FINAL PROJECT REPORT

A Safety Signal Analysis with Three-Level Hierarchical Mixture Model

Guanlin Zhang*

[1]Department of Biostatistics, University of Kansas Medical Center, Kansas City

Correspondence
*Email: gzhang2@kumc.edu

In this manuscript we follow the footsteps of others [1] and fit a three-level hierarchical mixture model into a dataset featuring adverse effects of trials across different body systems, and make inference on the comparison of treatment and control groups.

KEYWORDS:
Safety Signal, Three-Level Hierarchical model, Adverse Effects(AEs), Body Systems, Multiplicities;

1 | INTRODUCTION

To our example we are going to apply a three-level hierarchical mixture model, proposed by Berry & Berry [1]. The reference literature was initially produced in order to address the multiplicity issues when considering potentially drug-caused adverse effects across and within different body systems.

The idea is to model different adverse effects during drug trials both across and within the body systems with a Bayesian approach through a three-level hierarchical model, that way we could make posteior inferences on the AEs simultaneously without losing power as the frequentists perspecitve would on this matter. For details we refer to our main reference [1].

Here we are just going to introduce the model proposed by Berry&Berry.

2 | MODEL

Assuming in a drug trial we have control (total number N_c) and treatment (total number N_t) groups. Suppose there are B body systems and within body system b there are k_b types of AEs labeled A_{bj}, where $b = 1, \dots, B$ and $j = 1 \dots, k_b$. Of the N_c controls,

X_{bj} experience A_{bj} and of the N_T patients in the treatment group, Y_{bj} experience A_{bj}. The probabilities of experiencing A_{bj} are c_{bj} and t_{bj} for control and treatment patients, respectively [1]. We use logistic transformations:

$$\gamma_{bj} = \log\left(\frac{c_{bj}}{1-c_{bj}}\right) \text{ and } \theta_{bj} = \log\left(\frac{t_{bj}}{1-t_{bj}}\right) - \gamma_{bj}$$

So given the data X_{bj} and Y_{bj}, there is:

$$X_{bj} \sim \text{Bin}(c_{bj}, N_c) \text{ and } Y_{bj} \sim \text{Bin}(t_{bj}, N_t)$$

Here γ_{bj} and θ_{bj} are level 1 prior parameters and we assume:

$$\gamma_{bj} \sim N(\mu_{\gamma b}, \sigma^2_{\gamma}) \qquad \text{for } b = 1, \dots, B \text{ and } j = 1, \dots, k_b.$$

For θ_{bj} it represents the log odds ratio of AEs in treatment group versus in control group. It is modeled with a mixture prior distribution that

$$\theta_{bj} \sim \pi_b I_{[0]} + (1-\pi_b) N(\mu_{\theta b}, \sigma^2_{\theta b}) \qquad \text{for } b = 1, \dots, B; j = 1, \dots, k_b.$$

The above expression can also be equivalently phrased as:

$$\theta_{bj} = (1 - \text{mix}_{1,bj}) \times \text{mix}_{2,bj}$$

with

$$\text{mix}_{1,bj} \sim \text{Bernoulli}(\pi_b) \text{ and } \text{mix}_{2,bj} \sim N(\mu_{\theta b}, \sigma^2_{\theta b})$$

We take this form in our OpenBUGS coding later on.

Now for the second level prior parameters we have $\mu_{\gamma b}$, σ^2_{γ}, π_b, $\mu_{\theta b}$, and $\sigma^2_{\theta b}$. Assume:

$$\mu_{\gamma b} \sim N(\mu_{\gamma 0}, \tau^2_{\gamma 0}) \text{ for } b = 1, \dots, B \text{ and } \sigma^2_{\gamma} \sim \text{IG}(\alpha_{\sigma\gamma}, \beta_{\sigma,\gamma})$$

we also have

$$\pi_b \sim \text{Beta}(\alpha_\pi, \beta_\pi), b = 1, \dots, B$$

Here the π_b depends on the body systems, so the model considers different probabilities of having same AE effect between control and treatment group in different body systems.

We also have

$$\mu_{\theta b} \sim N(\mu_{\theta 0}, \tau^2_{\theta 0}) \text{ for } b = 1, \dots, B \text{ and } \sigma^2_{\theta b} \sim \text{IG}(\alpha_\theta, \beta_\theta).$$

Finally for the third level prior parameters, we have $\mu_{\gamma 0}$, $\tau^2_{\gamma 0}$, α_π, β_π, $\mu_{\theta 0}$ and $\tau^2_{\theta 0}$. We assume:

$$\mu_{\gamma 0} \sim N(\mu_{\gamma 00}, \tau^2_{\gamma 00}) \text{ and } \tau^2_{\gamma 0} \sim \text{IG}(\alpha_{\tau\gamma}, \beta_{\tau\gamma}).$$

For α_π and β_π as the hyper parameter of π_b in the beta distribution, we restrict them to be \$ > 1\$ and assume a left truncated exponential distribution:

$$\alpha_\pi \sim \frac{\lambda_\alpha \exp(-\alpha\lambda_\alpha)}{\exp(-\lambda_\alpha)} I_{[\alpha>1]} \text{ and } \beta_\pi \sim \frac{\lambda_\beta \exp(-\beta\lambda_\beta)}{\exp(-\lambda_\beta)} I_{[\beta>1]}$$

The reason of doing so is explained in detail, see (Berry & Berry 2004)[1].

For constants in the model it is assumed as $\mu_{\theta 00} = 0, \tau^2_{\theta 00} = 10$, $\alpha_\theta = 3$, $\beta_\theta = 1$, $\alpha_{\theta 0} = 3$, $\beta_{theta0} = 1$, $\alpha_{\tau\gamma} = 3$, $\beta_{\tau\gamma} = 1$, $\alpha_{\sigma\gamma} = 3$, $\beta_{\sigma\gamma} = 1$ and $\lambda_\gamma = \lambda_\beta = 1$.

Figure 1 provides an OpenBUGS doodle that helps with seeing the hierarchical structure. The variables inside eclipses are either stochastic or logical node, and those in rectangles are constants.

The notation we use in OpenBUGs is consistent with our reference literature, except that to avoid transform between precision and variance, everywhere where there is notation for variance in the paper, we use it as precision in our code. For example, when the paper mention $\tau^2_{00} = 10$, we put it as $\tau^2_{0,0} = 0.1$ in the OpenBUGs code

3 | ANALYSIS

We proceed the analysis with a few steps. We first fit the data presented by our reference[1] in OpenBUGs. In the meantime, to check consistency, we also utilize an R package to do the same job, and we compare between both results, and also with those

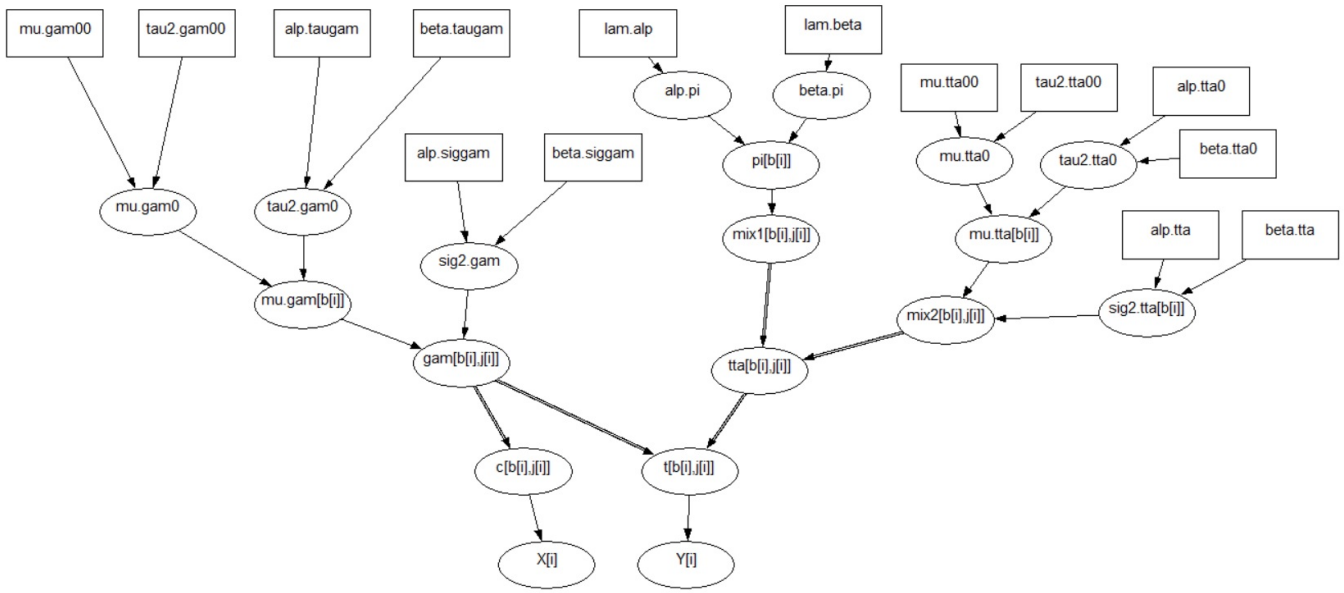

FIGURE 1 OpenBUGS doodle for the hierarchical model

presented by the paper. Then we fit the three-level hierarchical model to a different set of data in winBUGS, and we compare the results with an independent model.

3.1 | Fit data from Berry(2004)[1] by WinBUGS and R separately

The reference paper[1] present the data as shown in Figure 2 :

and the results are partially presented in the paper (we are mainly interested in the posterior $\theta > 0$ probability since it reflects drug effects on AEs) as shown in Figure 3

The original authors are particularly interested in the 4 AEs that are significant in the Fisher-exact test since we want to know how the Bayesian hierarchical model concludes differently. Table 1 presents the results:

AEs	Index	$\theta > 0$	Fisher exact p
Irritability	$\theta_{8,3}$	0.780	0.003
Diarrhea	$\theta_{3,4}$	0.231	0.029
Rash	$\theta_{10,4}$	0.19	0.021
Rash, measles /rub-like	$\theta_{10,6}$	0.126	0.039

TABLE 1 Fisher exact test vs Posterior probability

The key argument made by Berry & Berry is that the order of Fisher exact p values on these four AEs are not the same as the order of the posterior probability $\theta > 0$, and that is due to the reason that Fisher exact test only focus on individual AE while the hierarchicao model takes into account the body system factor.

			Treatment ($N_T = 148$)		Control ($N_C = 132$)		
b	j	Type of AE A_{bj}	Y_{bj}	Rate	X_{bj}	Rate	Fisher's exact p
1	1	Astenia/fatigue	57	0.385	40	0.303	0.167
1	2	Fever	34	0.230	26	0.197	0.561
1	3	Infection, fungal	2	0.014	0	0.000	0.500
1	4	Infection, viral	3	0.020	1	0.008	0.625
1	5	Malaise	27	0.182	20	0.152	0.525
3	1	Anorexia	7	0.047	2	0.015	0.179
3	2	Cendisiasis, oral	2	0.014	0	0.000	0.500
3	3	Constipation	2	0.014	0	0.000	0.500
3	4	Diarrhea	24	0.162	10	0.076	0.029*
3	5	Gastroenteritis	3	0.020	1	0.008	0.625
3	6	Nausea	2	0.014	7	0.053	0.089
3	7	Vomiting	19	0.128	19	0.144	0.730
5	1	Lymphadenopathy	3	0.020	2	0.015	1.000
6	1	Dehydration	0	0.000	2	0.015	0.221
8	1	Crying	2	0.014	0	0.000	0.500
8	2	Insomnia	2	0.014	2	0.015	1.000
8	3	Irritability	75	0.507	43	0.326	0.003*
9	1	Bronchitis	4	0.027	1	0.008	0.375
9	2	Congestion, nasal	4	0.027	2	0.015	0.375
9	3	Congestion, respiratory	1	0.007	2	0.015	0.603
9	4	Cough	13	0.088	8	0.061	0.497
9	5	Infection, upper respiratory	28	0.189	20	0.152	0.431
9	6	Laryngotracheobronchitis	2	0.014	1	0.008	1.000
9	7	Pharyngitis	13	0.088	8	0.061	0.497
9	8	Rhinorrhea	15	0.101	14	0.106	1.000
9	9	Sinusitis	3	0.020	1	0.008	0.625
9	10	Tonsillitis	2	0.014	1	0.008	1.000
9	11	Wheezing	3	0.020	1	0.008	0.625
10	1	Bite/sting	4	0.027	0	0.000	0.125
10	2	Eczema	2	0.014	0	0.000	0.500
10	3	Pruritis	2	0.014	1	0.008	1.000
10	4	Rash	13	0.088	3	0.023	0.021*
10	5	Rash, diaper	6	0.041	2	0.015	0.288
10	6	Rash, measles/rubella-like	8	0.054	1	0.008	0.039*
10	7	Rash, varicella-like	4	0.027	2	0.015	0.687
10	8	Urticaria	0	0.000	2	0.015	0.221
10	9	Viral exanthema	1	0.007	2	0.015	0.603
11	1	Conjunctivitis	0	0.000	2	0.015	0.221
11	2	Otitis media	18	0.122	14	0.106	0.711
11	3	Otorrhea	2	0.014	1	0.008	1.000

FIGURE 2 Data from Berry & Berry's paper

While I agree with the authors, our OpenBugs and R both return similar results to each other but different than those presented in the paper.

Particularly, we did our own OpenBugs code based on the model, while the R result actually comes from a package named "c212", which is developed by Dr Raymond Carragher from Strathclyde Institute of Pharmacy and Biomedical Sciences.

For OpenBugs we burned $1,000$ iterations and take the next $10,000$ updates, then we extract the coda information of θ and compute the mean number of θs of interest that is positive to estimate the posterior $P(\theta > 0)$. For R the package will also generate point estimate for θ and we compute the posterior $P(\theta > 0)$ in the same way.

Figure 4 and Figure 5 shows the trace and history plot from OpenBUGS iteration. Graphs show that our model converges all right.

Table 2 presents the comparison among the three for the four AEs:

			Post probability					Post probability	
b	j	Type of AE	$\theta = 0$	$\theta > 0$	b	j	Type of AE	$\theta = 0$	$\theta > 0$
1	1	Astenia/fatigue	0.762	0.211	9	4	Cough	0.906	0.062
1	2	Fever	0.827	0.122	9	5	Infection, respiratory	0.897	0.083
1	3	Infection, fungal	0.796	0.101	9	6	Bronchitis	0.898	0.047
1	4	Infection, viral	0.813	0.100	9	7	Pharyngitis	0.906	0.061
1	5	Malaise	0.826	0.116	9	8	Rhinorrhea	0.904	0.051
3	1	Anorexia	0.821	0.117	9	9	Sinusitis	0.903	0.051
3	2	Cendisiasis, oral	0.835	0.083	9	10	Tonsillitis	0.905	0.042
3	3	Constipation	0.812	0.101	9	11	Wheezing	0.907	0.050
3	4	**Diarrhea**	0.743	**0.231**	10	1	Bite/sting	0.859	0.087
3	5	Gastroenteritis	0.823	0.093	10	2	Eczema	0.860	0.070
3	6	Nausea	0.805	0.050	10	3	Pruritis	0.868	0.062
3	7	Vomiting	0.849	0.076	10	4	**Rash**	0.784	**0.190**
5	1	Lymphadenopathy	0.717	0.136	10	5	Rash, diaper	0.852	0.099
6	1	Dehydration	0.666	0.087	10	6	**Rash, measles/rub-like**	0.836	**0.126**
8	1	Crying	0.655	0.185	10	7	Rash, varicella-like	0.862	0.076
8	2	Insomnia	0.661	0.153	10	8	Urticaria	0.852	0.048
8	3	**Irritability**	0.214	**0.780**	10	9	Viral exanthema	0.855	0.055
9	1	Bronchitis	0.900	0.059	11	1	Conjunctivitis	0.721	0.079
9	2	Congestion, nasal	0.901	0.058	11	2	Otitis media	0.757	0.102
9	3	Congestion, respiratory	0.896	0.040	11	3	Otorrhea	0.749	0.121

FIGURE 3 Berry & Berry's paper results

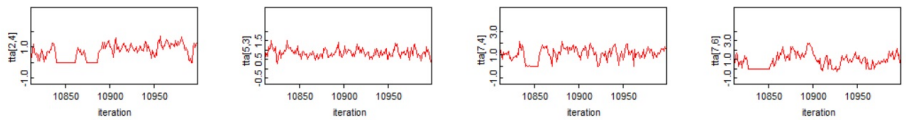

FIGURE 4 Trace Plot from OpenBugs

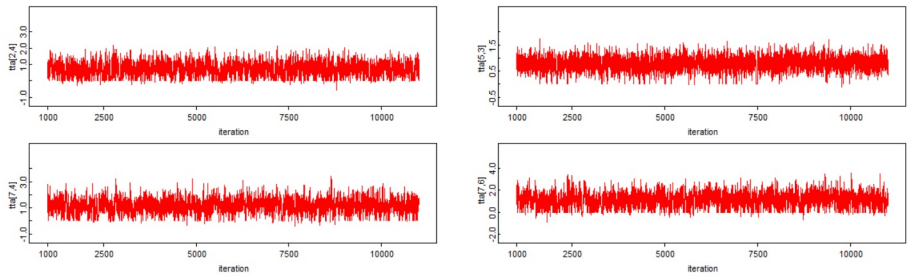

FIGURE 5 Trace Plot from OpenBugs

From table 2 we can conclude the following:

1. OpenBUGs and R return similar estiamtes on posterior probability of treatment effects on AEs, but different than the original reference paper. Our analysis suggests that for all four adverse effects, there is a significant effect from the treatment (high probability of $\theta > 0$).

2. Posterior probability of treatment effect on irritability is strong, consistent with Fisher exact test.In the mean time, due to body system effect the extent of treatment effect for these adverse effects are not completely in the same order as the p values from Fisher exact tests, although the difference is nearly ignorable.

We argue that since the original paper was published in 2004 and over the time the computational software has evolved tremendously, the result from the original paper may be subject to further verification.

		post $P(\theta > 0)$			
AEs	Type	Reference	OpenBugs	R	Fisher exact p
Irritability	$\theta_{8,3}$	0.78	0.978	0.984	0.003
Diarrhea	$\theta_{3,4}$	0.231	0.847	0.853	0.029
Rash	$\theta_{10,4}$	0.19	0.945	0.993	0.021
Rash, measles/rub-like	$\theta_{10,6}$	0.126	0.890	0.946	0.039

TABLE 2 OpenBugs and R results compare with reference

3.2 | Fit the new example with both OpenBUGs and R and compare with an independent model

Figure 6 presents a new set of data, similar to the one from our main reference[1]:

			Treatment (N_T=320)		Control (N_C=320)	
b	j	Type of SAE A_{bj}	Y_{bj}	Rate	X_{bj}	Rate
1	1	Arrhythmia	0	0.000	1	0.003
1	2	Increased BP	3	0.009	3	0.009
1	3	Other CB AEs	0	0.000	2	0.006
1	4	Pre-eclampsia	8	0.025	6	0.019
2	1	Emesis	0	0.000	1	0.003
2	2	Other GI AEs	0	0.000	1	0.003
3	1	Depression	0	0.000	1	0.003
3	2	Headache	0	0.000	1	0.003
3	3	Other HNMB AEs	1	0.003	0	0.000
4	1	Gestational Diabetes Mellitus	2	0.006	0	0.000
4	2	Other MAN AEs	0	0.000	1	0.003
5	1	Chorioamnionitis	0	0.000	2	0.006
5	2	Decreased Fetal Movement	1	0.003	2	0.006
5	3	Endomyometritis	1	0.003	0	0.000
5	4	Miscarriage	0	0.000	1	0.003
5	5	Other PD AEs	3	0.009	1	0.003
5	6	Postpartum Hemorrhage	1	0.003	0	0.000
5	7	Premature Delivery	5	0.016	5	0.016
5	8	Premature ROM	2	0.006	10	0.031
5	9	Preterm Contractions	2	0.006	4	0.013
6	1	Other RESP AEs	0	0.000	1	0.003
6	2	Shortness of Breath	0	0.000	1	0.003
7	1	Other UG AEs	1	0.003	0	0.000
7	2	Pyelnephritis	1	0.003	4	0.013
7	3	Urinary Tract Infection	1	0.003	1	0.003
7	4	Vaginal Bleeding	1	0.003	1	0.003
8	1	Abdominal Pain	1	0.003	2	0.006
8	2	Other BODY AEs	0	0.000	3	0.009
8	3	PD012	2	0.006	0	0.000
8	4	Pelvic Pain	1	0.003	0	0.000
8	5	Polyhydramnios	0	0.000	1	0.003

FIGURE 6 Safety Signal Data

We fit the data with winBUGS and R. For the winBUGS part this time we also fit an independent model, which means that every γ_{bj} and θ_{bj} has its own normal prior distribution. Figure 7 shows an OpenBUGS doodle version of the model:

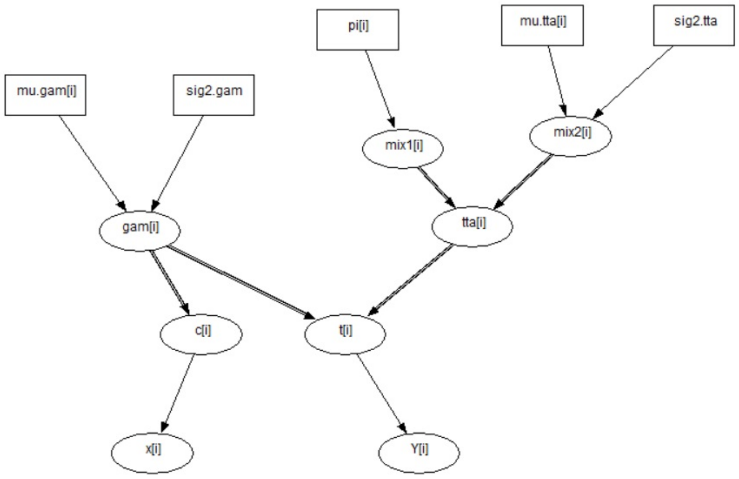

FIGURE 7 OpenBugs doodle for Independent Model

Table 3 compares the posterior probability $P(\theta > 0)$ between winBUGS results for hierarchical and independetn model, and also the R package output for the hierarchical model.

From Table 3 we conclude that when using hierarchical model we tends to be more conservative about claiming that there is a treatment effect on the adverse effects. Both OpenBUGS and R gave similar results to the posterior probability $P(\theta > 0)$.

However the independent model gives in general larger values of this probability, if we focus on comparing row by row. When using independent model, we are considering each AE without taking into consideration of body system, so this is essentially in the same spirit of making inference with fisher exact test on each individual adverse effect.

Take some observations for example, for "Other HNMB AEs", the hierarchical models give a posteior probability fairly small (0.046 from R and 0.053 from OpenBUGS), while the independent model gives a value up to 0.20.

Should we claim a treatment effect on this particular AE? In this case the posteior probability is pretty small even for independent model. So the answer is probably no.

Suppose the posterior probability is quite large for independent model, say pass over 0.5, but the value given by hierarchical model is pretty small, then we need to think over before making a decision. What seems to be a treatment effect might actually be caused by the body system itself.

4 | CONCLUSION

We have explored the three-level hierarchical model proposed by Berry & Berry, and tried to fit different datasets into the model. Both R and OpenBUGS give similar outcomes when it comes to estimating the posterior probabilities.

Although we did not show for every dataset in every model, the model generally converges quite well. (The R package does not provide convergence check).

	post $P(\theta > 0)$		
AE	R(Hierarchical)	OpenBUGS(Hierarchical)	OpenBUGS(Independent)
Arrhythmia	0.0207	0.0267	0.0496
Increased BP	0.0487	0.0591	0.0939
Other CB AEs	0.0136	0.0237	0.0227
Pre-eclampsia	0.0763	0.1102	0.1058
Emesis	0.0258	0.0264	0.0457
Other GI AEs	0.022	0.0249	0.0463
Depression	0.0229	0.0275	0.0464
Headache	0.0232	0.0301	0.0508
Other HNMB AEs	0.046	0.0529	0.2012
Gestational Diabetes Mellitus	0.0763	0.0961	0.2920
Other MAN AEs	0.0331	0.0341	0.0476
Chorioamnionitis	0.0061	0.0156	0.0240
Decreased Fetal Movement	0.0199	0.0281	0.0661
Endomyometritis	0.037	0.0413	0.1984
Miscarriage	0.0062	0.0194	0.0493
Other PD AEs	0.0634	0.0718	0.2238
Postpartum Hemorrhage	0.0309	0.0406	0.1871
Premature Delivery	0.0486	0.0671	0.0708
Premature ROM	0.0059	0.0109	0.0045
Preterm Contractions	0.0186	0.0331	0.0434
Other RESP AEs	0.0113	0.0232	0.0472
Shortness of Breath	0.0105	0.0251	0.0474
Other UG AEs	0.0466	0.0508	0.1887
Pyelnephritis	0.0254	0.0313	0.0272
Urinary Tract Infection	0.0466	0.0452	0.1047
Vaginal Bleeding	0.0402	0.0473	0.1046
Abdominal Pain	0.0335	0.0434	0.0616
Other BODY AEs	0.0125	0.0213	0.0118
PD012	0.0502	0.074	0.2885
Pelvic Pain	0.0414	0.0542	0.1897
Polyhydramnios	0.0198	0.0268	0.0483

TABLE 3 hierarchical model vs independent model, and winBUGS vs R

In the mean time, we particularly compared the results between fitting the hierarchical model and the independent model. We conclude that the hierarchical model is more conservative when it comes to attributing the adverse effects to the effect of treatment, due to the consideration of body system effect.

Finally, it seems that we could not reproduce the exact same estimate on the data provided in the reference paper. However both our OpenBUGS and R program return similar results to each other. We argue that the difference than the original publication might due to the change of computational software over a time period of more than a decade.

References

1. Berry Scott M., Berry Donald A.. Accounting for Multiplicities in Assessing Drug Safety: A Three-Level Hierarchical Mixture Model. *Biometrics.* 2004;:418-426.