*Biostatistics* (2007), **8**, 1, *pp.* 1–15
doi:10.1093/biostatistics/kxl000
Advance Access publication on May 15, 2006

# Bayesian Inference for Generalized Linear Mixed Models: Supplementary Material

YOUYI FONG

*Department of Biostatistics, University of Washington, Seattle, WA 98112, USA*

HÅVARD RUE

*The Norwegian University for Science and Technology, Trondheim, Norway*

JON WAKEFIELD*

*Departments of Statistics and Biostatistics, University of Washington, Seattle, WA 98112, USA*
jonno@u.washington.edu

## EXAMPLES

### *Overdispersion*

This example concerns data on the proportion of seeds that germinated on each of $m = 21$ plates arranged in a $2 \times 2$ factorial design with respect to seed variety and type of root extract. These data were presented by Crowder (1978). The sampling model is $Y_i|\boldsymbol{\beta}, p_i \sim \text{Binomial}(n_i, p_i)$ where, for plate $i$, $Y_i$ is the number of germinating seeds and $n_i$ is the total number of seeds (which range between 4 and 81), $i = 1, ..., m$. To account for between plate variability, Breslow and Clayton (1993) introduce plate-level random effects,

*To whom correspondence should be addressed.

and then fit main effects and interaction models:

$$\text{logit } p_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + b_i \tag{0.1}$$

$$\text{logit } p_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{1i} x_{2i} + b_i \tag{0.2}$$

with $b_i|\sigma^2 \sim_{iid} N(0, \sigma^2)$, and $x_{1i}$, $x_{2i}$ representing the seed variety and type of root extract for plate $i$. We fit each of these models and, following the procedure outlined in Section 4.2, assume a marginal Cauchy distribution for the random effects with the residual odds of germination across plates being [0.1,10] with probability 0.95, to give prior $\sigma^{-2} \sim \text{Ga}(0.5, 0.0164)$.

Breslow and Clayton (1993) evaluated the likelihood for $\boldsymbol{\beta}, \sigma$ using Gaussian quadrature, and Table 1 reports the maximum likelihood estimates (MLEs), along with asymptotic standard errors. These are compared with the posterior means and posterior standard deviations from INLA (we present standard deviations on the linear predictor scale for direct comparison with likelihood methods, in practice interval estimates and inference on a more interpretable scale are preferable). There is reasonably close correspondence between the ML and INLA results, though the posterior standard deviations of the fixed effects are on all occasions slightly larger than the asymptotic standard errors, which probably reflects that with $m = 21$ clusters a little accuracy is lost when using asymptotic inference.

| Variable | ML | INLA | ML | INLA |
|---|---|---|---|---|
| Intercept | $-0.389 \pm 0.166$ | $-0.389 \pm 0.173$ | $-0.548 \pm 0.167$ | $-0.549 \pm 0.182$ |
| Seed | $-0.347 \pm 0.215$ | $-0.351 \pm 0.221$ | $-0.097 \pm 0.278$ | $0.087 \pm 0.298$ |
| Extract | $1.029 \pm 0.205$ | $1.033 \pm 0.212$ | $1.337 \pm 0.237$ | $1.347 \pm 0.259$ |
| Interaction | — | — | $-0.811 \pm 0.385$ | $-0.821 \pm 0.414$ |
| $\sigma$ | $0.295 \pm 0.112$ | $0.315 \pm 0.125$ | $0.236 \pm 0.110$ | $0.278 \pm 0.117$ |

Table 1. ML and INLA summaries for the seeds data.

*A Mixed Model for the Log Odds Ratio*

We analyze data from Kneale (1971), following Breslow and Clayton (1993). The data concern the classification of childhood cancer deaths by age at death, year of birth and whether or not the mother reported being exposed to pelvic radiation during pregnancy. These data are combined with a set of matched controls. Following previous authors we arrange the data as a set of $m = 120$ $2 \times 2$ tables, one for each combination of age and birth year combination. We let $Y_{i0}$ represent the number of exposed of the $n_{i0}$ control individuals in stratum $i$, and $Y_{i1}$ the equivalent number of exposed amongst the $n_{i1}$ cases. We fit the model presented by Spiegelhalter et al. (1998) in which $Y_{ik}|p_{ik} \sim_{ind}$ Binomial$(n_{ik}, p_{ik})$, $k = 0, 1$, with

$$\text{logit } p_{i0} = \mu_i$$

$$\text{logit } p_{i1} = \mu_i + \theta_i$$

so that $\mu_i$ represent stratum effects, which are modeled as independent normals with large variances, and $\exp(\theta_i)$ is the odds ratio of interest. The log odds ratio is modeled as

$$\theta_i = \beta_0 + \beta_1 \text{year}_i + \beta_2(\text{year}_i^2 - 22) + b_i$$

with $b_i|\sigma^2 \sim_{iid} N(0, \sigma^2)$. We assume the residual odds lie in the range [0.1,10] with probability 0.9, and a log Cauchy marginal, to give a Ga(0.5,0.0164) prior for $\sigma^{-2}$.

The INLA results, alongside those using PQL, are presented in Table 2. For $\beta_0$ we once again see a slightly larger standard error under the Bayesian analysis.

| Variable | PQL | INLA |
|---|---|---|
| $\beta_0$ | $0.566 \pm 0.070$ | $0.569 \pm 0.075$ |
| $\beta_1$ | $-0.047 \pm 0.017$ | $-0.047 \pm 0.017$ |
| $\beta_2$ | $0.007 \pm 0.003$ | $0.007 \pm 0.003$ |
| $\sigma$ | $0.150 \pm 0.100$ | $0.170 \pm 0.066$ |

Table 2. PQL and INLA summaries for the Oxford childhood cancer data.

*Spatial Aggregation in Scottish Lip Cancer Rates*

The Scottish lip cancer data have been analyzed by a number of authors, and concern the incidence of male lip cancer in 56 counties of Scotland. The available data are the number of cases, $Y_i$, the expected number of cases, $n_i$, and the proportion of individuals who are employed in agriculture, fishing or forestry, $x_i$. We assume $Y_i|\mu_i \sim_{\text{ind}} \text{Poisson}(\mu_i)$ and fit four different models:

$$\log \mu_i = \log n_i + \beta_0 + v_i \tag{0.3}$$

$$\log \mu_i = \log n_i + \beta_0 + \beta_1 x_i/10 + v_i \tag{0.4}$$

$$\log \mu_i = \log n_i + \beta_0 + v_i + u_i \tag{0.5}$$

$$\log \mu_i = \log n_i + \beta_0 + \beta_1 x_i/10 + v_i + u_i \tag{0.6}$$

The unstructured random effects, $v_i$, are such that $v_i \sim_{iid} N(0, \sigma_v^2)$, and $u_i$ are assigned an intrinsic conditional autoregressive (ICAR) model. Specifically, $\text{E}[u_i|\{u_j : j \in \partial_i\}] = \overline{u}_i$ and $\text{Var}(u_i|\{u_j : j \in \partial_i\}, \sigma_u^2) = \sigma_u^2/m_i$, where $\partial_i$ represents the collection of neighbors of area $i$, $\overline{u}_i$ is the mean of the neighbors, and $m_i$ is the number of neighbors. The prior for $\beta_0$ is improper uniform. For the unstructured variability we assume that the residual relative risks lie in the interval $[0.2, 5]$ with probability 0.95 and assume $d = 2$ to to give the exponential prior distribution Ga(1,0.140) for $\sigma_v^{-2}$. We assume a zero mean normal prior with variance 1000 for $\beta_1$, and note that the above model does not aggregate correctly from a plausible individual-level model, see Wakefield (2007) for further discussion. We consider this model

for compatibility with the majority of previous analyses.

A prior for $\sigma_u^2$ is more difficult to specify, being a conditional variance. Using the formula given in Section 4.2 we evaluate the marginal variances with $\sigma_u^2 = 1$, which vary by area; these are plotted in Figure 1. The mean of these variances is 0.59 so that the average marginal variance is smaller than the conditional value (of 1). In this example, without other information being available, we would like the priors for the unstructured and spatial random effects to be roughly equivalent, in the sense of producing similar distributions for random effects in a generic area, $v$ and $u$. Since the latter vary by area we content ourselves with achieving this in a "typical" area. The median of the inverse Ga(1,0.140) distribution that was assumed for $\sigma_v^{-2}$ is 0.20, and so we take the prior for $\sigma_u^2$ as inverse Ga(1,0.20/0.59) where the 0.59 factor adjusts for the disparity between the conditional and marginal variances. This prior gives the marginal variances in Figure 1(b); comparison with panel (a) shows that the variances are larger, as desired. Simulating from this prior gives, across areas, the median of the 0.025 quantile of $\exp(u)$ as 0.18 and the median 0.975 quantile as 5.14. These summaries tie in reasonably well with the [0.2,5] range for $\exp(v)$. Figure 1, panels (c) and (d), plot the 0.025 and 0.975 quantiles of $\exp(u)$ versus the number of neighbors and we see that the areas with fewer neighbors have a far greater spread, as we would expect. Bernardinelli et al. (1995) describe a similar method based on simulations from the prior (for different values of $\sigma_u^2$), and evaluating the ratio of marginal to conditional variances. Figure 2 shows maps of realizations from the unstructured and spatial random effects, with two sets for each. One set results from taking the median value of $\sigma_v^2$ or $\sigma_u^2$ from the prior, and the other from taking the 0.95 point of the prior. Notice that the scales of the shading on the maps differs across the four panels. Perhaps the most striking feature of these maps is the difficulty in assessing visually the strength of smoothing.

Table 3 gives PQL and INLA results for models (0.3) and (0.4), and INLA results for models (0.5)

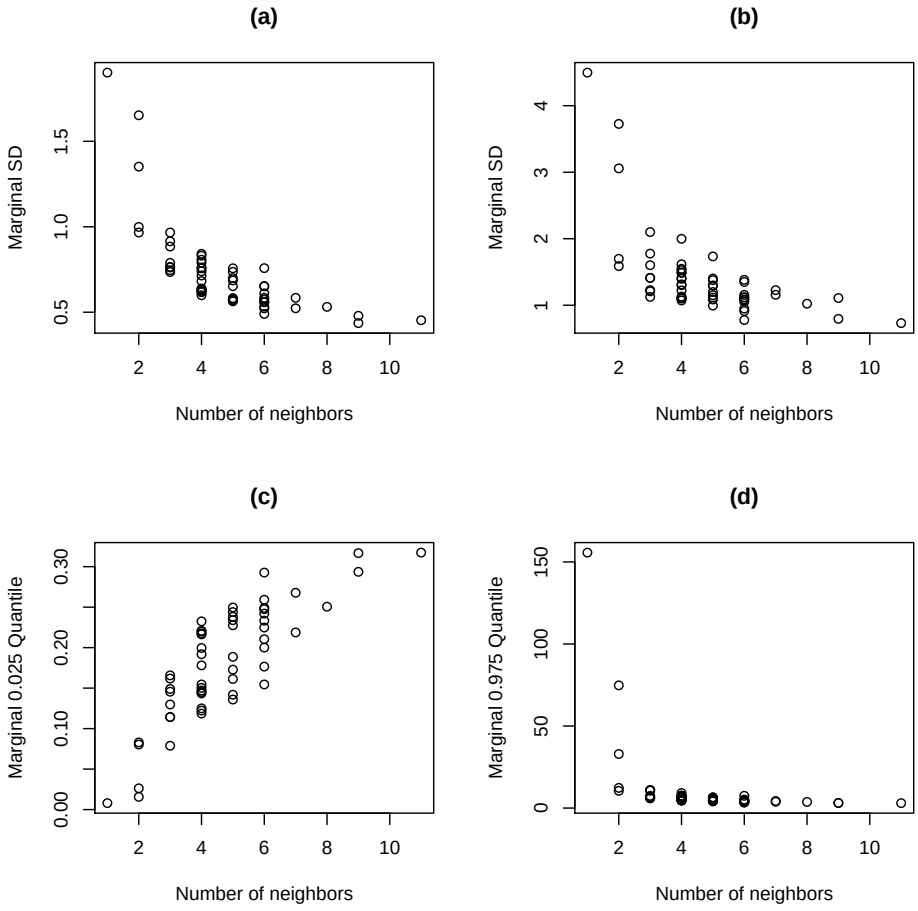


Fig. 1. (a) Marginal standard deviations from ICAR model with $\sigma_u^2 = 1$, versus number of neighbors, (b) marginal standard deviations from ICAR model with $\sigma_u^{-2} \sim \text{Ga}(1, 0.20/0.59)$, versus number of neighbors, (c) empirical 0.025 quantile of residual relative risk from ICAR model with $\sigma_u^{-2} \sim \text{Ga}(1, 0.20/0.59)$, versus number of neighbors, (d) empirical 0.975 quantile of residual relative risk from ICAR model with $\sigma_u^{-2} \sim \text{Ga}(1, 0.20/0.59)$, versus number of neighbors.

and (0.6); we could not fit the spatial models in standard software implementations of PQL. Breslow and Clayton (1993) considered models with spatial random effects only, but when the ICAR random effects are present we prefer to always include unstructured random effects, the reason being that since the ICAR model contains only a single parameter to govern both the spatial extent of dependence, and the strength of this dependence, there is no place for pure unstructured randomness to be accommodated (which can be a problem, particularly if there is negligible spatial dependence). INLA and PQL again give similar estimates and standard errors. When the spatial random effects are added in models (0.5) and (0.6) the coefficient associated with the covariate is greatly attenuated. This is a common phenomenon and occurs because the covariate has spatial structure, and the effect of any unmeasured confounders with spatial structure will be soaked into the estimate associated with the covariate $x$. When spatial random effects are included the spatial confounder effects can be accommodated in the $u_i$'s. This also explains the decrease in the estimate of $\sigma_u$ when $x$ is added to the model.

| | Model (0.3) | | Model (0.4) | | Model (0.5) | Model (0.6) |
|---|---|---|---|---|---|---|
| Variable | PQL | INLA | PQL | INLA | INLA | INLA |
| Intercept | $0.14 \pm 0.11$ | $0.08 \pm 0.12$ | $-0.44 \pm 0.16$ | $-0.49 \pm 0.16$ | 0.10±0.05 | -0.11±0.11 |
| $x$/10 | — | — | $0.68 \pm 0.14$ | $0.68 \pm 0.14$ | — | 0.26±0.12 |
| $\sigma_v$ | $0.76 \pm 0.09$ | $0.75 \pm 0.10$ | $0.60 \pm 0.08$ | $0.58 \pm 0.09$ | 0.05±0.03 | 0.05±0.03 |
| $\sigma_u$ | — | — | — | — | 0.75±0.11 | 0.69±0.11 |

Table 3. PQL and INLA summaries for four models fitted to the Scotland data.

As an example of how straightforward the implementation for GLMMs is, we present the R code for the Scotland example. Specifically, model (0.6), which includes a covariate and independent and spatial random effects, is specified by the following code:

```
data(Scotland)

Scotland$Region2 <-  Scotland$Region
```

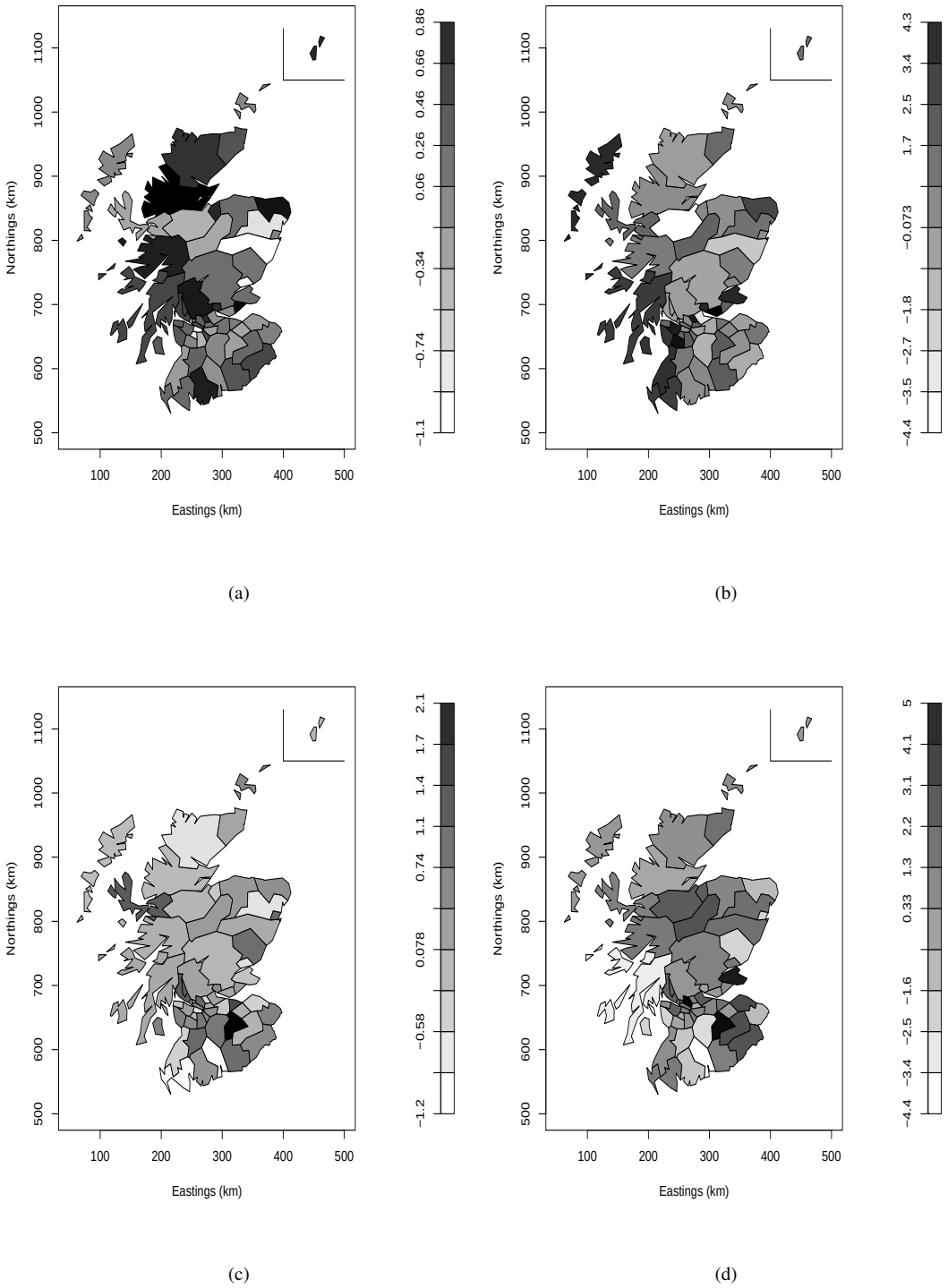


Fig. 2. Simulations from the prior for: (a) unstructured random effects with $\sigma_v$ set at the median, and (b) set at the 95% point of the prior, (c) spatial random effects with $\sigma_v$ set at the median, and (d) set at the 95% point of the prior.

```
scotlandmod4 = inla(Counts~1+I(X/10)+f(Region,model="iid",param=c(1,0.0014))+
f(Region2,model="besag",graph.file="scotland.graph",param=c(1,0.4/0.59)),
data=Scotland,family="poisson",E=E )
```

*Crossed Random Effects: The Salamander Data*

McCullagh and Nelder (1989) describe data on the success of matings between male and female salamanders of two population types (roughbutts, RB, and whitesides, WS). The experimental design is complex but involves three experiments having multiple pairings, with each salamander being involved in multiple matings, so that crossed random effects are required. The first experiment was in the summer of 1986, and the second and third were in the fall of 1986.

Let $Y_{ijk}$ denote the response (failure/success) for female $i$ and male $j$ in experiment $k$. There are 360 binary responses in total. For illustration we fit model C that was previously considered by Karim and Zeger (1992) and Breslow and Clayton (1993):

$$\text{logit } \Pr(Y_{ijk} = 1|\boldsymbol{\beta}, b^f_{ik}, b^m_{jk}) = \boldsymbol{x}_{ijk}\boldsymbol{\beta}_k + b^f_{ik} + b^m_{jk}$$

where $\boldsymbol{x}_{ijk}$ is a $1 \times 4$ vector representing the intercept and indicators for female WS, male WS, and male and female both WS, and $\boldsymbol{\beta}_k$ is the corresponding fixed effect (so that this model allows the fixed effects to vary by experiment). The model contains six random effects:

$$b^f_{ik} \sim_{iid} N(0, \sigma^2_{fk}), \quad b^m_{ik} \sim_{iid} N(0, \sigma^2_{mk}), \quad k = 1, 2, 3$$

one for each of males and females, and in each experiment. We assume that for each of the random effects the residual odds lies between 0.1 and 10 with probability 0.9, and that the marginal distribution of these odds is a log Student $t$ with 2 degrees of freedom so that Ga(1,0.622) priors are used for each of the six

precisions, $\sigma_{fk}^{-2}, \sigma_{mk}^{-2}$ for $k = 1, 2, 3$.

Table 4 gives results for the fixed effects and variance components, with the model fitted using both REML and INLA. Again we see the reduction in standard errors in the REML analysis. There is some attenuation of the Bayesian results here due to the INLA approximation strategy. This was revealed by comparison with a Bayesian analysis was carried out with `JAGS`. For example, for the first two variance components the `JAGS` posterior means were 1.42 and 0.85, as compared to the INLA results of 1.29 and 0.78, respectively.

| | Experiment 1 | | Experiment 2 | | Experiment 3 | |
|---|---|---|---|---|---|---|
| Variable | REML | INLA | REML | INLA | REML | INLA |
| Intercept | 1.34±0.62 | 1.48±0.72 | 0.57±0.67 | 0.56±0.71 | 1.02±0.65 | 1.07±0.73 |
| $WS_F$ | -2.94±0.88 | -3.26±1.01 | -2.46±0.93 | -2.51±1.02 | -3.23±0.83 | -3.39±0.92 |
| $WS_M$ | -0.42±0.63 | -0.50±0.73 | -0.77±0.72 | -0.75±0.75 | -0.82±0.86 | -0.85±0.94 |
| $WS_F \times WS_M$ | 3.18±0.94 | 3.52±1.03 | 3.71±0.96 | 3.74±1.03 | 3.82±0.99 | 4.03±1.05 |
| $\sigma_f$ | 1.25$^\star$ | 1.29±0.46 | 1.35$^\star$ | 1.38±0.50 | 0.59$^\star$ | 0.80±0.28 |
| $\sigma_m$ | 0.27$^\star$ | 0.78±0.29 | 0.96$^\star$ | 1.00±0.36 | 1.36$^\star$ | 1.46±0.48 |

Table 4. REML and INLA summaries for Salamander data. For the entries marked with a $^\star$ standard errors were unavailable.

## A Simulation Study

The Salamander example with binary data shows some inaccuracy. Hence we now examine the accuracy of INLA in the case of binomial data. Following the simulation study described in Zeger and Karim (1991) and Breslow and Clayton (1993) we assume $Y_{ij}|p_{ij} \sim_{ind}$ Binomial$(m, p_{ij})$ with $i = 1, ..., 100$ clusters, $j = 1, ..., 7$ observations per cluster, and with $m$ varying across simulations, and taken to be one of 1, 2,

4 and 8. Further, two models are examined:

$$\text{logit } p_{ij} = \beta_0 + \beta_1 t_j + \beta_2 x_i + \beta_3 t_j x_i + b_{0i} \tag{0.7}$$

$$\text{logit } p_{ij} = \beta_0 + \beta_1 t_j + \beta_2 x_i + \beta_3 t_j x_i + b_{0i} + b_{1i} t_j \tag{0.8}$$

so that in the second model the slope varies by cluster. The sampling times are $\{-3, -2, -1, 0, 1, 2, 3\}$, with $x_i = 0$ for half the sample, and 1 for the remainder. In model (0.7), $b_{0i} \sim_{iid} N(0, \sigma_0^2)$, with $\sigma_0^2 = 1$, while in model (0.8) $(b_{0i}, b_{1i})$ are zero mean bivariate normal with diagonal variance-covariance matrix, $\mathbf{\Sigma}$, with $\sigma_{00}^2 = 0.5$, $\sigma_{11}^2 = 0.25$. The values of the $\beta$ parameters are given in Table 5.

Due to computational overheads, simulation studies are rarely reported for Bayesian methods; INLA allows such studies to be quickly carried out, but care must be taken with prior specification, so as to not favor one method over another. For model (0.7) we assume the prior $\sigma_{00}^{-2} \sim \text{Ga}(0.5, 0.0164)$ which has 2.5%, 50%, 97.5% quantiles of (0.0065,0.072,33.4) with 86% of the prior mass to the left of 1 (the value used in the simulation). For model (0.8) we assume $\mathbf{\Sigma}^{-1}$ follow a Wishart distribution with 3 degrees of freedom, and diagonal scale parameter with diagonal elements 0.17 and 0.025. The latter yields 2.5%, 50% and 97.5% points of (0.023,0.12,3.1) for $\sigma_0^2$ and (0.0034,0.018,0.50) for $\sigma_{11}^2$. These priors have 84% and 72% of the mass the left of the values that were used to simulate the data. We note also that we are evaluating the frequentist accuracy of Bayesian summaries.

Table 5 gives the mean parameter estimates over 100 simulations. For model (0.7) fixed effect estimation is unbiased, but for $m = 1$ (and to a lesser extent $m = 2$) the estimate of $\sigma_{00}^2$ is attenuated. For model (0.8) there is some attenuation of the fixed effects in general, and more serious attenuation of the variances, along with poor estimation of $\rho$. With reference to Table 1 of Breslow and Clayton (1993) we note that INLA provides more accurate estimation than PQL.

| $m$ | $\sigma^2_{00}$ | $\sigma^2_{11}$ | $\rho$ | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ |
|---|---|---|---|---|---|---|---|
| Truth | 1.00 | — | — | -2.50 | 1.00 | -1.00 | -0.50 |
| 1 | 0.67 | — | — | -2.46 | 0.98 | -1.01 | -0.48 |
| 2 | 0.91 | — | — | -2.52 | 1.00 | -1.03 | -0.48 |
| 4 | 0.97 | — | — | -2.53 | 1.00 | -1.00 | -0.50 |
| 8 | 0.98 | — | — | -2.53 | 1.01 | -0.98 | -0.50 |
| Truth | 0.50 | 0.25 | 0.00 | -2.50 | 1.00 | -1.00 | -0.50 |
| 1 | 0.29 | 0.11 | 0.18 | -2.42 | 0.97 | -0.90 | -0.48 |
| 2 | 0.27 | 0.17 | 0.43 | -2.40 | 0.96 | -0.97 | -0.50 |
| 4 | 0.32 | 0.21 | 0.33 | -2.44 | 0.98 | -0.98 | -0.51 |
| 8 | 0.40 | 0.23 | 0.14 | -2.49 | 0.99 | -0.97 | -0.51 |

Table 5. Mean values (over 100 simulations) of parameter estimates in the binomial simulation study. The top and bottom parts of the table give the results for models (0.7) and (0.8), respectively.

Table 6 gives the average posterior standard deviations along with the empirical standard deviations of the posterior means (across simulations). This table may be compared directly with Table 2 of Breslow and Clayton (1993) which evaluated the PQL method. We see a good correspondence between the standard deviations for model (0.7), while for model (0.8) we see some discrepancy for $m = 1$.

We also analyzed the simulated data with MCMC (using the `JAGS` program), and with $200,000$ iterations being performed. Table 7 gives the average of the posterior means over the 100 simulations. We see that for small $m$ there is underestimation of the variances, due to the strong influence of the prior (recall that our priors were such that the simulated value occur at the 0.72 to 0.86 quantile of the prior). The $\beta$ parameters are well estimated.

Tables 8 and 9 summarize the accuracy of INLA as compared to MCMC. For a generic parameter $\theta$ each entry corresponds to $(\mathrm{E}[\theta^{inla}|\boldsymbol{y}] - \mathrm{E}[\theta^{mcmc}|\boldsymbol{y}])/\mathrm{sd}(\theta^{mcmc}|\boldsymbol{y})$, where we transform the variance components to the whole real line to make the summaries more meaningful. Hence we are examining the size of the error relative to the "true" posterior standard deviation. We see that the error in approximation is around 30% of the (true) posterior standard deviation for $m = 1$, but by $m = 4$, at least in the scenarios

| $m$ | | $\sigma^2_{00}$ | $\sigma^2_{11}$ | $\rho$ | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ |
|---|---|---|---|---|---|---|---|---|
| 1 | Sim | 0.58 | — | — | 0.32 | 0.14 | 0.48 | 0.22 |
| | Est | 0.50 | — | — | 0.31 | 0.13 | 0.46 | 0.20 |
| 2 | Sim | 0.42 | — | — | 0.25 | 0.09 | 0.37 | 0.14 |
| | Est | 0.37 | — | — | 0.24 | 0.09 | 0.36 | 0.14 |
| 4 | Sim | 0.27 | — | — | 0.20 | 0.06 | 0.33 | 0.09 |
| | Est | 0.29 | — | — | 0.20 | 0.07 | 0.29 | 0.10 |
| 8 | Sim | 0.22 | — | — | 0.18 | 0.05 | 0.29 | 0.07 |
| | Est | 0.25 | — | — | 0.17 | 0.05 | 0.25 | 0.07 |
| 1 | Sim | 0.16 | 0.11 | 0.30 | 0.29 | 0.17 | 0.43 | 0.25 |
| | Est | 0.30 | 0.11 | 0.52 | 0.28 | 0.14 | 0.43 | 0.21 |
| 2 | Sim | 0.15 | 0.10 | 0.32 | 0.22 | 0.12 | 0.35 | 0.16 |
| | Est | 0.19 | 0.09 | 0.38 | 0.20 | 0.11 | 0.31 | 0.16 |
| 4 | Sim | 0.19 | 0.07 | 0.30 | 0.17 | 0.10 | 0.25 | 0.13 |
| | Est | 0.18 | 0.07 | 0.30 | 0.17 | 0.10 | 0.24 | 0.14 |
| 8 | Sim | 0.17 | 0.05 | 0.23 | 0.13 | 0.08 | 0.22 | 0.11 |
| | Est | 0.15 | 0.06 | 0.21 | 0.14 | 0.09 | 0.20 | 0.12 |

Table 6. Average estimated posterior deviations, and empirical standard deviation (over 100 simulations) of the posterior means in the binomial simulation study. The top and bottom parts of the table give the results for models (0.7) and (0.8), respectively.

| $m$ | $\sigma^2_{00}$ | $\sigma^2_{11}$ | $\rho$ | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ |
|---|---|---|---|---|---|---|---|
| Truth | 1.00 | — | — | -2.50 | 1.00 | -1.00 | -0.50 |
| 1 | 0.90 | — | — | -2.55 | 1.02 | -1.05 | -0.49 |
| 2 | 1.03 | — | — | -2.56 | 1.02 | -1.04 | -0.48 |
| 4 | 1.04 | — | — | -2.54 | 1.01 | -1.01 | -0.50 |
| 8 | 1.02 | — | — | -2.53 | 1.01 | -0.98 | -0.50 |
| Truth | 0.50 | 0.25 | 0.00 | -2.50 | 1.00 | -1.00 | -0.50 |
| 1 | 0.36 | 0.14 | 0.31 | -2.49 | 1.00 | -0.93 | -0.49 |
| 2 | 0.31 | 0.19 | 0.47 | -2.43 | 0.97 | -0.99 | -0.50 |
| 4 | 0.35 | 0.22 | 0.33 | -2.45 | 0.98 | -0.98 | -0.51 |
| 8 | 0.43 | 0.24 | 0.13 | -2.49 | 1.00 | -0.97 | -0.51 |

Table 7. Average of posterior means from MCMC analysis of 100 simulated datasets. The top and bottom parts of the table give the results for models (0.7) and (0.8), respectively.

| $m$ | $\log \sigma_{00}^{-2}$ | $\log \sigma_{11}^{-2}$ | logit $[(1+\rho)/2]$ | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ |
|---|---|---|---|---|---|---|---|
| 1 | 0.35 | — | — | 0.26 | -0.25 | 0.08 | 0.04 |
| 2 | 0.29 | — | — | 0.13 | -0.14 | 0.04 | 0.01 |
| 4 | 0.17 | — | — | 0.06 | -0.09 | 0.01 | 0.01 |
| 8 | 0.10 | — | — | 0.02 | -0.05 | 0.01 | 0.00 |
| 1 | 0.22 | 0.30 | -0.28 | 0.20 | -0.17 | 0.07 | -0.01 |
| 2 | 0.17 | 0.15 | -0.13 | 0.09 | -0.06 | 0.04 | -0.04 |
| 4 | 0.18 | 0.08 | 0.00 | 0.02 | -0.00 | 0.01 | -0.01 |
| 8 | 0.17 | 0.07 | 0.06 | 0.01 | -0.01 | 0.01 | 0.00 |

Table 8. Comparison between INLA and MCMC. For a generic parameter $\theta$ each entry corresponds to $(\mathrm{E}[\theta^{inla}|\boldsymbol{y}] - \mathrm{E}[\theta^{mcmc}|\boldsymbol{y}])/\mathrm{sd}(\theta^{mcmc}|\boldsymbol{y})$. The top and bottom parts of the table give the results for models (0.7) and (0.8), respectively.

| $m$ | $\log \sigma_{00}^{-2}$ | $\log \sigma_{11}^{-2}$ | logit $[(1+\rho)/2]$ | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ |
|---|---|---|---|---|---|---|---|
| 1 | 1.12 | — | — | 0.84 | 0.87 | 0.86 | 0.89 |
| 2 | 1.07 | — | — | 0.91 | 0.91 | 0.91 | 0.92 |
| 4 | 1.00 | — | — | 0.95 | 0.94 | 0.94 | 0.95 |
| 8 | 0.99 | — | — | 0.98 | 0.97 | 0.97 | 0.97 |
| 1 | 1.01 | 1.05 | 1.00 | 0.85 | 0.84 | 0.90 | 0.91 |
| 2 | 1.02 | 1.09 | 1.02 | 0.91 | 0.93 | 0.93 | 0.96 |
| 4 | 1.02 | 1.02 | 1.02 | 0.95 | 0.96 | 0.96 | 0.98 |
| 8 | 1.02 | 1.00 | 1.02 | 0.97 | 0.98 | 0.97 | 0.99 |

Table 9. Ratio of posterior variances as estimated by INLA, as compared to those estimated by MCMC. The top and bottom parts of the table give the results for models (0.7) and (0.8), respectively.

considered here, the approximation is reasonable. This is confirmed by Table 9 in which we examine the ratio of posterior variances $\mathrm{Var}(\theta^{inla}|\boldsymbol{y})/\mathrm{Var}(\theta^{mcmc}|\boldsymbol{y})$. The posterior variance for the transformed variance components are overestimated, but only slightly, while the posterior variances for the regression coefficients are underestimated.

## REFERENCES

Bernardinelli, L., Clayton, D.G., and Montomoli, C. (1995). Bayesian estimates of disease maps: How important are priors? *Statistics in Medicine 14*, 2411–32.

Breslow, N.E. and Clayton, D.G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association 88*, 9–25.

Crowder, M.J. (1978). Beta-binomial ANOVA for proportions. *Applied Statistics 27*, 24–37.

Karim, M.R. and Zeger, S.L. (1992). Generalized linear models with random effects: Salamander mating revisited. *Biometrics 48*, 631–644.

Kneale, G.W. (1971). Problems arising in estimating from retrospective survey data the latent period of juvenile cancers initiated by obstetric radiography. *Biometrics 27*, 563–590.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models, Second Edition*. London: Chapman and Hall.

Spiegelhalter, DJ., Thomas, A., and Best, NG. (1998). WinBUGS User Manual, version 1.1.1. Cambridge, UK.

Wakefield, J.C. (2007). Disease mapping and spatial regression with count data. *Biostatistics 8*, 158–183.

Zeger, S.L. and Karim, M.R. (1991). Generalized linear models with random effects: A Gibbs sampling approach. *Journal of the American Statistical Association 86*, 79–86.