Spatio-temporal modeling using B-spline for disease mapping:  
Analysis of childhood cancer trends

Mahmoud Torabi a and Rhonda J. Rosychuk a∗

Department of Pediatrics, University of Alberta, Edmonton, Alberta, Canada
(Received 00 Month 200x; in final form 00 Month 200x)

To analyze childhood cancer diagnoses in the province of Alberta, Canada during 1983-2004, we construct a generalized linear mixed model for the analysis of geographic and temporal variability of cancer rates. In this model, spatially correlated random effects and temporal components are adopted. The interaction between space and time is also accommodated. Spatio-temporal models that use conditional autoregressive smoothing across the spatial dimension and B-spline over the temporal dimension are considered. We study the patterns of incidence rates over time and identify areas with consistently high rate estimates as areas for further investigation. We apply the method of penalized quasi-likelihood to estimate the model parameters. We illustrate this approach using a yearly data set of childhood cancer diagnoses in the province of Alberta, Canada during 1983-2004.

Keywords: conditional autoregressive; disease mapping; generalized linear mixed model; geographic epidemiology; penalized quasi-likelihood

1. Introduction

Childhood cancer differ from adult cancers in terms of type and distribution. Leukemias, brain and other nervous system tumours, lymphomas (lymph node cancers), bone cancers, soft tissue sarcomas, kidney cancers, eye cancers, and adrenal gland cancers, are the most common types of cancer in children, while skin, prostate, breast, lung, and colorectal cancers are the most common cancers in adults [7]. Childhood cancers also differ from adult cancers based on biological, clinical and environmental features, growth rates, and treatment responses. While in most cases the causes of childhood cancers are unknown, the causes of adult cancers are environmental, occupational and lifestyle factors such as diet, alcohol and smoking [20, 23].

In North America, childhood cancer is the most common cause of death from disease in the pediatric population (one year of age through adolescence) [4, 39]; more deaths than asthma, diabetes, cystic fibrosis and AIDS combined in Canada [4]. With such an impact, it is important to identify trends in childhood cancer incidence that may suggest further epidemiological studies to identify risk factors and identify any changes in important factors. Trends may occur over space and time and the focus of our paper is to examine geographical and temporal variations in annual number of childhood cancer diagnoses during 1983 to 2004 in the western Canadian province of Alberta.

*Corresponding author. Email: rhonda.rosychuk@ualberta.ca
As part of routine monitoring of geographic areas by public health agencies, disease rate maps are often used to examine trends over space and time. These agencies need to have reliable maps that are methodologically sound. The statistical and epidemiological literature has devoted considerable attention to this important issue since regional maps of morbidity and mortality rates over time are useful tools in determining spatial and temporal patterns of disease. Identifying regions with substantively different rates may be suggestive of region-level characteristics that could be responsible for the geographic pattern of disease rates. These characteristics could be further examined to determine any causal relationship with the disease.

To analyze disease rates, it is customary to use Poisson regression with the assumption that the rates in nearby regions are independent and the mean and variance of responses are equal. However, extra-Poisson variation may result if (unknown) causal factors are not included in the model. To have an appropriate model, we may want to incorporate a certain degree of spatial correlation, depending on how smoothly the omitted factors vary across the regions. This issue was first addressed by Clayton and Kaldor [10] by incorporating spatially correlated random effects in mixed models to account for the extra-Poisson variation. There are many different ways to perform inference in mixed models. With advances in computational power, one may want to use Markov chain Monte Carlo (MCMC) methods such as the Gibbs sampler [2, 3, 9, 22, 46, 48]. The method of penalized quasi-likelihood (PQL) may also be used for inference in generalized linear mixed models (GLMMs). This method was first proposed by Breslow and Clayton [5] and they also provided an example of the use of PQL for estimation in mapping studies. As another alternative method, one may want to use generalized estimating equation (GEE) approach proposed by Liang and Zeger [28] and Prentice and Zhao [36], to analyzing the longitudinal data using the generalized linear models.

Researchers have also focused on temporal smoothing of rates. For instance, an autoregressive (AR) model for temporal count data was used by Zeger [50]. Waller et al. [48] and Knorr-Held [22] developed the existing hierarchical Bayes spatial models to account for temporal effects and spatio-temporal interactions. Spatio-temporal models that use AR local smoothing across the spatial dimension and cubic B-spline smoothing over the temporal dimension, were proposed by MacNab and Dean [32] and Silva et al. [40]. In some contexts such as infectious diseases, malaria [30] and influenza-related mortality [21] have been noted to have spatio-temporal as well as seasonal effects. A new methodology was recently developed by Torabi and Rosychuk [45] to handle seasonal effects in the context of spatio-temporal modeling of disease rates.

We applied a comprehensive GLMM-based model to account for the spatio-temporal analysis of risks in our childhood cancer data set (Section 2). In this model, the spatially correlated random effects and temporal effects are included. The well-known conditional autoregressive (CAR) model [3] is used for the spatial random effects and B-spline approaches for the temporal effects. The interaction between space and time is also accommodated in the model. We employ the penalized quasi-likelihood method for parameter estimates (Section 3). In Section 4, we apply the GLMM model on childhood cancer incidence in the province of Alberta, Canada during 1983-2004. Section 5 gives a short general discussion and concluding remarks.
2. Methodology

Let $y_{it}$ denote the diseased cases for the geographic region $i$ in time $t$, and let $e_{it}$ be the corresponding expected number for $i = 1, ..., I; t = 1, ..., T$. When the expected number of disease cases varies by important strata, say age group and gender, adjustments can be made. Without loss of generality, the expected number of disease cases, $e_{it}$, is adjusted by age group (0-4, 5-9, 10-14, 15-19 years in our data set) and gender (male, female). The conditional expectation of the $y_{it}$ given the random effects is denoted by $\mu_{it}^c$. A generalized and flexible Poisson model for $\mu_{it}^c$ is given by

$$\mu_{it}^c = \exp(\log e_{it} + m + S(t) + \eta_i + \theta_{it}),$$

where $m$ is the overall mean ratio over the time and region, $\eta_i$ represents spatially correlated random effects, and $\theta_{it}$ is the interaction between spatial and temporal effects. For account the fixed temporal effects, $S(t)$ takes the flexible form of a cubic B-spline without an intercept with three inner knots [18]. The B-spline is provided without an intercept due to an overall mean of ratio $m$ in our model. A cubic B-spline, $S(t)$, is given by

$$S(t) = \beta_1 B_1(t) + \beta_2 B_2(t) + \beta_3 B_3(t) + \beta_4 B_4(t),$$

where $(\beta_l, B_l)$ denote the coefficients and basis functions of the B-spline, respectively $(l = 1, ..., 4)$ [17, 18].

The usual CAR model is used to capture the spatial random effects $\eta_i$. A variety of CAR models may also be used by taking a collection of mutually compatible conditional distributions $p(\eta_i | \eta_{-i}), i = 1, ..., I$ where $\eta_{-i} = \{ \eta_j : j \neq i, j \in \partial_i \}$ and $\partial_i$ refers a set of neighbours for the $i$-th area [3]. We consider the following general model for the spatial effects $\eta_i$,

$$\eta = (\eta_1, ..., \eta_I)^t \sim N(0, \Sigma_{\eta}),$$

$$\Sigma_{\eta} = \sigma_{\eta}^2 P^{-1},$$

$$P = \lambda_{\eta} D + (1 - \lambda_{\eta}) I_I,$$

where $\sigma_{\eta}^2$ is the spatial dispersion parameter, $\lambda_{\eta}$ measures the spatial autocorrelation, $0 \leq \lambda_{\eta} \leq 1$, and $I_I$ is an identity matrix of dimension $I$. The neighbourhood matrix $D$ has its $i$-th diagonal element equal to the number of neighbours of the corresponding area ($\#\partial_i$), and the off-diagonal elements in each row equal $-1$ if the corresponding areas are neighbours and zero otherwise [27, 31, 32]. Neighbours can be defined in various ways, depending on the context of the analysis, but one popular definition that we use is the set of areas that have common borders. When the spatial correlation $\lambda_{\eta}$ is one, we have $\eta_i | \eta_{-i} \sim N(\bar{\eta}, \sigma_{\eta}^2 / \#\partial_i)$, where $\bar{\eta}$ is the mean of the random effects in the neighbourhood of the $i$-th area. A more general form of the spatial random effects $\eta_i$ is given by

$$\eta_i | \eta_{-i} \sim N\left(\frac{\sum_{j \in \partial_i} w_{ij} \eta_j}{\sum_{j \in \partial_i} w_{ij}}, \frac{\sigma_{\eta}^2}{\sum_{j \in \partial_i} w_{ij}}\right),$$

where $(w_{ij})$ are the weights in the neighbourhood of the $i$-th area.
where $w_{ij}$ is the user-specified weights linking areas $i$ and $j$ [25, 31, 32, 47].

The interaction effect of space and time $\theta_{it}$ may be defined in many different ways. One way to define $\theta_{it}$ is

$$\theta = (\theta_{11}, ..., \theta_{IT})' \sim N(0, \sigma^2_\theta \mathbf{K}_\theta^{-1}),$$

where $\sigma^2_\theta$ measures the dispersion between the space and time effects and $\mathbf{K}_\theta$ is a pre-specified structure matrix. The model (1) reduces to model with the main effects space and time when $\theta_{it} = 0$. Hence, the variation that is not explained by the main effects is captured by $\theta$. It was suggested by Clayton [8] to specify $\mathbf{K}_\theta$ as the Kronecker product of the structure matrices of those main effects, which are assumed to interact. Alternatively, one may define $\theta_{it}$ as $\theta_{it}$ or $S_i(t)$ depending on the nature of data set [32, 40], where $\theta_i$ is a fixed parameter or an CAR model, and $S_i(t)$ is a cubic B-spline for specific region $i$. In this paper, we use $\theta = (\theta_{11}, ..., \theta_{IT})' \sim N(0, \sigma^2_\theta \mathbf{K}_\theta^{-1})$ which was found useful in our exploration of the data.

The model (1) can be written as a GLMM, $E(y|\nu) = g(\text{offset} + \mathbf{X}\beta + \mathbf{Z}\nu)$, where $y = (y_{11}, ..., y_{IT})'$, $g(\cdot) = \exp(\cdot)$, $\beta = (\beta_1, ..., \beta_4)'$ is the vector of the fixed parameters, and $\mathbf{X}$ and $\mathbf{Z}$ are the known $N \times p$ and $N \times h$ matrices of full rank, $(N = IT, p = 5, h = I + IT)$, and $\nu$ is independently distributed with mean 0 and covariance matrix $\mathbf{\Sigma}_\nu$ depending on variance parameters $\zeta = (\sigma^2_\nu, \lambda_\nu, \sigma^2_\theta)'$, where $\mathbf{\Sigma}_\nu = \text{diag}(\sigma^2_\nu, \sigma^2_\theta \mathbf{K}_\theta)$. The design matrix $\mathbf{X} = \text{col}_{1 \leq i \leq I}(\mathbf{X}_i)$ corresponds to the fixed effects and has dimension $N \times p$. The design matrix $\mathbf{Z} = \text{col}_{1 \leq i \leq I}(\mathbf{Z}_i)$ for the random effects has dimension $N \times h$, where $\mathbf{Z}_i = (\mathbf{Z}_{i1}, \mathbf{Z}_{i2}), i = 1, ..., I$. The matrix $\mathbf{Z}_{i1}$ has dimension $T \times I$ where the corresponding $i$-th column is one, elsewhere 0; $\mathbf{Z}_{i2}$ has dimension $T \times IT$ which depends on the structure of $\mathbf{K}_\theta$.

There are various methods to estimate the fixed parameters $\beta$ and variance components $\zeta$ such as penalized quasi-likelihood (PQL) [5, 6, 15, 29, 31, 47], marginal quasi-likelihood (MQL) [5, 43], generalized estimating equations (GEEs) [28, 36], estimating functions [49], hierarchical likelihood [26], Bayesian analysis using MCMC [2, 3, 9, 19], and the EM algorithm [33].

3. Estimation of model parameters

We study the approximate method PQL for inference in GLMM. PQL estimation is an approximate inference technique proposed by Breslow and Clayton [5] for GLMMs which uses weighted least-squares algorithms for estimation of fixed model parameters along with likelihood equations from an approximate normal model for estimating variance components. Breslow and Clayton [5] used ideas proposed by Tierney and Kadane [44] and Barndorff-Nielson and Cox [1] on Laplace methods for integral approximations. To estimate the fixed effects using an iterative weighted least-squares algorithm in the normal mixed effects, following Breslow and Clayton [5], we define a working response vector $y^* = \mathbf{X}\beta + \mathbf{Z}\nu + (y - \mu^c)/\mu^c$ where $\mu^c = (\mu^c_{11}, ..., \mu^c_{IT})$. The associated normal theory model is then given by

$$y^* = \mathbf{X}\beta + \mathbf{Z}\nu + \varepsilon,$$
where $\varepsilon \sim N(0, W^{-1})$, $W = \text{diag}(\mu_{itk}^2)$. Using a starting value of the variance component, $\varsigma^{(0)}$, and $\beta$ and $\nu$, the $\hat{\beta}$ and $\hat{\nu}$ are obtained iteratively using

$$
\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y^*
$$

and

$$
\hat{\nu} = \Sigma_\nu Z'V^{-1}(y^* - X\hat{\beta}),
$$

where $V = W^{-1} + Z\Sigma_\nu Z'$. The estimated asymptotic variance of $\hat{\beta}$ is then given by $(X'V^{-1}X)^{-1}$.

To estimate the variance components, $\hat{\varsigma}$, we update $\varsigma$ given $\hat{\beta}$ and $\hat{\nu}$ from (4) and (5), by taking one Newton step towards a new value $\varsigma^{(1)}$:

$$
\varsigma^{(1)} = \varsigma^{(0)} + I_\varsigma^{-1}S_\varsigma,
$$

where $I_\varsigma$ and $S_\varsigma$ are evaluated at $\varsigma = \varsigma^{(0)}$, and the $k$-th component of $S_\varsigma$ is given by

$$
S_\varsigma k = \frac{1}{2} \left[ (y^* - X\beta)'V^{-1}\frac{\partial V}{\partial \varsigma_k}V^{-1}(y^* - X\beta) - \text{tr}(P\frac{\partial V}{\partial \varsigma_k}) \right],
$$

with $P = V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}$ and $I_{\varsigma k}$ is the corresponding Fisher information matrix with elements

$$
I_{\varsigma k} = \frac{1}{2}\text{tr}\left(P\frac{\partial V}{\partial \varsigma_k}P\frac{\partial V}{\partial \varsigma_l}\right).
$$

The asymptotic variance of $\hat{\varsigma}$ is then given by $I_\varsigma^{-1}$. Using the updated estimate $\varsigma^{(1)}$ from (6), we solve $\hat{\beta}$ and $\hat{\nu}$ from (4) and (5), and then use (6) to get the updated value for variance components, $\varsigma^{(2)}$. This process is repeated until both the fixed model parameters, $\beta$, and variance components, $\varsigma$, converge.

4. Application

4.1 Data analysis

We use a yearly data set of cancer diagnoses in children (age $\leq 19$) in the western Canadian province of Alberta during the 1983-2004 fiscal years. The population of Alberta increased during the study period from 2.4 million in 1983 to 3.2 million in 2004 with the average population of children around 800,000. During the last study year, the province consisted of nine Regional Health Authorities that were responsible for the delivery of health care services. These nine regions were further sub-divided into 70 sub-Regional Health Authorities (sRHAs) and these sRHAs are the geographic unit used in our model and all data were linked to these geographic boundaries. The number of cancer cases totaled 3,684 over the study period with mean and median number of yearly cases per sRHA of 2.4 and 2.0 (range 0 to 13), respectively. The sRHA population sizes vary from 2,390 to 29,060, with mean and median number of 11,670 and 10,140, respectively. The largest population belongs to region 50, while region 53 has the least population.
We report the model (1) which was found useful in our exploration of the data. In our application, \( I = 70, T = 22 \), and the design matrix \( X = \text{col}_{1 \leq i \leq I}(X_i) \) corresponds to the fixed effects with dimension \( N \times p \), where \( N = 1,540 \) and \( p = 5 \).

In our model, \( X_i \) is the same for all regions. We define \( X_i \) with dimension \( 22 \times 5 \) as

\[
X_i = \begin{pmatrix}
1 & B_1(1) & B_2(1) & B_3(1) & B_4(1) \\
... & ... & ... & ... & ...
\end{pmatrix},
\]

where \( B_i(t), i = 1, \ldots, 4 \), are easily obtained by using the B-spline command \( \text{bs}(.) \) in the statistical software package \( R \) [37]. The design matrix \( Z = \text{col}_{1 \leq i \leq I}(Z_i) \) for the random effects has dimension \( N \times h \), where \( h = 1,610 \).

As a first step, we fit the simple Poisson model, \( y_{it} \sim \text{Poisson}(e_{it}) \), independently. It clearly fails to provide an adequate fit to the data, where the Pearson goodness-of-fit statistic is 3,178 with 1,539 degrees of freedom.

We also investigated the use of a simple linear trend to model the fixed temporal effects. The estimated coefficient of the linear temporal term was 0.028 with standard error 0.004. For the B-spline smoothing of the fixed temporal effects, we used a cubic B-spline with three inner knots. These results indicated a general increase in cancer ratios over the period. Figure 1 shows the estimates of the overall ratios of cancer cases for crude ratios over regions, simple linear trend, and cubic B-spline, where the crude ratio for time \( t \) is \( \sum_{i=1}^{I} y_{it} / \sum_{i=1}^{I} e_{it} \).

We fitted the model

\[
\mu_{it} = \exp(\log e_{it} + m + S(t) + \eta_i),
\]

and resulted that the temporal effects had more contribution into the model than spatial random effects. Hence, we took \( (\theta_{i1}, \ldots, \theta_{iT})' \) as random walk with first degree, \( \text{RW}(1) \), for each \( i = 1, \ldots, I \) [22]. We then fitted the model (1) to the data set of cancer cases using the PQL method (Table 1). Most of B-spline coefficient are clearly significant at the 5% level, indicating the existence of the fixed temporal effects. The estimated spatial dispersion parameter is 0.02 (standard error 0.01) with spatial autocorrelation 0.33 (standard error 0.25). The estimate of the temporal dispersion for the interaction effects is 0.01, with standard error 0.004.

Table 1. Parameter estimates and standard errors, spatio-temporal mixed model, yearly childhood cancer cases, 1983-2004.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficients of the fixed effects</th>
<th>Variance components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Standard error</td>
</tr>
<tr>
<td>( m )</td>
<td>-0.360</td>
<td>0.092</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.029</td>
<td>0.126</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>0.487</td>
<td>0.161</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>0.332</td>
<td>0.151</td>
</tr>
</tbody>
</table>
Formal tests for overdispersion may be used to indicate its presence in the model [12, 14, 41]. We applied the parametric bootstrap approach proposed by Sinha [41] to evaluate the significance of variance components and resulted that in terms of significance, there is more overdispersion in spatial effects than temporal effects.

For a diagnostic analysis, we calculated the deviance residual [34] as

$$d_{it} = \text{sign}(y_{it} - \hat{\mu}_{it}^c)[2\{y_{it} \log(y_{it} / \hat{\mu}_{it}^c) - y_{it} + \hat{\mu}_{it}^c\}]^{1/2},$$

where

$$\text{sign}(z) = \begin{cases} 1 & z > 0 \\ 0 & z = 0 \\ -1 & z < 0 \end{cases}.$$

Figure 2 gives the residuals versus predicteds diagnostic plot. It is clear from Figure 2 that there is no serious lack of fit in our model.

To have better understanding of the estimated spatial risk profile, we obtained the adjusted relative risk $\exp(\eta_i + \theta_{it})$ which is automatically calibrated on a common base for the temporal effects [22]. Figure 3 plots the estimated spatial effects based on the fitted model, where the regional risk of cancer cases corresponds to four selected years. For confidentiality, the years are not shown, however, the order of
the maps increased with time. The overall spatial pattern suggests that sRHAs with relatively high cancer ratios are clustered in the south-central part of the province. In general, the spatial pattern does not change much over time. Some parts of the two largest population centres, shown as inserts in Figure 3, have the highest estimated cancer ratios. There are also some parts of the region between these centres with relatively high estimated cancer ratios. More investigation may be warranted to explore the reasons for seemingly higher ratios of cancer cases in these regions compared to other parts of the province.

We also provide the childhood cancer ratio estimates for four selected health regions in Figure 4. The crude ratio estimates are defined by $y_{it}/e_{it}$, whereas other estimate is $\exp(\eta_i + \theta_{it})$. As shown, the crude ratio estimates are variable in these sub-regions. It is generally true that a certain trend in the log ratio estimates over time for a region would suggest that the underlying ratio of cancer cases in that region also has the same pattern relative to the provincial average.

4.2 Simulation study on the performance of PQL estimates

We conducted a simulation study to evaluate the performance of PQL estimates using the scenario similar to our childhood cancer data. More specifically, data are generated from the model (1) with the parameters close to those obtained in the analysis of the childhood cancer data; $m, \beta_l(l = 1, ..., 4)$, and $\sigma^2_{\eta}, \lambda_{\eta}, \sigma^2_\theta$ are listed
Figure 3. Childhood cancer adjusted relative risk for the spatial effects of the regional cancer risks for four selected years; Alberta childhood cancer data, 1983-2004. Major urban areas are provided as inserts.

in Table 2. The neighborhood structure and the population sizes are exactly as for the childhood cancer data. Estimates were obtained using PQL analyses of 1,000 data sets generated from the mixed Poisson model (1).

Table 2 presents the bias values of the fixed cubic B-spline parameters and the variance components parameters, as well as the standard deviation of the estimated
parameters and mean values of the estimated standard errors. The estimates are reasonably unbiased, and it seems that their standard errors are estimated quite accurately. Table 3 and Figure 5 assess the adequacy of the normal approximation to the distribution of the standardized estimators. Table 3 reports coverage probabilities of confidence intervals obtained using the usual normal approximation. It seems that $qq$-plots of the standardized estimates of model parameters show good agreement of their distribution with the standard normal, where we only show the $qq$-plots of $\beta_4$ and $\sigma^2$. The use of the normal approximation for constructing confidence intervals seems satisfactory for inference in the mean parameters, while this approach may not be as good for inference concerning variance components; see [13, 15, 24] for more details. Overall, it seems that PQL provides good point estimates and standard errors for this data analysis, and the use of the normal approximation for inference concerning the mean parameter is reasonable, but not as good for variance components.
Figure 5. Normal probability plots of the standardized estimators of $\beta_4$ (a) and the variance component $\sigma^2_\eta$ (b). Quantiles of the standard normal are plotted on the x-axis.

Table 2. Mean values of biases and standard errors, and simulated standard errors of PQL estimates based on 1,000 simulated data sets.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Bias</th>
<th>Standard error (X100)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simulated</td>
<td>PQL</td>
</tr>
<tr>
<td>$m = -0.40$</td>
<td>-0.0061</td>
<td>9.23</td>
</tr>
<tr>
<td>$\beta_1 = 0.03$</td>
<td>0.0032</td>
<td>12.59</td>
</tr>
<tr>
<td>$\beta_2 = 0.50$</td>
<td>-0.0051</td>
<td>16.19</td>
</tr>
<tr>
<td>$\beta_3 = 0.30$</td>
<td>-0.0043</td>
<td>15.11</td>
</tr>
<tr>
<td>$\beta_4 = 0.60$</td>
<td>0.0012</td>
<td>10.43</td>
</tr>
<tr>
<td>$\sigma^2_\eta = 0.02$</td>
<td>-0.0003</td>
<td>1.04</td>
</tr>
<tr>
<td>$\lambda_\eta = 0.33$</td>
<td>0.0001</td>
<td>25.10</td>
</tr>
<tr>
<td>$\sigma^2_\theta = 0.01$</td>
<td>-0.0001</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Table 3. Coverage probabilities of confidence intervals of the model parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Confidence coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.90</td>
</tr>
<tr>
<td>$m$</td>
<td>0.894</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.905</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.906</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.896</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>0.901</td>
</tr>
<tr>
<td>$\sigma^2_\eta$</td>
<td>0.889</td>
</tr>
<tr>
<td>$\lambda_\eta$</td>
<td>0.895</td>
</tr>
<tr>
<td>$\sigma^2_\theta$</td>
<td>0.901</td>
</tr>
</tbody>
</table>

5. Conclusion

We have presented and illustrated a model for spatio-temporal analysis with focus on the mapping of area level disease rates over time. We selected the form and terms in the model to provide a good fit for our childhood cancer incidence data. The model accommodated a cubic B-spline for fixed temporal effects, CAR models for the spatial random effects, and RW models for interaction random effects. The
model adjusted for the main spatial and temporal effects as well as the interaction between them. A class of GLMMs was used to fit the model. We used the PQL method to estimate the model parameters.

It is important to note that there are inferential difficulties with PQL inference in that the relevance of asymptotic standard errors for specific effects is not well established. As indicated by Lin and Breslow [29] and Breslow and Lin [6], the bias corrections are essential for the PQL analysis of data with low numbers of population at risk. Our simulation results indicated that PQL estimators have little bias for the correlated model considered, except when means are very low yielding sparse data with an abundance of zero counts.

We adjusted our expected number of disease by two important factors, age and gender. The model can be easily extended to include some covariates directly into the model which may be required for some applications. However, we could not include such covariates because of low total numbers of cancer cases in our data set. To ensure valid inference, one needs to have more observations in order to include the covariates directly in the model.

Overall, it was suggested by the model estimates that the childhood incidence ratios were increasing over time. It is worthwhile to note that in the United States and Europe the overall rates of childhood cancer have been also increasing since the 1970s [11, 16, 35, 38, 42]. We also showed that the relatively high cancer ratio estimates were mainly located in the south-central part of the province. These findings may represent real increases or may be indicative of different distributions of important covariates that are unmeasured and unadjusted for in our modeling. Further investigation may be needed to explore these findings.

Acknowledgements

Dr. Torabi is supported by a Full-time Fellowship Award from Alberta Heritage Foundation for Medical Research (AHFMR). Dr. Rosychuk is supported by AHFMR as a Health Scholar. The authors thank the Alberta Cancer Registry for providing the cancer data and Alberta Health and Wellness for providing the population data. Disclaimer: The interpretations, conclusions and opinions expressed in this paper are those of the authors and do not necessarily reflect the position of the Alberta Cancer Board. This study is based in part on data provided by Alberta Health and Wellness. The interpretation and conclusions contained herein are those of the researchers and do not necessarily represent the views of the Government of Alberta. Neither the Government nor Alberta Health and Wellness express any opinion in relation to this study.

References

REFERENCES


REFERENCES