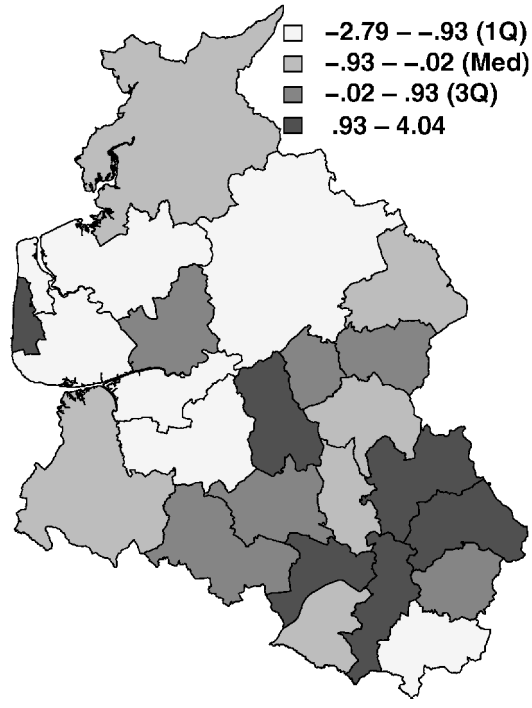

Figure 7. Townsend Deprivation Index (averaged within districts).

5. DISCUSSION

We have presented an attempt to model spatial variation in survival rates of AML patients in northwest England. Our finding is that there is good evidence of variation in survival across the region after allowing for covariates. The smoothing and shrinkage consequent to the spatial modeling approach lead to results that intuitively seem more reasonable than those obtained under the simpler, but perhaps naive, method of allocating fixed effects to each district (compare Figs. 2 and 4). There are concentrated regions of high and low risk with systematic variation of the order of $\pm 20\%$ in hazard. These differences are small when contrasted with the estimated amount of individual heterogeneity, but large in practical terms given that high numbers of patients can be affected. The task now (although not solely for statisticians) is to try to identify causes of this variation.

As mentioned, we have no information on treatment center. Center-to-center variation in survival rates can occur (Andersen, Klein, and Zhang 1999), and it is quite possible that local differences in treatment policy may explain at least some of the variability in the AML data. On the other hand, there is also a case for the argument that local population or environmental factors may explain observed center-to-center variation. The Townsend index used as a covariate is designed

to capture the level of deprivation or affluence for the enumeration district for each patient. We have no further information on social or other characteristics that could influence prognosis once AML is diagnosed. If such characteristics can be assumed to vary smoothly in space, then there should be more association between survival rates at centers that are relatively close together than at those that are far apart. Here the hierarchic gamma frailty model of Section 3, unlike the more traditional CAR model, may be useful if treatment center is known, because we can postulate a shared frailty effect for all patients treated at a single center but allow these frailties to be associated in space.

The correlation structure taken in Sections 3 and 4 for R assumed decreasing association with distance between locations. Provided that all correlations are nonnegative, such a restriction is not necessary, and any valid correlation matrix R can be used. Moreover, there is no requirement that the data be spatially referenced, the multivariate and hierarchic gamma frailty models also may be useful for familial or pedigree studies, for instance, with correlations depending on the closeness of genetic relationship. In current work we investigate a method for introducing time-dependent frailty in recurrent event studies, by discretizing the time axis, having one frailty for each interval but allowing positive serial correlation between intervals.

We do not consider our estimation methods definitive. The point location MCMC method is very time-consuming with our large sample size. The multivariate gamma model of Section 4, however, has an advantage over the hierarchic model of Section 3, because several parameters can be consistently estimated under a gamma frailty model ignoring spatial location, as in Section 2. Had we preferred, we could have chosen to use marginal estimates of ξ and $\boldsymbol{\beta}$ as well as cumulative baseline hazard A_0, although we had no difficulty estimating all terms simultaneously. On the other hand, the hierarchic model is applicable when there is a shared frailty, component as in the district-based analyses, whereas the multivariate gamma model does not extend in this way because of the inclusion of part of the nugget effect in Z_S.

A final comparison of the hierarchic model and the multivariate gamma model is that the former is less time-consuming to implement than the latter. However, the hierarchic model is based on what may be unnatural district boundaries in terms of potential risk factors, and treatment centers are not necessarily linked to districts. Furthermore, the averaging over districts and demonstration of existence of spatial correlations between the mean district effects can lead to the so-called ecological fallacy, whereby the same effect is mistakenly assumed to apply at the individual level. Neither of these issues is of concern under the multivariate gamma approach.

We have adopted MCMC procedures with frailties considered as missing data. Other methods are possible, although a drawback is that the joint frailty density is in practice unmanageable. A closed-form expression for the joint survival distribution can be obtained from the Laplace transform given in the Appendix, which might be exploited in future work. We will be very interested in alternative approaches to this problem.

APPENDIX: MULTIVARIATE GAMMA DERIVATION

Suppose that $Y_1, Y_2, \ldots, Y_m$ are iid n-variate Gaussian variables with standard marginals and correlation matrix C. Then (see, e.g., Krishnamoorthy and Parthasarthy 1951; Johnson and Kotz 1972, pp. 220–224) the vector $Z^* = (Z_1^*, Z_2^*, \ldots, Z_n^*)'$ with elements

$$Z_i^* = \sum_{j=1}^{m} Y_{ij}^2, \qquad i = 1, 2, \ldots, n$$

has a multivariate chi-squared distribution with Laplace transform (LT)

$$\mathcal{L}^*(a) = E[\exp\{-a'Z^*\}] = |I + 2\,\mathrm{diag}(a)C|^{-m/2},$$

where $a = (a_1, a_2, \ldots, a_n)'$, $a_i \in \mathbb{R}$. Marginal distributions are χ_m^2 or, in other words, $\Gamma(m/2, 1/2)$. Although the joint density for Z^* is intractable, it is easy to use the LT in the bivariate case to show that the correlation structure is simply obtained as the elementwise squares of C, that is, $\mathrm{corr}(Z_i^*, Z_j^*) = c_{ij}^2$.

An LT like the foregoing, but with $m/2$ replaced with $\xi > 0$, defines a proper multivariate gamma distribution for all $\xi > 0$ in the bivariate case (Vere-Jones 1967) and in higher dimensions when $c_{ij} = c$ (Moran and Vere-Jones 1969). Otherwise, necessary and sufficient conditions for the LT to define a proper multivariate distribution for all $\xi > 0$ were derived by Griffiths (1984) and given in an alternative form by Bapat (1989). These conditions are not satisfied for the spatial correlation structures R of interest in this work. However, provided that there is a number $\tau\xi < \xi_0 < \xi/\tau$ such that $2/\xi_0$ is a positive integer, we can use the foregoing LT to construct a distribution satisfying P1 and P2.

First, take $c_{ij} = (\tau\xi_0 r_{ij}/\xi)^{1/2}$ and $m = 2/\xi_0$. Then take $Z_S = \xi_0 Z^*/2$, which has $\Gamma(1/\xi_0, 1/\xi_0)$ marginals. Next, define Z_I to be an n-vector of mutually independent $\Gamma(1/\xi - 1/\xi_0, 1/\xi_0)$ variables, noting that $\xi < \xi_0$ so the shape parameter is positive. Then $Z = \xi(Z_S + Z_I)/\xi_0$ satisfies properties P1 and P2, as required. This construction is particularly straightforward when $\xi_0 = 2$, so that $m = 1$ and Z_S is then the elementwise square of a single Gaussian vector Y, with standard marginals and correlations equal to the square roots of those for Z_S. A disadvantage is that we have a constraint that $\xi \le \xi_0$, which implies $\xi \le 2$ for $\xi_0 = 2$. The limit 2 for ξ seems reasonably high in comparison with the marginal likelihood estimate of Section 2, and hence we judged it worthwhile to run the chain at this choice ξ_0. In other situations, it may be necessary to use other values of ξ_0 and have more than one latent Gaussian variable.

[Received February 2001. Revised August 2002.]

REFERENCES

Andersen, P. K., Klein, J. P., and Zhang, M. J. (1999), "Testing for Centre Effects in Multi-Centre Survival Studies: A Monte Carlo Comparison of Fixed and Random Effects Tests," *Statistics in Medicine*, 18, 1489–1500.

Bapat, R. B. (1989), "Infinite Divisibility of Multivariate Gamma Distributions and M-Matrices," *Sankyā*, Ser. A, 51, 73–78.

Breslow, N. E. (1974), "Covariance Analysis of Censored Survival Data," *Biometrics*, 30, 89–99.

Cassileth, P. A., Lynch, E., Hines, J. D., Oken, M. M., Mazz, J. J., Bennett, J. M., McGlave, P. B., Edelstein, M., Harrington, D. P., and O'Connell, M. J. (1992). "Varying Intensity of Postremission Therapy in Acute Myeloid Leukemia," *Blood*, 79, 1924–1930.

Cressie, N. (1993), *Statistics for Spatial Data* (rev. ed.), New York: Wiley.

Diggle, P. J. (1983), *Statistical Analysis of Spatial Point Patterns*, London: Academic Press.

Diggle, P. J., Tawn, J. A., and Moyeed, R. A. (1998), "Model-Based Geostatistics," *Applied Statistics*, 47, 299–350.

Elliott, P., Wakefield, J. C., Best, N. C., and Briggs, D. J. (eds.) (2000), *Spatial Epidemiology: Methods and Applications*, Oxford, U.K.: Oxford University Press.

Estey, E. H., Shen, Y., and Thall, P. F. (2001), "Effect of Time to Complete Remission on Subsequent Survival and Disease-Free Survival Time in AML, RAEB-t, and RAEB," *Blood*, 95, 72–77.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (eds.) (1996), "Introducing Markov Chain Monte Carlo," in *Markov Chain Monte Carlo in Practice*, London: Chapman and Hall, pp. 1–20.

Gorst, D. (1995), "North West Leukaemia Register," Technical Report 1995, North West Leukaemia Register.

Griffiths, R. C. (1984), "Characterization of Infinitely Divisible Multivariate Gamma Distributions," *Journal of Multivariate Analysis*, 15, 13–20.

Hastings, W. K. (1970), "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, 57, 97–109.

Henderson, R., and Oman, P. (1999), "Effect of Frailty on Marginal Regression Estimates in Survival Analysis," *Journal of the Royal Statistical Society*, Ser. B, 61, 367–379.

Hougaard, P. (2000), *Analysis of Multivariate Survival Data*, New York: Springer-Verlag.

Johnson, N. L., and Kotz, S. (1972), *Continuous Multivariate Distributions*, New York: Wiley.

Krishnamoorthy, A. S., and Parthasarathy, M. (1951), "A Multivariate Gamma-Type Distribution," *Annals of Mathematical Statistics*, 22, 549–557. corr. (1960), 31, 229.

Korsgaard, R., and Andersen, A. H. (1998), "The Additive Genetic Gamma Frailty Model," *Scandinavian Journal of Statistics*, 25, 255–269.

Matérn, B. (1986), Spatial Variation (2nd ed.), Berlin: Springer-Verlag.

Moran, P. A. P., and Vere-Jones, D. (1969), "The Infinite Divisibility of Multivariate Gamma Distributions," *Sankhyā*, Ser. A, 31, 191–194.

Petersen, J. H. (1998), "An Additive Frailty Model for Correlated Life Times," *Biometrics*, 54, 646–661.

Petersen, J. H., Andersen, P. K., and Gill, R. D. (1996), "Variance Components Models for Survival Data," *Statistica Neerlandica*, 50, 193–197.

Pickles, A. R., and Crouchley, R. (1994), "Generalizations and Applications of Frailty Models for Survival and Event Data," *Statistical Methods in Medical Research*, 3, 263–278.

Schoch, C., Haferlach, T., Haase, D., Fonatsch, C., Loffler, H., Schlegelberger, B., Staib, P., Sauerland, M. C., Heinecke, A., Buchner, T., Hiddemann, W., and German AML Cooperative Study Group. (2001), "Patients with de novo Acute Myeloid Leukaemia and Complex Karyotype Aberrations Show a Poor Prognosis Despite Intensive Treatment: A Study of 90 Patients," *British Journal of Haematology*, 112, 118–126.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002), "Bayesian Measures of Model Complexity and Fit," *Journal of the Royal Statistical Society*, Ser. B, 64, 583–640.

Townsend, P., Phillimore, P., and Beattie, A. (1988), *Health and Deprivation: Inequality and the North*, London: Croom Helm.

Vere-Jones, D. (1967), "The Infinite Divisibility of a Bivariate Gamma Distribution," *Sankhyā*, Ser. A, 29, 421–422.

Yashin, A. I., Vaupel, J. W., and Iachine, I. A. (1995), "Correlated Individual Frailty: An Advantageous Approach to Survival Analysis of Bivariate Data," *Mathematical Population Studies*, 5, 145–159.

Zahl, P. H. (1997), "Frailty Modelling for the Excess Hazard," *Statistics in Medicine*, 16, 1573–1585.