

Poisson-Gamma and Spatial-Temporal Models: with Application to Cervical Cancer in Kenya's Counties

Joseph Kuria Waitara¹, Gregory Kerich¹, John Kihoro², Anne Korir³

¹School of Sciences and Aerospace Studies, Moi University, Eldoret, Kenya

²School of Computing and Mathematics, The Co-operative University of Kenya, Nairobi, Kenya

³National Cancer Registry, Nairobi, Kenya

Email address:

jkjoseph834@gmail.com (J. K. Waitara), kerichgregoryy@gmail.com (G. Kerich), kihoro.jm@cuk.ac.ke (J. Kihoro),

annkorir@yahoo.com (A. Korir)

To cite this article:

Joseph Kuria Waitara, Gregory Kerich, John Kihoro, Anne Korir. Poisson-Gamma and Spatial-Temporal Models: With Application to Cervical Cancer in Kenya's Counties. *American Journal of Theoretical and Applied Statistics*. Vol. 10, No. 3, 2021, pp. 158-166.

doi: 10.11648/j.ajtas.20211003.14

Received: June 1, 2021; **Accepted:** June 18, 2021; **Published:** June 26, 2021

Abstract: In Africa, Cancer is an emerging health problem where in 2012 new cancer cases were about 847,00 and around 519,00 deaths, three quarters of those deaths occurred in sub-Saharan region. In 2018, cancer was ranked as the third leading cause of deaths in Kenya after infectious and cardiovascular diseases. In 2018 cancer incidences were estimated to be 47,887 new cancer cases and 32,987 deaths. According to data from World Health Organization in 2020, cervical cancer is the second most prevalent cancer among women while breast cancer is the first. In this study, data collected by the Nairobi Cancer Registry (NCR) was used to produce spatial-temporal distribution of the cervical cancer in counties in Kenya. The results showed that counties where data was available among them Embu, Meru, Machakos, Mombasa, Nyeri, Kiambu, Kakamega, Nairobi and Bomet respectively had high risk of cervical cancer. Availability of county-based estimates and spatial-temporal distribution of cervical cancer cases will aide development of targeted county strategies, enhance early detection, promote awareness and implementation of universal coverage of major control interventions which will be crucial in reducing and halting the rising burden of the cancer cases in Kenya. In counties where data was not available the model showed relative risks for cervical cancer disease was minute but it was present, therefore spatial temporal models are very appropriate to estimate relative risks of diseases even when there is a small sample (and possibly without a sample) in a given area by borrowing information from other neighboring regions.

Keywords: Small Area Estimation, Spatial Temporal, Integrated Nested Laplace Approximation, Generalized Linear Mixed Models

1. Introduction

Cancer arises when normal cells transforms into tumour cell in various stages from pre-cancerous lesion to a malignant tumour. According to the International Agency for Research on Cancer (IARC), in 2018 the global new cancer cases was estimated to be 18.1 million and approximately 9.6 million deaths. In Africa, Cancer is an emerging health problem where in 2012 new cancer cases were about 847,00 and 519,00 deaths, three quarters of those deaths occurred in sub-Saharan region. In 2018, cancer was ranked as the third leading cause of deaths in Kenya after infectious and cardiovascular diseases.

The annual incidence of cancer in Kenya was estimated to be 47,887 new cancer cases, with an annual mortality 32,987 in 2018 [6].

Among women cervical cancer is the fourth prevalent cancer worldwide [27]. According to data from World Health Organization in 2020, in Kenya breast cancer in the most prevalent followed by cervical cancer among women [1]. Human papillomavirus (HPV) infection which is transmitted through direct contact is the cause of almost all cervical cancers [9]. According to Schiffman et al. [25] sexually transmitted HPV genotypes, notably HPV16 cause virtually all cervical cancers world-wide if not controlled immunologically or by screening. The control strategies for

cervical cancer includes early screening, vaccination against HPV, treatment of pre-cancerous lesions, diagnosis and treatment of invasive cervical cancer and palliative care [28].

Cervical cancer screening aides in detection of abnormalities which can be treated and pre cancers which may progress into actual cancer thus reducing cervical cancer incidences, deaths and morbidity related to treatment [25].

Spatial temporal components of events associated with time and location can be utilized to reveal aspects related to where and when the events occurred. Cervical cancer is an event associated with time and space (location) therefore analysis can be conducted to determine the trends, spread and patterns to develop ways to halt its spread [20]. This research project presented a spatial temporal approach to analyze spatial dynamics of the cervical cancer cases in Kenya over two years (2015 and 2016) in ten counties namely Bomet, Embu, Kakamega, Kiambu, Machakos, Meru, Mombasa, Nairobi, Nakuru, and Nyeri County.

Generating spatial or spatial-temporal maps by mapping cancer rates may help to identify distribution of cervical cancer in different geographical locations. Identification of

areas with high disease burden helps in prioritization of control efforts and interventions leading to change of risk behaviours [12].

The study main aim was to determine cervical cancer hotspots by density analysis, and evaluate the trend of spatial correlation for the distribution of cases in Kenya from counties with available data.

2. Materials and Methods

The study focused on cervical cancer data collected from different counties for a period of two years (2015 and 2016) by Kenya National Cancer Registry, which is a national Population-Based Cancer Registry (PBCR) that provides high quality cancer surveillance data which includes incidences, mortality, and survival information of cancer patients.

Kenya has area of 582,650 km square and is divided into 47 counties (See Figure 1) and has population of 47.5 Million people as per the census conducted in 2019 by Kenya National Bureau of Statistics (KNBS).



Figure 1. Kenya Administrative Units.

3. Methodology

According to Chandra [7] Small Area Estimation is a twofold problem. First is how to obtain reliable estimates based on data obtained from small areas where some have empty samples and others with very small samples. The second question is how estimation error is to be assessed. Survey data can be used to estimate indicators of small areas with small samples sizes or empty samples. Direct estimators of indicators in small areas might have large sampling errors, their estimation is improved by introducing regression models which provides a relation between independent and dependent variables of interest. Small area estimation is applied in modelling data from this form of non-planned domains (small areas) [2]. Small Area Estimation methods are divided into area and unit level small area models [22].

Gómez-Rubio et al. [11] noted that, fitting Bayesian models using available in-samples data (samples where data is available) can provide reliable estimates for off-samples areas (where data was not available), area level covariates can be included in calculation of the estimates when available. The estimates maybe less accurate as compared when the survey data is available for each area, but they are reasonable since they have lower bias. To model large-scale spatial patterns and with very sparse data, spatial random effects are included at regional level.

In this study the Bayesian Hierarchical Generalized Linear Mixed Models (BHGLMMs) which are used in small area estimation because of their ability to incorporate multiple levels of model dependencies [8] were considered.

Posterior distributions of model parameters are estimated using Bayes method through combination of observed data, prior distributions and the available covariates [13]. Strong spatial autocorrelation is generally exhibited in small area level data

According to Lawson [15], introduction of spatially structured random effects in the model residual spatial autocorrelation is accounted. Conditional autoregressive priors first proposed by Besag et al. [4] are applied in modelling of spatially structured random effects. The equations in the paper were written using Math Type software [19].

3.1. Generalized Linear Mixed Model

In Generalized Linear Mixed Models (GLMMs), the major assumption is that the response variable Y_i comes from an exponential family of the form $Y_i / \theta_i, \phi_i \sim p(\cdot)$.

The exponential family is defined as,

$$p(y_i / \theta_i, \phi_i) = \exp\left(\frac{y_i \theta_i - b(\theta_i)}{a(\phi_i)} + c(y_i, \phi_i)\right) \quad (1)$$

for $i = 1, \dots, n$ observations and θ_i is the scalar canonical parameter.

Equation (2) is obtained by linking the mean $\mu_i = E(y_i / \beta f^i(\cdot)), \phi_i$ through monotonic link function

$g(\cdot)$. The link takes different forms depending on the model

applied for example \log for Poisson model.

$$g(\mu_i) = \eta_i = \beta_0 + \sum_{i=1}^n f_i(\mu_i) + \sum_{k=1}^m \beta_k x_{ki} + \varepsilon_i \quad (2)$$

If the form of the functions $f_i(\cdot)$ is varied, it can accommodate different models such as, spatial, spatial-temporal, hierarchical regression and time series. x 's are covariates and ε_i 's are error terms.

3.1.1. Poisson-Gamma Model

A typical Poisson model cannot model extra variance, Poisson-gamma (PG) model is also known as Negative Binomial model which incorporates gamma distributed random-effects is used as an alternative. Poisson-gamma (PG) model, has the following two-level formulation:

$$y_i \sim \text{Poisson}(E_i \theta_i);$$

$$\theta_i \sim \text{Gamma}(\alpha, \beta)$$

In first level the assumption is that the random variable y_i (count data) has Poisson distribution or can be written as $y_i \sim \text{Poisson}(e_i \mu_i \theta_i)$ with probability density function:

$$g(y_i | e_i \mu_i \theta_i) = \frac{e^{-(e_i \mu_i \theta_i)} (e_i \mu_i \theta_i)^{y_i}}{y_i!}, y_i = 0, 1, \dots \quad (3)$$

where $\underline{\mu}_i = \mu(x_i, \beta)$ is regression model

$x_i = (x_{1i}, x_{2i}, \dots, x_{pi})^T$ is a vector of covariates and

$\underline{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^T$ are regression coefficients.

In the second stage, θ_i has Gamma distribution or $\theta_i \sim \text{iidGamma}(\alpha, \alpha)$ where the probability density function is:

$$k(\theta_i) = \frac{\alpha^\alpha}{\Gamma(\alpha)} e^{-\alpha \theta_i} \theta_i^{\alpha-1}, \theta_i > 0 \quad (4)$$

based on equation above the joint probability density function is obtained as follows:

$$h(y_i, \theta_i) = \frac{e^{-(e_i \mu_i \theta_i)} (e_i \mu_i \theta_i)^{y_i}}{y_i!} \frac{\alpha^\alpha}{\Gamma(\alpha)} e^{-\alpha \theta_i} \theta_i^{\alpha-1}, y_i = 0, 1, \dots; \theta_i > 0 \quad (5)$$

The marginal probability density function can be obtained as follows

$$m(y_i) = \int_0^\infty h(y_i, \theta_i) d\theta_i$$

$$= \binom{y_i + \alpha - 1}{\alpha - 1} \left(\frac{\alpha}{e_i \mu_i + \alpha} \right)^\alpha \left(1 - \frac{\alpha}{e_i \mu_i + \alpha} \right)^{y_i} \quad (6)$$

The distribution of equation is negative binomial with mean and variance for y_i given as follows:

$$E(Y_i | \beta, \alpha) = e_i \mu_i \text{ and } Var(Y_i | \beta, \alpha) = e_i \mu_i \left(1 + \frac{e_i \mu_i}{\alpha} \right)$$

The posterior distribution for θ_i is estimated as follows

$$\begin{aligned} \pi(\theta_i | y_i, \beta, \alpha) &= \frac{h(y_i, \theta_i)}{m(y_i)} \\ &= \frac{(e_i \mu_i + \alpha)^{y_i + \alpha}}{\Gamma(y_i + \alpha)} e^{-(e_i \mu_i + \alpha)} (\theta_i)^{y_i + \alpha - 1}, \theta_i > 0 \end{aligned} \quad (7)$$

The posterior distribution for θ_i is Gamma based on equation 7 or can be written as:

$$\theta_i | y_i, \beta, \alpha \sim \text{Gamma}(y_i + \alpha, e_i \mu_i + \alpha)$$

The Bayes estimate for θ_i , generates the posterior mean and posterior variance a, as follows:

$$\widehat{\theta}_i^B(\beta, \alpha) = E_B(\theta_i | y_i, \beta, \alpha) = \frac{(y_i + \alpha)}{(e_i \mu_i + \alpha)} \quad (8)$$

and

$$Var_B(\theta_i | y_i, \alpha, v) = \frac{(y_i + \alpha)}{(e_i \mu_i + \alpha)^2} \quad (9)$$

Spatial correlation of the data is not taken in to account in this model since it only introduces a spatially-unstructured over-dispersion [21]. Due to the stated disadvantage and inability to include covariates Poisson-Gamma model has been criticized and has shown to be inferior to Conditional Autoregressive (CAR) convolution models which are more complex [16].

3.1.2. Conditional Autoregressive (CAR) Models

CAR models have been extensively used to model spatial data in various areas such as in epidemiology, agriculture, demography, economy and image analysis as model for both latent and observed variables

To obtain latent Gaussian models, Gaussian priors are assigned to $\beta_0, f_i(\cdot), \beta_k$ and ε_i in equation 2. This can be represented as $\Theta = (\beta_k, s, f_i, s, \dots)$. where Θ is unobserved multivariate Gaussian random variable, whose density $\pi(\Theta / \varphi)$ is controlled by a vector of hyper parameters Φ [17], $\Phi = (\phi_1, \phi_2)$ may not follow Gaussian distribution [18].

The key components of the Latent Gaussian model are, the likelihood of the data $\pi(y / \Theta)$, the Gaussian density of the random vector Θ , $\pi(\Theta / \Phi)$ and the prior distribution of the parameter vector $\pi(\Phi)$.

The posterior is therefore defined as

$$\pi(\Theta, \Phi / y) \propto \pi(\Theta) \pi(\Theta / \Phi) \prod_{i=1} \pi(y_i / x_i, \Phi) \quad (10)$$

The posterior marginal's for x_i and posterior marginal's for Φ or some Φ_j can be obtained by applying integrated Nested Laplace approximation (INLA) [17].

3.2. Integrated Nested Laplace Application (INLA) Methodology

This is an appropriate inference based method for approximating the posterior marginal's of the latent Gaussian field $\pi(x_i / y), i = 1, \dots, n$ in three steps.

The posterior marginal's of the latent effects Θ are written as

$$\pi(x_i / y) = \int \pi(x_i / \Phi, y) \pi(\Phi / y) d\Phi \quad (11)$$

$$\pi(\Phi_i / y) = \int \pi(\Phi / y) d\Phi_{-j} \quad (12)$$

The posterior marginal's $\tilde{\pi}(x_i / y)$ and $\tilde{\pi}(\Phi_i / y)$ can be approximated using the Laplace approximation. The first approximation to $\pi(\Phi / y)$ using Gaussian distributions is constructed as follows;

$$\pi(\Phi / y) = \frac{\pi(\Theta, \Phi, y)}{\tilde{\pi}_G(\Theta / y)} \Big|_{\Theta = \Theta^*(\Phi)} \quad (13)$$

where $\tilde{\pi}_G(\Theta / \Phi, y)$ is a Gaussian approximation to the full conditional of Θ and $\Theta^*(\Phi)$ is the mode of the full conditional for, for a given value of Φ . It involves locating the mode of $\tilde{\pi}(\Phi / y)$ which is used to integrate out the uncertainty with respect to Φ when approximating the posterior marginal of x_i .

The posterior marginals of the latent field are supposed to start from $\tilde{\pi}_G(x_i / \Phi, y)$ and approximate the density of $x_i / \Phi, y$ with the Gaussian marginal derived from $\tilde{\pi}_G(\Theta / \Phi, y)$ i.e

$$\tilde{\pi}(x_i / \Phi, y) = N(x_i; (\Phi), \delta_{ii}^2(\Phi)) \quad (14)$$

The marginals of the interest can be computed using numerical integration over a multidimensional grid of values of Φ

$$\tilde{\pi}(x_i / y) = \sum_k \tilde{\pi}(x_i / \Phi_k, y) \tilde{\pi}(\Phi_k / y) \Delta_k \quad (15)$$

where the sum is over the values of Φ with area weights Δ_k [24].

The first step in INLA computation involves approximating the posterior marginal of Φ by using Laplace approximation in equation (13).

The second step involves computing the Laplace

approximation of $\tilde{\pi}(x_i / \Phi, y)$ for selected values of Φ which improves the Gaussian approximation in equation (10)

$$\tilde{\pi}_{LA}(x_i / \Phi, y) \propto \frac{\pi(\Theta, \Phi, y)}{\tilde{\pi}_{GG}(\Theta_{-i} / x_i, \Phi, y)} \Big|_{\Theta_{-i} = \Theta_{-i}^*(x_i, \Theta)} \quad (16)$$

where $\tilde{\pi}_{GG}(\Theta_{-i} / x_i, \Phi, y)$ is a Gaussian approximation to $\Theta_{-i} / x_i, \Phi, y$ around its mode $\Theta_{-i}(x_i, \Phi)$.

An improved version of $\tilde{\pi}_{LA}(x_i / \Phi, y)$ known as Simplified Laplace approximation was developed by Rue et al. [23]. It involves a series of expansion of $\tilde{\pi}_{LA}(x_i / \Phi, y)$ around $x_i = \mu_i(\Phi)$ which corrects for skewness and location and it is also less computationally expensive [23].

The third step involves combining steps 1 and 2 using numerical integration in equation 13 [23].

Cervical Cancer Distribution

Generalized Linear Mixed Model with spatial and temporal random effects and assuming that the response variable was generated by a Poisson process (count data) was used to model cervical cancer cases.

Poisson-gamma (PG) model, has the following two-level formulation:

$$y_i \sim \text{Poisson}(E_i \theta_i);$$

$$\theta_i \sim \text{Gamma}(\alpha, \beta)$$

Where y_i represents observed cervical cancers cases for each county and E_i are the expected cervical cancer cases.

In convolution model, the Poisson regression model is stated as:

$$y_i \sim \text{Poisson}(\mu_i)$$

$$\mu_i = \exp(X_i \beta + \text{offset}_i)$$

Where y_i is the observed cervical cancer cases, X_i 's are the covariates (when available) for the i^{th} observation and offset term is the expected i^{th} population per county.

In this study cervical cancer cases y_i 's were modelled as shown in equation (17) below.

$$g(\mu_i) = \beta_0 + \sum_j \beta_j X_{ij} + f_{trend}(time) + f_{str}(S_i) + f_{unstr}(S_i) \quad (17)$$

Where β_j 's are coefficients, X_{ij} is vector of the covariates, f_{trend} is trend component, f_{str} and f_{unstr} are structured and unstructured spatial effects of the county.

- i. $g(\cdot)$ is a monotonic link function
- ii. β_0 an overall intercept term.
- iii. β_j 's are coefficients
- iv. Correlated random time effects, f_{trend} , to account for time dependence, were modelled via first order random walk [14, 15]. The assumption is values in the

latest year depend on values of the previous year in a specific county. The correlated temporal random effect, f_{trend} , which has a random walk prior distribution, with precision, $\tau_{\phi 1}$ was assigned Gamma (1, 0.001) prior.

- v. The spatial effects $f_{str}(S_i)$, estimated at county level where households were located. The spatial effects by county to account for strong spatial autocorrelation, and were modelled via normal conditionally autoregressive priors (CAR) [4].
- vi. Where $i = 1, \dots, m$, and $j = 1, \dots, T$, are counties and years respectively;

$$u_i | u_j, \tau_u \sim N\left(\frac{1}{m} \sum_{j \in \delta_i} w_{ij} u_j, \frac{1}{n_{\delta_i} \tau_u}\right) \quad i \neq j \quad (18)$$

where, τ_u is the conditional precision of spatial random effects and δ_i is the neighborhood of the i^{th} region, n_{δ_i} is the

number of neighbours, $\sum_{j=1}^m w_{ij}$, and the spatial weight, ω_{ij}

equals 1 if counties i and j are neighbours and zero otherwise. Spatial distribution is utilized to determine the number of neighbours and it is assumed that each county has not less than one neighbour. $\tau_u \sim \text{Gamma}(1, 0.001)$ prior was assigned as the conditional precision for the spatially structured random effect [5].

The Gamma (α, β) density is defined as:

$$\pi(\tau) = \frac{\beta^\alpha}{(\alpha - 1)} \tau^{(\alpha - 1)} \exp(-\beta \tau), \text{ for } \tau > 0, \text{ where } \alpha > 0, \text{ the shape parameter, and } \beta > 0, \text{ the inverse scale parameter.}$$

$f_{unstr}(S_i)$ un-correlated random effects by county were assigned a Normal prior, $f_{unstr} \sim N(0, \frac{1}{\tau_v})$, with precision, τ_v assigned Gamma (1, 0.001) prior.

3.3. Model Selection Criteria

Deviance Information Criteria

The Deviance Information Criterion (DIC) [26] similar to Akaike Information Criterion (AIC) and is applied in hierarchical Bayesian models. It is defined for improper priors and provides the effective number of parameters in the Bayesian models.

The deviance is

$$D(x, \theta) = -2 \sum_{i \in I} \log \pi(y_i | x_i, \theta) + \text{constant} \quad (19)$$

The model with smaller DIC value [26] was used to calculate the relative risks. The calculated relative risks are subsequently mapped to geographical areas to produce spatial

temporal maps.

4. Results

Data in this study was analyzed using spatial temporal model R-packages. The packages contain functions for Generalized Linear Mixed Model (GLMM) and INLA methodology.

4.1. Spatial Temporal Maps

This section displays the distribution of notified cases, Standard Incidence Ratios (SIR) map, for all counties with notified cervical cancer cases over two years (2015-2016). Where ($SIR_i > 1$) indicates the risk of cervical cancer is higher, equal ($SIR_i = 1$) or ($SIR_i < 1$) lower risk than that which is expected from the standard population.

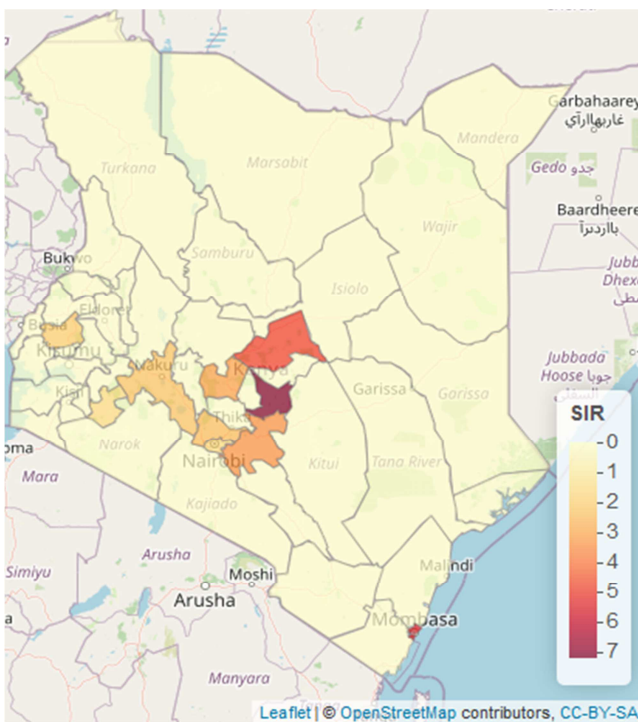


Figure 2. Standardized Incidence Ratios (SIR).

Clearly, from Figure 2 in most counties there was greater risk of cervical cancer cases than expected from the standard population since all counties where data was available had a SIR value greater than 1. Bomet=1.59, Embu =7.13, Kakamega =2.02, Kiambu =2.42, Machakos =3.44, Meru=4.82, Mombasa =5.51, Nairobi=1.66, Nakuru =2.26, Nyeri =3.07.

The deep purple areas indicated higher risks ($SIR > 1$) while the light shaded areas are low risk ($SIR < 1$). The highest burden of cervical cancer cases was in Embu, Mombasa, Meru, Machakos and Nyeri counties respectively.

4.2. Assessing the Presence of Spatial Autocorrelation

Spatial-autocorrelation in spatial-temporal models is modelled via random effects. Therefore, assessment of presence of spatial autocorrelation was conducted by computing the residuals of a fitted simple Poisson log-linear

model.

Table 1. Spatial autocorrelation estimate.

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	8.09e-11	0.0306	4.15	1

Table 2. Over dispersion test.

z-value	p-value	Dispersion parameter
4.15	0.0306	31.202

The over dispersion parameter is equal to 31.202 indicating a relatively high overdispersion, the Poisson model has equal mean and variance assumption.

Moran's I statistic was computed and permutation test for each year of the data to check for spatial-autocorrelation in the model residuals.

The null hypothesis is; there is no spatial autocorrelation while the alternative is; there is positive spatial autocorrelation. The estimated Moran's I statistic was 0.0399 and the p-value was 0.2104 > 0.05, suggesting there was no unexplained spatial autocorrelation in the residuals.

Table 3. Moran I statistic.

Moran I Statistic	p-value
0.0399	0.2104

4.3. Poisson-gamma Model

A Poisson-gamma (Negative-Binomial) model that takes care of over dispersion model and zero inflated variables was fitted. The dispersion parameter for random effect was 1.8692 which was close to 1 and the Akaike Information Criterion (AIC) was 262.9605.

Table 4. Poisson-Gamma estimate.

	Estimate	Std. Error	t-value	Pr(> z)
(Intercept-Fixed effects)	-0.3178	0.3731	-0.852	0.399
Summary of random effects estimates				
Baringo	0.0144	2.5729		
Bomet	2.1385	0.4115		
Bungoma	0.0058	2.5734		
...				

Relative risk estimates for counties where data was available were presented in Table 5. In counties where data was not available the relative risks ranged from 0.01 to 0.06 (Lamu).

Table 5. Relative risks for cervical cancer Poisson-Gamma model.

County	Relative Risk
Bomet	2.14
Embu	9.89
Kakamega	2.78
Kiambu	3.40
Machakos	4.76
Meru	6.48
Mombasa	7.41
Nairobi	2.28
Nakuru	2.19
Nyeri	4.28

The relative risks in Table 5 revealed that Embu county had the highest risk of cervical cancer, followed by Mombasa, Meru, Machakos, Nyeri, Kiambu, Kakamega, Nairobi, Nakuru and Bomet county respectively.

4.4. Spatio-temporal Modelling

Standardized Incidence Ratios (SIRs) sometimes can be useful, but in areas with rare diseases or with small (possibly empty) samples SIRs may be misleading and not very reliable as measure of risk since expected counts may be low as compared to actual observations.

Therefore, estimating disease risks using models which borrow information from neighbouring areas and incorporate covariates information in shrinking or smoothing of extreme values based on small samples is appropriate [10].

The parametric formulation for spatial- temporal models was introduced by Bernardinelli et al. [3], with assumption that the linear predictor can be written as:

$$n_{it} = \alpha + u_i + v_i + \beta t \tag{20}$$

u_i is the spatial structured and v_i unstructured components with a main linear trend β , which represents the overall time effect.

In R-INLA, the first model was specified as follows:

formula1 <-y~1+f (Area, model="bym", graph= Kenya.adj) + f (Area1, Year, model= "iid") + Year...Model 1

Area in Model 1 represents the spatial structured effect while Area1 represents spatial unstructured component.

This specification assumes a linear effect of time for each area (i). The parameters estimated by INLA are

$\theta = \{\alpha, \beta, u, v\}$, while the hyper-parameters are $\psi = \{\tau_u, \tau_v\}$.

The assumption of linearity in the i can be realized using a dynamic non parametric formulation for the linear predictor

$$\eta_{it} = \alpha + u_i + v_i + \gamma_t \tag{21}$$

Here, α , u_i and v_i have the same parameterisation as in (20); however, the term γ_t denotes the temporally structured effect, where it is modelled using a first order random walk. In this formulation parameters estimated by INLA are $\theta = \{\alpha, \xi, v, \gamma\}$ and hyper-parameters are represented by $\psi = \{\tau_u, \tau_v, \tau_\gamma\}$.

This second model was specified in R-INLA as:

formula2 <-y~1+f (Area, model="bym", graph=Kenya.adj)+ f(year, model= "rw1")+ f(Area1, model="iid")...Model 2

A third model expanding Model 1 to explain the time trend of cervical cancer cases and to include space-time interaction was specified as follows:

$$\eta_{it} = \alpha + u_i + v_i + \gamma_t + \delta_{it} \tag{22}$$

In this model parameters estimated by INLA are $\theta = \{\alpha, \xi, v, \gamma, \delta\}$ and hyper-parameters are represented by

$$\psi = \{\tau_u, \tau_v, \tau_\gamma, \tau_\delta\}.$$

The third model R-INLA code was formulated as follows:
formula3 <-y~1+f (Area, model="bym", graph=Kenya.adj) + f (year, model= "rw1")+ f(Area. Year, model="iid")...(Model 3)

The DIC values for the three models are presented in Table 6. Despite the added complexity the third model with dynamic time trend and space-time interaction had a smaller DIC value indicating a that the model fitted well to the data and was the most appropriate and the relative risks were obtained from the model.

Table 6. Spatial-temporal models Deviance Information Criterion (DIC) for defined Equations (13)-(15).

Model	\bar{D}	PD	DIC
Model 1	157.0350	39.2626	196.2976
Model 2	174.4258	30.1654	204.5911
Model 3	158.1371	35.6865	193.8236

The disease risks are usually presented as relative risks in Poisson models. Posterior relative risk distributions greater than 1 indicates an elevated risk of the disease.

Table 7. Relative risks for cervical cancer Poisson-Gamma model.

County	Relative Risk
Bomet	1.53
Embu	7.92
Kakamega	2.06
Kiambu	2.80
Machakos	3.48
Meru	4.43
Mombasa	5.12
Nairobi	1.63
Nakuru	1.95
Nyeri	3.33

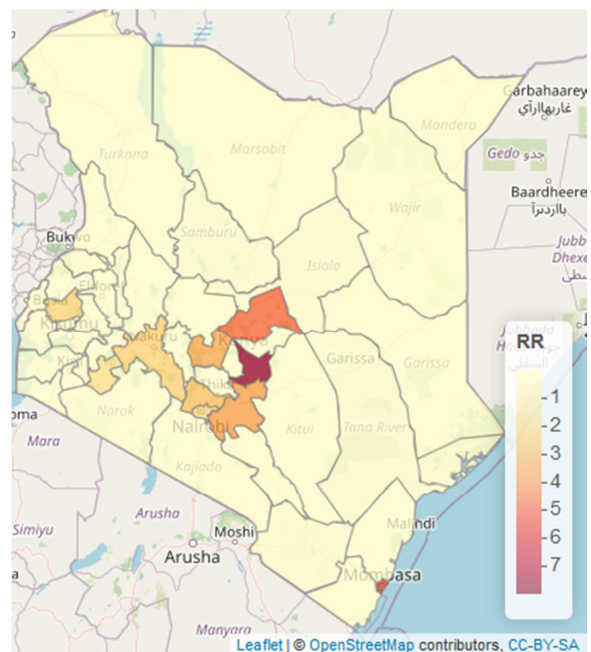


Figure 3. Distribution of the county specific relative risks of cervical cancer in the spatial-temporal model.

The relative risks in Table 7 indicates that Embu county had the highest risk of cervical cancer, followed by Mombasa, Meru, Machakos, Nyeri, Kiambu, Kakamega, Nakuru, Nairobi and Bomet county respectively.

In Figures 3-4 elevated risks are manifested by values greater than 1 in some parts of the country, and a posterior probabilities greater than 0.8.

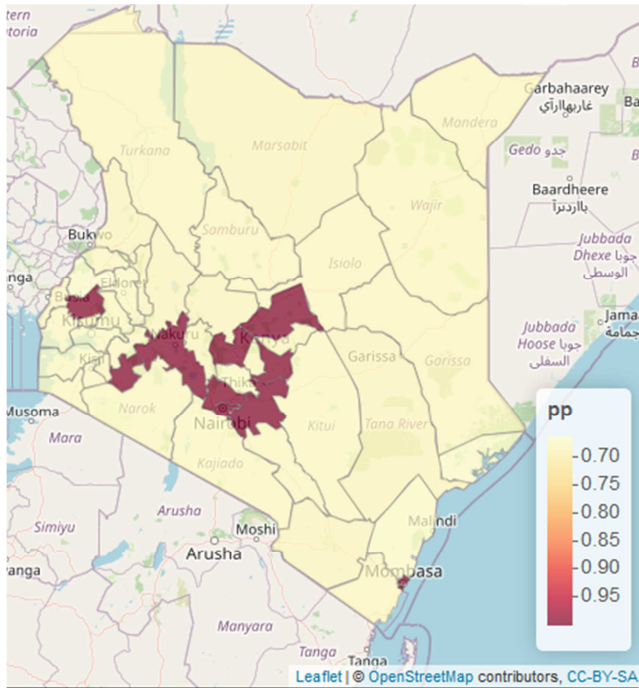


Figure 4. Map of the uncertainty for the spatial effect: $p(\theta_i > 1|y)$.

5. Discussion and Conclusion

The results show that counties where data was available among them Embu, Mombasa, Meru, Machakos and Nyeri counties had very high risk of cervical cancer. The national and county institutions can utilize spatial temporal tools to identify various cancer hot spots and improve screening and treatment facilities based on specific cancer case.

In counties where data was not available the model showed relative risks of cervical cancers was not high but the risk was present, therefore spatial temporal models are very appropriate to estimate relative risks of diseases even when there is a small sample (and possibly an empty sample) in a given area by borrowing information from other neighboring regions. Based on the study findings we recommend the counties with high relatives to create awareness, provide screening services and provide vaccines to the groups which are at higher risk of cervical cancer.

Despite success of this study, the biggest impediment in spatial temporal study is non-availability of adequate county data which will provide more insight on the distribution of cervical cancer cases in Kenya. Therefore the National Cancer Registry in collaboration with counties health departments should work closely to enhance cancer data collection. This will facilitate research and inform the appropriate measures to be implemented in mitigation of the increase of cancer cases.

References

- [1] The Global Cancer Observatory, (2021, March 2021): <https://gco.iarc.fr/today/data/factsheets/populations/404-kenya-fact-sheets.pdf>.
- [2] Roberto Benavent and Domingo Morales. Multivariate fay-herriot models for small area estimation. *Computational Statistics & Data Analysis*, 94: 372–390, 2016.
- [3] L Bernardinelli, D Clayton, C Pascutto, C Montomoli, M Ghislandi, and M Songini. Bayesian analysis of space-time variation in disease risk. *Statistics in medicine*, 14 (21-22): 2433–2443, 1995.
- [4] Julian Besag, Jeremy York, and Annie Mollié. Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics*, 43 (1): 1–20, 1991.
- [5] Roger Bivand. Creating neighbours. 2019.
- [6] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L Siegel, Lindsey A Torre, and Ahmedin Jemal. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68 (6): 394–424, 2018.
- [7] Hukum Chandra. Overview of small area estimation techniques. *Indian Agricultural Statistics Research Institute*, pages 75–88, 2003.
- [8] Avital Cnaan, Nan M Laird, and Peter Slasor. Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Statistics in medicine*, 16 (20): 2349–2380, 1997.
- [9] Elizabeth TH Fonham, Andrew MD Wolf, Timothy R Church, Ruth Etzioni, Christopher R Flowers, Abbe Herzig, Carmen E Guerra, Kevin C Oeffinger, Ya-Chen Tina Shih, Louise C Walter, et al. Cervical cancer screening for individuals at average risk: 2020 guideline update from the american cancer society. *CA: A Cancer Journal for Clinicians*, 70 (5): 321–346, 2020.
- [10] Alan E Gelfand, Peter Diggle, Peter Guttorp, and Montserrat Fuentes. *Handbook of spatial statistics*. CRC press, 2010.
- [11] Virgilio Gómez-Rubio, Nicky Best, Sylvia Richardson, Guangquan Li, and Philip Clarke. Bayesian statistics small area estimation. 2010.
- [12] Marie-Josèphe Horner, Sean F Altekruse, Zhaohui Zou, Louise Wideroff, Hormuzd A Katki, and David G Stinchcomb. Us geographic distribution of prevaccine era cervical cancer screening, incidence, stage, and mortality. *Cancer Epidemiology and Prevention Biomarkers*, 20 (4): 591–599, 2011.
- [13] Diba Khana, Lauren M Rossen, Holly Hedegaard, and Margaret Warner. A bayesian spatial and temporal modeling approach to mapping geographic variation in mortality rates for subnational areas with r-inla. *Journal of data science: JDS*, 16 (1): 147, 2018.
- [14] Leonhard Knorr-Held and G Rasser. Bayesian detection of clusters and discontinuities in disease maps: Simulations. (revised, june 1999). 1999.

- [15] Andrew B Lawson. *Bayesian disease mapping: hierarchical modeling in spatial epidemiology*. Chapman and Hall/CRC, 2013.
- [16] Andrew B Lawson, AB Biggeri, Dankmar Böhning, Emmanuel Lesaffre, Jean-François Viel, ALLAN Clark, PETER Schlattmann, and Fabio Divino. Disease mapping models: an empirical evaluation. disease mapping collaborative group. *Statistics in medicine*, 19 (17): 2217–41, 2000.
- [17] Sara Martino and Håvard Rue. Implementing approximate bayesian inference using integrated nested laplace approximation: A manual for the inla program. *Department of Mathematical Sciences, NTNU, Norway*, 2009.
- [18] Thiago G Martins, Daniel Simpson, Finn Lindgren, and Håvard Rue. Bayesian computing with inla: new features. *Computational Statistics & Data Analysis*, 67: 68–83, 2013.
- [19] Mathtype. <https://www.wiris.com/en/mathtype/office-tools/>. April 2021.
- [20] Kirumba John Mwangi. *Use of GIS in mapping of cancer prevalence a case study of Uasin Gishu County*. PhD thesis, UNIVERSITY OF NAIROBI, 2014.
- [21] Thomas Neyens, Christel Faes, and Geert Molenberghs. A generalized poisson-gamma model for spatially overdispersed data. *Spatial and spatio-temporal epidemiology*, 3 (3): 185–194, 2012.
- [22] JNK Rao. Some new developments in small area estimation. *Journal of the Iranian Statistical Society*, 2 (2): 145–169, 2003.
- [23] Håvard Rue, Sara Martino, and Nicolas Chopin. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, 71 (2): 319–392, 2009.
- [24] Håvard Rue and Sara Martino. Approximate bayesian inference for hierarchical gaussian markov random field models. *Journal of statistical planning and inference*, 137 (10): 3177–3192, 2007.
- [25] M Schiffman, J Doorbar, N Wentzensen, S de Sanjose, C Fakhry, BJ Monk, MA Stanley, and S Franceschi. vol. 2. 2016. carcinogenic human papillomavirus infection. *Nature Reviews. Disease Primers*. [Abstract][Google Scholar], page 16086, 2016.
- [26] David J Spiegelhalter, Nicola G Best, Bradley P Carlin, and Angelika Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the royal statistical society: Series b (statistical methodology)*, 64 (4): 583–639, 2002.
- [27] Chiara Vidoni, Letizia Vallino, Alessandra Ferraresi, Eleonora Secomandi, Amreen Salwa, Menaka Chinthakindi, Alessandra Galetto, Danny N Dhanasekaran, and Ciro Isidoro. Epigenetic control of autophagy in women’s tumors: role of non-coding rnas. *Journal of Cancer Metastasis and Treatment*, 7, 2021.
- [28] WHO (2020, March2020), Human papillomavirus (HPV) and cervical cancer: [https://www.who.int/en/news-room/fact-sheets/detail/human-papillomavirus-\(hpv\)-and-cervical-cancer](https://www.who.int/en/news-room/fact-sheets/detail/human-papillomavirus-(hpv)-and-cervical-cancer).