



# Bayesian Disease Mapping: A Literature Review With an Application, Using WinBugs Software

Richard R. Southey<sup>1</sup>, Lizanne Raubenheimer<sup>1,\*</sup> and Sarah E. Radloff<sup>1</sup>

<sup>1</sup>Department of Statistics, Rhodes University, Makhanda (Grahamstown), South Africa

**Abstract** In this paper we review the improper and proper conditional autoregressive (CAR) models. A set of spatially correlated Gaussian random effects are assumed. The CAR model which contains components from both the uncorrelated heterogeneity (UH) and correlated heterogeneity (CH) models, have been applied to the South African acute pericarditis 2014 data set, where a Poisson model is used. Acute pericarditis is caused by an inflammation of the pericardium in the heart. The data set has been used to examine whether there is a significant difference between the proper conditional autoregressive prior and the intrinsic conditional autoregressive prior for the correlated heterogeneity component in a convolution model. Both the hyperpriors of the precision for the uncorrelated heterogeneity components were modelled by using the Jeffreys' prior. The sensitivity of the hyperprior has also been investigated. The deaths from this disease in 2014 in South Africa have been considered, where disease maps and the relative risk of acquiring and dying from acute pericarditis have been investigated, as well as the standardised mortality ratio (SMR). The convergence and burn-in period of the models were assessed by the Brook-Gelman-Rubin (BGR) diagnostic. The deviance information criterion (DIC) was used to assess the models.

**Keywords** Acute Pericarditis, Bayesian Statistics, Conditional Autoregressive Model, Disease Mapping, Standardised Mortality Ratio.

AMS 2010 subject classifications 62F15, 62C10, 92B15

DOI: 10.19139/soic-2310-5070-1395

## 1. Introduction

Disease mapping is a useful tool in the analysis and detection of patterns in diseases and involves mapping characteristics of a disease onto a region where the disease occurs. More recently, disease mapping has increased in popularity as the computational power of computers has improved along with the software used to produce disease maps. The method is popular as it is a good way to visualise the characteristics of a disease. There are many different applications of disease mapping. The application which will be considered in this paper is the analysis of the relative risk of contracting and dying from a disease. In this paper, the disease which will be considered is acute pericarditis in South Africa in 2014. A Bayesian approach will be used in this paper. The Bayesian approach requires a prior distribution to be used along with the likelihood function to derive the posterior distribution, a conjugate prior and Jeffreys' prior will be used as the prior distributions and will be combined with the Poisson likelihood to derive the posterior distributions.

Lawson et al. [15] mention that the earliest application of disease mapping was done by Snow in 1854, where a map of addresses was put together for cholera victims relating to the proximity of the

---

\*Correspondence to: Lizanne Raubenheimer (Email: L.Raubenheimer@ru.ac.za). Department of Statistics, Rhodes University, PO Box 94, Makhanda (Grahamstown), South Africa, 6193.

water supply. Recently, disease mapping has experienced substantial growth as it is a vital tool for disease prevention. Present-day disease mapping can incorporate geographic information systems as well as many statistical applications and procedures together. In disease mapping, confounders arise because of the nature of populations. Confounders are defined by Lesaffre & Lawson [16] as known explanatory variables which account for the influence of variables affecting the study population. The confounders are considered in a disease mapping sense via the introduction of random effects into the model. Other variables to be considered are deviation indices. These are variables which are related to a range of “poverty-related explanatory variables.” Disease mapping has also spread across many topics including ecological, geographical, health and spatial studies.

Lawson et al. [15] separate the main activities of disease mapping and list them as disease mapping, disease clustering and ecological analysis. For the disease mapping activity, which is mainly used in the health sector, an estimate of the “true relative risk” is found over a region or map. This application is mainly used for disease prevention and resource allocation. Disease clustering is used to determine where diseases are clustered most in a region. Lawson et al. [15] state that this application is mainly used in environmental assessments. An ecological analysis is primarily used in epidemiological research to provide a “geographical distribution of a disease in relation to explanatory covariates, usually at an aggregated spatial level.”

There are two situations which are most used in the geographic spread of a disease, Lawson et al. [15]:

- a study has a fixed period for which it takes place
- the location of the observations of interest are recorded. This type of study is called a case-event analysis. Often the locations are street addresses or postal codes, and the privacy of information is a limiting factor in these types of studies because of the sensitive nature of the information which is collected. To address the limitations of the case-event analysis study, the tract count analysis study may be used. The tract count analysis study mentioned is based on the premise that the counts of the observations of interest within a small region are often recorded instead of using the locations of street addresses. The regions are arbitrary and may be used for census information.

For this paper, we downloaded the map shape files which are freely available from a number of websites including the Global Administrative Areas available at <http://gadm.org/country>. The shape files were then imported into Quantum Geographic Information System (QGIS)<sup>®</sup>[2016] available at <http://www.qgis.org/en/site/> and imported into WinBUGS<sup>®</sup> developed by Lunn et al. [17]. These maps are then used as the disease maps for the study and provide a graphical representation which simplifies the analysis of the diseases which are analysed. The disease maps used in this paper are used to analyse the posterior expected exceedance probability (PP) and relative risk (RR).

Diggle et al. [8] considered an application to the mapping of environmental count data. Jahan et al. [11] proposed a hierarchical Bayesian meta-analysis model which analysed the point and interval estimates from an online atlas, where they modelled the published cancer incidence estimates available as part of the online Australian Cancer Atlas. Wakefield et al. [24] looked into an application to the mapping of cancer risk in the United Kingdom. Pascutto et al. [20] discussed a number of issues dealing with the analysis of disease mapping, where they considered the mapping of larynx cancer across electoral wards in the North West Thames region of the United Kingdom, making use of Bayesian methods.

A comprehensive review of the main classes of hierarchical models that have been used for Bayesian disease mapping has been considered in Best et al. [4], where they considered extensions to model the joint spatial distribution of health indicators. Anderson et al. [1] proposed a two-stage solution to the problem of identifying the spatial extent of high-risk clusters, where the first stage is a spatially adjusted hierarchical agglomerative clustering algorithm and the second stage fits a separate Poisson log-linear model to the study data for each cluster structure. Coly et al. [6] modelled rare infectious diseases where they proposed some specific and generic variations of the current methodology. They used a specific spatio-temporal link in order to better model the stronger spatial and temporal dependencies due to the transmission of the disease. They have used Bayesian hierarchical models on bovine tuberculosis data.

The model will be discussed in Section 2, followed by the data analysis and a sensitivity analysis in Section 3. The discussion and conclusion will be given in Section 4.

## 2. The Model

Suppose that there exists a set of spatially correlated Gaussian random effects, denoted by  $S_1, S_2, \dots, S_N$ , for the  $N$  regions in the area under the study and suppose that the joint distribution of the random effects are

$$\mathbf{S} \sim MVN(\boldsymbol{\mu}, \omega\Sigma),$$

where  $\mathbf{S}$  is a vector containing  $S_i$ ,  $i = 1, 2, \dots, N$  and is distributed as a multivariate normally distributed random variable of  $N$ -dimension,  $\boldsymbol{\mu}$  is a mean vector of  $1 \times N$  dimension,  $\omega$  is a strictly non-negative value which governs the variability of  $S_i$  and  $\Sigma$  is a  $N \times N$  positive definite matrix as in Spiegelhalter et al. [22]. The “between-region” covariance matrix  $\omega\Sigma$  is expressed in Osei [19] as

$$\omega\Sigma = \omega(I - \rho W)^{-1}M, \quad (1)$$

where  $I$  is an identity matrix of  $N \times N$  dimension,  $M$  is a diagonal matrix of  $N \times N$  dimension containing values, denoted by  $m_{ii}$ , which are proportional to a conditional covariance  $S_i|S_j$  and  $W$  is a weight matrix of  $N \times N$  dimension with values, denoted by  $w_{ij}$ . Where  $w_{ij}$  is representing the spatial correlation between region  $i$  and  $j$  and  $\rho$  is a representation of the amount of spatial dependence.

The joint multivariate Gaussian model is expressed in Spiegelhalter et al. [22] as

$$S_i | \mathbf{S}_{-i} \sim N\left(\mu_i + \sum_{j=1}^N \rho w_{ij}(S_j - \mu_j), \omega m_{ii}\right), \quad (2)$$

where  $\mathbf{S}_{-i}$  denotes the whole vector  $\mathbf{S}$  excluding the value of  $S_i$ . When there are no explanatory variables then  $\mu_i = x_i'\beta = 0$  and (2) can be written as

$$S_i | \mathbf{S}_{-i} \sim N\left(\sum_{j=1}^N \rho w_{ij}S_j, \omega m_{ii}\right). \quad (3)$$

Theorem 1. When there are no explanatory variables then the joint distribution of the random effects are

$$\mathbf{S} \sim MVN(\mathbf{0}, \omega\Sigma),$$

where  $\omega\Sigma$  is given in (1).

Proof

The Brook expansion developed by Brook [5] can be written as:

$$\frac{\pi(\mathbf{z})}{\pi(\mathbf{x})} = \prod_{i=1}^N \frac{\pi(z_i | z_1, z_2, \dots, z_{i-1}, x_{i+1}, \dots, x_N)}{\pi(x_i | z_1, z_2, \dots, z_{i-1}, x_{i+1}, \dots, x_N)},$$

thus

$$\begin{aligned}
\frac{\pi(\mathbf{S})}{\pi(\mathbf{0})} &= \prod_{i=1}^N \frac{\pi(S_i | S_1, S_2, \dots, S_{i-1}, 0_{i+1}, \dots, 0_N)}{\pi(0_i | S_1, S_2, \dots, S_{i-1}, 0_{i+1}, \dots, 0_N)} \\
&= \prod_{i=1}^N \frac{\exp\left\{-\frac{1}{2\omega m_{ii}} \left(S_i - \rho \sum_{j<i} w_{ij} S_j - \rho \sum_{j>i} 0_j\right)^2\right\}}{\exp\left\{-\frac{1}{2\omega m_{ii}} \left(0_i - \rho \sum_{j<i} w_{ij} S_j - \rho \sum_{j>i} 0_j\right)^2\right\}} \\
&= \prod_{i=1}^N \exp\left\{-\frac{1}{2\omega m_{ii}} \left[\left(S_i - \rho \sum_{j<i} w_{ij} S_j\right)^2 - \left(\rho \sum_{j<i} w_{ij} S_j\right)^2\right]\right\} \\
&= \prod_{i=1}^N \exp\left\{-\frac{1}{2\omega m_{ii}} \left[S_i^2 - 2S_i \rho \sum_{j<i} w_{ij} S_j + \left(\rho \sum_{j<i} w_{ij} S_j\right)^2 - \left(\rho \sum_{j<i} w_{ij} S_j\right)^2\right]\right\} \\
&= \prod_{i=1}^N \exp\left\{-\frac{1}{2\omega m_{ii}} \left[S_i^2 - 2\rho S_i \sum_{j<i} w_{ij} S_j\right]\right\} \\
&= \exp\left\{-\frac{1}{2\omega \sum_{i=1}^N m_{ii}} \left[\sum_{i=1}^N S_i^2 - 2\rho \sum_{i=1}^N \sum_{j<i} S_i w_{ij} S_j\right]\right\}.
\end{aligned}$$

Since  $W$  is symmetric then  $2\rho \sum_{i=1}^N \sum_{j<i} S_i w_{ij} S_j = \rho \sum_{i=1}^N \sum_{j=1}^N S_i w_{ij} S_j$  and thus:

$$\begin{aligned}
\exp\left\{-\frac{1}{2\omega \sum_{i=1}^N m_{ii}} \left[\sum_{i=1}^N S_i^2 - 2\rho \sum_{i=1}^N \sum_{j<i} S_i w_{ij} S_j\right]\right\} &= \exp\left\{-\frac{1}{2\omega \sum_{i=1}^N m_{ii}} \left[\sum_{i=1}^N S_i^2 - \rho \sum_{i=1}^N \sum_{j=1}^N S_i w_{ij} S_j\right]\right\} \\
&= \exp\left\{-\frac{1}{2\omega} \left[\sum_{i=1}^N S_i^2 - \rho \sum_{i=1}^N \sum_{j=1}^N S_i w_{ij} S_j\right] \sum_{i=1}^N m_{ii}^{-1}\right\} \\
&= \exp\left\{-\frac{1}{2} \left[S' (I - \rho W) (\omega M)^{-1} S\right]\right\} \\
\therefore S &\sim MVN(\mathbf{0}, \omega (I - \rho W)^{-1} M) \\
\therefore S &\sim MVN(\mathbf{0}, \omega \Sigma).
\end{aligned}$$

□

The use of the covariance matrix  $\Sigma$  requires the matrices  $W$  and  $M$ , and the parameter  $\rho$  to be specified. The CAR model requires  $\Sigma$  to be symmetric positive definite such that the following conditions are met, as outlined in Spiegelhalter et al. [22]:

1. The matrix  $\Sigma$  must be symmetric such that  $w_{ij} m_{jj} = w_{ji} m_{ii}$ .
2. When  $\rho = 0$  there is no spatial dependence.
3. The value of  $\rho$  must be in-between  $\rho_{min}$  and  $\rho_{max}$  such that  $\rho_{min}^{-1}$  and  $\rho_{max}^{-1}$  are lowest and highest eigenvalues of  $M^{-\frac{1}{2}} W M^{\frac{1}{2}}$ .
4.  $Var(S_j | S_i) = \omega m_{ii} > 0$  such that  $m_{ii} > 0$ .
5. Since spatial dependence is mostly positive,  $\rho$  is constrained to be between 0 and  $\rho_{max}$ .

### 2.1. Improper CAR (ICAR) Model

This model’s foundation dates back to 1987 when Künsch [12] developed lattice models for intrinsic autoregressions. These autoregressions were based on a two-dimensional lattice and could be applied to intrinsic models for which stationarity is only assumed when parameters change value. These models were built based on spatial distributions and permit the use of a “singular normal joint distribution”, Lawson [14].

Besag et al. [3] show that the ICAR model is a CAR model where the covariance matrix  $\Sigma$  is semi-definite, whereas the CAR model has a positive definite covariance matrix. The ICAR model creates weights given by

$$w_{ij} = \begin{cases} \frac{1}{ne_i} & \text{if regions } i \text{ and } j \text{ are adjacent,} \\ 0 & \text{otherwise,} \end{cases}$$

where  $w_{ii}$  is also 0 and let  $m_{ii} = \frac{1}{ne_i}$  where  $ne_i$  is the number of neighbours of region  $i$ . The ICAR model also requires

$$\rho = 1 = \rho_{max}.$$

Since  $\rho = 1$ ,  $m_{ii} = \frac{1}{ne_i}$  and  $w_{ij} = \frac{1}{ne_i}$  and since there are no explanatory variables in both data sets then (3) will change to the following equation for the ICAR model

$$S_i | \mathbf{S}_{-i} \sim N\left(\sum_{j \in ne_i} \frac{S_j}{ne_i}, \frac{\omega}{ne_i}\right).$$

WinBUGS® requires unnormalised weights and sets  $Z_{ij} = 1$  for  $i$  and  $j$  adjacent regions and 0 otherwise and let  $w_{ij} = \frac{Z_{ij}}{Z_{ik}}$  where  $Z_{ik} = \sum_{j=1}^N w_{ij}$ .

A constraint which requires the random effects in the model sum to 0 is proposed by Besag & Kooperberg [2] and includes an additional intercept term with a  $U(-\infty, \infty)$  prior which is location invariant and gives the same result as an unconstrained parameterisation without an additional intercept term, this is modelled as  $\beta_0$  in the (4). The prior distribution for the parameter  $\omega$  must be included in the model and be modelled in terms of precision as  $\tau = \frac{1}{\omega}$ . This prior is very sensitive because the posterior variance in the random effect will be affected by the prior variance, Spiegelhalter et al. [22]. The prior distribution is usually modelled by a gamma distribution where the parameters have to be selected sensibly as a large prior variance places most of the prior focus away from 0. Various suggestions are available for the parameters of the gamma distribution. The one proposed by Besag & Kooperberg [2] has parameters of  $\alpha = 0.5$  and  $\beta = 0.005$ .

The model that is applied to both data sets is the CAR model which contains components from both the UH and CH models and is implemented in WinBUGS®. The model has the following parameters

$$\begin{aligned} y_i | \mu_i &\sim Pois(\mu_i) \text{ where } i = 1, 2, \dots, N \text{ and } \mu_i = e_i \theta_i \\ L(\mu_i | y_i) &= \prod_{i=1}^N \frac{\mu_i^{y_i} \exp(-\mu_i)}{y_i!} \\ E(y_i | \mu_i) &= e_i \theta_i = \mu_i \\ \log(\mu_i) &= \log(e_i) + \beta_0 + v_i + S_i, \end{aligned} \tag{4}$$

where  $\theta_i$  is the relative risk in the  $i^{th}$  region,  $e_i$  is the expected number of events taking place in the  $i^{th}$  region,  $v_i \sim N(\mu = 0, \phi)$  is the uncorrelated heterogeneity,  $S_i$  is the ICAR model prior and  $\beta_0$  is the intercept term due to the requirement that the random effects in the model must sum to 0. The model focuses on making inference on the relative risk of each region. The following link is assumed

$$\log(\mu_i) = \eta_i,$$

then

$$\mu_i = e_i^{\eta_i},$$

which results in the Bayesian linear model given by

$$\eta_i = x_i' \beta + \beta_0 + v_i + S_i.$$

The relative risk is modelled by

$$\theta_i = \exp[\beta_0 + v_i + S_i],$$

where  $x_i' \beta$  is not modelled as there are no explanatory variables in the data.

The prior distribution for the vector  $\mathbf{S}$  is written by Lawson [14] as

$$\pi(\mathbf{S} | \omega) \propto \frac{1}{\omega^{\frac{N}{2}}} \exp \left\{ -\frac{1}{2\omega} \sum_i \sum_{j \in ne_i} (S_i - S_j)^2 \right\},$$

which is an application of a ‘‘Markov random field’’. There are various methods for weighting schemes for regions.

The zero-mean Gaussian prior distribution was assigned by Besag et al. [3] to the uncorrelated heterogeneity  $v_i$  in each region. This prior is written as

$$\pi(\mathbf{v}) \propto \phi^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2\phi} \sum_{i=1}^N v_i^2 \right\},$$

where  $\omega$  and  $\phi$  have the following priors

$$\omega \sim Gam(\alpha_1, \beta_1)$$

$$\phi \sim Gam(\alpha_2, \beta_2).$$

The posterior distribution has the following form

$$\begin{aligned} \pi(\mathbf{S}, \mathbf{v}, \omega, \phi | y_i) &\propto \prod_{i=1}^N \left\{ \frac{\exp(-e_i \theta_i) (e_i \theta_i)^{y_i}}{y_i!} \right\} \times \frac{1}{\omega^{\frac{N}{2}}} \exp \left\{ -\frac{1}{2\omega} \sum_i \sum_{j \in ne_i} (S_i - S_j)^2 \right\} \\ &\times \phi^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2\phi} \sum_{i=1}^N v_i^2 \right\} \times Gam(\alpha_1, \beta_1) \times Gam(\alpha_2, \beta_2). \end{aligned}$$

The advantage of using such a model is shown in Lawson [14] as conditional moments

$$\begin{aligned} E(S_i | \mathbf{S}_{-i}) &= \bar{S}_i \\ var(S_i | \mathbf{S}_{-i}) &= \frac{\omega}{ne_i} \\ [S_i | \mathbf{S}_{-i}] &\sim N\left(\bar{S}_i, \frac{\omega}{ne_i}\right), \end{aligned}$$

where  $\bar{S}_i = \sum_{j \in ne_i} \frac{S_j}{ne_i}$ , the average calculated over the districts of the  $i^{th}$  region.

An advantage of using this model is that the conditional moments are easy to calculate and are functions of the regions in the study. The justification for using both uncorrelated heterogeneity (UH) and correlated

heterogeneity (CH) components in the model is provided in Lawson [14]. The CH and UH components must be used in the model since unobserved effects within the region under study can arise in different forms. The UH effect is included in the model to model uncorrelated additional variation. In the absence of prior information, there is little justification not to include both effects, and it is straightforward to include both effects in the form of an additive random effect in the model. These effects may not be identified, but focus may be on the “total effect of the unobserved confounding”, then the sum of these effects can be identified and forms part of the model.

The intraclass correlation coefficient (ICC) may be calculated from the variance components of the UH and CH components as follows.

Let  $\phi_\nu$  and  $\phi_S$  denote the variance of the UH and CH components respectively, then the “relative variance contribution/ intraclass correlation” is given by  $\frac{\phi_\nu}{\phi_\nu + \phi_S}$ , as in Lawson [14]. This “relative variance contribution” is only useful when the components may be identified.

The WinBUGS<sup>®</sup> function for the ICAR model is the *car.normal* function. The following parameters are required for the ICAR model as in Spiegelhalter et al. [22]

$$S[1 : N] \sim \text{car.normal}(\mathbf{adj}[], \mathbf{weights}[], \mathbf{num}[], \omega),$$

where:

- **adj**[] is a vector containing the identification numbers of the neighbours of each region and is generated using the adjacency matrix in the mapping tool menu in GeoBUGS<sup>®</sup>[2003].
- **weights**[] is a vector containing the unnormalised weights for each pair of regions and having the same length as the **adj**[] vector. This is generated as a vector of 1’s by assigning  $m_{ij} = 1$  ( $w_{ij} = \frac{1}{ne_i}$ ) if regions  $i$  and  $j$  are adjacent and 0 otherwise. This is the standard CAR model proposed by Besag et al. [3].
- **num**[] is a vector the same length as the number of regions in the study and contains the number of neighbours ( $ne_i$ ) for each region. This vector is created using the adjacency matrix in the mapping tool menu in GeoBUGS<sup>®</sup>.
- $\omega$  is a value which is the precision of the CAR prior or the inverse scale parameter when using the Laplace prior. This may be assigned a gamma prior distribution in the model.
- The intercept term  $\beta_0$  is assigned a flat prior distribution and must be used in a CAR model that has random effects.

## 2.2. Proper CAR (PCAR) Model

Another form of the Gaussian Markov random field (GMRF) model is the PCAR model. Suppose  $\rho$  is as defined earlier and to ensure definiteness of  $\Sigma$ , the covariance matrix,  $\rho$  must lie in the interval  $\rho_{min}$  and  $\rho_{max}$  such that  $\rho_{min}^{-1}$  and  $\rho_{max}^{-1}$  are the lowest and highest eigenvalues of  $M^{-\frac{1}{2}}WM^{\frac{1}{2}}$ . Thus, the range of  $\rho$  is a function of the eigenvalues of the matrix  $M^{-\frac{1}{2}}WM^{\frac{1}{2}}$ . This is different to the ICAR model where  $\rho$  is fixed, as explained in Spiegelhalter et al. [22]. The distribution which arises is thus proper assuming that the properties of  $M$  and  $W$  are met. A straight forward uniform distribution,  $U(\rho_{min}, \rho_{max})$ , may be given as a hyperprior for  $\rho$ , as mentioned by Besag & Kooperberg [2]. The biggest advantage of using this type of model is that the results will be similar to those produced by the fully specified Gaussian covariance model, because the model does not have to perform a matrix inversion during sampling. The model may also be used as a data likelihood as it is proper and the variance and correlation parameters are specified, and this is in contrast to the ICAR model which is improper and therefore cannot be used as a data likelihood. There are two types of specifications which will be considered in this paper. The specifications are for the elements of the matrices  $M$  and  $W$  and the parameter  $\rho$ . The first specification is made by Besag et al. [3] and is used in the ICAR model, this specification is

$$w_{ij} = \begin{cases} \frac{1}{ne_i} & \text{if regions } i \text{ and } j \text{ are adjacent,} \\ 0 & \text{otherwise,} \end{cases}$$

and

$$m_{ii} = \frac{1}{ne_i},$$

$$\rho = 1 = \rho_{max}.$$

The second specification, which is used in the PCAR model, is proposed by Cressie & Chan [7] and Stern & Cressie [23]

$$w_{ij} = \begin{cases} \left(\frac{e_i}{e_j}\right)^{\frac{1}{2}} & \text{if regions } i \text{ and } j \text{ are adjacent,} \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

and

$$m_{ii} = \frac{1}{e_i}, \quad (6)$$

$$\rho \in (\rho_{min}, \rho_{max}),$$

where  $e_i$  is the expected number of events at the  $i^{th}$  region.

The model with no trend is then written as:

$$S_i | \mathbf{S}_{-i} \sim N\left(\rho \sum_{j \in ne_i} \left(\frac{e_i}{e_j}\right)^{\frac{1}{2}} S_j, \frac{\omega}{e_i}\right).$$

The WinBUGS<sup>®</sup> function for the PCAR model is the `car.proper` function. The following parameters are required for the PCAR model, as is copied from Spiegelhalter et al. [22]

$$\mathbf{S}[1 : N] \sim \text{car.proper}(\mathbf{mu}[], \mathbf{W}[], \mathbf{adj}[], \mathbf{num}[], \mathbf{M}[], \text{omega}, \rho),$$

where:

- $\mathbf{mu}[]$  is a vector containing the means for each region, this may be given a prior distribution, defined in the data or calculated within the model.
- $\mathbf{adj}[], \text{omega}$  and  $\mathbf{num}[]$  are the same as in the ICAR model.
- $\mathbf{W}[]$  is a vector of normalised weights of each pair of regions. The ICAR model used unnormalised weights.
- $\mathbf{M}[]$  is a vector containing the diagonal values of  $M$  denoted by  $m_{ii}$  and has length  $N$ .
- $\rho$  is a value which specifies the strength of the spatial dependence and is constrained by the highest and lowest eigenvalues of  $M^{-\frac{1}{2}}WM^{\frac{1}{2}}$ .

The bounds are created by using the parameters of the PCAR model as

$$\min(\mathbf{W}[], \mathbf{adj}[], \mathbf{num}[], \mathbf{M}[]) \text{ and } \max(\mathbf{W}[], \mathbf{adj}[], \mathbf{num}[], \mathbf{M}[]).$$

### 2.3. Residuals

A common technique used in model goodness-of-fit is the analysis of the residuals of the model. The residuals represent the difference between the observed value and the expected value, this difference is known as the deviance, as defined by Lunn et al. [18]. Residuals may be used to analyse various short falls in the model, some of which are given in Lunn et al. [18] as: autocorrelation, the shape of the distribution and the fit of the model. The general case of a residual is defined as

$$r_{Ri} = y_i - \hat{y}_i, \quad (7)$$



where  $\hat{y}_i$  is the fitted value of the model.

The standardised form of the Pearson residual is well-known and defined as

$$r_{pi} = \frac{y_i - E(y_i|\theta)}{\sqrt{\text{var}(y_i|\theta)}},$$

where  $y_i$  is the  $i^{\text{th}}$  observed value,  $E(y_i|\theta)$  is the expected value of the  $i^{\text{th}}$  observed value given the parameter  $\theta$  and  $\text{var}(y_i|\theta)$  is the variance of the  $i^{\text{th}}$  observed value given the parameter  $\theta$ . The Pearson residual,  $r_p$ , has a mean of 0 and a variance of 1 and thus the residuals are assumed to occur between -2 and 2, as explained by Lunn et al. [18].

The Bayesian residual is defined by Lawson [14] as

$$r_{Bi} = y_i - \frac{1}{G} \sum_{g=1}^G E(y_i|\theta_i^{(g)}),$$

where  $\theta_i^{(g)}$  is a vector of values resulting from the sampling of the posterior distribution and  $E(y_i|\theta_i^{(g)})$  is the expected value resulting from the posterior distribution.

In the case of a model with a Poisson likelihood with mean parameter  $e_i\theta_i$ , the Bayesian residual can be approximated by

$$r_{Bi} = y_i - \frac{1}{G} \sum_{g=1}^G e_i\theta_i,$$

when a constant region rate is applied to the model. Another parameterisation is to use the posterior expected value,  $e_i\hat{\theta}_i$ , as the fitted value and then the Bayesian residual is calculated in Lawson [14] as

$$r_{Bi} = y_i - e_i\hat{\theta}_i,$$

which follows from (7).

The deviance residual is defined by Lunn et al. [18] as

$$r_{di} = \text{sign}_i \sqrt{D_i}, \quad (8)$$

where  $\text{sign}_i$  is the  $i^{\text{th}}$  sign of the result of  $y_i - \hat{y}_i$  and  $D_i$  is the  $i^{\text{th}}$  standardised deviance defined as

$$D(\boldsymbol{\theta}) = \sum_{i=1}^N D_i(\boldsymbol{\theta}) = -2\log[\pi(y|\boldsymbol{\theta})] + 2\log[\pi(y|\hat{\theta}_i(y))], \quad (9)$$

where  $\hat{\theta}_i(y)$  is the ‘‘saturated estimate’’ as in Lunn et al. [18].

Theorem 2. The standardised deviance for the Poisson distribution when  $y_i \sim \text{Poi}(\theta_i)$  is

$$D(\boldsymbol{\theta}) = 2 \sum_{i=1}^N \left[ y_i \log\left(\frac{y_i}{\hat{\theta}_i}\right) - (y_i - \hat{\theta}_i) \right]. \quad (10)$$

Proof

Let  $\hat{\theta}_i$  denote the  $i^{\text{th}}$  maximum likelihood estimate of the fitted model and let  $\tilde{\theta}_i = y_i$  under the saturated model. The likelihood of the fitted model is then

$$L(\hat{\theta}_i | y_i) = \prod_{i=1}^N \frac{\exp(-\hat{\theta}_i) \hat{\theta}_i^{y_i}}{y_i!}$$

and the likelihood of the saturated model is

$$L(\tilde{\theta}_i | y_i) = \prod_{i=1}^N \frac{\exp(-\tilde{\theta}_i) \tilde{\theta}_i^{y_i}}{y_i!}.$$

The log likelihood of the fitted model is then

$$l(\hat{\theta}_i | y_i) = \sum_{i=1}^N \left[ y_i \log(\hat{\theta}_i) - \hat{\theta}_i - \log(y_i!) \right]$$

and the log likelihood of the saturated model is

$$l(\tilde{\theta}_i | y_i) = \sum_{i=1}^N \left[ y_i \log(\tilde{\theta}_i) - \tilde{\theta}_i - \log(y_i!) \right].$$

Substituting the log likelihoods into (9) results in

$$\begin{aligned} D(\boldsymbol{\theta}) &= -2 \sum_{i=1}^N \left[ y_i \log(\hat{\theta}_i) - \hat{\theta}_i - \log(y_i!) \right] + 2 \sum_{i=1}^N \left[ y_i \log(\tilde{\theta}_i) - \tilde{\theta}_i - \log(y_i!) \right] \\ &= 2 \sum_{i=1}^N \left[ \log(y_i!) - y_i \log(\hat{\theta}_i) + \hat{\theta}_i + y_i \log(\tilde{\theta}_i) - \tilde{\theta}_i - \log(y_i!) \right] \\ &= 2 \sum_{i=1}^N \left[ \hat{\theta}_i + y_i \log\left(\frac{\tilde{\theta}_i}{\hat{\theta}_i}\right) - \tilde{\theta}_i \right]. \end{aligned}$$

Since  $\tilde{\theta}_i = y_i$  under the saturated model, then the deviance is

$$\begin{aligned} D(\boldsymbol{\theta}) &= 2 \sum_{i=1}^N \left[ \hat{\theta}_i + y_i \log\left(\frac{y_i}{\hat{\theta}_i}\right) - y_i \right] \\ &= 2 \sum_{i=1}^N \left[ y_i \log\left(\frac{y_i}{\hat{\theta}_i}\right) - (y_i - \hat{\theta}_i) \right]. \end{aligned}$$

□

### 3. Data Analysis

The data was collected via the civil registration system which is part of the Department of Home Affairs. Data collection is important to analyse the current health system within the country. These death notifications are collected by the Department of Home Affairs and are processed fortnightly by Statistics South Africa which produces statistical releases based on the data collected.

The South African data for acute pericarditis (Classification number: I30) is given in Table 1 along with the expected values, which are calculated based upon the province-wide incidence rate. The disease map for the observed values is given in Figure 1. Imazio et al. [10] state that pericarditis is caused by an inflammation of the pericardium in the heart. The pericardium is a sac which surrounds the heart and contains many blood vessels and is made up of visceral and parietal layers, Lange & Hillis [13]. These layers are distanced by the pericardial cavity and contain between 15 to 50 ml of fluid. Pericarditis may cause no pain at all and may be lethargic which is common in patients who have tuberculosis, Lange &

Hillis [13]. In other cases, pain may be severe and occur suddenly without warning. Pain often resonates behind the breast bone and changes depending on the position of the patient. Patients may experience relief from pain by shifting location or sitting upward. Pain may often spread to different parts of the body including the arms, neck and shoulders.

Diagnosis of acute pericarditis is challenging and is thought to be a process of exclusion, Imazio et al. [10]. Lange & Hillis [13] state that in 90% of cases, the cause of acute pericarditis is unknown or thought to be a viral disease. In the other 10% of cases, the cause of acute pericarditis is believed to be “transmural myocardial infarction” along with other infections and associated with a “dissecting aortic aneurysm”.

Table 1. The observed number of deaths and expected values of acute pericarditis (Classification number: I30) in South Africa by province in 2014.

Province	Number of Observed Deaths	Number of Expected Deaths
Eastern Cape	3	2.0108547
Free State	1	0.8256759
Gauteng	2	3.8264741
KwaZulu-Natal	1	3.1686066
Limpopo	2	1.6682253
Mpumalanga	0	1.2530874
North-West	3	1.0892270
Northern Cape	1	0.3456705
Western Cape	3	1.8121786
Total	16	16

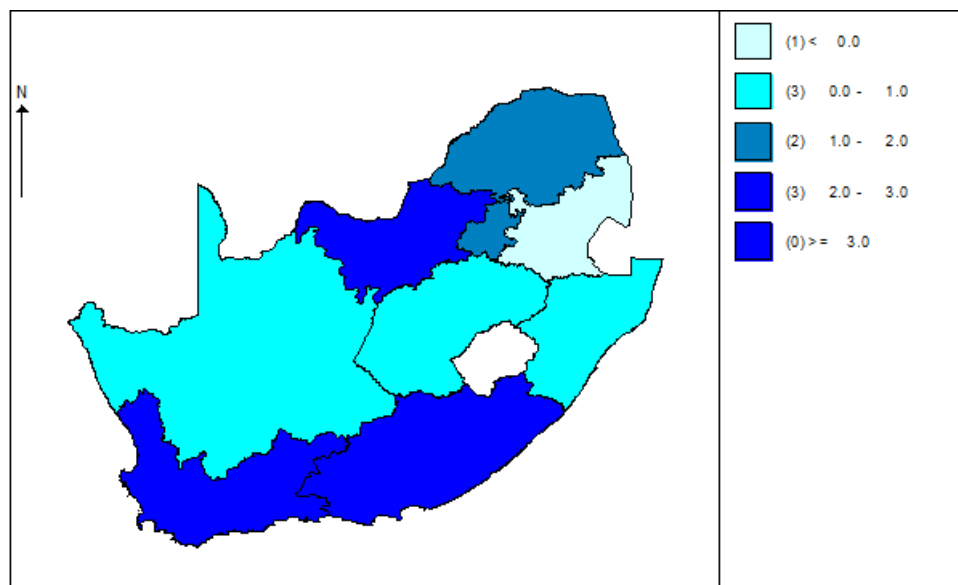


Figure 1. The acute pericarditis deaths in South Africa by province in 2014.

### 3.1. Results for ICAR model

The ICAR model is run for the South African data given in Table 1. The uncorrelated heterogeneity (UH) component is modelled with a zero mean Gaussian distribution with a hyperprior distribution for the precision as a gamma distribution with parameters  $\alpha = 0.5$  and  $\beta = 0.0005$ , respectively, which is the Jeffreys' prior. The precision of the ICAR prior distribution is modelled using a gamma hyperprior distribution with parameters  $\alpha = 0.5$  and  $\beta = 0.0005$ , respectively. These parameters,  $\alpha = 0.5$  and  $\beta = 0.0005$ , are proposed by [?] and are the parameters of the Jeffreys' prior. The model is run for  $n_T = 50000$  iterations and a further  $n_A = 20000$  for the calculation of the DIC. The convergence and burn-in period of the model are assessed by the Brook-Gelman-Rubin (BGR) diagnostic and trace plots. The BGR is a tool to check for convergence and used to get an idea of the burn-in period.

There are four chains run for the deviance and relative risk in all provinces. These chains are run for  $n_T = 50000$  iterations using different initial values. The convergence of the model is assessed by trace plots, and kernel densities of the relative risks and deviance. The model fit is assessed by a Q-Q plot of the deviance residuals. The deviance in both of the South African models is the saturated deviance which is calculated as the sum of the deviance of each province.

Table 2. The statistics from the BGR for the ICAR model.

Iteration Range	Unnormalised		Normalised as plotted		BGR ratio
	of pooled chains	mean within chain	of pooled chains	mean within chain	
251–500	5.217	5.316	0.9814	1.0000	0.9814
501–1000	3.977	3.967	0.7480	0.7461	1.0030
751–1500	3.881	3.846	0.7300	0.7233	1.0090
1001–2000	3.713	3.743	0.6985	0.7040	0.9921
1251–2500	3.976	4.000	0.7478	0.7524	0.9939
1501–3000	4.277	4.309	0.8046	0.8105	0.9927
1751–3500	4.097	4.158	0.7706	0.7821	0.9853
2001–4000	4.118	4.150	0.7745	0.7807	0.9921
2251–4500	4.090	4.076	0.7693	0.7667	1.0030
2501–5000	4.050	4.050	0.7617	0.7618	0.9999
2751–5500	3.838	3.834	0.7220	0.7212	1.0010
3001–6000	3.792	3.791	0.7132	0.7130	1.0000
3251–6500	3.888	3.885	0.7313	0.7308	1.0010
3501–7000	3.908	3.903	0.7351	0.7342	1.0010
3751–7500	3.864	3.860	0.7267	0.7261	1.0010
4001–8000	3.902	3.920	0.7340	0.7374	0.9953
4251–8500	3.918	3.904	0.7369	0.7342	1.0040
4501–9000	3.844	3.832	0.7230	0.7209	1.0030
4751–9500	3.852	3.826	0.7244	0.7197	1.0070
5001–10000	3.833	3.814	0.7210	0.7174	1.0050
5251–10500	3.862	3.843	0.7265	0.7229	1.0050
5501–11000	3.916	3.905	0.7366	0.7346	1.0030
5751–11500	4.000	3.996	0.7524	0.7517	1.0010
6001–12000	4.012	4.012	0.7547	0.7546	1.0000

Table 2 shows the results of the BGR convergence diagnostic applied to the deviance and relative risks of the model. From the BGR ratio, it can be seen that the model reaches convergence between 501-1000 iterations as this is the first time that the BGR ratio reaches 1. Gelman & Hill [9] proposed that a BGR

ratio of 1.1 is an “acceptable limit” for convergence. The burn-in period may also be determined from Table 2. Although interpretation may be slightly subjective, the burn-in period is between iterations 5751-11500 as the statistics in the table become more stable around these iterations. The burn-in period is therefore set at  $N_B = 12000$  as this is above 11500 and ensures stability in the statistics.

Table 3. An analysis of the burn-in period for the ICAR model.

Province	Percentage
Eastern Cape	0.7517
Free State	0.5317
Gauteng	0.7660
KwaZulu-Natal	1.0253
Limpopo	0.5530
Mpumalanga	0.9240
North-West	1.2775
Northern Cape	1.0941
Western Cape	1.2102

The proportion of the MC error to the standard deviation of the relative risks of the ICAR model are used to assess whether the burn-in period of  $N_B = 12000$  is sufficient. This proportion, expressed as a percentage in Table 3, has been calculated for all the relative risks of the provinces and all of the percentages are below 5% which indicates that the burn-in period of  $N_B = 12000$  is sufficient.

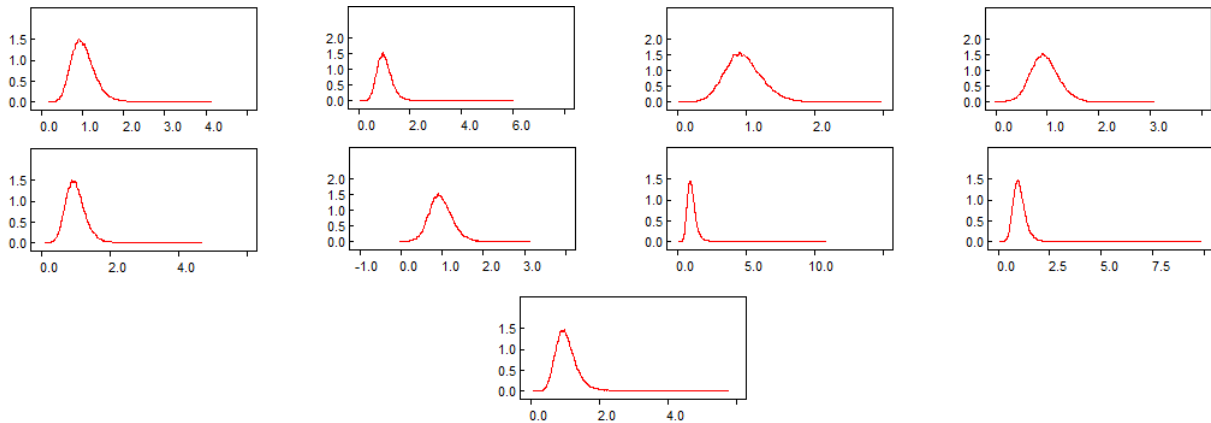


Figure 2. The density kernel for the relative risks of the South African data for the ICAR model from top left: Eastern Cape, Free State, Gauteng, KwaZulu-Natal, Limpopo, Mpumalanga North-West, Northern Cape and Western Cape.

The density kernels for the relative risks are shown in Figure 2. The density kernel plot gives an estimate of the shape of the marginal posterior distribution for each relative risk. The densities have a heavy tail. All of the densities seem to have a maximum around 1.

The MAPE and MSPE of the ICAR model are relatively low values and are similar to those calculated for the PCAR model.

The quantile ratio (QR) is calculated for the relative risks and is used to assess the between-province variation in the relative risks. The QR shows the extent of the spread of the “empirical distribution” of the relative risks as in Lunn et al. [18]. The 60% QR, denoted by QR60, is calculated for the model by using the *ranked()* function in WinBUGS<sup>®</sup>, which will rank the relative risks of the provinces and

Table 4. The model diagnostics of the South African data for the ICAR model.

Statistic	Mean	Standard Deviation	MC Error	95% Credible Interval	
				2.5%	97.5%
MAPE	1.461	0.3972	0.002155	0.7778	2.333
MSPE	3.774	2.343	0.01192	1.111	9.889
$p_V$	0.9569	0.09267	0.001765	0.6207	0.9993
$p_E$	0.02647	0.07265	0.001765	2.416E-4	0.2711
QR60	1.321	0.7269	0.0145	1.031	3.063
$p_S$	0.01667	0.06089	0.001737	6.883E-5	0.2245
Saturated Deviance	10.44	1.928	0.02203	6.451	14.95
DIC	32.074				
$p_D$	1.45				
$\bar{D}$	29.174				
$\hat{D}$	30.624				

then the calculation involves estimating the “exponentiated difference” between the relative risks at the provinces with 80% and 20% quantiles which are then ranked. The code to calculate QR60 is:  $QR60 < -ranked(rr[], 8)/ranked(rr[], 2)$ .

The QR is calculated as 1.321 which indicates that there is a relatively low heterogeneity in the relative risks of the middle 60% of the provinces in the model.

The amount of spatial to unstructured variation in the model is assessed. The variation contributed by the unobserved spatial factors is 1.667% of the total variation in the model, this is denoted by  $p_S$  in Table 4. The variation contributed by the unobserved non-spatial factors, denoted by  $p_V$ , is 2.647% of the total variation, with the other 95.69%,  $p_E$ , of the total variation is due to the expected values of the provinces.

The DIC is 32.074, resulting from the model being run for an additional  $n_A = 20000$  iterations. The posterior mean of the deviance,  $\bar{D}$ , is 29.174.

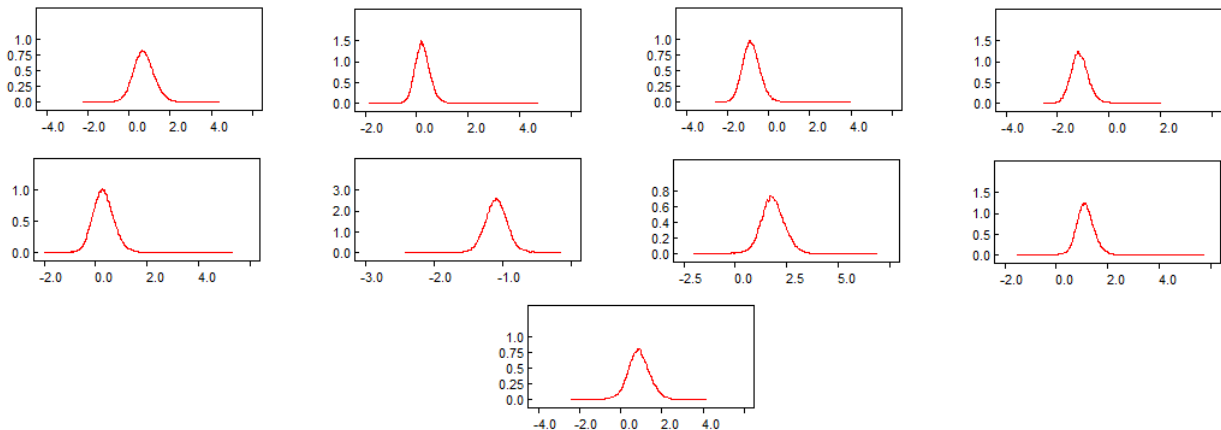


Figure 3. The density kernels for the standardised Bayesian residuals of the South African data ICAR model from top left: Eastern Cape, Free State, Gauteng, KwaZulu-Natal, Limpopo, Mpumalanga, North-West, Northern Cape and Western Cape.

Lawson [14] proposed two features of residuals in a model when the model fits the data well. The first feature is that the residuals should be symmetric and centred around 0. Although the centring around 0 is

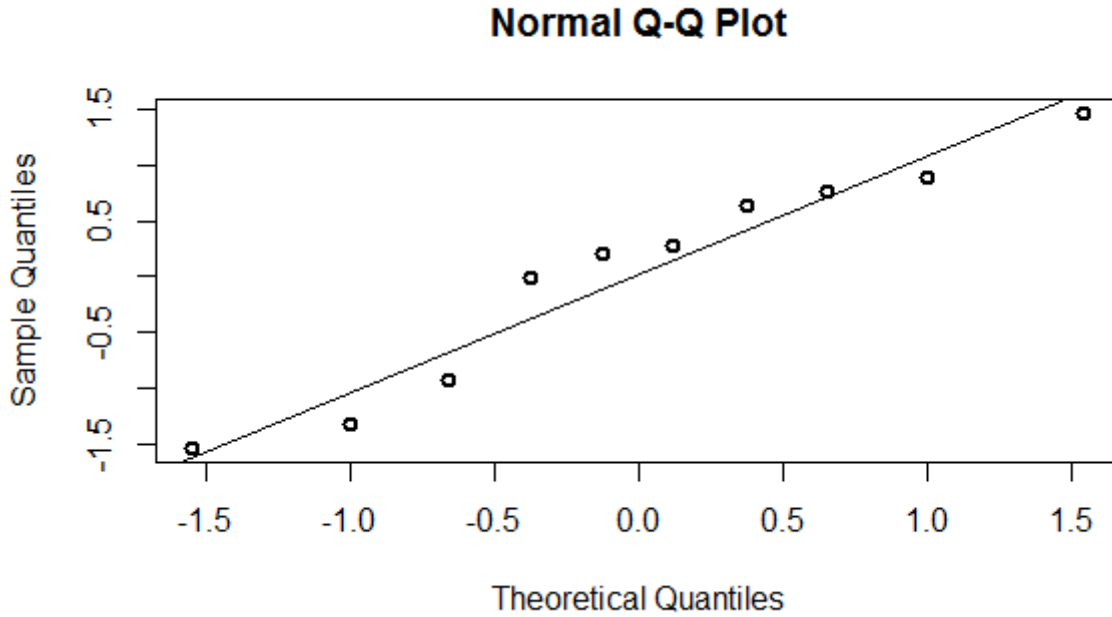


Figure 4. Q-Q plot of the deviance residuals of the ICAR model.

an approximation due to the nature of the simulation in the model. The second feature is that the residuals must show a random pattern and show no particular structure. Figure 3 shows the density kernels for the residuals in each province in the South African ICAR model. Based on the Figure, Eastern Cape, Free State and Limpopo appear to be approximately centred around 0. KwaZulu-Natal, Mpumalanga and North-West have centres which are further away from 0. An approximate symmetry of the kernel densities is shown in Eastern Cape, Free State, North-West and Western Cape which may indicate a good model fit for these provinces. The symmetry of the other provinces appears to be skewed.

Table 5. The statistics from the Bayesian residuals of the South African data for the ICAR model.

Province	Bayesian Residual			95% Credible Interval	
	Mean	Standard Deviation	MC error	Interval	
				2.5%	97.5%
Eastern Cape	0.9258	0.6194	0.004785	-0.4678	1.8980
Free State	0.1650	0.2568	0.001421	-0.3893	0.5735
Gauteng	-1.7100	1.0520	0.007847	-3.9750	0.1699
KwaZulu-Natal	-2.0370	0.8872	0.008782	-3.9250	-0.4344
Limpopo	0.3113	0.5091	0.002780	-0.8099	1.1320
Mpumalanga	-1.2150	0.3600	0.003093	-1.9900	-0.5715
North-West	1.8380	0.4216	0.005140	0.9020	2.3880
Northern Cape	0.6370	0.1316	0.001392	0.3584	0.8140
Western Cape	1.0820	0.6464	0.008066	-0.4070	2.0000

The Q-Q plot of the deviance residuals for the ICAR model shows a reasonable fit for most of the provinces. Most of the provinces show adequate model fit although some provinces are further away from

the straight line than other provinces are. Based on the two features proposed by Lawson [14], the model fits the data adequately.

Table 6. The deviance residuals of the ICAR model.

Province	Deviance Residual			95% Credible Interval	
	Mean	Standard Deviation	MC error	2.5%	97.5%
Eastern Cape	0.6448	0.4425	0.002901	-0.2644	1.477
Free State	0.2039	0.2787	0.001473	-0.3455	0.740
Gauteng	-0.9309	0.5047	0.004340	-1.8940	0.096
KwaZulu-Natal	-1.3240	0.4469	0.004640	-2.1610	-0.397
Limpopo	0.2725	0.3928	0.002003	-0.5093	1.033
Mpumalanga	-1.5430	0.2334	0.002103	-2.0010	-1.074
North-West	1.4730	0.4009	0.003986	0.6212	2.184
Northern Cape	0.8958	0.2338	0.001869	0.4195	1.315
Western Cape	0.7723	0.4586	0.004218	-0.2000	1.606

KwaZulu-Natal and Mpumalanga have the lowest SMR values out of all the provinces. There are 1 and 0 observed deaths with expected number of deaths of 3.1686 and 1.253 for KwaZulu-Natal and Mpumalanga, respectively. A low number of observed deaths and relatively high expected values contribute to the low values for SMR. The high expected value for KwaZulu-Natal results from the large population of the province at 10694434, which is the second highest population of a province in South Africa after Gauteng. The combination of a large population size and a small number of observed deaths result in a small SMR value. Mpumalanga has the sixth largest population size of all the provinces in South Africa but since the observed number of deaths is 0, the SMR is also approximately 0. The highest SMR values are at the North-West and Northern Cape with 3 and 1 observed deaths and expected number of deaths of 1.0892 and 0.3457 respectively. The low expected number of deaths with respect to the number of observed deaths result in a high SMR value. The low expected number of deaths is a result of a small population size in these provinces. The Northern Cape and North-West have population sizes of 1166680 and 3676274, respectively, which are the smallest and third smallest populations, respectively. Therefore the relative risk is a better estimate of the risk of a disease than the SMR, as the relative risk is not a function of the expected values whereas SMR is calculated based on the expected values, and thus is influenced by the population size of the province.

Table 7. The SMR of the South African data for the ICAR model.

Province	SMR			95% Credible Interval	
	Mean	Standard Deviation	MC Error	2.5%	97.5%
Eastern Cape	1.4920	0.007775	0.00004145	1.477	1.508
Free State	1.2110	0.008173	0.00004347	1.196	1.228
Gauteng	0.5226	0.003961	0.00002114	0.515	0.530
KwaZulu-Natal	0.3155	0.006860	0.00003661	0.302	0.329
Limpopo	1.1990	0.003791	0.00002021	1.192	1.207
Mpumalanga	0.0008	0.025410	0.0001354	-0.052	0.047
North-West	2.7560	0.051330	0.0002733	2.660	2.861
Northern Cape	2.910	0.180500	0.0009529	2.606	3.309
Western Cape	1.6560	0.011500	0.00006131	1.634	1.679



Table 7 shows the results from the MCMC simulation with  $n_T = 50000$  iterations. The SMR for Mpumalanga is 0.0008 which is expected as the observed number of deaths is 0. The SMR would usually be 0 in this province as the observed value is 0 but because the SMR has a prior distribution, this is not the case although the approximation is very close. The Northern Cape, which has the lowest expected number of deaths in the model, has the highest SMR value which is attributed to the low expected number of deaths.

Table 8. The relative risk of the South African data for the ICAR model.

Province	Relative Risk			95% Credible Interval	
	Mean	Standard Deviation	MC Error	2.5%	97.5%
Eastern Cape	1.0320	0.3113	0.002340	0.5526	1.7410
Free State	1.0110	0.3138	0.001669	0.5164	1.6990
Gauteng	0.9677	0.2756	0.002111	0.4833	1.5650
KwaZulu-Natal	0.9547	0.2808	0.002879	0.4382	1.5590
Limpopo	1.0110	0.3065	0.001695	0.5183	1.7010
Mpumalanga	0.9666	0.2894	0.002674	0.4478	1.5810
North-West	1.0710	0.3997	0.005106	0.5587	1.9510
Northern Cape	1.0520	0.3901	0.004268	0.5375	1.8940
Western Cape	1.0590	0.3582	0.004335	0.5547	1.8780

KwaZulu-Natal, Mpumalanga and Gauteng have the lowest relative risk of all the provinces and are all below 0.97. This result is different from the SMR values as Mpumalanga had the lowest SMR value but does not have the lowest relative risk value which is now KwaZulu-Natal. Although KwaZulu-Natal had the second smallest SMR value. Western Cape, Northern Cape and North-West have high relative risk values with the highest associated with North-West at 1.071. The SMR was the highest at the Northern Cape which is now the third highest for the relative risk. All of the relative risks are not far away from 1 which is regarded as the “null risk value” and resulted in a low PP. It can thus be said that the risk of contracting and dying from acute pericarditis in South Africa is low when using the ICAR model.

Table 9. The posterior exceedance probability of the South African data for the ICAR model.

Posterior Expected Exceedance Probability			
Province	Mean	Standard Deviation	MC Error
Eastern Cape	0.4925	0.4999	0.002652
Free State	0.4690	0.4990	0.002522
Gauteng	0.4235	0.4941	0.003162
KwaZulu-Natal	0.4098	0.4918	0.003744
Limpopo	0.4701	0.4991	0.002553
Mpumalanga	0.4235	0.4941	0.003305
North-West	0.5131	0.4998	0.003337
Northern Cape	0.5025	0.5000	0.003108
Western Cape	0.5084	0.4999	0.003646

The PP is the posterior expected exceedance probability calculated as the proportion of relative risk values which surpass the threshold of 1. It can be seen that the mean PP is below 0.52 for all provinces. The highest values of the PP occur at Western Cape, Northern Cape and North-West. This is expected as these 3 provinces have the highest SMR and relative risk values.

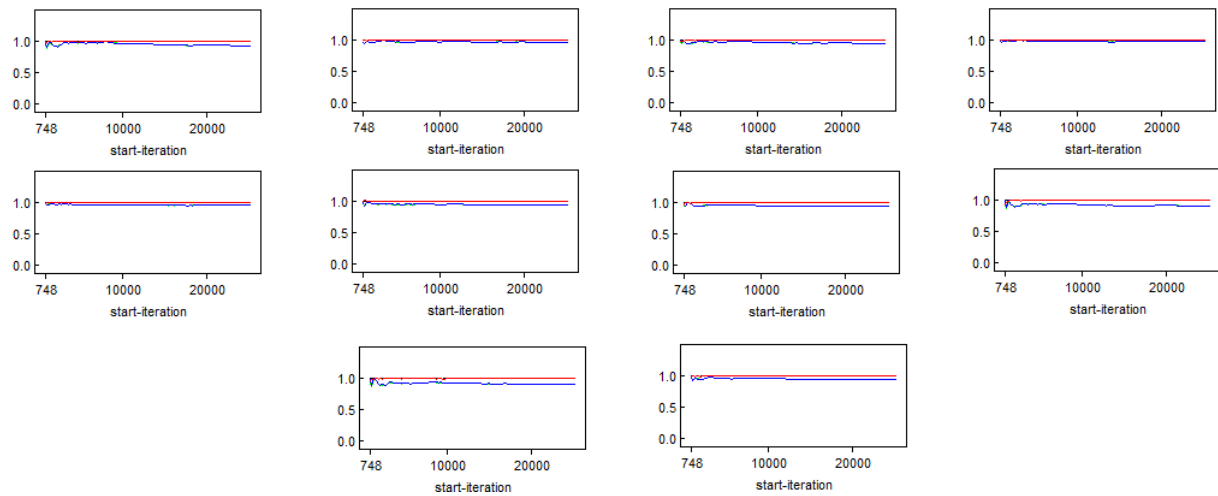


Figure 5. The BGR plots of the PCAR model from top left: Deviance, Eastern Cape, Free State, Gauteng, KwaZulu-Natal, Limpopo, Mpumalanga, North-West, Northern Cape and Western Cape.

### 3.2. Results for PCAR model

The PCAR model is given in Section 2.2 and is also based on the data in Table 1. The precision for the PCAR prior distribution is assigned a gamma hyperprior distribution with parameters of  $\alpha = 0.5$  and  $\beta = 0.0005$ , respectively, which are the parameters of the Jeffreys' prior. The UH component is assigned a zero mean Gaussian distribution with the hyperprior of the precision modelled using a gamma distribution with parameters of  $\alpha = 0.5$  and  $\beta = 0.0005$ , respectively, which are again the parameters of the Jeffreys' prior. These are the same parameters as the ICAR model in Section 3.1. Model assessment is performed in the same way as the ICAR model in Section 3.1. The BGR plots for the deviance and the relative risk in all provinces for the PCAR model are given in Figure 5. The model is run for  $n_T = 50000$  iterations and four chains are used with four different sets of initial values, as with the ICAR model in Section 3.1.

The implementation of the PCAR model is based on the method by Spiegelhalter et al. [22]. The PCAR model is implemented in WinBUGS<sup>®</sup> by firstly calculating  $m_{ii}$  which are the elements proportional to a conditional covariance  $S_i | S_j$  and are contained in the vector  $\mathbf{M}$ . The elements  $m_{ii}$  are calculated using (6) which is written in WinBUGS<sup>®</sup> as:  $M[i] < -1/e[i]$ . The next step is to calculate the cumulative sum of the neighbours vector. The first element of the cumulative sum is 0 which is implemented as:  $csum[1] < -0$  and the rest of the cumulative sum is calculated by starting at the second element using a *loop* as

$$\text{for } (i \text{ in } 2 : (N + 1)) \quad \{ \\ csum[i] < - \quad sum(num[1 : (i - 1)]),$$

where  $\mathbf{num}$  is the vector of the number of neighbours for each province. The next step is to determine an index matrix, denoted by  $pick[]$ , which will contain the index of the elements of the weight vector  $\mathbf{W}$  and will have  $N$  columns and number of rows equal to the number of elements in  $\mathbf{W}$ . The  $pick[]$  matrix is calculated such that the  $i^{th}$  column contains a value of 1 in all of the  $J$  rows such that,  $csum[i] < J \leq csum[i + 1]$  and 0 otherwise. This is done such that,  $\mathbf{W}$  at the elements of  $W[(csum[i] + 1) : csum[i + 1]]$ , are the set of weights  $w_{ij}$  corresponding to the spatial correlation between provinces  $i$  and  $j$ . The code for the  $pick[]$  matrix is

$$\text{for}(k \text{ in } 1 : sumNumNeigh) \quad \{ \\ \text{for}(i \text{ in } 1 : N) \quad \{ \\ pick[k, i] < - \quad step(k - csum[i] - epsilon) * step(csum[i + 1] - k),$$

where *epsilon* is equal to 0.0001.

The vector  $\mathbf{W}$  can now be calculated using the *inprod()* function in WinBUGS<sup>®</sup>. The *inprod()* function is used to calculate the inner product of two vectors. The *inprod()* function is used to determine which province is associated with which element in  $\mathbf{W}$ . This is done using the inner product of the  $k^{th}$  row of the *pick[]* matrix and the expected value, which is the denominator in Equation 5, given as

$$w_{ij} = \begin{cases} \left(\frac{\epsilon_i}{\epsilon_j}\right)^{\frac{1}{2}} & \text{if regions } i \text{ and } j \text{ are adjacent} \\ 0 & \text{otherwise.} \end{cases}$$

Thus the *inprod()* function results in 0 and 1 elements which will determine which provinces are adjacent. The last part of the calculation of the vector  $\mathbf{W}$  is to take the square root of the proportion of the expected values of the adjacent provinces. The element  $k$  of the vector  $\mathbf{W}$  is thus calculated as:

$$W[k] < -\text{sqrt}(e[\text{adj}[k]]/\text{inprod}(e[], \text{pick}[k, ])).$$

The PCAR prior can now be implemented using

$$S[1 : N] \sim \text{car.proper}(mu[], W[], \text{adj}[], \text{num}[], M[], \text{omega}, \rho),$$

where  $mu[]$  is assigned a prior distribution as:  $alpha \sim \text{dnorm}(0, 0.0001)$  and  $omega$  is assigned a Jeffreys' prior and  $\rho$  has a uniform distribution which is bounded in WinBUGS<sup>®</sup> as

$$\begin{aligned} \text{rho.min} &< - \text{min.bound}(W[], \text{adj}[], \text{num}[], M[]) \\ \text{rho.max} &< - \text{max.bound}(W[], \text{adj}[], \text{num}[], M[]). \end{aligned}$$

Table 10. The statistics from the BGR of the PCAR model.

Iteration Range	Unnormalised		Normalised as plotted		BGR ratio
	of pooled chains	mean within chain	of pooled chains	mean within chain	
748–995	3.944	4.007	0.9579	0.9732	0.9842
996–1490	3.679	3.728	0.8935	0.9055	0.9868
1243–1985	4.028	4.046	0.9783	0.9827	0.9956
1491–2480	3.934	3.992	0.9554	0.9696	0.9854
1738–2975	3.814	3.826	0.9262	0.9292	0.9968
1986–3470	3.840	3.817	0.9326	0.9270	1.0060
2233–3965	3.794	3.768	0.9213	0.9150	1.0070
2481–4460	3.928	3.899	0.9538	0.9469	1.0070
2728–4955	3.928	3.918	0.9540	0.9515	1.0030
2976–5450	4.083	4.048	0.9915	0.9830	1.0090
3223–5945	4.063	4.028	0.9867	0.9782	1.0090
3471–6440	3.998	3.976	0.9708	0.9657	1.0050
3718–6935	4.048	4.046	0.9830	0.9825	1.0010
3966–7430	4.070	4.050	0.9883	0.9836	1.0050
4213–7925	4.033	4.038	0.9795	0.9806	0.9989
4461–8420	4.020	4.033	0.9762	0.9795	0.9966
4708–8915	4.012	4.035	0.9743	0.9799	0.9943
4956–9410	4.057	4.072	0.9852	0.9889	0.9963
5203–9905	4.063	4.069	0.9867	0.9882	0.9985
5451–10400	3.998	3.999	0.9709	0.9711	0.9998
5698–10895	4.001	3.996	0.9718	0.9703	1.0010
5946–11390	4.054	4.041	0.9844	0.9815	1.0030
6193–11885	4.063	4.060	0.9868	0.9859	1.0010
6441–12380	4.062	4.055	0.9866	0.9848	1.0020

The PCAR model seems to show convergence slower than the ICAR model. The blue line is also closer to the red line in the PCAR model and seems more stable than in the ICAR model.

Convergence for the PCAR model is only reached at iterations 1986-3470, which is slower than the ICAR model which reached convergence at iterations 501-1000. From Table 10, it can be seen that stability is reached at iterations 5946-11390 and thus the burn-in period is again set to  $N_B = 12000$ . This is the same burn-in period as the ICAR model in which stability was reached at iterations 5751-11500.

Table 11. An analysis of the burn-in period of the PCAR model.

Province	Percentage
Eastern Cape	0.7048
Free State	0.6042
Gauteng	0.7695
KwaZulu-Natal	0.8801
Limpopo	0.5417
Mpumalanga	0.7727
North-West	1.2028
Northern Cape	1.0193
Western Cape	0.7466

As with the ICAR model, and keeping with the rule of thumb proposed by Woodward [25], the proportion of the MC error to the standard deviation of the relative risk is calculated. All of these proportions are well below 5% and thus, it can be concluded that the burn-in period of  $N_B = 12000$  is sufficient, which is the same conclusion as in the ICAR model.

The PCAR model was run with a single chain of  $n_T = 50000$  iterations with a burn-in period of  $N_B = 12000$ . Trace plots show the model converged for the relative risks in all of the provinces.

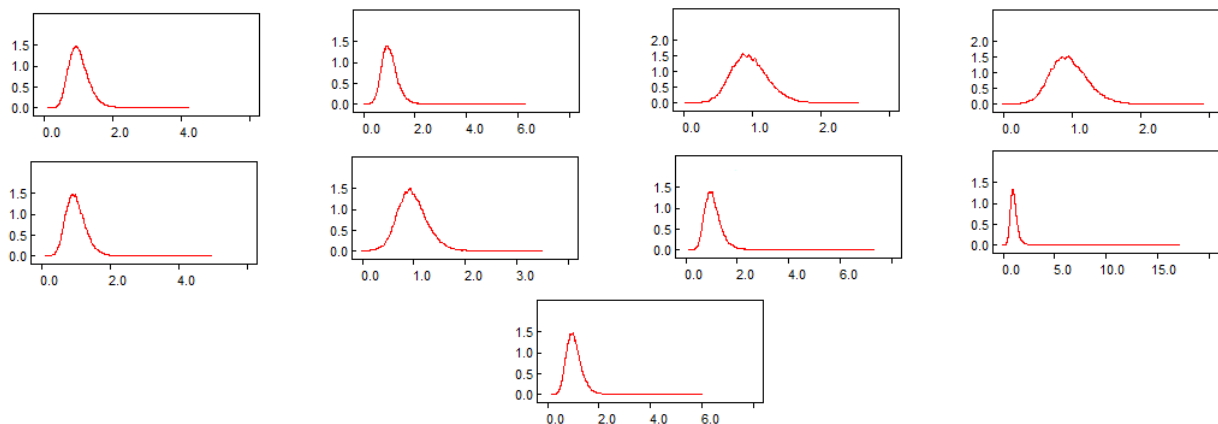


Figure 6. The density kernels for the relative risk of the South African data for the PCAR model from top left: Eastern Cape, Free State, Gauteng, KwaZulu-Natal, Limpopo, Mpumalanga, North-West, Northern Cape and Western Cape.

The density kernels show a maximum of approximately 1 for the relative risks in all the provinces. All of the provinces show a heavy tail towards the right.

Table 12. The model diagnostics of the South African data for the PCAR model.

Statistic	Mean	Standard Deviation	MC Error	95% Credible Interval	
				2.5%	97.5%
MAPE	1.461	0.399	0.002174	0.7778	2.333
MSPE	3.766	2.345	0.01171	1.111	9.889
$p_E$	0.9449	0.1033	0.002736	0.6222	0.9989
$p_V$	0.02744	0.07497	0.001854	2.37E-4	0.2663
QR60	1.379	0.7671	0.01596	1.039	2.967
$p_S$	0.02768	0.07518	0.001969	2.174E-4	0.25
Saturated Deviance	10.51	1.868	0.01758	6.885	14.99
DIC	32.139				
$p_D$	1.441				
$\hat{D}$	29.257				
$D$	30.698				

The MAPE is the same for both models, while the MSPE is higher in the ICAR model with a difference of 0.008. The quantile ratio, QR60, is 1.379 for the PCAR model and indicates that there is relatively low heterogeneity in the risk across the middle 60% of the provinces. The QR is higher in the PCAR model than in the ICAR model with a difference of 0.058.

The amount of variation attributed to the unobserved spatial factors,  $p_S$ , is 2.768% of the total variation in the model. This compares to 1.667% in the ICAR model. The variation attributed to the unobserved non-spatial factors,  $p_V$ , is 2.744% of the total variation in the model which is higher than the 2.647% in the ICAR model. The other 94.49% of the variation in the model is due to the expected values of the provinces.

The DIC is calculated based on an additional  $n_A = 20000$  iterations, as in the ICAR model, and resulted in a DIC value of 32.139. This is slightly higher than the ICAR model, which has a DIC value of 32.074. The difference in the saturated deviance between the two models is 0.07 over  $n_T = 50000$  iterations, which is not a significant difference.

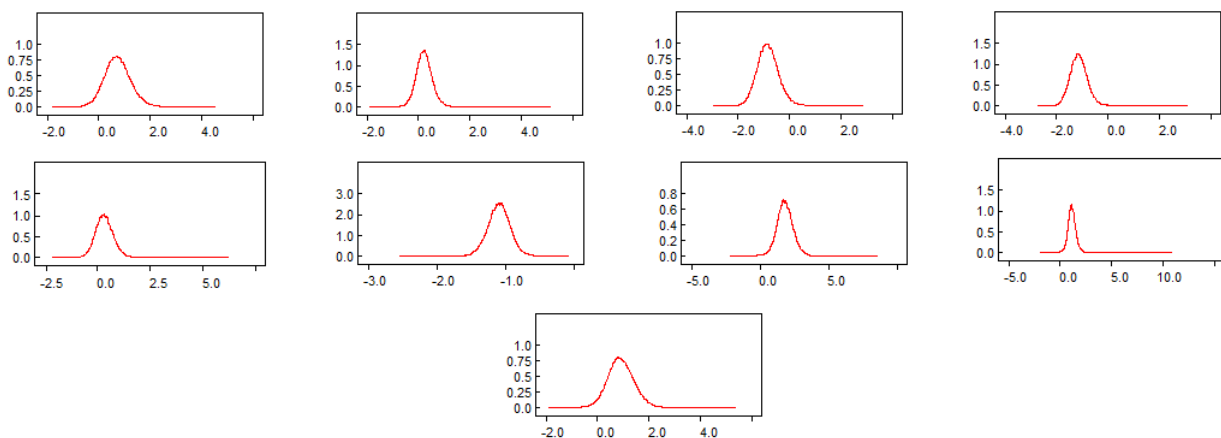


Figure 7. The density kernels of the standardised Bayesian residuals of the South African data for the PCAR model from top left: Eastern Cape, Free State, Gauteng, KwaZulu-Natal, Limpopo, Mpumalanga, North-West, Northern Cape and Western Cape.

Based on Figure 7, Eastern Cape, Free State and Limpopo appear to have centres closer to 0. KwaZulu-Natal, Mpumalanga, North-West and Western Cape have centres which are further away from 0. An approximate symmetry of the kernel densities is shown in Eastern Cape, Free State, North-West and Western Cape which may indicate a good model fit for these provinces. The symmetry of the other provinces appears to be skewed, with heavy tails occurring at Free State, KwaZulu-Natal, Limpopo and Mpumalanga.

Table 13. The statistics from the Bayesian residuals of the South African data for the PCAR model.

Province	Bayesian Residual			95% Credible Interval	
	Mean	Standard Deviation	MC error	Interval	
				2.5%	97.5%
Eastern Cape	0.9285	0.6277	0.004333	-0.5052	1.911
Free State	0.1575	0.2869	0.002026	-0.4725	0.5889
Gauteng	-1.707	1.032	0.007437	-3.943	0.0962
KwaZulu-Natal	-2.049	0.8697	0.007079	-3.925	-0.5027
Limpopo	0.3108	0.5125	0.003164	-0.8222	1.133
Mpumalanga	-1.224	0.3787	0.002789	-2.05	-0.5658
North-West	1.819	0.4437	0.005242	0.776	2.396
Northern Cape	0.6269	0.1843	0.001834	0.254	0.8335
Western Cape	1.128	0.5761	0.004475	-0.1782	2.011

Table 14. The deviance residuals of the PCAR model.

Province	Deviance Residual			95% Credible Interval	
	Mean	Standard Deviation	MC error	Interval	
				2.5%	97.5%
Eastern Cape	0.654	0.4386	0.002821	-0.2408	1.485
Free State	0.2029	0.2972	0.001971	-0.4014	0.773
Gauteng	-0.9363	0.4918	0.003991	-1.886	0.04419
KwaZulu-Natal	-1.336	0.431	0.003684	-2.158	-0.4744
Limpopo	0.2737	0.3891	0.002281	-0.4986	1.017
Mpumalanga	-1.547	0.2388	0.001986	-2.017	-1.081
North-West	1.463	0.4145	0.004286	0.5604	2.195
Northern Cape	0.8957	0.2719	0.002036	0.3275	1.371
Western Cape	0.8055	0.4293	0.00291	-0.07843	1.616

There are small differences between the Bayesian residuals in the PCAR and ICAR models. The biggest difference between the Bayesian residuals of the two models occurs at the Western Cape with the Bayesian residuals having values of 1.082 and 1.128 for the ICAR and PCAR models, respectively. This is a difference of 0.046, which is not a significant difference. The other provinces all have very small differences between the two models.

As with the Bayesian residuals, the deviance residuals for both models are also very similar.

The model seems to fit the data adequately as the normal Q-Q plot shows that most of the points are close to the straight line. There is a province at -1.5 which is the furthest away from the straight line than all the other provinces and indicates the model does not fit well to this province. This is different to the Q-Q plot of the ICAR model.

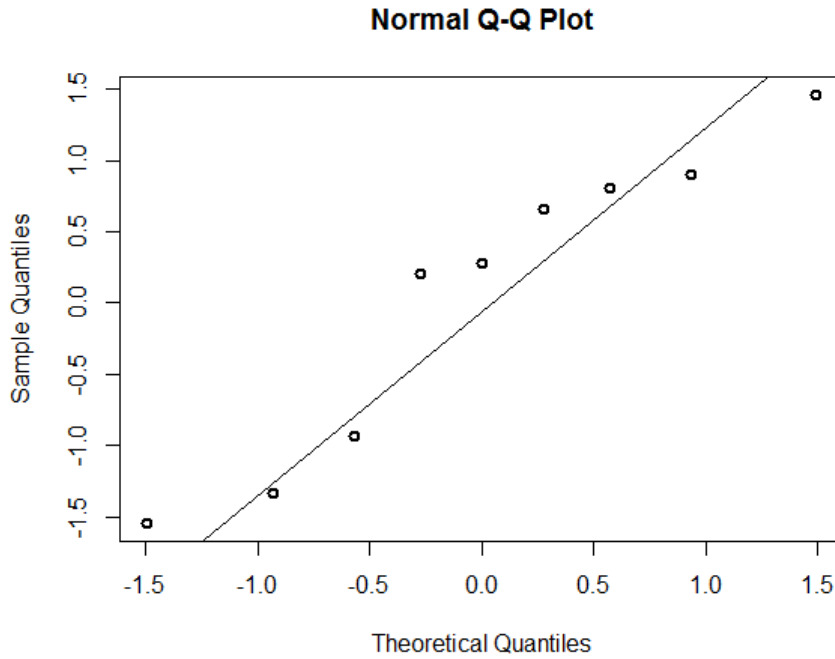


Figure 8. Q-Q plot of the deviance residuals of the PCAR model.

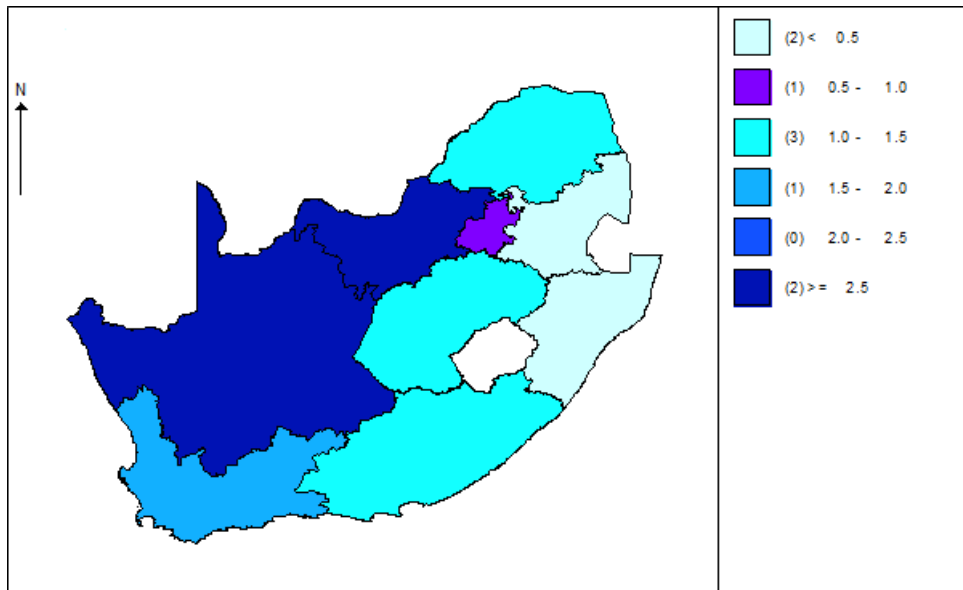


Figure 9. The SMR of the South African data for the PCAR model.

The PCAR model produces the same figure as the ICAR model for the SMR. As in the ICAR model, KwaZulu-Natal and Mpumalanga have the lowest SMR values. The highest values of the SMR are again

North-West and the Northern Cape. The same conclusions can be drawn from the PCAR model that were drawn from the ICAR model.

Table 15. The SMR of the South African data for the PCAR model.

Province	SMR			95% Credible Interval	
	Mean	Standard Deviation	MC Error	2.5%	97.5%
Eastern Cape	1.4920	0.007741	0.00004084	1.4770	1.5080
Free State	1.2110	0.008138	0.00004283	1.1970	1.2280
Gauteng	0.5226	0.003944	0.00002082	0.5148	0.5302
KwaZulu-Natal	0.3155	0.006830	0.00003606	0.3019	0.3286
Limpopo	1.1990	0.003775	0.00001991	1.1920	1.2070
Mpumalanga	0.0007	0.025300	0.00013340	-0.0522	0.0467
North-West	2.7560	0.051110	0.00026930	2.6610	2.8600
Northern Cape	2.9100	0.179800	0.00094090	2.6070	3.3080
Western Cape	1.6560	0.011450	0.00006040	1.6340	1.6790

The SMR values in Table 15 give almost the same means as the ICAR model, with the only difference in the SMR value of Mpumalanga. In the ICAR model, the mean SMR value at Mpumalanga is 0.0007581 whereas in the PCAR model the value is 0.0007134. The mean SMR value at Mpumalanga is slightly higher than the PCAR model with a difference of only 0.0000447. Other slight differences do occur in the standard deviations and MC errors between the two models. The standard deviations are slightly different for the two models, with the PCAR model having a slightly lower standard deviation for all the provinces. The MC error, as with the standard deviation, is slightly lower for the PCAR model than for the ICAR model.

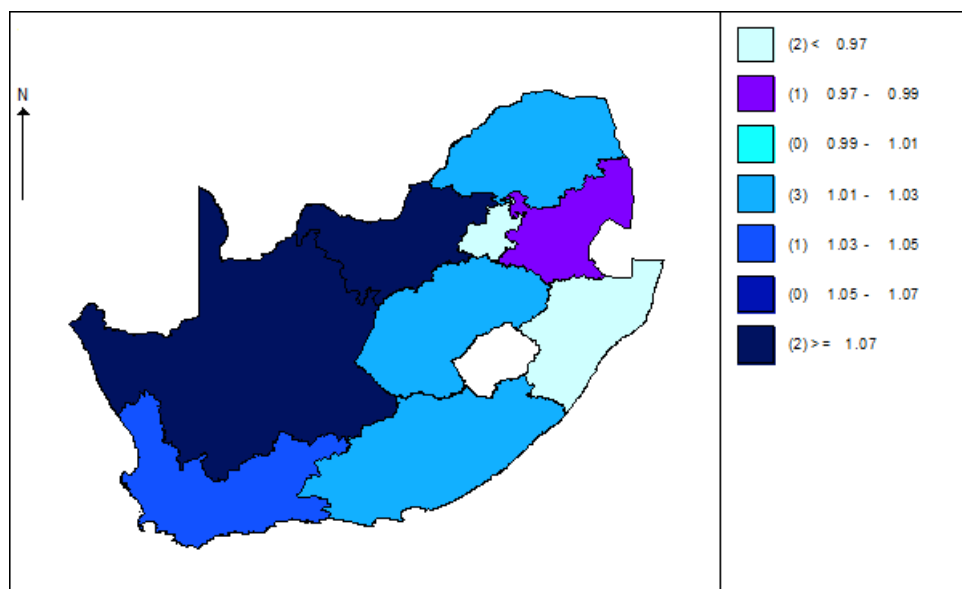


Figure 10. The relative risk of the South African data for the PCAR model.



Table 16. The relative risk of the South African data for the PCAR model.

Province	Relative Risk			95% Credible Interval	
	Mean	Standard Deviation	MC Error	2.5%	97.5%
Eastern Cape	1.0300	0.3130	0.002206	0.5434	1.712
Free State	1.0240	0.3522	0.002128	0.4984	1.805
Gauteng	0.9696	0.2720	0.002093	0.4994	1.560
KwaZulu-Natal	0.9632	0.2786	0.002452	0.4759	1.547
Limpopo	1.0130	0.3085	0.001671	0.5267	1.692
Mpumalanga	0.9757	0.3032	0.002342	0.4595	1.628
North-West	1.0830	0.4018	0.004833	0.5536	2.017
Northern Cape	1.0750	0.5235	0.005336	0.4881	2.136
Western Cape	1.0360	0.3204	0.002392	0.5457	1.760

The relative risk is slightly higher for the ICAR model than the PCAR model in provinces Gauteng and the Western Cape. All the other provinces have a slightly lower relative risk in the ICAR model. The highest difference is 0.023. The same can be concluded for the PCAR model as was concluded for the ICAR model, in that the relative risk of contracting and dying from acute pericarditis in South Africa in 2014 has a relatively low risk.

The PP for the PCAR model shows that there is a slight difference in the PCAR model compared to the ICAR model. The biggest difference occurs at the Western Cape where the ICAR model has a slightly higher value than the PCAR model with the difference being 0.0141. There are three provinces where the PP for the ICAR model is lower than the PP for the PCAR model, these occur at Gauteng, KwaZulu-Natal and Mpumalanga. The standard deviation for the PP in the ICAR model is lower or equal to the standard deviation for the PP in the PCAR model in 7 provinces. There are slight differences in the standard deviations and MC errors in the two models.

The PP shows that there is no usual or unexpected clustering of the disease in the PCAR model when the exceedance probability,  $\hat{P}$ , of interest is set at 0.9.

The credible intervals for the PP are not given in Table 17 as all of the credible intervals have a lower bound of 0 and an upper bound of 1.0.

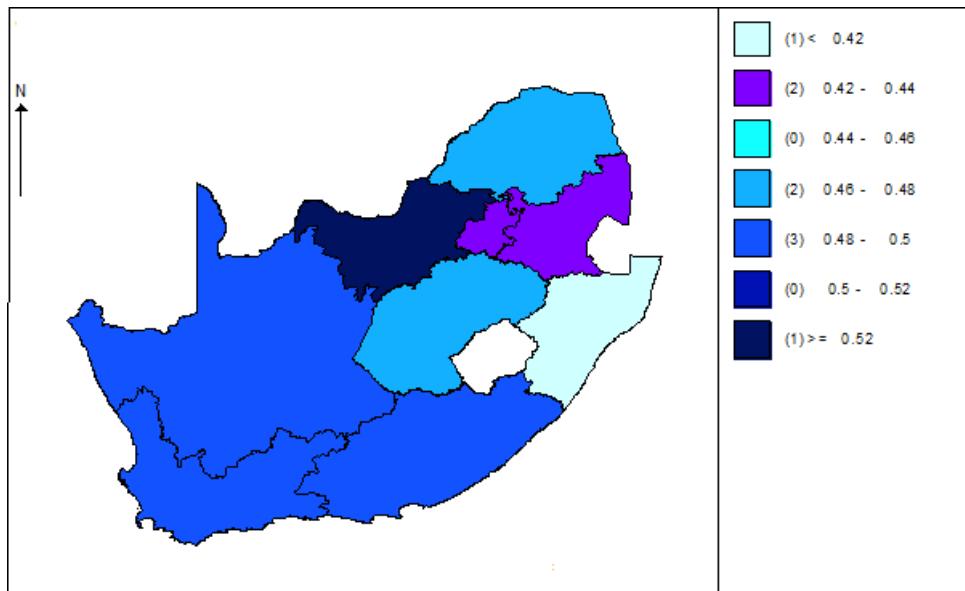


Figure 11. The posterior exceedance probability of the South African data for the PCAR model.

Table 17. The posterior exceedance probability of the South African data for the PCAR model.

Posterior Expected Exceedance Probability			
Province	Mean	Standard Deviation	MC Error
Eastern Cape	0.4887	0.4999	0.002814
Free State	0.4750	0.4994	0.002656
Gauteng	0.4257	0.4944	0.003089
KwaZulu-Natal	0.4146	0.4927	0.003332
Limpopo	0.4693	0.4991	0.002616
Mpumalanga	0.4332	0.4955	0.003117
North-West	0.5265	0.4993	0.003634
Northern Cape	0.4914	0.4999	0.002407
Western Cape	0.4943	0.5000	0.002807

### 3.3. Sensitivity analysis

In this Section an investigation will be done to see how sensitive the ICAR and PCAR models are, to the various specifications for the hyperparameters. Vague priors and a conjugate subjective prior will be used. A summary of the various priors are given in Table 18. The choice of which prior distribution will give the best results is controversial and is often subjective. Four priors will be considered in this section. The parameters of the conjugate prior distribution for the South African data are estimated from previous studies dating back from 2010 until 2013 for which, the mean and variance are calculated for each province and summed to yield the total mean and variance for 2010-2013. The parameters of the prior are then estimated by using the mean and variance of a gamma distribution and solving for the two parameters,  $\alpha$  and  $\beta$ .

Table 18. Summary of various priors.

	ICAR	PCAR
Prior 1 <sup>†</sup>	<i>Gamma</i> (0.5, 0.0005)	<i>Gamma</i> (0.5, 0.0005)
Prior 2	<i>Gamma</i> (13.787, 0.766)	<i>Gamma</i> (13.787, 0.766)
Prior 3	<i>Gamma</i> (0.5, 0.0001)	<i>Gamma</i> (0.5, 0.0001)
Prior 4	<i>Gamma</i> (0.1, 0.00001)	<i>Gamma</i> (0.1, 0.00001)

The DIC and  $p_D$  of the various models are given in Table 19, the DIC has been used to assess the sensitivity of the model to the choice of prior. From Table 19 it can be seen that the DIC values are very similar for the various priors, the DIC for the subjective prior is slightly less than for the three vague priors.

Table 19. DIC and  $p_D$  of the various models.

	ICAR		PCAR	
	DIC	$p_D$	DIC	$p_D$
Prior 1	32.074	1.45	32.139	1.441
Prior 2	31.913	1.835	31.919	1.846
Prior 3	32.070	1.318	32.117	1.275
Prior 4	32.221	1.694	32.287	1.749

#### 4. Conclusion

The results from the ICAR and PCAR models showed that it was possible to use either model and still obtain very similar results for the PCAR and ICAR models for acute pericarditis in South Africa in 2014. Both the ICAR and PCAR models also showed that there was a low risk of contracting and dying from acute pericarditis in all provinces in 2014. The ICAR model seems to be the more popular model among researchers. It is thus recommended that the ICAR model is used. The work done in this paper illustrated to the user how WinBUGS could be applied in Bayesian disease mapping.

We found the disadvantages of the PCAR as:

- The implementation of the PCAR model was more complex than the ICAR model.
- The WinBUGS<sup>®</sup> code for the PCAR model was clumsy.
- The PCAR model had a longer computation time than the ICAR model. This became evident when the number of parameters and number of regions in the model increased. Some of the computations may be done in RStudio<sup>®</sup> and the result of these computations may be inserted into WinBUGS<sup>®</sup> as data as a temporary fix for the longer computation time.
- The PCAR model only worked for models with a small number of regions and a small number of parameters.
- The choice of initial values in the PCAR model may lead to a trap error in WinBUGS<sup>®</sup>, thus the specification of initial values in the PCAR model was more difficult than in the ICAR model.
- An incorrect specification of the bounds of the parameter  $\rho$ , the amount of spatial dependence, in the PCAR model resulted in incorrect results.

The complexity of diseases made it difficult to model a contagious disease as the independence of the data would be broken. The likelihood is based on the foundation that the data is independent thus an investigation into the methods and techniques required to perform such an analysis would be very

<sup>†</sup>Prior used in the model in the data analysis

interesting. Another issue, requiring further investigation, would be the analysis of a disease which results in a higher number of deaths. The Poisson distribution is used when there are a small number of deaths as a large number of deaths create a lot of overdispersion. It would be interesting to investigate which distribution would fit the data well in that case. Defining a model which can handle multiple diseases at once and a mapping technique to map the output of such a model.

#### Acknowledgements

The authors would like to thank the editor and referees for the helpful comments and suggestions to improve the paper. The work of the first and second author work is based on the research supported by the National Research Foundation (NRF). Any opinion, finding and conclusion or recommendation expressed in this material is that of the author(s) and the NRF does not accept any liability in this regard.

#### REFERENCES

- [1] Anderson, A., Lee, D., & Dean, N. 2014. Identifying clusters in Bayesian disease mapping. *Biostatistics*, 15, 457 – 469.
- [2] Besag, J., & Kooperberg, C. 1995. On Conditional and Intrinsic Autoregressions. *Biometrika*, 82(4), 733–746.
- [3] Besag, J., York, J., & Mollié, A. 1991. Bayesian Image Restoration, with Two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1), 1–20.
- [4] Best, N., Richardson, S., & Thomson, A. 2005. A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research*, 14, 35 – 59.
- [5] Brook, D. 1964. On the Distinction Between the Conditional Probability and the Joint Probability Approaches in the Specification of Nearest-Neighbour Systems. *Biometrika*, 51(3/4), 481–483.
- [6] Coly, S., Garrido, M., Abrial, D., & Yao, A-F. 2021. Bayesian hierarchical models for disease mapping applied to contagious pathologies. *PLoS ONE*, 16, 1 – 28.
- [7] Cressie, N. A., & Chan, N. H. 1989. Spatial modeling of regional variables. *Journal of the American Statistical Association*, 84(406), 393–401.
- [8] Diggle, P. J., Tawn, J. A., & Moyeed, R. A. 1998. Modelbased geostatistics. *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 47, 299 – 350.
- [9] Gelman, A., & Hill, J. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Analytical Methods for Social Research. Cambridge University Press.
- [10] Imazio, M., Cecchi, E., Demichelis, B., Ierna, S., Demarie, D., Ghisio, A., Pomari, F., Coda, L., Belli, R., & Trincherò, R. 2007. Indicators of Poor Prognosis of Acute Pericarditis. *Circulation*, 115(21), 2739–2744.
- [11] Jahan, F., Duncan, E. W., Cramb, S. M., Baade, P. D., & Mengersen, K. L. 2000. Augmenting disease maps: a Bayesian meta-analysis approach. *Royal Society Open Science*, 7, 1 – 27.
- [12] Künsch, H. R. 1987. Intrinsic Autoregressions and Related Models on the Two-Dimensional Lattice. *Biometrika*, 74(3), 517–524.

- [13] Lange, R. A., & Hillis, L. D. 2004. Acute Pericarditis. *New England Journal of Medicine*, 351(21), 2195–2202.
- [14] Lawson, A. B. 2009. *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*. CRC Press.
- [15] Lawson, A. B., Browne, W. J., & Rodeiro, C. L. Vidal. 2003. *Disease Mapping with WinBUGS and MLwiN*. John Wiley & Sons, Inc.
- [16] Lesaffre, Emmanuel, & Lawson, Andrew B. 2012. *Bayesian Biostatistics*. John Wiley & Sons, Ltd.
- [17] Lunn, D. J., Thomas, A., Best, N. G., & Spiegelhalter, D. J. 2000. WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 325–337.
- [18] Lunn, D. J., Jackson, C., Best, N. G., Thomas, A., & Spiegelhalter, D. J. 2012. *The BUGS Book: A Practical Introduction to Bayesian Analysis*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press: Taylor & Francis Group.
- [19] Osei, P. P. 2010. *Statistical Methods for Disease Mapping*. Ph.D. thesis, African Institute for Mathematical Sciences (AIMS), 6 Melrose Road, Muizenberg, Cape Town, South Africa 7950.
- [20] Pascutto, C., Wakefield, J. C., Best, N. G., Richardson, S., Bernardinelli, L., Staines, A., & Elliott, P. 2000. Statistical issues in the analysis of disease mapping data. *Statistics in Medicine*, 19, 2493 – 2519.
- [21] Quantum GIS Development Team. 2016 (February). Quantum GIS Geographic Information System. 2.14 LTR Essen edn. Open Source Geospatial Foundation Project.
- [22] Spiegelhalter, D. J., Thomas, A., & Best, N. G. 2003. *WinBUGS Version 1.4 User Manual*. Imperial College and Medical Research Council, United Kingdom.
- [23] Stern, H. S., & Cressie, N. A. C. 1999. Inference for extremes in disease mapping. In: *Disease Mapping and Risk Assessment for Public Health*, Chapter 5, Lawson, A. B., Biggeri, A., Boehning, D., Lesaffre, E., Viel, J. F. & Bertollini, R. (eds). New York: Wiley.
- [24] Wakefield, J. C., Best, N. G., & Waller, L. A. 2000. Bayesian approaches to disease mapping. Pages 104 – 127 of: and Elliott, P., C., Wakefield J., Best, N. G., & J., Briggs D. (eds), *Spatialepidemiology: methods and applications*. Oxford: Oxford University Press.
- [25] Woodward, P. 2016. *Bayesian Analysis Made Simple: AN Excel GUI for WinBUGS*. Chapman & Hall/CRC Biostatistics Series. CRC Press.