Lip Cancer Example

- Consider the spatial disease mapping model:

$$\begin{aligned} Y_i \mid \mu_i &\overset{ind}{\sim} Po\left(E_i\, e^{\mu_i}\right) \;, \quad \text{where} \\ Y_i &= \text{observed disease count,} \\ E_i &= \text{expected count (known), and} \\ \mu_i &= \mathbf{x}_i'\boldsymbol{\beta} + \theta_i + \phi_i \end{aligned}$$

Lip Cancer Example

- Consider the spatial disease mapping model:

$$\begin{aligned} Y_i \,|\, \mu_i &\overset{ind}{\sim} Po\left(E_i\, e^{\mu_i}\right) \ , \quad \text{where} \\ Y_i &= \text{observed disease count,} \\ E_i &= \text{expected count (known), and} \\ \mu_i &= \mathbf{x}_i' \boldsymbol{\beta} + \theta_i + \phi_i \end{aligned}$$

- The $\mathbf{x}_i$ are explanatory spatial covariates; typically $\boldsymbol{\beta}$ is assigned a flat prior.

Lip Cancer Example

- Consider the spatial disease mapping model:

$$\begin{aligned} Y_i \mid \mu_i &\overset{ind}{\sim} Po\left(E_i \, e^{\mu_i}\right) \;, \quad \text{where} \\ Y_i &= \text{observed disease count,} \\ E_i &= \text{expected count (known), and} \\ \mu_i &= \mathbf{x}_i' \boldsymbol{\beta} + \theta_i + \phi_i \end{aligned}$$

- The $\mathbf{x}_i$ are explanatory spatial covariates; typically $\boldsymbol{\beta}$ is assigned a flat prior.
- Note the mean structure also contains two sets of random effects! The first, θ_i, capture heterogeneity among the regions via

$$\theta_i \overset{iid}{\sim} N(0 \,,\, 1/\tau_h) \;.$$

Lip Cancer Example

- The second set, ϕ_i, capture regional clustering via a conditionally autoregressive (CAR) prior,

$$\phi_i \,|\, \phi_{j \neq i} \sim N(\bar{\phi}_i \, , \, 1/(\tau_c m_i)) \ ,$$

where m_i is the number of “neighbors” of region i, and $\bar{\phi}_i = m_i^{-1} \Sigma_{j \in \partial_i} \, \phi_j$.

Lip Cancer Example

- The second set, ϕ_i, capture regional clustering via a conditionally autoregressive (CAR) prior,

$$\phi_i \,|\, \phi_{j \neq i} \sim N(\bar{\phi}_i \; , \; 1/(\tau_c m_i)) \; ,$$

where m_i is the number of “neighbors” of region i, and $\bar{\phi}_i = m_i^{-1} \Sigma_{j \in \partial_i} \, \phi_j$.

- The CAR prior is translation invariant, so typically we insist $\Sigma_{i=1}^I \phi_i = 0$ (imposed numerically after each MCMC iteration). Still, Y_i cannot inform about θ_i or ϕ_i, but only about their sum $\eta_i = \theta_i + \phi_i$.

Lip Cancer Example

- The second set, ϕ_i, capture regional clustering via a conditionally autoregressive (CAR) prior,

$$\phi_i \,|\, \phi_{j \neq i} \sim N(\bar{\phi}_i \,,\, 1/(\tau_c m_i)) \;,$$

  where m_i is the number of "neighbors" of region i, and $\bar{\phi}_i = m_i^{-1} \Sigma_{j \in \partial_i} \phi_j$.

- The CAR prior is translation invariant, so typically we insist $\Sigma_{i=1}^{I} \phi_i = 0$ (imposed numerically after each MCMC iteration). Still, Y_i cannot inform about θ_i or ϕ_i, but only about their sum $\eta_i = \theta_i + \phi_i$.

- Making the reparametrization from $(\boldsymbol{\theta}, \boldsymbol{\phi})$ to $(\boldsymbol{\theta}, \boldsymbol{\eta})$, we have the joint posterior

$$p(\boldsymbol{\theta}, \boldsymbol{\eta} | \mathbf{y}) \propto L(\boldsymbol{\eta}; \mathbf{y}) p(\boldsymbol{\theta}) p(\boldsymbol{\eta} - \boldsymbol{\theta}).$$

Lip Cancer Example

- This means that

$$p(\theta_i \mid \theta_{j \neq i}, \boldsymbol{\eta}, \mathbf{y}) \propto p(\theta_i)\, p(\eta_i - \theta_i \mid \{\eta_j - \theta_j\}_{j \neq i}) \ .$$

Lip Cancer Example

- This means that

$$p(\theta_i \mid \theta_{j\neq i}, \boldsymbol{\eta}, \mathbf{y}) \propto p(\theta_i)\, p(\eta_i - \theta_i \mid \{\eta_j - \theta_j\}_{j\neq i}) \;.$$

- Since this distribution is free of the data $\mathbf{y}$, the θ_i are Bayesianly unidentified (and so are the ϕ_i).

Lip Cancer Example

- This means that

$$p(\theta_i \mid \theta_{j \neq i}, \boldsymbol{\eta}, \mathbf{y}) \propto p(\theta_i)\, p(\eta_i - \theta_i \mid \{\eta_j - \theta_j\}_{j \neq i})\ .$$

- Since this distribution is free of the data $\mathbf{y}$, the θ_i are Bayesianly unidentified (and so are the ϕ_i).
- BUT this does not preclude Bayesian learning about θ_i; this would instead require

$$p(\theta_i | \mathbf{y}) = p(\theta_i)\ .$$

[Stronger condition: data have no impact on the marginal (not conditional) posterior.]

Lip Cancer Example

- Dilemma: Though unidentified, the θ_i and ϕ_i are interesting in their own right, as is

$$\psi = \frac{sd(\boldsymbol{\phi})}{sd(\boldsymbol{\theta}) + sd(\boldsymbol{\phi})} ,$$

where $sd(\cdot)$ is the empirical marginal standard deviation function. Can we specify vague but proper prior values τ_h and τ_c that

Lip Cancer Example

- Dilemma: Though unidentified, the θ_i and ϕ_i are interesting in their own right, as is

$$\psi = \frac{sd(\boldsymbol{\phi})}{sd(\boldsymbol{\theta}) + sd(\boldsymbol{\phi})} ,$$

  where $sd(\cdot)$ is the empirical marginal standard deviation function. Can we specify vague but proper prior values τ_h and τ_c that
    - lead to acceptable convergence behavior, and

Lip Cancer Example

- Dilemma: Though unidentified, the θ_i and ϕ_i are interesting in their own right, as is

$$\psi = \frac{sd(\boldsymbol{\phi})}{sd(\boldsymbol{\theta}) + sd(\boldsymbol{\phi})} \ ,$$

  where $sd(\cdot)$ is the empirical marginal standard deviation function. Can we specify vague but proper prior values τ_h and τ_c that
  - lead to acceptable convergence behavior, and
  - still allow Bayesian learning?

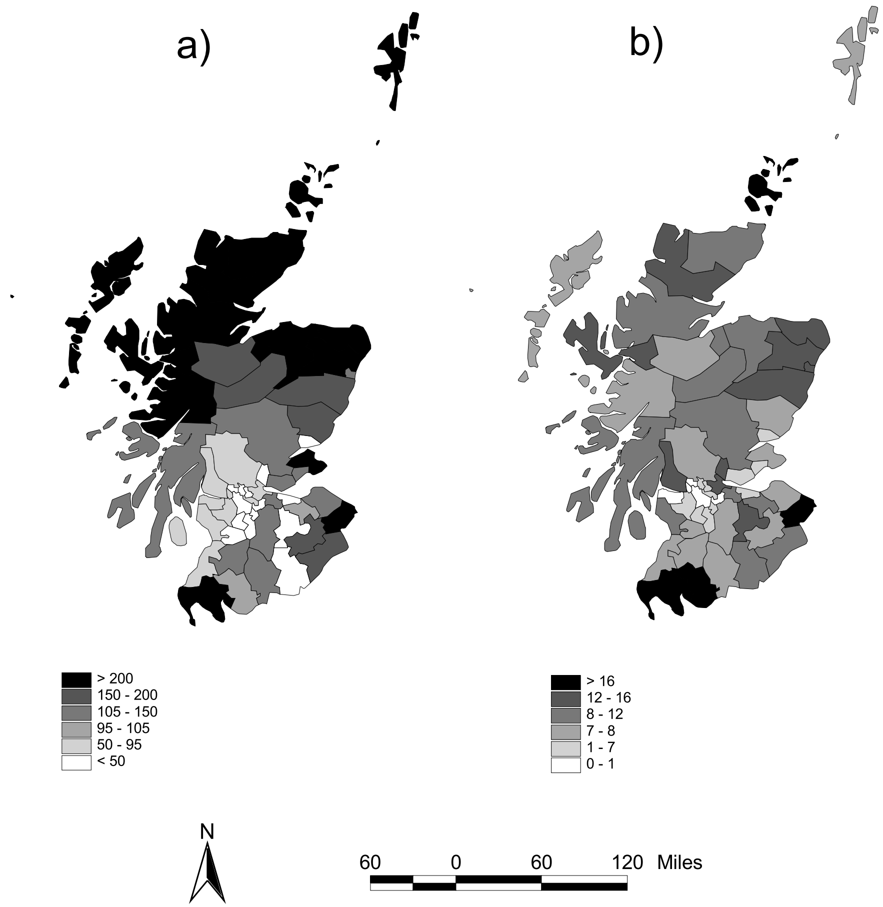
Lip Cancer Example

- Dilemma: Though unidentified, the θ_i and ϕ_i are interesting in their own right, as is

$$\psi = \frac{sd(\boldsymbol{\phi})}{sd(\boldsymbol{\theta}) + sd(\boldsymbol{\phi})} ,$$

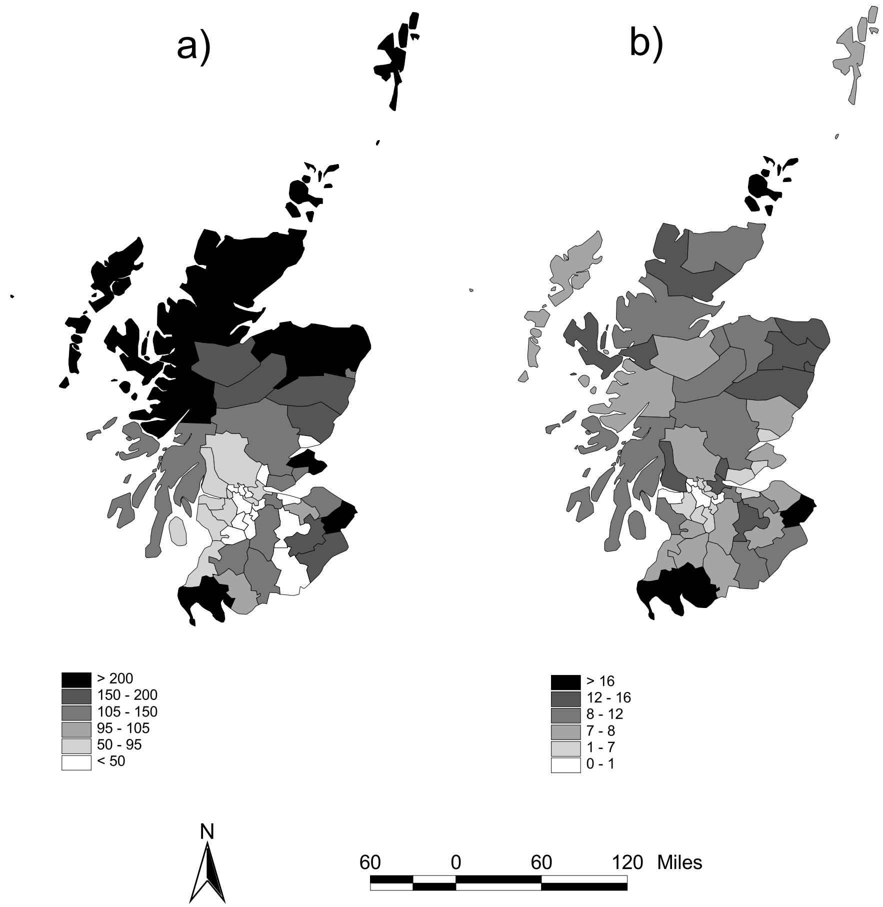
  where $sd(\cdot)$ is the empirical marginal standard deviation function. Can we specify vague but proper prior values τ_h and τ_c that
  - lead to acceptable convergence behavior, and
  - still allow Bayesian learning?
- Tricky to specify a “fair” prior balance between heterogeneity and clustering (e.g., one for which $\psi \approx 1/2$) since θ_i prior is specified marginally while the ϕ_i prior is specified conditionally.

Dataset: Scottish lip cancer data



a) $SMR_i = 100Y_i/E_i$, standardized mortality ratio for lip cancer in $I = 56$ districts, 1975–1980

Dataset: Scottish lip cancer data



a) $SMR_i = 100 Y_i / E_i$, standardized mortality ratio for lip cancer in $I = 56$ districts, 1975–1980

b) one covariate, x_i = percentage of the population engaged in agriculture, fishing or forestry (AFF)

WinBUGS code to fit this model

```
model {
   for (i in 1 : regions) {
      O[i] ~ dpois(mu[i])
      log(mu[i]) <- log(E[i]) + beta*aff[i]/10 + phi[i] + theta[i]
      theta[i] ~ dnorm(0.0,tau.h)
      eta[i] <- theta[i] + phi[i]
   }
   phi[1:regions] ~ {\red car.normal}(adj[], weights[], num[], tau.c)

   beta ~ dnorm(0.0, 1.0E-5)  # vague prior on covariate effect

   tau.h ~ dgamma(1.0E-3,1.0E-3)  #  ''fair'' prior from Best et al.
   tau.c ~ dgamma(1.0E-1,1.0E-1)  #   (1999, Bayesian Statistics 6)

   sd.h <- sd(theta[]) # marginal SD of heterogeneity effects
   sd.c <- sd(phi[])   # marginal SD of clustering (spatial) effects
   psi <- sd.c / (sd.h + sd.c)
}

(See WinBUGS Map manual for DATA and INITS for this example)
```

Lip Cancer Results

	posterior for ψ			posterior for β		
priors for τ_c, τ_h	mean	sd	l1acf	mean	sd	l1acf
G(1.0, 1.0), G(3.2761, 1.81)	.57	.058	.80	.43	.17	.94
G(.1, .1), G(.32761, .181)	.65	.073	.89	.41	.14	.92
G(.1, .1), G(.001, .001)	.82	.10	.98	.38	.13	.91

	posterior for η_1			posterior for η_{56}		
priors for τ_c, τ_h	mean	sd	l1acf	mean	sd	l1acf
G(1.0, 1.0), G(3.2761, 1.81)	.92	.40	.33	–.96	.52	.12
G(.1, .1), G(.32761, .181)	.89	.36	.28	–.79	.41	.17
G(.1, .1), G(.001, .001)	.90	.34	.31	–.70	.35	.21

- AFF covariate is significantly $\neq 0$ under all 3 priors

Lip Cancer Results

priors for τ_c, τ_h	posterior for ψ mean	sd	l1acf	posterior for β mean	sd	l1acf
G(1.0, 1.0), G(3.2761, 1.81)	.57	.058	.80	.43	.17	.94
G(.1, .1), G(.32761, .181)	.65	.073	.89	.41	.14	.92
G(.1, .1), G(.001, .001)	.82	.10	.98	.38	.13	.91

priors for τ_c, τ_h	posterior for η_1 mean	sd	l1acf	posterior for η_{56} mean	sd	l1acf
G(1.0, 1.0), G(3.2761, 1.81)	.92	.40	.33	–.96	.52	.12
G(.1, .1), G(.32761, .181)	.89	.36	.28	–.79	.41	.17
G(.1, .1), G(.001, .001)	.90	.34	.31	–.70	.35	.21

- AFF covariate is significantly $\neq 0$ under all 3 priors
- convergence is slow for ψ and β, but rapid for η_i (and μ_i)

Lip Cancer Results

	posterior for ψ			posterior for β		
priors for τ_c, τ_h	mean	sd	l1acf	mean	sd	l1acf
G(1.0, 1.0), G(3.2761, 1.81)	.57	.058	.80	.43	.17	.94
G(.1, .1), G(.32761, .181)	.65	.073	.89	.41	.14	.92
G(.1, .1), G(.001, .001)	.82	.10	.98	.38	.13	.91
	posterior for η_1			posterior for η_{56}		
priors for τ_c, τ_h	mean	sd	l1acf	mean	sd	l1acf
G(1.0, 1.0), G(3.2761, 1.81)	.92	.40	.33	–.96	.52	.12
G(.1, .1), G(.32761, .181)	.89	.36	.28	–.79	.41	.17
G(.1, .1), G(.001, .001)	.90	.34	.31	–.70	.35	.21

- AFF covariate is significantly $\neq 0$ under all 3 priors
- convergence is slow for ψ and β, but rapid for η_i (and μ_i)
- Excess variability in the data is mostly due to clustering ($E(\psi|\mathbf{y}) > .50$), but the posterior distribution for ψ does not seem robust to changes in the prior.