1. Run the BUGS model(s) of example 'Seeds' from WinBUGS manual, examples volume I. Compute the same model using hierarchical centering as explained in the example.

2. Below is a Poisson log-linear model for the Danish lung cancer data with STZ constraints. Compute posterior estimates for the intercept which now represents overall mean. Then replace the age and city main effects ('fixed/constant effects') by random ('varying') effects $u_i \sim \mathrm{N}(0, \sigma_a^2)$ and $v_j \sim \mathrm{N}(0, \sigma_c^2)$ ($i = 1, \ldots, 6, j = 1, \ldots, 4$). Priors for variance components could be $\sigma_a \sim \mathrm{U}(0, 1000)$ and $\sigma_c \sim \mathrm{U}(0, 1000)$, and a flat prior for $\mu \sim \mathrm{N}(0, 0.001)$. The means are then $\log(\lambda_k) = \mu + u_{i_k} + v_{j_k}$, $k = 1, \ldots, 24$. This is identifiable since σ_a and σ_c will not allow infinitely large positive or negative random effects. (In comparison, the main effects had basically flat prior). Parameter μ represents the overall mean in random effects model, comparable to the intercept in STZ ANOVA model. These random effects are unstructured. Construct structured random effects for age groups by assuming $u_{j+1} \sim \mathrm{N}(u_j, \sigma_a^2)$. From the random effects model we could generate predictions for 'another Danish city' by generating $v^* \sim \mathrm{N}(0, \sigma_c^2)$. And if we had completely missing age group in all cities, this could be predicted by generating u^* either as an unstructured random effect (describing variation between age groups) or as a structured random effect (describing variation between successive age groups). The latter would draw information both from the previous and the next age group since u_i^* would depend on the previous u_{i-1} whereas the next u_{i+1} would depend on u_i^*.

`http://www.sci.usq.edu.au/staff/dunn/Datasets/glms/poisson/danishlc.html`

```
model{
for(i in 1:24){
cases[i] ~ dpois(mu[i]); group[i] <- i
mu[i] <- pop[i]*4*lambda[i] # lambda = incidence per year
LA[i] <- lambda[i]/100000 # LA = incidence per 100000 per year
log(lambda[i]) <- inprod(alpha[],X[i,])
X[i,1] <- 1
for(k in 2:6){X[i,k] <- equals(age[i],k)-equals(age[i],1)} # STZ
for(k in 2:4){X[i,k+5] <- equals(city[i],k)-equals(age[i],1)}# STZ
}
for(k in 1:9){alpha[k] ~  dnorm(0,0.001);A[k] <- exp(alpha[k])}
}
# inits:
list(alpha=c(0,0,0,0,0,0,0,0,0))
```

```
cases[] pop[]   age[]   city[]
11      3059    1       1
11      800     2       1
11      710     3       1
10      581     4       1
11      509     5       1
10      605     6       1
13      2879    1       2
6       1083    2       2
```

```
15      923     3       2
10      834     4       2
12      634     5       2
2       782     6       2
4       3142    1       3
8       1050    2       3
7       895     3       3
11      702     4       3
9       535     5       3
12      659     6       3
5       2520    1       4
7       878     2       4
10      839     3       4
14      631     4       4
8       539     5       4
7       619     6       4
END
```

3. Assume the true disease incidence varies over a number of geographical regions i, e.g. $i = 1, \ldots, 100$. To make your own data, generate these true values λ_i from Gamma$(5, 1000)$. Then, for each generated λ_i, generate the actual disease count X_i for each region from Poisson$(N_i\lambda_i)$, assuming the exposure (population count) is $N_i = i \times 10$. Now, you know the 'true values', and you have 'observed data'. Compute the observed incidences $\hat{\lambda}_i = X_i/N_i$ and plot them as a function of population size. Observe how these behave when N_i is small. Compute the posterior distribution of all λ_i assuming the prior $\lambda_i \sim$ Gamma(α, β), $\alpha \sim$ Exp(0.01), $\beta \sim$ Exp(0.01). (You might also try the log-linear model, employing normal prior density). Compute the posterior medians and 95% CIs and compare them with the observed incidences, $\hat{\lambda}_i$, as a function of population size. In similar applications, it might be interesting to study urban-rural effects. The 'true incidence' λ_i might then be a function of population size. (Simulated data could be done for this too). When estimating $\log(\lambda_i)$ we could then assume a prior $\log(\lambda_{i+1}) \sim \mathrm{N}(\log(\lambda_i), \sigma^2)$ to incorporate smoothness over urban-rural continuum. This would help estimating λ_i also when the corresponding X_i might be missing.

```
model{
for(i in 1:100){
N[i] <- i*10
x[i] ~ dpois(par[i]);  xpred[i] ~ dpois(par[i])
par[i] <- lambda[i]*N[i]
lambda[i] ~ dgamma(a,b)
laobs[i] <- x[i]/N[i]
}
a ~ dexp(0.01); b~dexp(0.01)
}
list(x=c(0,0,3,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,3,1,1,0,1,
0,0,2,0,1,3,0,3,3,2,3,2,1,1,5,3,2,4,2,4,3,5,3,3,2,2,0,4,
2,2,2,3,2,2,7,5,1,1,0,3,3,3,3,10,4,0,5,6,1,2,1,6,2,3,5,
 3,1,1,7,8,7,2,6,9,4,3,3,3,3,7,8,10,4,2,8,2,3))
#,lambdatrue=c(0.006189,0.005392,0.007482,0.002101,0.00404,0.004104,
0.002443,0.003522,0.00304,0.003879,0.006207,0.005146,
0.007665,0.006171,0.006467,0.004278,0.004247,0.003849,
0.006132,0.006311,0.004918,0.007915,0.006253,0.0028,
0.00269,0.002432,0.004463,0.00487,0.004416,0.005386,
0.003095,0.004779,0.003521,0.003467,0.005463,0.002086,
0.001188,0.004572,0.006427,0.007555,0.006153,0.002609,
0.005903,0.005654,0.008799,0.007592,0.004678,0.003258,
0.003312,0.002201,0.005663,0.002479,0.004547,0.00375,
0.003309,0.002688,0.006507,0.009398,0.01047,0.006518,
0.001765,0.002879,0.004183,0.003772,0.003067,0.002954,
0.01264,0.006287,0.002936,0.006267,0.00689,0.003587,
0.002659,0.003542,0.004375,0.001684,0.003801,0.006977,
0.002438,0.001452,0.003746,0.007303,0.01207,0.003909,
0.002182,0.006481,0.006271,0.004222,0.006121,0.002793,
0.003799,0.004475,0.008524,0.007171,0.006329,0.003326,
0.001944,0.006952,0.003755,0.003938) )
```

Note that the easiest way to generate the data is to use R or other similar program and then import it to WinBUGS. (Alternatively, define just the data model with your given parameters in WinBUGS, run for some iterations, and select the values from some single iteration step, using 'coda'-button, and copy-paste).

4. Compute the Eyes example from WinBUGS examples Vol II. The data are also given below. Since the model assumes a mixture of two distributions, take a look at the data histogram first to see if this is plausible model. In R you can do this by `hist(x)` or by computing a nonparametric density estimate `plot(density(x))`. Check also the suggested other parametrization, and the possibility of having no members in the second component distribution.

```
x<-c(29.0, 30.0, 32.0, 33.1, 33.4, 33.6, 33.7, 34.1, 34.8,
    35.3, 35.4, 35.9, 36.1, 36.3, 36.4, 36.6, 37.0, 37.4, 37.5,
    38.3, 38.5, 38.6, 39.4, 39.6, 40.4, 40.8, 42.0, 42.8, 43.0,
    43.5, 43.8, 43.9, 45.3, 46.2, 48.8, 48.7, 48.9, 49.0, 49.4,
    49.9, 50.6, 51.2, 51.4, 51.5, 51.6, 52.8, 52.9, 53.2)
```