Lecture 9

Further hierarchical models

Outline

There is huge scope for elaborating the basic hierarchical models discussed in the previous lecture to reflect additional structure and complexity in the data, e.g.

- Adding covariates at different levels of the hierarchy
- Adding further levels to the hierarchy (patients within wards within hospitals, pupils within schools within local authorities, . . .)
- Adding non-nested (cross-classified) levels (patients within GPs crossed with hospitals, . . .)
- Repeated observations on some/all units (longitudinal data)
- Modelling temporal or spatial structure in data, . . .

In this lecture, we will discuss:

- Hierarchical models for count data and including covariates
- Hierarchical models for longitudinal data
- Cross-classified models

Hierarchical models for count data: Disease mapping

- In disease mapping, we are interested in modelling counts of disease cases collected on each of a number of geographical areas within a study region
- Here we consider data on the observed number of cases of childhood leukaemia, y_i, diagnosed in a 10 year period in each of $i = 1, \ldots, 879$ areas (electoral wards) in London (data from Thames Cancer Registry)
- Using national age/sex-standardised reference rates for leukaemia and Census population counts, we can also calculate the expected number of cases, E_i, in each area
- Assume a Poisson likelihood for the disease count in each area:

$$y_i \sim \text{Poisson}(\mu_i); \quad \mu_i = \lambda_i E_i; \quad i = 1, \ldots, 879$$

- We have 879 *distinct* relative risk parameters λ_i
- What prior should we specify for each λ_i?

Different modelling assumptions

Identical parameters

Assume $\lambda_i = \lambda$ for all i and assign a prior

$$\lambda \sim \text{Gamma}(a, b)$$

with *specified values* of a and b, e.g.

$$\lambda \sim \text{Gamma}(1, 1)$$

$\rightarrow$ conjugate Poisson-gamma model

Independent parameters

Assume independent vague priors for each relative risk, e.g.

$$\lambda_i \sim \text{Gamma}(0.1, 0.1), \quad i = 1, \ldots, 879$$

$\rightarrow$ This will give estimates of the posterior mean for $\lambda_i \approx y_i / E_i$, which is the MLE (also termed standardised morbidity ratio, SMR)

Different modelling assumptions (continued)

Similar (exchangeable) parameters

Specify a hierarchical random effects prior:

$$\lambda_i \sim \text{Gamma}(a, b), \qquad i = 1, \ldots, 879$$

where a and b are *unknown parameters* to also be estimated

$\rightarrow$ assign hyperprior distributions to a and b

$\rightarrow$ what is a suitable hyperprior for these parameters?

A more flexible hierarchical prior for the relative risks

- A gamma random effects prior for the λ_i is mathematically convenient, but may be restrictive:
  - covariate adjustment (regression) is difficult
  - no possibility for allowing spatial correlation between risks in nearby areas
- A normal random effects prior for $\log \lambda_i$ is more flexible:

$$\begin{aligned} y_i &\sim \text{Poisson}(\mu_i = \lambda_i E_i) \\ \log \lambda_i &= \alpha + \theta_i \\ \theta_i &\sim \text{Normal}(0, \sigma^2) \end{aligned}$$

- Need to specify hyperprior distributions for

  σ^2 (between-area variance), e.g. $\sigma^{-2} \sim \text{Gamma}(0.001, 0.001)$

  α (mean log relative risk), e.g. $\alpha \sim \text{Normal}(0, 10000)$

Parameter Interpretation

- θ_i are the random effects
- $\lambda_i = \exp(\alpha + \theta_i) =$ relative risk in area *i* compared to expected risk based on age and sex of population
- θ_i can also be thought of as a latent variable which captures the effects of unknown or unmeasured area level covariates
- If these area level covariates are spatially structured (e.g. environmental effects), our model for θ_i should allow for this (i.e. replace normal random effects distribution by spatial distribution — not covered in this course)
- The variance of the random effects (σ^2) reflects the amount of extra-Poisson variation in the data

Ranking in hierarchical models

- Recent trend in UK towards ranking 'institutional' performance e.g. schools, hospitals or areas
- Rank of a point estimate is a highly unreliable summary statistic
  - would like measure of uncertainty about rank
- Bayesian methods provide posterior interval estimates for ranks
- For the leukemia example, at each MCMC iteration, ranking sampled values of $\lambda_1, \ldots, \lambda_{879}$ gives sample from posterior distribution of ranks for each area
- See Goldstein and Spiegelhalter (1996) for further discussion on ranking

BUGS contains 'built-in' options for ranks:

- `Rank` option of `Inference` menu monitors the rank of the elements of a specified vector
- `rank(x[],i)` returns the rank of the ith element of `x`
- `ranked(x[],i)` returns the value of the ith-ranked element of `x`

Quantile ratios to summarise level 2 variability

- Unclear how to define or calculate the VPC for generalised linear hierarchical models
- Alternative summary of variability between units in a hierarchical model is to rank the random effects and calculate the difference or ratio between two units at opposite extremes
- For the leukemia example, suppose we consider the 5^{th} and 95^{th} percentiles of the area relative risk distribution
  - let $\lambda_{5\%}$ denote the relative risk of leukemia for the area ranked at the 5^{th} percentile
  - let $\lambda_{95\%}$ denote the relative risk of leukemia for the area ranked at the 95^{th} percentile
  - then $\text{QR}_{90} = \frac{\lambda_{95\%}}{\lambda_{5\%}}$ = ratio of relative risks of leukemia between the top and bottom 5% of areas
- Using MCMC, we can calculate the ranks, and hence the QR_{90}, at each iteration, and hence obtain a posterior distribution for QR_{90}

BUGS code

```
model {
 for(i in 1 : N) {
  Y[i] ~ dpois(mu[i])
  log(mu[i]) <- log(E[i]) + alpha + theta[i]
  theta[i] ~ dnorm(0, tau) # area random effects
  lambda[i] <- exp(alpha + theta[i]) # area relative risk
 }
 # Priors:
 alpha ~ dnorm(0, 0.0001) # vague prior on overall intercept
 tau ~ dgamma(0.5, 0.0005) # precision of area random effects
 sigma <- 1/sqrt(tau)      # between-area sd of random effects

 # 90% quantile ratio for area relative risks
 QR90 <- ranked(lambda[],835)/ranked(lambda[],44)

 #rank
 for(i in 1 : N) {
  rank.lambda[i] <- rank(lambda[], i) # rank of area i
 }
}
```

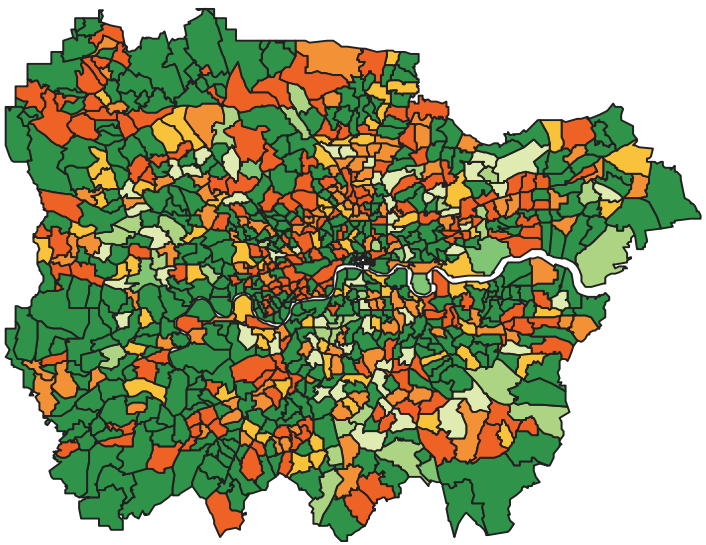
Results for childhood leukaemia example

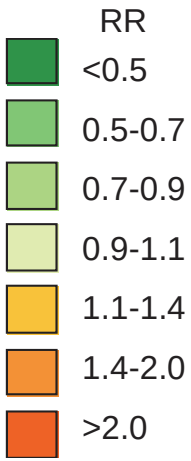
Parameters of interest:

- $e^{\alpha+\theta_i}$ (`lambda[i]`) = relative risk of leukaemia in area *i* relative to expected (see map)
- σ (`sigma`) = between-area standard deviation of log relative risk of leukaemia
  - posterior mean and 95% interval = 0.46 (0.34, 0.62)
- QR_{90} (`QR90`) = 4.7 (95% interval 2.9 to 7.5)
  - so 4.7-fold variation in relative risk of leukemia between top and bottom 5% of areas

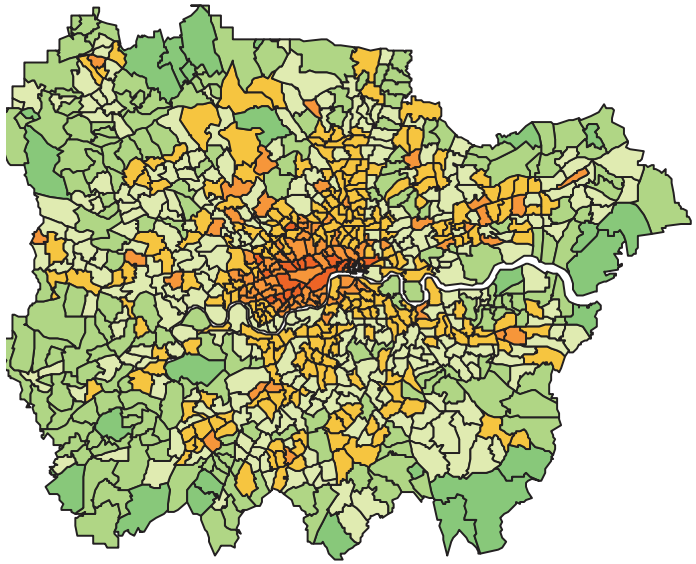
Maps of estimated area-specific RR of leukaemia

SMR

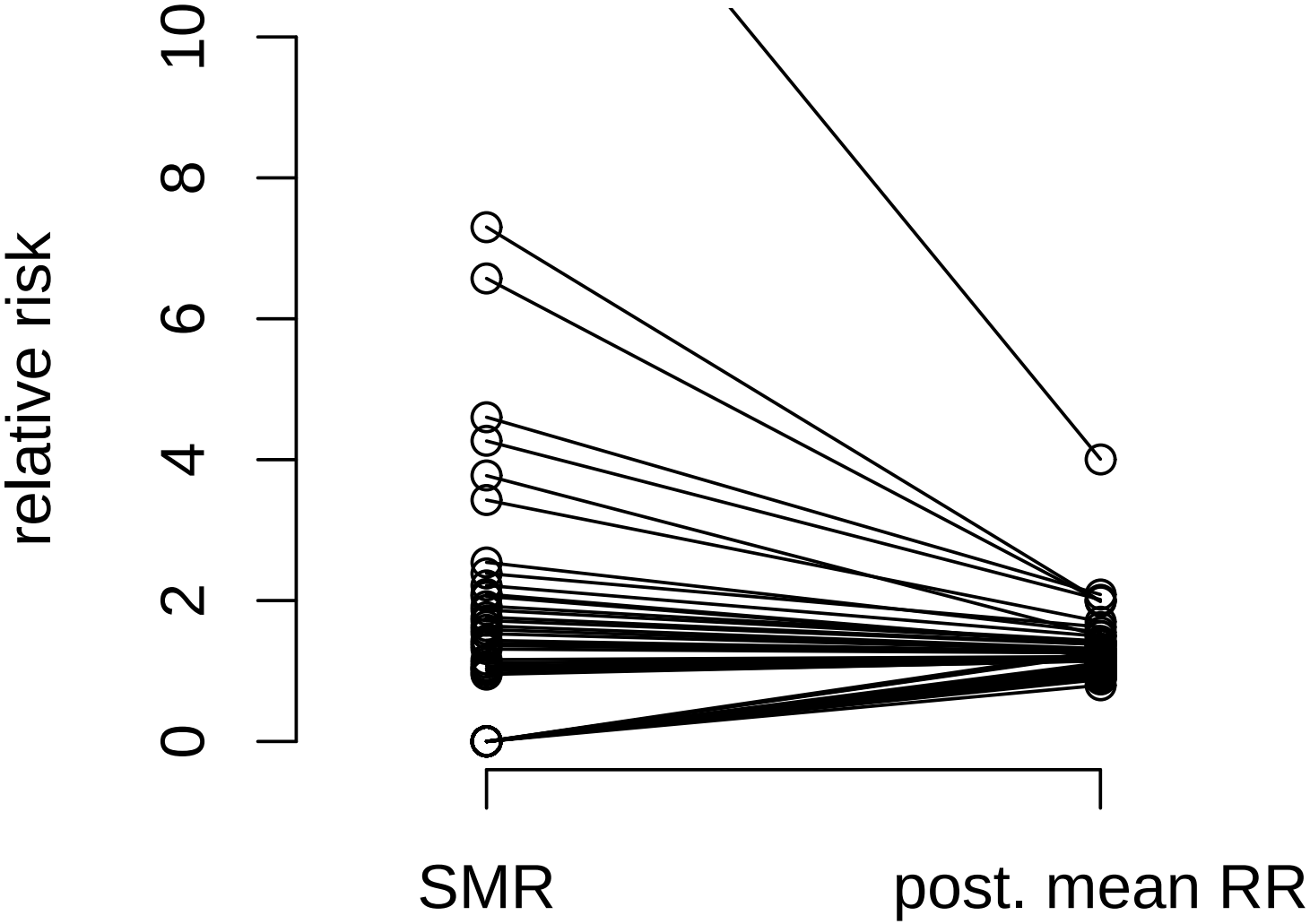




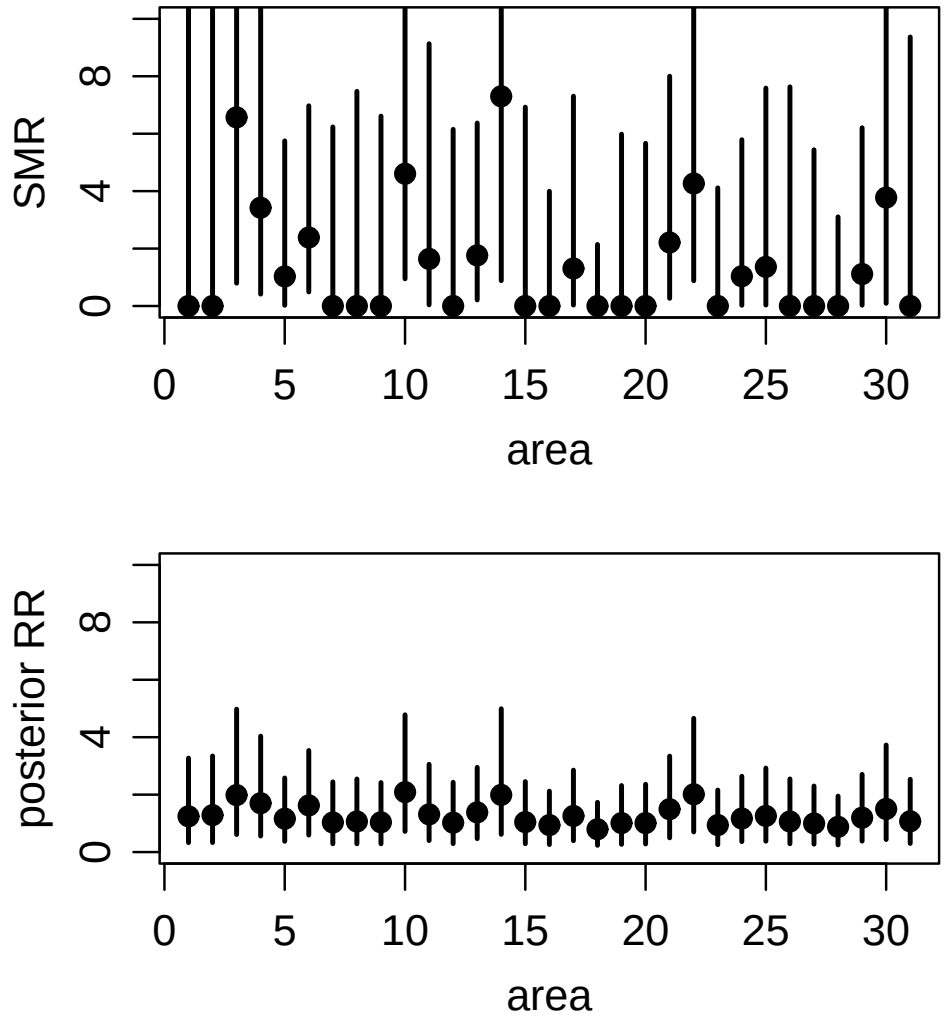
Smoothed RR



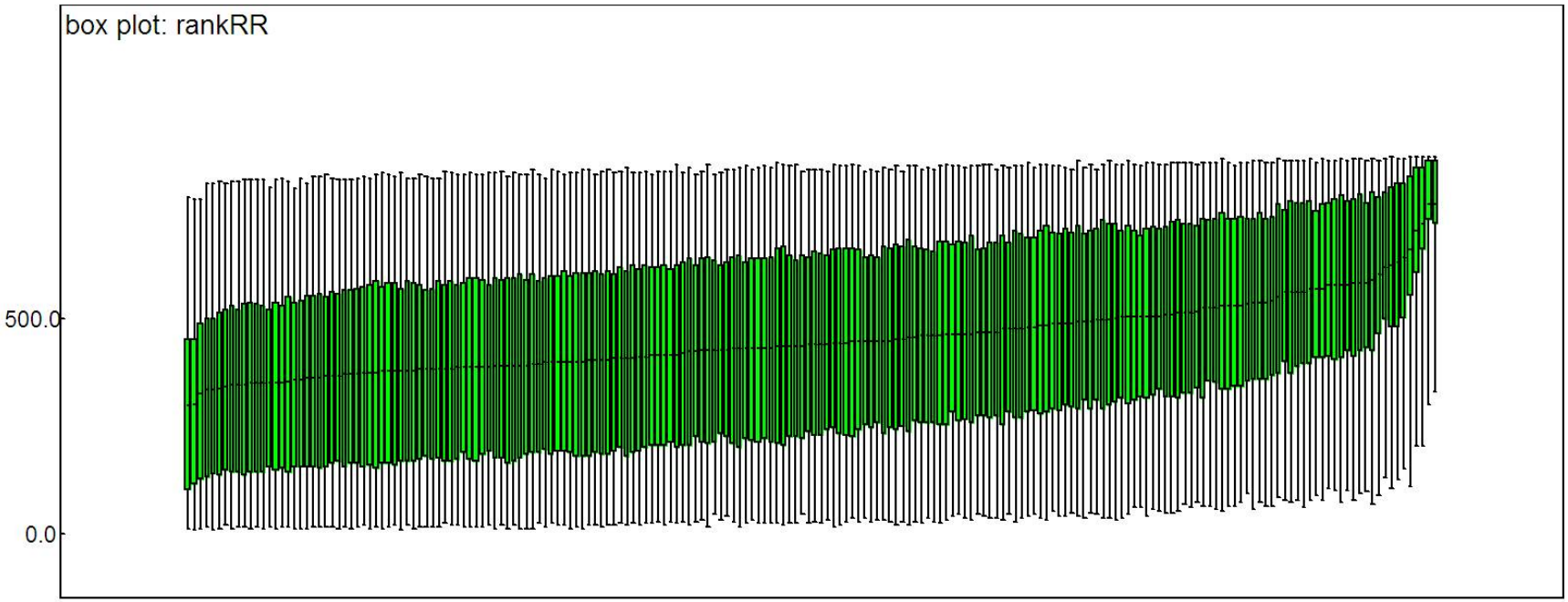
SMR versus posterior mean RR for selected areas



Point estimate and 95% interval for relative risk in selected areas

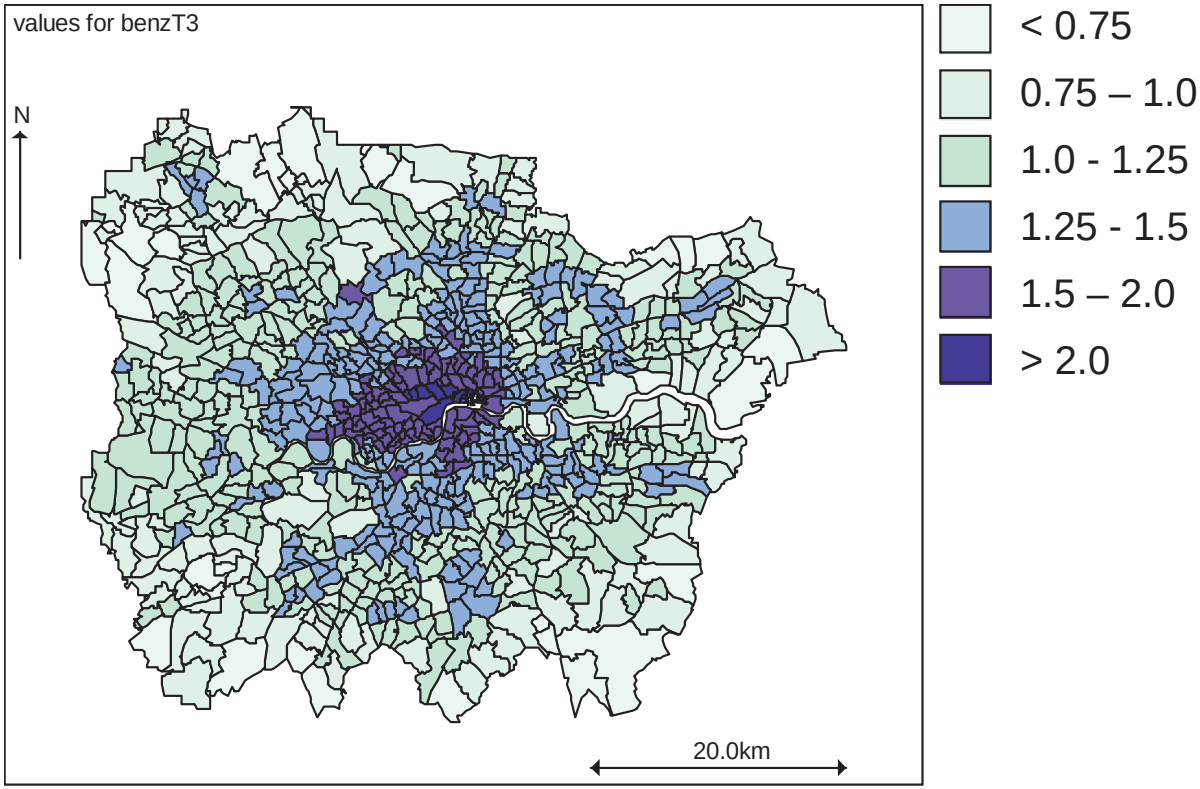


Posterior distribution of area ranks



Map of mean benzene levels per ward

(tonnes per annum on cube root scale)



- Can we explain some of the variation in risk of leukaemia by environmental exposure to benzene?

Including covariates in hierarchical models

- Let X_i = average benzene emissions (tonnes per annum) in ward i (following cube-root transformation to reduce skew)
- Include X as a covariate in the hierarchical model:

$$\begin{aligned} y_i &\sim \text{Poisson}(E_i\lambda_i); \quad i = 1, \ldots, 873 \\ \log \lambda_i &= \alpha + \beta X_i + \theta_i \\ \theta_i &\sim \text{Normal}(0, \sigma^2) \\ \alpha, \beta, \sigma^2 &\sim \text{vague priors} \end{aligned}$$

Extract from BUGS code

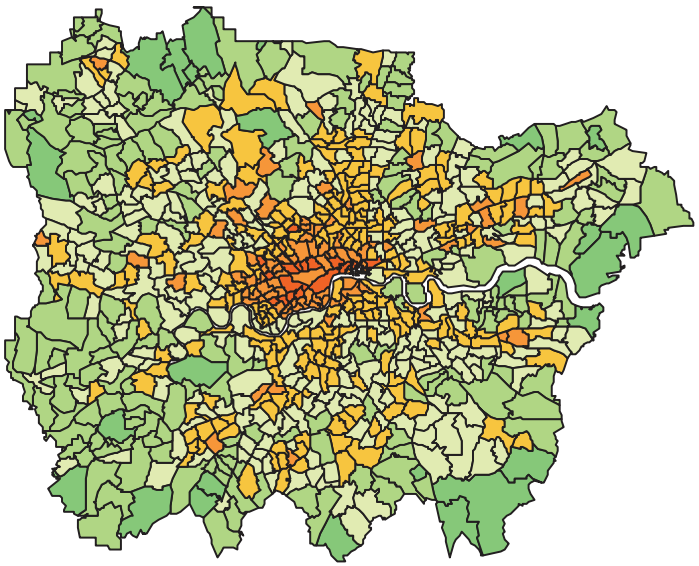
```
for(i in 1 : N) {
 Y[i] ~ dpois(mu[i])
 log(mu[i]) <- log(E[i]) + alpha + beta*X[i] + theta[i]
 theta[i] ~ dnorm(0, tau) # area random effects
 lambda[i] <- exp(alpha + beta*X[i] + theta[i]) # area RR
 residRR[i] <- exp(theta[i]) # unexplained area residual RR
}
# Priors:
alpha  ~ dnorm(0, 0.0001) # vague prior on overall intercept
beta ~ dnorm(0, 0.0001) # vague prior on regression coefficient
RR.benz <- exp(beta) # RR per unit increase in X (benzene)

tau ~ dgamma(0.5, 0.0005) # precision of area random effects
sigma <- 1/sqrt(tau) # between-area sd of random effects

# 90% quantile ratio for area relative risks
QR90 <- ranked(lambda[],835)/ranked(lambda[],44)
# 90% quantile ratio for area residual relative risks
residQR90 <- ranked(residRR[],835)/ranked(residRR[],44)
```

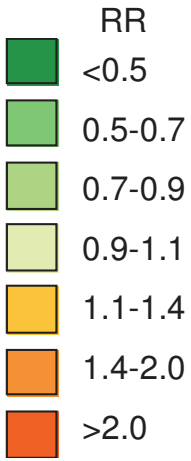
Results

- e^{β} (`RR.benz`) = RR of leukaemia associated with unit increase in cube root benzene emissions in area of residence = 2.23 (1.64, 2.96)
- Residual 90% quantile ratio (`residQR90`) indicates that there is a 3.9-fold (95% CI 1.8 to 5.0-fold) variation in residual relative risk between the top and bottom 5% of areas after adjusting for effects of benzene
  - Compare with estimate of QR_{90} = 4.7 from model without benzene
- λ_i (`lambda`) = RR of leukaemia in area *i* relative to London average (see map)
- e^{θ_i} (`residRR`) = residual relative risk of leukaemia in area *i* relative to London average after adjusting for effects of benzene (see map)

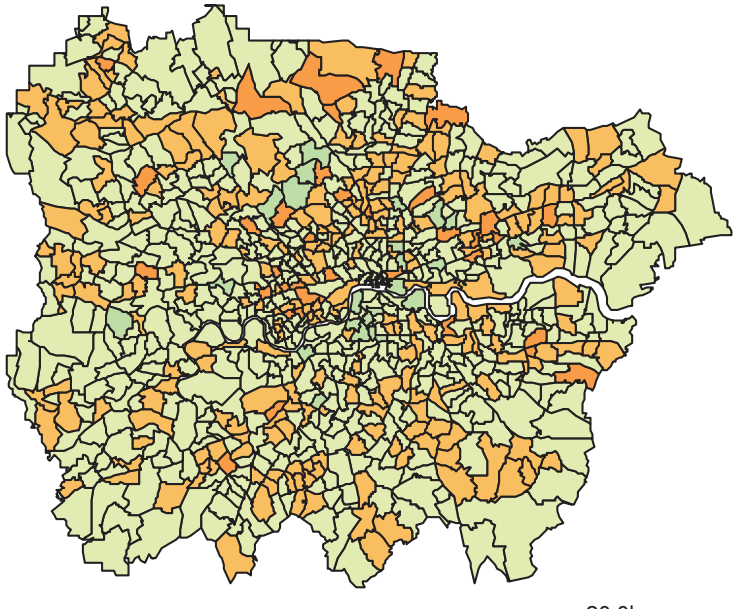
Maps of area-specific RR of leukaemia

Smoothed RR





Smoothed residual RR
after adjusting for benzene



Longitudinal data

- Arise in studies where individual (or units) are measured repeatedly over time
- For a given individual, observations over time will be typically dependent
- Longitudinal data can arise in various forms:
  - continuous or discrete response
    - discrete response can be binary/binomial, categorical or counts
  - equally spaced or irregularly spaced
  - same or different time points for each individual
  - with or without missing data
  - many or few time points, T
  - many or few individuals or units, n

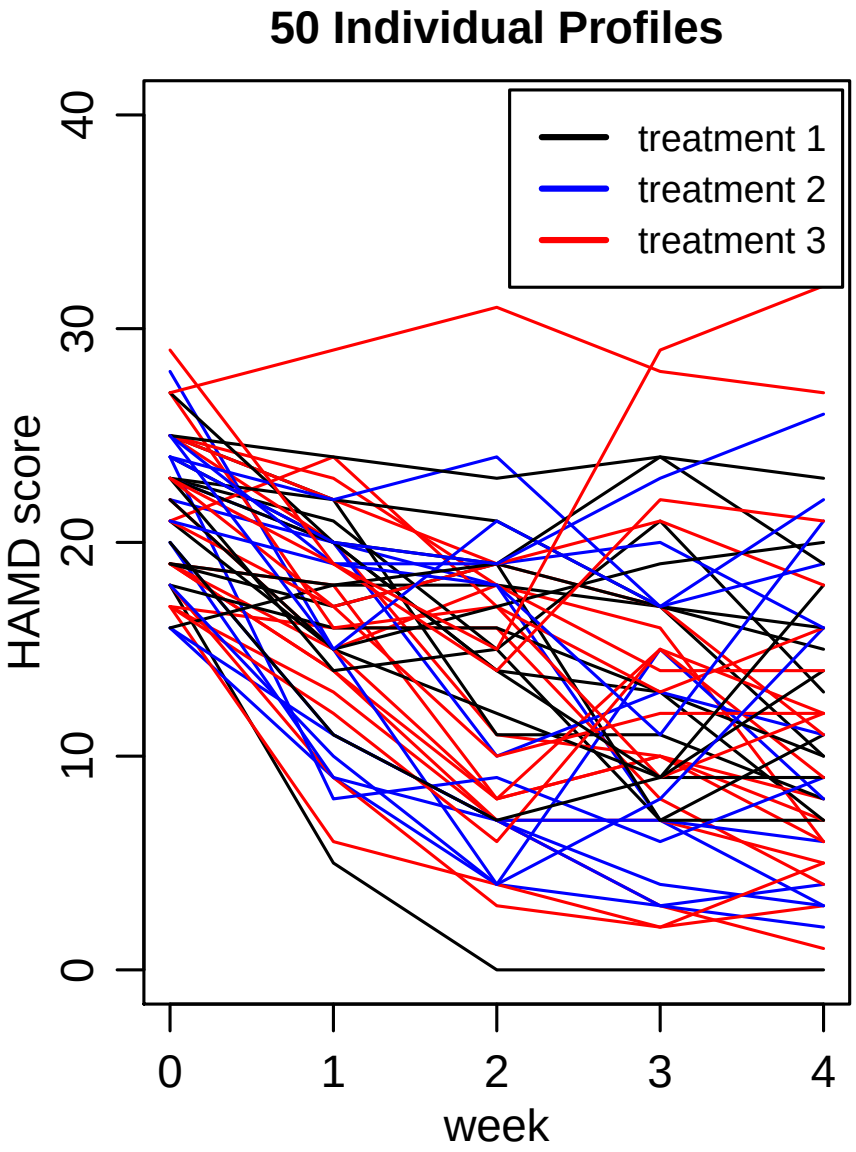
Analysing longitudinal data

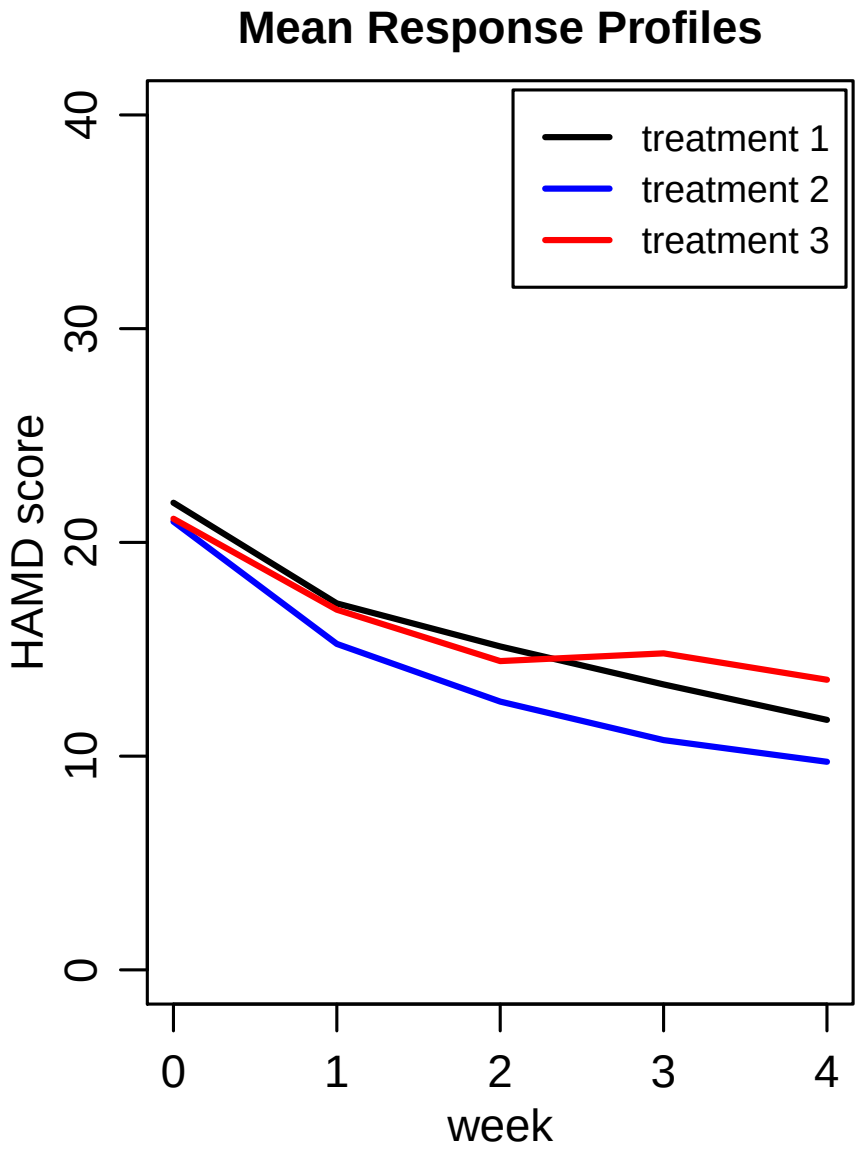
- There are many different ways to analyse longitudinal data
- The key feature of longitudinal data is the need to account for the dependence structure of the data
- Two common methods:
  - random effects (hierarchical) models
  - autoregressive models
- Here, we will focus on random effects models

HAMD Example: antidepressant clinical trial

- 6 centre clinical trial, comparing 3 treatments of depression
- 367 subjects randomised to one of 3 treatments
- Subjects rated on Hamilton depression score (HAMD) on 5 weekly visits
  - week 0 before treatment
  - weeks 1-4 during treatment
- HAMD score takes values 0-50
  - the higher the score, the more severe the depression
- Subjects drop out from week 2 onwards, but for now we
  - ignore the subjects who dropped out
  - analyse the 246 complete cases
- Data was previously analysed by Diggle and Kenward (1994)

HAMD Example: data





HAMD Example: objective

- Study objective: are there any differences in the effects of the 3 treatments on the change in HAMD score over time?
- The variables we will use are:
  - y: Hamilton depression (HAMD) score
  - t: treatment
  - w: week
- For simplicity we will
  - ignore any centre effects
  - assume linear relationships
- The models we will consider are:
  - a non-hierarchical model (standard linear regression)
  - a hierarchical model with random intercepts
  - a hierarchical model with random intercepts and random slopes

HAMD Example: a Bayesian (non-hierarchical) linear model (LM)

- Specification:
  - probability distribution for responses:

$$y_{iw} \sim \text{Normal}(\mu_{iw}, \sigma^2)$$

y_{iw} = the HAMD score for individual i in week w (weeks 0,. . . ,4)

  - linear predictor: $\mu_{iw} = \alpha + \beta_{treat(i)} w$

$treat(i)$ = the treatment indicator of individual i, so it can take values 1, 2 or 3

w = the week of the visit, takes value 0 for visit before treatment and values 1-4 for follow-up visits

- In this model no account is taken of the repeated structure (observations are nested within individuals)
- Assume vague priors for all parameters:

$$\begin{aligned} \alpha, \beta_1, \beta_2, \beta_3 &\sim \text{Normal}(0, 10000) \\ \frac{1}{\sigma^2} &\sim \text{Gamma}(0.001, 0.001) \end{aligned}$$

HAMD Example: a Bayesian hierarchical linear model

- Modify LM to allow a separate intercept for each individual:

$$\begin{aligned} y_{iw} &\sim \text{Normal}(\mu_{iw}, \sigma^2) \\ \mu_{iw} &= \alpha_i + \beta_{treat(i)} w \end{aligned}$$

  We are assuming that *conditionally* on α_i, $\{y_{iw}, w = 0, \ldots, 4\}$ are independent

- Assume that all the $\{\alpha_i\}$ follow a *common* prior distribution, e.g.

$$\alpha_i \sim \text{Normal}(\mu_\alpha, \sigma_\alpha^2) \quad i = 1, \ldots, 246$$

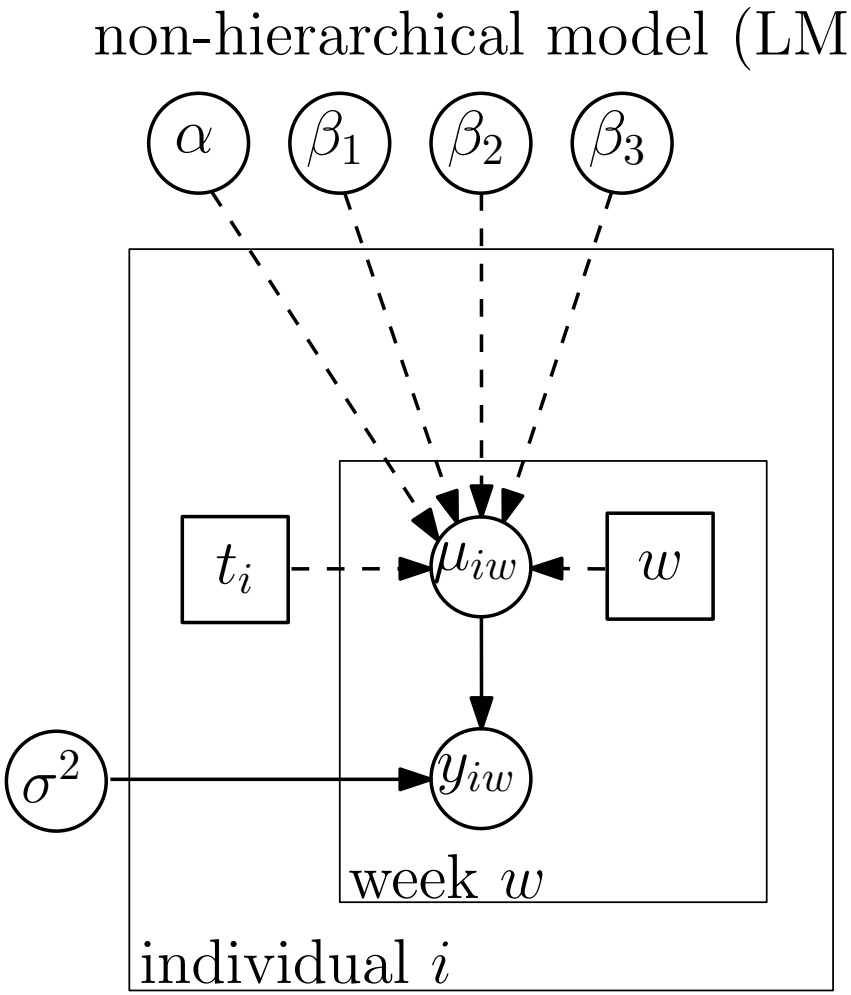
  Here we are assuming exchangeability between all the individuals

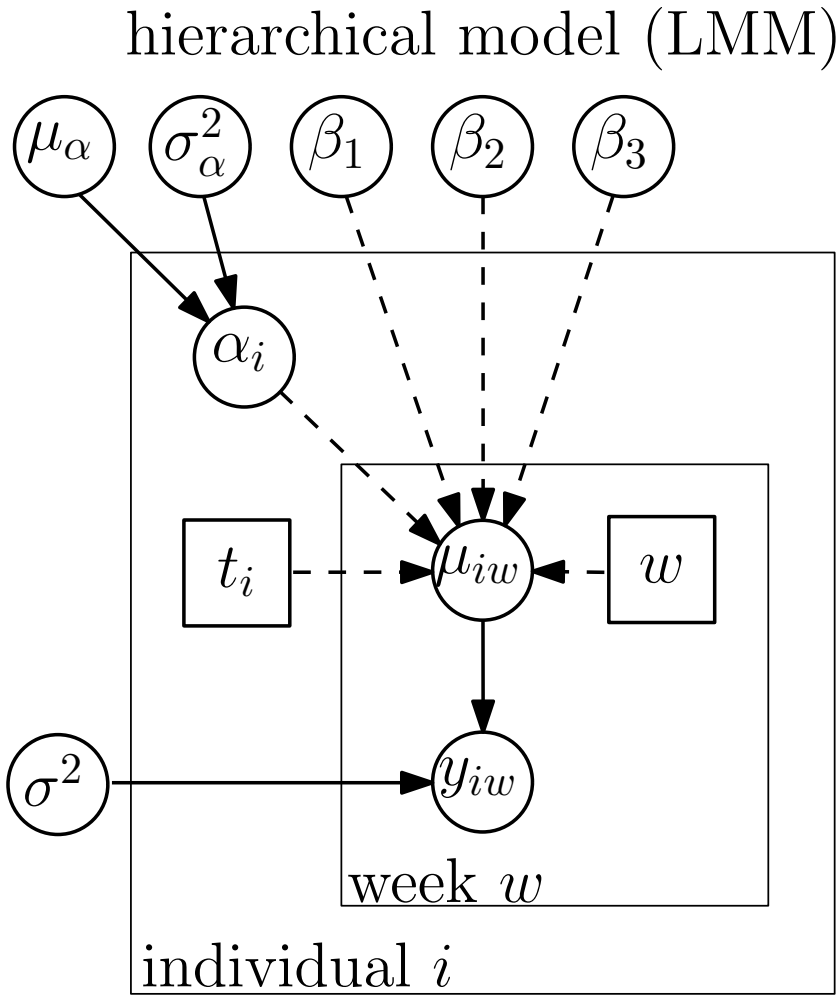
- We may then assume vague priors for the *hyperparameters* of the population distribution:

$$\begin{aligned} \mu_\alpha &\sim \text{Normal}(0, 10000) \\ \sigma_\alpha &\sim \text{Uniform}(0, 100) \end{aligned}$$

- This is an example of a *Hierarchical LM* or *Linear Mixed Model (LMM)* or *Random Intercepts* model

HAMD Example: DAGs for LM and LMM





t_i represents the treatment indicator of individual i

HAMD Example: WinBUGS code for LM and LMM

Part of WinBUGS code for non-hierarchical model:

```
for (i in 1:N) { # N individuals
  for (w in 1:W) { # W weeks
    hamd[i,w]~dnorm(mu[i,w],tau)
    mu[i,w]<-alpha+beta[treat[i]]*(w-1)
  }
}
# specification of priors ....
```

Part of WinBUGS code for hierarchical model:

```
for (i in 1:N) { # N individuals
  for (w in 1:W) { # W weeks
    hamd[i,w]~dnorm(mu[i,w],tau)
    mu[i,w]<-alpha[i]+beta[treat[i]]*(w-1)
  }
  alpha[i]~dnorm(alpha.mu,alpha.tau) # random intercepts
}
# specification of priors ....
```

HAMD Example: WinBUGS code for priors

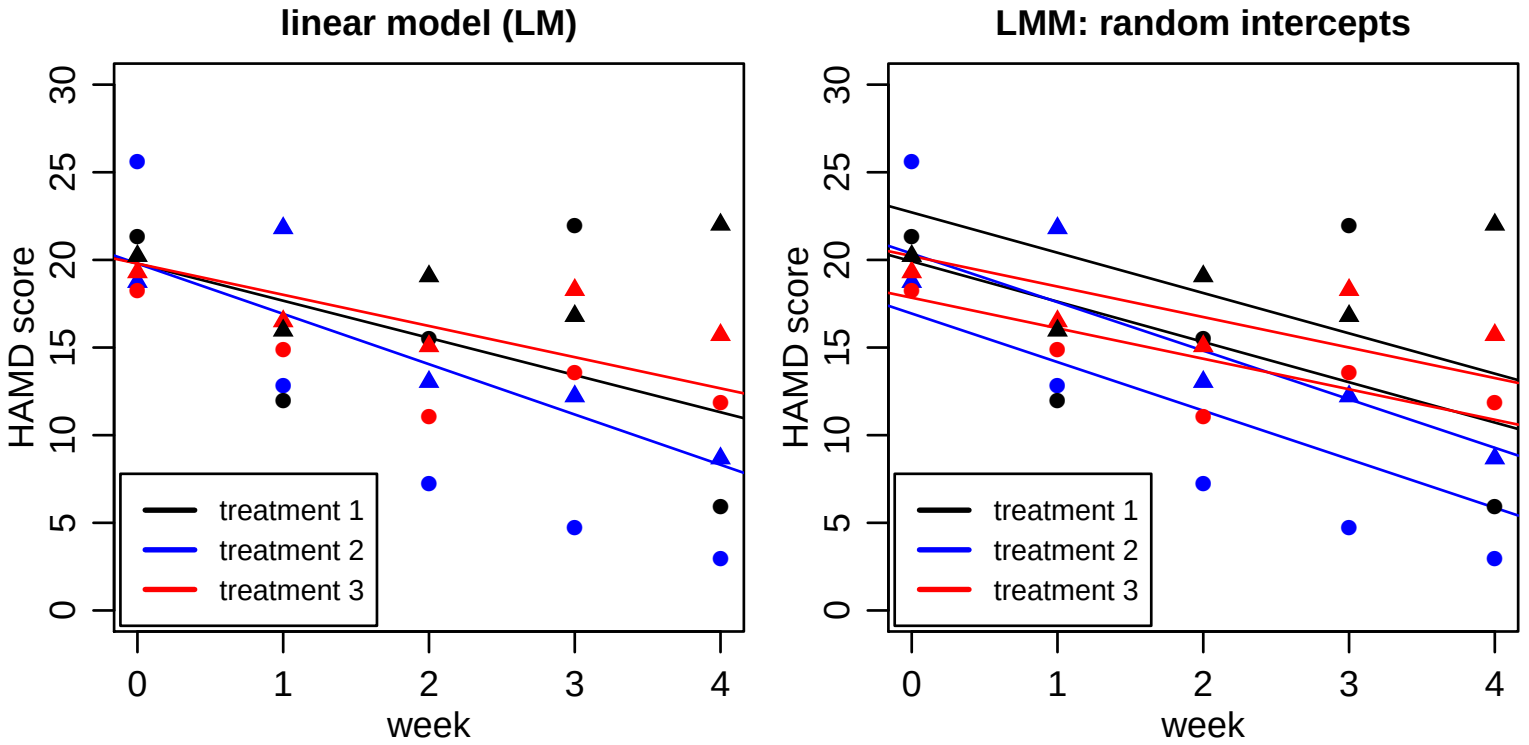
Prior specification for non-hierarchical model:

```
alpha~dnorm(0,0.00001)
for (t in 1:T){  # T treatments
  beta[t]~dnorm(0,0.00001)
  }
tau~dgamma(0.001,0.001)
sigma.sq<-1/tau  # Normal errors
```

Prior specification for hierarchical model:

```
alpha.mu~dnorm(0,0.00001)
alpha.sigma~dunif(0,100)
alpha.sigma.sq<-pow(alpha.sigma,2)
alpha.tau<-1/alpha.sigma.sq
for (t in 1:T){  # T treatments
  beta[t]~dnorm(0,0.00001)
  }
tau~dgamma(0.001,0.001)
sigma.sq<-1/tau  # Normal errors
```

HAMD Example: LM and LMM fitted lines



circles and triangles represent scores for 6 individuals (2 for each treatment)

- LM:
  - 3 regressions lines fitted, 1 for each treatment
  - each treatment has the same intercept, but a different slope
- LMM:
  - each individual has a different regression line
  - but for each treatment, individuals have the same slope

HAMD Example: results for LM and LMM

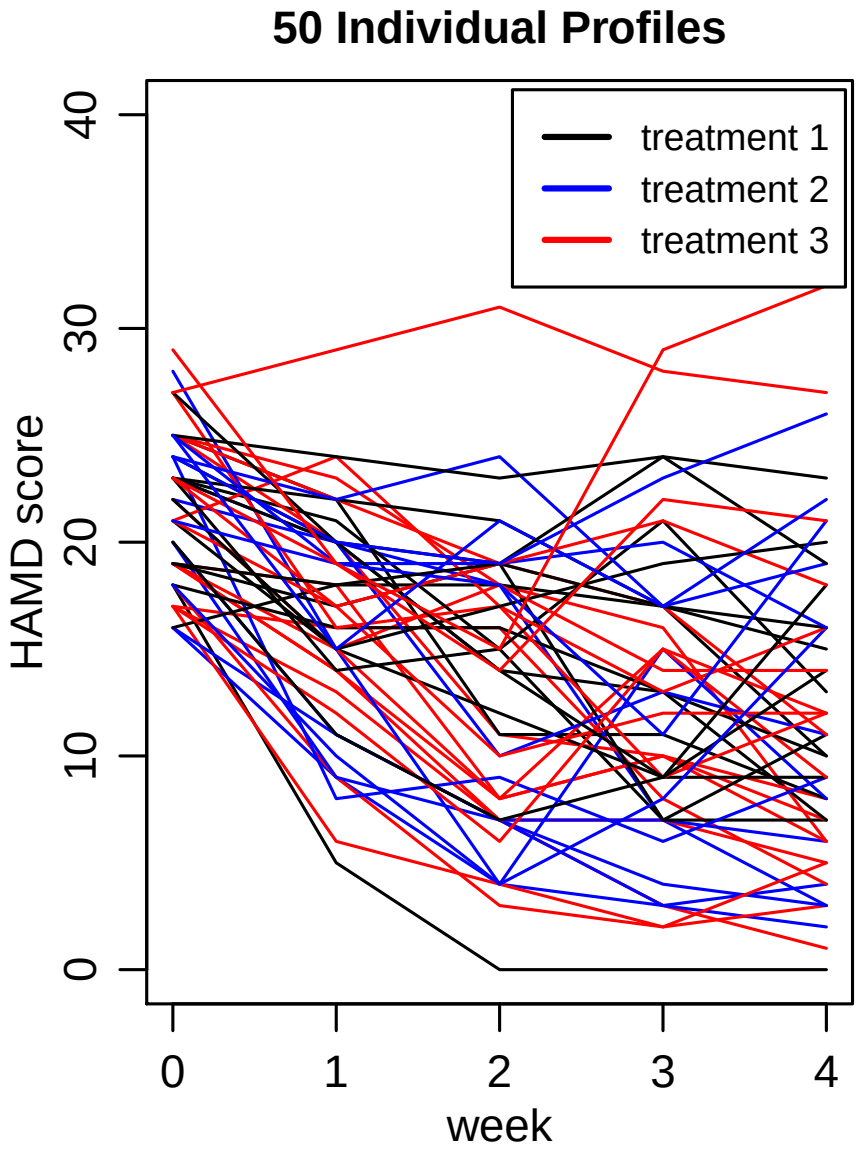
Table: posterior mean (95% credible interval) for the non-hierarchical and hierarchical models fitted to the HAMD data

	non-hierarchical model			hierarchical model	
α	19.8	(19.2,20.4)	μ_α	19.8	(19.1,20.5)
			σ^2_α	17.6	(14.0,21.9)
β_1	-2.1	(-2.4,-1.8)	β_1	-2.3	(-2.6,-2.0)
β_2	-2.9	(-3.2,-2.6)	β_2	-2.8	(-3.0,-2.5)
β_3	-1.8	(-2.1,-1.5)	β_3	-1.7	(-2.0,-1.5)
σ^2	35.4	(32.6,38.5)	σ^2	18.2	(16.6,19.8)

Note

- the variability in the intercept in the hierarchical model
- how the residual variance (σ^2) is reduced when random effects are incorporated

HAMD Example: revisiting the data



The plot of the raw data

- indicates that separate intercepts are appropriate
- also suggests including separate slopes

So we add random slopes to the hierarchical model

HAMD Example: adding random slopes

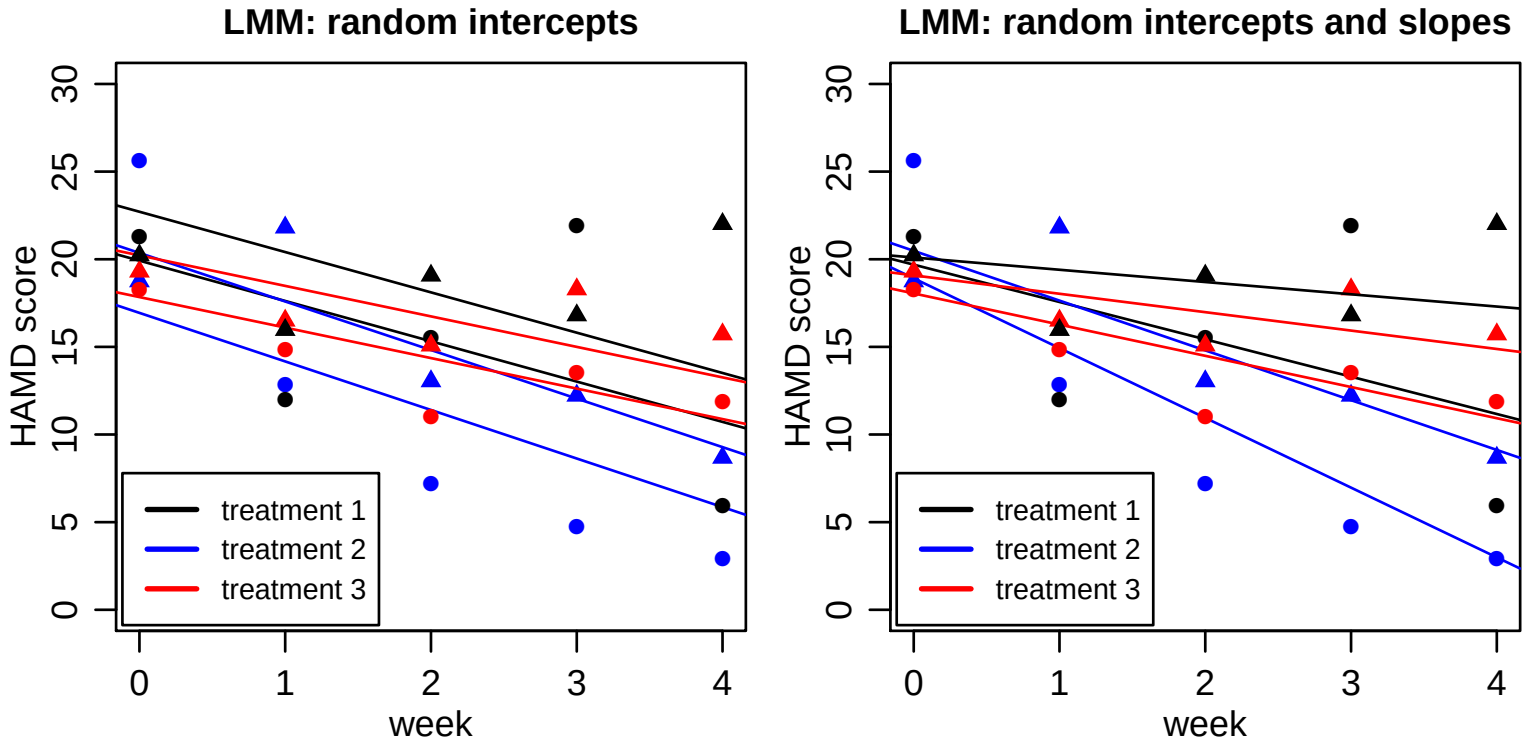
- Modify LMM to allow a separate slope for each individual:

$$\begin{aligned} y_{iw} &\sim \text{Normal}(\mu_{iw}, \sigma^2) \\ \mu_{iw} &= \alpha_i + \beta_{(treat(i),i)} w \end{aligned}$$

- As for the $\{\alpha_i\}$, assume that the $\{\beta_{(1,i)}\}$,$\{\beta_{(2,i)}\}$ & $\{\beta_{(3,i)}\}$ follow *common* prior distributions with vague priors on their *hyperparameters*

```
for (i in 1:N) {  # N individuals
  for (w in 1:W) {  # W weeks
    hamd[i,w]~dnorm(mu[i,w],tau)
    mu[i,w]<-alpha[i]+beta[treat[i],i]*(w-1)
  }
  alpha[i]~dnorm(alpha.mu,alpha.tau)
  for (t in 1:T){beta[t,i]~dnorm(beta.mu[t],beta.tau[t])}
}
# Priors
for (t in 1:T){  # T treatments
  beta.mu[t]~dnorm(0,0.00001)
  beta.sigma[t]~dunif(0,100)
  beta.sigma.sq[t]<-pow(beta.sigma[t],2)
  beta.tau[t]<-1/beta.sigma.sq[t]
}        # specification of other priors as before ....
```

HAMD Example: random intercepts and slopes



circles and triangles represent scores for 6 individuals (2 for each treatment)

- LMM with random intercepts only:
  - each individual has a different regression line
  - but for each treatment, only intercept varies by individual
- LMM with random intercepts and random slopes:
  - now intercepts and slopes both vary
  - better fit for each individual

HAMD Example: results comparison

Table: posterior mean (95% credible interval) for the non-hierarchical and hierarchical models fitted to the HAMD data

linear model			hierarchical model 1*			hierarchical model 2†	
α	19.8	(19.2,20.4)	μ_α	19.8	(19.1,20.5)	μ_α	19.8 (19.2,20.4)
			σ^2_α	17.6	(14.0,21.9)	σ^2_α	11.1 (8.4,14.4)
β_1	-2.1	(-2.4,-1.8)	β_1	-2.3	(-2.6,-2.0)	μ_{β_1}	-2.3 (-2.7,-1.9)
						$\sigma^2_{\beta_1}$	2.0 (1.2,3.0)
β_2	-2.9	(-3.2,-2.6)	β_2	-2.8	(-3.0,-2.5)	μ_{β_2}	-2.8 (-3.2,-2.4)
						$\sigma^2_{\beta_2}$	1.2 (0.5,2.0)
β_3	-1.8	(-2.1,-1.5)	β_3	-1.7	(-2.0,-1.5)	μ_{β_3}	-1.7 (-2.1,-1.4)
						$\sigma^2_{\beta_3}$	1.9 (1.1,2.9)
σ^2	35.4	(32.6,38.5)	σ^2	18.2	(16.6,19.8)	σ^2	14.4 (13.0,15.9)
p_D	5		p_D	207		p_D	314
DIC	7882		DIC	7263		DIC	7082

* random intercepts only

† random intercepts and random slopes

HAMD Example: interpretation of results

- Study objective: are there any differences in the effects of the 3 treatments on the change in HAMD score over time?
- So we are particularly interested in the differences in the slope parameters, i.e.
  - $\beta_1 - \beta_2$, $\beta_1 - \beta_3$ and $\beta_2 - \beta_3$ or
  - $\mu_{\beta_1} - \mu_{\beta_2}$, $\mu_{\beta_1} - \mu_{\beta_3}$ and $\mu_{\beta_2} - \mu_{\beta_3}$ for models with random slopes
- To monitor these contrasts, add the following lines of BUGS code

```
# Calculate contrasts
contrasts[1]<-beta[1]-beta[2]
contrasts[2]<-beta[1]-beta[3]
contrasts[3]<-beta[2]-beta[3]
```

or

```
contrasts[1]<-beta.mu[1]-beta.mu[2] ...
```

HAMD Example: contrasts

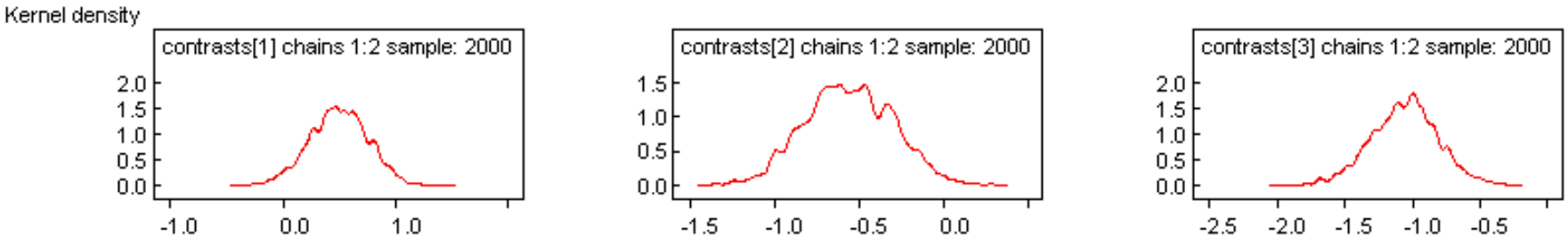
Table: posterior mean (95% credible interval) for the contrasts (treatment comparisons) from models fitted to the HAMD data

treatments	linear model		hierarchical 1*		hierarchical 2†	
1 v 2	0.8	(0.4,1.1)	0.5	(0.1,0.8)	0.5	(0.0,1.0)
1 v 3	-0.3	(-0.7,0.0)	-0.6	(-0.9,-0.2)	-0.6	(-1.1,0.0)
2 v 3	-1.1	(-1.4,-0.8)	-1.0	(-1.4,-0.7)	-1.1	(-1.6,-0.6)

* random intercepts only

† random intercepts and random slopes

Density plots for hierarchical 2



Cross-classified random effects models

- Straightforward to extend basic hierarchical model to include non-nested random effects structures, e.g.
  - THM measurements cross-classified within zones and years
  - pupils cross-classified within primary and secondary schools
- Easiest to formulate cross-classified models in BUGS using nested index notation (see example)

Example: Schools – exam scores cross-classified by primary and secondary school

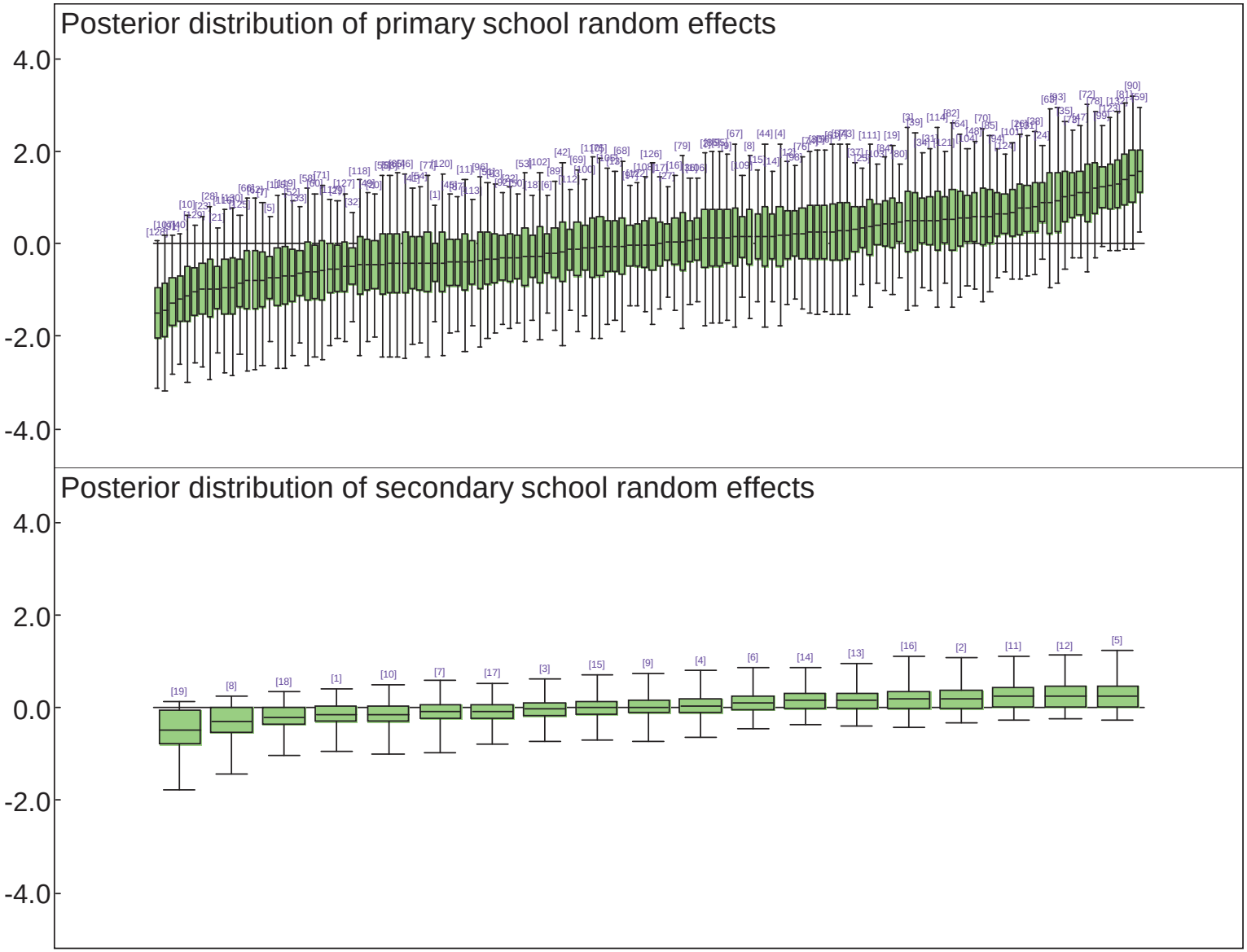
- These data were obtained from the MLwiN website `www.mlwin.com/softrev/2lev-xc.html`
- We use a random sample of 800 children who attended 132 primary schools and 19 secondary schools in Scotland
- The following variables were used
  - `Y` exam attainment score of pupils at age 16
  - `VRQ` verbal reasoning score taken on secondary school entry
  - `SEX` pupil's gender (0 = boy, 1 = girl)
  - `PID` primary school identifying code
  - `SID` secondary school identifying code
- **Model 1:** Normal hierarchical model with independent random effects for primary school and secondary school
- **Model 2:** Verbal reasoning score + gender included as 'fixed' covariate effects (but note that in Bayesian framework, 'fixed' effect coefficients are still assigned prior distributions)

BUGS model code (Model 2)

```
for(i in 1:Nobs) {
 Y[i] ~ dnorm(mu[i], tau.e)
 mu[i] <- alpha + beta[1]*SEX[i] + beta[2]*VRQ[i] +
                  theta.ps[PID[i]] + theta.ss[SID[i]]
}
### random effects distributions
 for(j in 1:Nprim) { theta.ps[j] ~ dnorm(0, tau.ps) } # primary
 for(k in 1:Nsec) { theta.ss[k] ~ dnorm(0, tau.ss) } # secondary
### priors on regression coefficients and variances
 tau.e ~ dgamma(0.001, 0.001)
 sigma2.e <- 1/tau.e                # residual error variance
 tau.ps ~ dgamma(0.001, 0.001)
 sigma2.ps <- 1/tau.ps              # between primary school var.
 tau.ss ~ dgamma(0.001, 0.001)
 sigma2.ss <- 1/tau.ss              # between secondary school var.
 alpha ~ dnorm(0, 0.000001)         # intercept
 for(q in 1:2) {beta[q] ~ dnorm(0, 0.000001)} # regression coeff.
### percentage of total variance explained
 VPC.ps <- sigma2.ps/(sigma2.e+sigma2.ps+sigma2.ss) # primary
 VPC.ss <- sigma2.ss/(sigma2.e+sigma2.ps+sigma2.ss) # secondary
```

Results

Parameters	Model 1		Model 2	
α	5.53	(5.17, 5.88)	5.85	(5.59, 6.10)
β_1 (sex)	–	–	0.23	(-0.08, 0.53)
β_2 (VRQ)	–	–	0.16	(0.15, 0.17)
$\sigma^2_{[e]}$	8.18	(7.35, 9.10)	4.49	(4.03, 5.00)
$\sigma^2_{[ps]}$	1.12	(0.43, 1.98)	0.36	(0.08, 0.70)
$\sigma^2_{[ss]}$	0.19	(0.10, 0.82)	0.02	(0.0007, 0.12)
VPC_{ps}	11.8%	(4.7%, 19.8%)	7.4%	(1.5%, 13.8%)
VPC_{ss}	2.0%	(0.1%, 8.3%)	0.4%	(0.01%, 2.4%)
DIC	4008		3514	
p_D	58.0		43.8	


Posterior distribution of primary school random effects
4.0
2.0
0.0
-2.0
-4.0
Posterior distribution of secondary school random effects
4.0
2.0
0.0
-2.0
-4.0
[19]
[8]
[18]
[1]
[10]
[7]
[17]
[3]
[15]
[9]
[4]
[6]
[14]
[13]
[16]
[2]
[11]
[12]
[5]