Brazilian Journal of Probability and Statistics (2001), 15, pp. 33–47.
©Associação Brasileira de Estatística

META ANALYSIS: A BAYESIAN APPROACH

Jorge Alberto Achcar[1], **Viviane Carla Fortulan**[1] and **Josmar Mazucheli**[2]

[1] *Universidade de São Paulo, ICMC, Caixa Postal 668, 13560-970, São Carlos, SP, Brazil.*
E-mails: jorge@icmc.sc.usp.br and vivifor@zaz.com.br

[2] *Departamento de Estatística, Universidade Estadual de Maringá, PR, Brazil, Graduate student at COPPE/UFRJ/PO.*
E-mail: josmar@uem.br

Summary
In this paper, we consider the use of Bayesian methods in meta-analysis using Markov Chain Monte Carlo (MCMC) methods. Usually the proposed models considered for meta- analysis assume many parameters and the Bayesian approach with MCMC is a reasonable alternative to combine information from independent studies. Some numerical illustrations are considered with medical data from different studies.

Key Words: Bayesian analysis; MCMC methods; meta-analysis.

1 Introduction

Meta-analysis refers to statistical methods for combining results from independent studies in order to draw overall conclusions. This approach is useful when the number of studies on some research is small, and usually these studies have small sample sizes.

The term meta-analysis was introduced by Glass(1976) as the analysis of the results of statistical analyses for the purpose of drawing general conclusions. Since the introduction of the term, the use of statistical methods to combine the results of replicated research studies has become widespread in education, psychology, and the biomedical sciences (see for example, Hedges, 1992).

Meta-analysis has been used by social researchers for many years (see for example, Rosenthal, 1978; Glass, McGraw and Smith, 1981; Hedges,

1983, 1992; Hunter, Schmidt and Jackson, 1982; or Light and Pillemer, 1984).

In medical research, each year, over 8000 clinical trials are performed only in the United States (See DuMouchel, 1996). Some early applications of meta-analysis to combine results from different medical studies are given by L'Abbe, Detsky, O'Rourke, (1987); Sacks et al, (1988) and Bulpitt, (1988).

In this paper, we consider the use of Bayesian hierarchical models in meta-analysis. Bayesian formulation offers a natural mechanism for describing and explaining heterogeneity between individual studies, as well as between groups of studies, or double- blind and open studies. Bayesian methods also offer a convenient framework for incorporating available prior information to improve the precision of inference.

Bayesian calculations are carried out using Markov Chain Monte Carlo (MCMC) methods such as Gibbs Sampling (see for example, Gelfand and Smith, 1990) and Metropolis-Hastings algorithm (see for example, Roberts and Smith, 1993).

The use of MCMC methods has been used by many statisticians in meta-analysis (see for example, Larose and Dey, 1997).

An outline of the paper is as follows. In Section 2, we consider a Bayesian reanalysis of a data set introduced by Van Belle (1992) assuming a modified form for the Bayesian hierarchical model considered by Larose and Dey (1997). In Section 3, we consider a reanalysis of two data sets related to the use of estrogen and the incidence of breast and endometrium cancers assuming a different model as considered by Larose and Dey (1995, 1997). Section 4 gives some conclusions.

2 Use of Bayesian methods: an application with a new antiepileptic drug

Van Belle (1992) considered fifteen independent reports of clinical trials of progabide, a new antiepileptic drug. The response variable was defined to be the number y_i of improved patients in each study considering n_i patients. The aims of this study is to verify the efficacy of progabide considering the same drug in two distinct groups: "open" studies that permits investigator knowledge of the treatment regime, while "closed" studies are double-blind controlled clinical trials. The data for the N equals to 15 studies is presented in Table 1.

In Table 1, we have the lists of study number, the study number from the original set, the total number (n_i) of patients, the number of patients improved (y_i), the observed proportion of patients improved $\hat{p}_i$ and the logit of the observed proportion improved $\hat{\theta}_i = \log[\hat{p}_i/(1-\hat{p}_i)]$. The first six were open studies, and the last nine were closed studies.

Larose and Dey (1997), considered for a meta-analysis of the data of Table 1, a Bayesian hierarchical model given by,

$$\begin{aligned} y_i &\sim b(n_i, p_i), \quad i = 1, 2, \ldots, N; \\ \theta_i &\sim N(\mu + \delta_k, \sigma_k^2), \quad k = 1, 2; \\ \sigma_k^2 &\sim IG(a_k, b_k), \quad k = 1, 2; \ a_k \text{ and } b_k \text{ known}; \\ \mu &\sim N(0, c); \ c \text{ known}; \\ \delta_k &\sim N(0, d); \ k = 1, 2; \ d \text{ known}, \end{aligned} \tag{2.1}$$

Table 1
Data from 15 independent studies

	Study	Original	n_i	y_i	$\hat{p}_i$	$\hat{\theta}_i$
Open studies	1	1	30	17	0.5666	0.2680
	2	7	16	8	0.5000	0.0000
	3	9	69	41	0.5942	0.3814
	4	10	23	13	0.5652	0.2623
	5	11	42	32	0.7619	1.1631
	6	15	151	90	0.5960	0.3888
Closed studies	7	2	20	5	0.2500	-1.0986
	8	3	20	9	0.4500	-0.2007
	9	4	17	3	0.1765	-1.5402
	10	5	15	7	0.4667	-1.1334
	11	6	18	8	0.4444	-0.2233
	12	8	17	9	0.5294	-0.1177
	13	12	19	1	0.0526	-2.8910
	14	13	51	12	0.2353	-1.1786
	15	14	59	17	0.2881	-0.9046

where $b(n_i, p_i)$ denotes a binomial distribution, $N(\mu, \sigma^2)$ denotes a normal distribution and $IG(a, b)$ denotes an inverse gamma distribution with parameters a and b. They also assume prior independence among the parameters; y_i is the number of patients improved for study i, n_i is the number of patients who completely study i, p_i is the parameter of percentage improved for study i, and $\theta_i = \log[p_i/(1-p_i)]$, and $k = 1$ for the open studies, and $k = 2$ for the closed studies.

The likelihood function for $p_1, \ldots, p_N$ is given by

$$L(p_1, \ldots, p_N) = \prod_{i=1}^{N} \binom{n_i}{y_i} p_i^{y_i} (1 - p_i)^{n_i - y_i}. \tag{2.2}$$

In the parametrization $\theta_i = \log[p_i/(1-p_i)]$, the likelihood function for $\theta_1, \ldots, \theta_N$ is given by

$$L(\theta_1, \ldots, \theta_N) \propto \prod_{i=1}^{N} \frac{e^{\theta_i y_i}}{(1+e^{\theta_i})^{n_i}}, \tag{2.3}$$

where $N = N_1 + N_2$, N_1 is the number of open studies and N_2 is the number of closed studies.

Combining (2.1) with (2.3), we get the joint posterior distribution for $\theta_1, \ldots, \theta_N$, $\sigma_1^2, \sigma_2^2, \mu, \delta_1$ and δ_2.

Samples of the joint posterior distribution for $\theta_1, \ldots, \theta_N$, $\sigma_1^2, \sigma_2^2, \mu, \delta_1$ and δ_2 are obtained from the full conditional distribution using Gibbs with Metropolis-Hastings algorithm (see Larose and Dey, 1997).

In the original parametrization, we usually have difficulties to get the convergence for the Gibbs Sampling algorithm, especially with the parameters μ and δ_k, $k = 1, 2$ (parameters are not identifiable).

In this way, we consider a reanalysis of the data set of Table 1 assuming the reparametrization $\xi_k = \mu + \delta_k$, $k = 1, 2$.

With this reparametrization, the conditional distributions for the Gibbs Sampling algorithm are given by,

$$\begin{aligned}
\pi(\xi_1|\sigma^2, \theta, D) &\propto N\left[\left(\frac{\frac{\sigma_1^2}{2c} + N_1\overline{\theta}_1}{\frac{\sigma_1^2}{c} + N_1}\right), \left(\frac{\sigma_1^2}{\frac{\sigma_1^2}{c} + N_1}\right)\right], \\
\pi(\xi_2|\sigma^2, \theta, D) &\propto N\left[\left(\frac{\frac{\sigma_2^2}{2d} + N_2\overline{\theta}_2}{\frac{\sigma_2^2}{d} + N_2}\right), \left(\frac{\sigma_2^2}{\frac{\sigma_2^2}{d} + N_2}\right)\right], \\
\pi(\sigma_k^2|\xi, \theta, D) &\propto IG\left[\left(\frac{N_k}{2} + a_k\right), \left(\left(\frac{1}{b_k} + \frac{1}{2}\sum_{i=1}^{N_k}(\theta_i - (\xi_k))^2\right)^{-1}\right)\right], \\
\pi(\theta_i|\xi, \sigma^2, \theta_{(i)}, D) &\propto e^{\theta_i y_i - \frac{(\theta_i - (\xi_k))^2}{2\sigma_k^2}}(1+e^{\theta_i})^{-n_i},
\end{aligned} \tag{2.4}$$

where $\overline{\theta}_k = \sum_{i=1}^{N_k} \theta_i/N_k$, $k = 1, 2$, $i = 1, 2, \ldots, N$ and $N = N_1 + N_2$.

Observe that we need to use the Metropolis-Hastings algorithm to generate the variables θ_i, $i = 1, 2, \ldots, N$.

Considering $a_k = 0.001$, $b_k = 1000$ $(k = 1, 2)$, $c = 10$ and $b = 10$ for the prior distributions given in (2.1), we generated 5 separate Gibbs chains each of which ran for 10000 iterations. For each parameter, we selected the 3070th, 3080th, ... iterations which for 5 chains yields a sample of size 3500. In Table 2, we have Monte Carlo estimates for the posterior means of the parameters based on the 3500 Gibbs samples generated from the conditional posterior distributions (2.4). We monitored the convergence

of the Gibbs Samples using the Gelman and Rubin (1992) method based on the analysis of variance technique to determine if further iterations are needed.

We also have in Table 2, 95% credible intervals for the parameters.

Table 2

Posterior mean and 95% credible intervals of parameters

Parameters	Posterior means	95% credible intervals
ξ_1	0.4499	(-0.1047 ; 1.0002)
ξ_2	-0.8396	(-1.7487 ; 0.0762)
σ_1^2	0.0575	(0.0007 ; 0.3804)
σ_2^2	0.3313	(0.0012 ; 1.6004)
p_1	0.5962	(0.4697 ; 0.7125)
p_2	0.5807	(0.3963 ; 0.7082)
p_3	0.6009	(0.5062 ; 0.6969)
p_4	0.5960	(0.4606 ; 0.7149)
p_5	0.6721	(0.5573 ;0.8137)
p_6	0.6017	(0.5298 ; 0.6703)
p_7	0.2729	(0.1283 ; 0.4359)
p_8	0.3973	(0.2365 ; 0.5949)
p_9	0.2333	(0.0860 ; 0.3999)
p_{10}	0.4003	(0.2242 ; 0.6261)
p_{11}	0.3892	(0.2195 ; 0.6017)
p_{12}	0.4386	(0.2560 ; 0.6539)
p_{13}	0.1599	(0.0273 ; 0.3307)
p_{14}	0.2531	(0.1493 ; 0.3661)
p_{15}	0.2913	(0.1918 ; 0.4000)

From the results of Table 2, we observe that $\sigma_1^2 < \sigma_2^2$, indicating that the open studies form a more homogeneous group than the closed studies.

From the results of Table 2, we also observe that for the open studies, the posterior means for $p_1, \ldots, p_6$ are in the interval (0.58 ; 0.68) and for the closed studies, the posterior means for $p_7, \ldots, p_{15}$ are in the interval (0.15 ; 0.44), that is, the average improvement of patients using progabide in the open studies is better than 0.5, and the average improvement of patients using progabide in the closed studies is less than 0.5. In other words: the open studies support the efficacy of progabide; the closed studies support the reverse hypothesis.

These results are in close agreement with the results obtained by Larose and Dey (1997).

3 Use of estrogen and the incidence of breast and endometrium cancers

In the last years, medical researchers are concerned that estrogen stimulation may contribute to the development of endometrial and breast cancer (see for example, Adami, 1992; Kaufman, 1991; or Brinton and Schairer, 1993).

To investigate if estrogen stimulation is generating endometrium and breast cancers, investigators have turned to meta-analysis, since there are many independent investigators in this medical research project. Besides the use or not of estrogen, usually there are many risk factors like duration of therapy, age at menopause, age at menarche, age at first child, obesity among many others.

The relative-risk of users of estrogen with respect to endometriun or breast cancer could be calculated for independent medical research centers. Relative risk (see for example, Fleiss, 1973) is a measure of the strength of association between estrogen and cancer given by

$$RR = \frac{ad}{bc}, \tag{3.1}$$

where a is the number of exposed cases; b is the number of exposed controls; c is the number of unexposed cases and d is the number of unexposed controls.

Observe that if $RR > 1$, we have evidence that the chance of cancer is larger for users of estrogen.

3.1 A study on the incidence of uteri and ovary cancer

In Table 3, we have a data set introduced by Larose and Dey (1997) from 8 different medical centers on the rates of endometrial cancer.

In Table 3, we observe that for some laboratories there are missing information for the covariates. To analyze the data set of Table 3, we consider a meta-analysis under the Bayesian approach. A Bayesian analysis of this data set was introduced by Larose and Dey (1997), assuming an extension of the hierarchical variance components model of DuMouchel and Harris (1983).

To reanalyze the data set of Table 3, we consider a different model given by,

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{w}\delta + \varepsilon, \tag{3.2}$$

Table 3
Data on endometrial cancer

Studies	Response	Duration			Type of estrogen		Obesity	
	$\log(RR)$	< 12	$[12, 60)$	≥ 60	Conjugate	Other	No	Yes
1	2.08	22	55	56	137	40	186	129
2	-0.06	68	12	11	68	134	584	126
3	0.76	49	70	121	334	122	405	146
4	1.58	18	25	39	74	10	-	-
5	0.87	11	17	49	-	-	722	140
6	0.71	57	85	39	-	-	991	188
7	1.24	26	61	98	-	-	596	190
8	1.99	8	17	27	54	45	-	-

where $\mathbf{y}$ is a $(n \times 1)$ vector of responses given by $\mathbf{y} = \log(RR)$; $\mathbf{X}$ is a $(n \times 3)$ covariate matrix given by,

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3), \tag{3.3}$$

where $\mathbf{x}_1$ is the covariate vector denoting the proportion of duration of treatment larger or equal to 60 months; $\mathbf{x}_2$ is the covariate vector denoting the proportion of conjugate estrogen users; $\mathbf{x}_3$ is the covariate vector denoting the proportion of obesity in each study; $\beta' = (\beta_1, \beta_2, \beta_3)$ is a vector of regression parameters; δ is a dummy variable (measuring the effects of two different groups) and $\mathbf{w}$ is a $(n \times 1)$ vector given by $\mathbf{w}' = (w_1, \ldots, w_n)$, where $w_i = 1(0)$, $i = 1, 2, \ldots, n$, indicating complete information (or not) for the covariates, and ε is a vector of error with elements ε_i, $i = 1, 2, ..., n$.

Let us assume that the random variables ε_i, $i = 1, 2, \ldots, n$, are independent with same $N(0, \sigma^2)$ distribution.

Associated to model (3.2) with the data of Table 3, we have,

$$\mathbf{X} = \begin{pmatrix} 0.42 & 0.77 & 0.41 \\ 0.12 & 0.33 & 0.18 \\ 0.50 & 0.73 & 0.26 \\ 0.47 & 0.88 & - \\ 0.64 & - & 0.16 \\ 0.21 & - & 0.16 \\ 0.53 & - & 0.24 \\ 0.52 & 0.54 & - \end{pmatrix}; \quad \mathbf{w} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{3.4}$$

For the missing observations in $\mathbf{X}$, we can consider the average of the complete observations in each vector $\mathbf{x}_j$, $j = 1, 2, 3$; or to use a simulation approach to obtain these missing observations.

The likelihood function for β, δ and σ^2 is given by,

$$L(\beta,\delta,\sigma^2) = \frac{1}{(\sigma^2)^{n/2}}\, e^{\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_{1i}\beta_1 - x_{2i}\beta_2 - x_{3i}\beta_3 - w_i\delta)^2\right\}}. \tag{3.5}$$

Assuming prior independence, consider the following prior densities for the parameters,

$$\begin{aligned} \delta &\sim N(0,a^2);\ a \text{ known;} \\ \beta_m &\sim N(0,b^2);\ b \text{ known}; m=1,2,3; \\ \sigma^2 &\sim IG(c,d);\ c,d \text{ known.} \end{aligned} \tag{3.6}$$

The conditional distributions for the Gibbs Sampling algorithm are given by,

$$\begin{aligned} \sigma^2|\beta,\delta,\mathbf{y} &\sim IG\left[\frac{n}{2}+c; d+\frac{1}{2}\sum_{i=1}^{n}(y_i - x_{1i}\beta_1 - x_{2i}\beta_2 - x_{3i}\beta_3 - w_i\delta)^2\right], \\ \beta_m|\beta_{(m)},\delta,\sigma^2,\mathbf{y} &\sim N\left[\frac{k_{m2}}{k_{m1}}; \frac{1}{k_{m1}}\right]; m=1,2,3; \beta_{(m)} = (\beta_j,\beta_k); \\ & \quad j\neq k;\ k\neq m;\ m,j,k=1,2,3; \\ \delta|\beta;\sigma^2,\mathbf{y} &\sim N\left[\frac{r_2}{r_1}; \frac{1}{r_1}\right], \end{aligned} \tag{3.7}$$

where

$$\begin{aligned} k_{j1} &= \frac{1}{b^2} + \frac{b_{jj}}{\sigma^2}; j=1,2,3; \\ k_{12} &= \frac{a_1}{\sigma^2} - \frac{b_{12}\beta_2}{\sigma^2} - \frac{b_{31}\beta_3}{\sigma^2} - \frac{c_1\delta}{\sigma^2}; \\ k_{22} &= \frac{a_2}{\sigma^2} - \frac{b_{12}\beta_1}{\sigma^2} - \frac{b_{23}\beta_3}{\sigma^2} - \frac{c_2\delta}{\sigma^2}; \\ k_{32} &= \frac{a_3}{\sigma^2} - \frac{b_{31}\beta_1}{\sigma^2} - \frac{b_{23}\beta_2}{\sigma^2} - \frac{c_3\delta}{\sigma^2}; \\ r_1 &= \frac{1}{a^2} + \frac{e}{\sigma^2}; \\ r_2 &= \frac{f}{\sigma^2} - \frac{c_1\beta_1}{\sigma^2} - \frac{c_2\beta_2}{\sigma^2} - \frac{c_3\beta_3}{\sigma^2}; \\ a_j &= \sum_{i=1}^{n} y_i x_{ji},\ j=1,2,3; \quad b_{jk} = \sum_{i=1}^{n} x_{ji}x_{ki},\ j,k=1,2,3; \\ c_j &= \sum_{i=1}^{n} x_{ji}w_i,\ j=1,2,3; \quad f = \sum_{i=1}^{n} y_i w_i \text{ and } e = \sum_{i=1}^{n} w_i^2\ . \end{aligned}$$

Assuming the model (3.2) for the data set of Table 3 and the prior distributions (3.6) with $a = 2$, $b = 2$, $c = 2.5$ and $d = 2$, we generated from the conditional posterior distributions (3.7), 2 separate Gibbs chains each of which ran for 50000 iterations. For each chain, we considered every 10th draw, and we finally obtained a sample of size 10000. We monitored the convergence of the Gibbs samples using the Gelman and Rubin (1992) method. In Table 4, we have posterior summaries and the estimated potential scale reductions $\hat{R}$ (see Gelman and Rubin, 1992) for all parameters. The considered number of iterations were sufficient for approximate convergence ($\sqrt{\hat{R}} < 1.1$ for all parameters).

Table 4
Posterior summaries (endometrial cancer)

Parameters	Posterior means	95% credible intervals	$\sqrt{\hat{R}}$
β_1	0.21986	(-1.9711; 2.3031)	1.0237
β_2	-1.16245	(-3.0216; 0.6801)	1.0018
β_3	9.71747	(4.9275; 14.4461)	1.0026
δ	-1.19423	(-1.9843; -0.4283)	1.0011
5	0.38427	(0.1967; 0.7705)	1.0046

From the results in Table 4, we observe that the covariates x_1 (duration of treatment) and x_2 (type of estrogen) do not indicate significative effects in the relative risk (the 95% credible interval include zero), but x_3 (obesity) indicate a significative effect in the relative risk. We also observe that there is a significative difference between the two groups (covariates with missing observations and without missing observations). These results are similar to the results obtained by Larose and Dey (1995),except that they did not find that obesity has a significative effect.

3.2 Incidence of breast cancer

In Table 5, we have a data set relative to different studies on the incidence of breast cancer and the use of estrogen (see Larose and Dey, 1995).

Also assuming the model (3.2), with $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_7)$ a $(n \times 7)$ matrix of covariates, where $\mathbf{x}_1$ is the vector denoting the proportion of women that gave birth to at least 3 children; $\mathbf{x}_2$ is the vector denoting the proportion of women that gave birth to the first child with at least 30 years old; $\mathbf{x}_3$ is a vector denoting the proportion of women that had the first menarche with at least 15 years old; $\mathbf{x}_4$ is a vector denoting the proportion of women with surgical menopause; $\mathbf{x}_5$ is a vector denoting the proportion of women with family history of cancers; $\mathbf{x}_6$ is a vector denoting the proportion of women

Table 5
Data on breast cancer

Studies	Response	Number of children			Age at 1st child			Age at Menarche		
	$\log(RR)$	0	1-2	≥ 3	< 20	20-29	≥ 30	< 12	12-14	≥ 15
1	0.0245	345	683	539	155	848	219	287	1031	242
2	0.0746	576	1266	1230	387	2556	376	590	2098	408
3	0.0827	204	731	770	113	1116	271	294	1194	201
4	-0.0419	331	981	1678	721	1740	136	666	1987	352
5	0.0185	359	733	844	146	1120	311	317	1361	250
6	0.0172	210	440	422	81	613	168	194	691	181
7	0.1793	-	-	-	-	-	-	202	717	403
8	0.1487	112	298	340	53	433	121	172	467	109
9	0.3338	197	411	473	-	-	-	-	-	-

Studies	Type of menopause		Family history of cancer		Duration of treatment			Age at menopause		
	Natur.	Surgic.	Posit.	Negat.	≤ 1	1-9	≥ 10	< 40	40-49	≥ 50
1	600	244	130	1438	49	108	45	42	243	370
2	2134	961	335	2761	187	259	105	571	1156	1239
3	1274	432	244	1416	177	142	29	164	695	808
4	1913	1101	279	1639	186	632	137	-	-	-
5	2435	1002	498	1440	589	1025	536	304	879	761
6	-	-	85	982	115	308	89	168	378	378
7	942	380	-	-	100	211	177	126	629	557
8	-	-	92	656	-	-	-	96	276	376
9	1081	201	70	1011	63	103	41	-	-	-

that are using estrogen for at least 10 years; and $\mathbf{x}_7$ is a vector denoting the proportion of women that entered at menopause with at least 50 years old. We also consider a vector $(n \times 1)$ where $w_i = 1(0)$, $i = 1, \ldots, n$, indicates complete information (or not) for the covariates.

In (3.2), the matrix $\mathbf{X}$ is given by,

$$\mathbf{X} = \begin{pmatrix} 0.34 & 0.18 & 0.15 & 0.29 & 0.08 & 0.22 & 0.56 \\ 0.40 & 0.11 & 0.13 & 0.31 & 0.11 & 0.19 & 0.42 \\ 0.45 & 0.18 & 0.12 & 0.25 & 0.15 & 0.08 & 0.48 \\ 0.56 & 0.05 & 0.12 & 0.36 & 0.14 & 0.14 & - \\ 0.44 & 0.20 & 0.13 & 0.29 & 0.26 & 0.25 & 0.39 \\ 0.39 & 0.19 & 0.17 & - & 0.08 & 0.17 & 0.41 \\ - & - & 0.30 & 0.29 & - & 0.36 & 0.42 \\ 0.45 & 0.20 & 0.15 & - & 0.12 & - & 0.50 \\ 0.44 & - & - & 0.16 & 0.06 & 0.20 & - \end{pmatrix}. \tag{3.8}$$

Assuming the same prior distributions (3.6) for the parameters and the same considerations on the missing values of covariates considered in Section 3.1, the conditional distributions for the Gibbs Sampling algorithm are given by,

$$\begin{aligned} \sigma^2|\beta, \delta, \mathbf{y} &\sim IG\left[\frac{n}{2} + c; d + \frac{1}{2}\sum_{i=1}^{n}(y_i - x_{1i}\beta_1 - x_{2i}\beta_2 - x_{3i}\beta_3 \right. \\ &\quad \left. - x_{4i}\beta_4 - x_{5i}\beta_5 - x_{6i}\beta_6 - x_{7i}\beta_7 - w_i\delta)^2\right]; \\ \beta_m|\beta_{(m)}, \delta, \sigma^2, \mathbf{y} &\sim N\left\{\frac{k_{m2}}{k_{m1}}; \frac{1}{k_{m1}}\right\}; \; m = 1, 2, \ldots, 7; \\ \delta|\beta, \sigma^2, \mathbf{y} &\sim N\left[\frac{r_2}{r_1}; \frac{1}{r_1}\right], \end{aligned} \tag{3.9}$$

where

$$\begin{aligned} k_{j1} &= \frac{1}{b^2} + \frac{b_{jj}}{\sigma^2}; \quad j = 1, 2, \ldots, 7; \\ k_{12} &= \frac{a_1}{\sigma^2} - \frac{b_{12}\beta_2}{\sigma^2} - \frac{b_{13}\beta_3}{\sigma^2} - \frac{b_{14}\beta_4}{\sigma^2} - \frac{b_{15}\beta_5}{\sigma^2} - \frac{b_{16}\beta_6}{\sigma^2} - \frac{b_{17}\beta_7}{\sigma^2} - \frac{c_1\delta}{\sigma^2}; \\ k_{22} &= \frac{a_2}{\sigma^2} - \frac{b_{12}\beta_1}{\sigma^2} - \frac{b_{23}\beta_3}{\sigma^2} - \frac{b_{24}\beta_4}{\sigma^2} - \frac{b_{25}\beta_5}{\sigma^2} - \frac{b_{26}\beta_6}{\sigma^2} - \frac{b_{27}\beta_7}{\sigma^2} - \frac{c_2\delta}{\sigma^2}; \\ k_{32} &= \frac{a_3}{\sigma^2} - \frac{b_{13}\beta_1}{\sigma^2} - \frac{b_{23}\beta_2}{\sigma^2} - \frac{b_{34}\beta_4}{\sigma^2} - \frac{b_{35}\beta_5}{\sigma^2} - \frac{b_{36}\beta_6}{\sigma^2} - \frac{b_{37}\beta_7}{\sigma^2} - \frac{c_3\delta}{\sigma^2}; \\ k_{42} &= \frac{a_4}{\sigma^2} - \frac{b_{14}\beta_1}{\sigma^2} - \frac{b_{23}\beta_2}{\sigma^2} - \frac{b_{34}\beta_3}{\sigma^2} - \frac{b_{45}\beta_5}{\sigma^2} - \frac{b_{46}\beta_6}{\sigma^2} - \frac{b_{47}\beta_7}{\sigma^2} - \frac{c_4\delta}{\sigma^2}; \end{aligned}$$

$$
\begin{aligned}
k_{52} &= \frac{a_5}{\sigma^2} - \frac{b_{15}\beta_1}{\sigma^2} - \frac{b_{25}\beta_2}{\sigma^2} - \frac{b_{35}\beta_3}{\sigma^2} - \frac{b_{45}\beta_4}{\sigma^2} - \frac{b_{56}\beta_6}{\sigma^2} - \frac{b_{57}\beta_7}{\sigma^2} - \frac{c_5\delta}{\sigma^2};\\
k_{62} &= \frac{a_6}{\sigma^2} - \frac{b_{16}\beta_1}{\sigma^2} - \frac{b_{26}\beta_2}{\sigma^2} - \frac{b_{36}\beta_3}{\sigma^2} - \frac{b_{46}\beta_4}{\sigma^2} - \frac{b_{56}\beta_5}{\sigma^2} - \frac{b_{67}\beta_7}{\sigma^2} - \frac{c_6\delta}{\sigma^2};\\
k_{72} &= \frac{a_7}{\sigma^2} - \frac{b_{17}\beta_1}{\sigma^2} - \frac{b_{27}\beta_3}{\sigma^2} - \frac{b_{37}\beta_3}{\sigma^2} - \frac{b_{47}\beta_4}{\sigma^2} - \frac{b_{57}\beta_5}{\sigma^2} - \frac{b_{67}\beta_6}{\sigma^2} - \frac{c_7\delta}{\sigma^2};\\
r_1 &= \frac{1}{a^2} + \frac{e}{\sigma^2};\\
r_2 &= \frac{f}{\sigma^2} - \frac{c_1\beta_1}{\sigma^2} - \frac{c_2\beta_2}{\sigma^2} - \frac{c_3\beta_3}{\sigma^2} - \frac{c_4\beta_4}{\sigma^2} - \frac{c_5\beta_5}{\sigma^2} - \frac{c_6\beta_6}{\sigma^2} - \frac{c_7\beta_7}{\sigma^2};\\
a_j &= \sum_{i=1}^{n} y_i x_{ji};\ j = 1, 2, \ldots, 7; \quad b_{jk} = \sum_{i=1}^{n} x_{ji} x_{ki};\ \ j, k = 1, 2, \ldots, 7;\\
c_j &= \sum_{i=1}^{n} x_{ji} w_i;\ j = 1, 2, \ldots, 7; \quad f = \sum_{i=1}^{n} y_i w_i; \quad e = \sum_{i=1}^{n} w_i^2\ .
\end{aligned}
$$

Assuming the prior distribution (3.6) with $a = 0.2$, $b = 0.1$, $c = 0.25$ and $d = 2$, we generated from (3.9), 2 separate Gibbs chains each of which ran for 50000 iterations. Taking every 10th draw, we obtained a final sample of size 10000.

In Table 6, we have the posterior means and 95% credible intervals for all parameters. We also have in Table 6, the estimated potential scale reductions (see Gelman and Rubin, 1992) for all parameters.

Table 6
Posterior summaries (breast cancer)

Parameters	Posterior means	95% credible intervals	$\sqrt{\hat{R}}$
β_1	1.3189	(-0.1175; 2.7868)	1.0007
β_2	0.4012	(-1.3759; 2.1904)	1.0030
β_3	-0.1612	(-2.3994; 2.0503)	1.0043
β_4	-1.7132	(-3.3587; -0.0361)	1.0009
β_5	-1.2338	(-3.2577; 0.8856)	0.9999
β_6	0.8967	(-0.7646; 2.4733)	1.0009
β_7	-0.2172	(-1.6351; 1.2790)	1.0011
δ	0.0738	(-0.1735; 0.3173)	0.9998
σ	0.1225	(0.05780; 0.2631)	1.0057

From the results in Table 6, we observe that the covariates x_1, x_2, x_3, x_5, x_6 and x_7 not present significative effects in the response $y = \log(RR)$,

where RR is the relative risk of cancer among the users of estrogen, but x_4 (type of menopause) presents a significative effect on the response. These results are in similar with the obtained results of Larose and Dey (1995).

4 Some conclusions

The use of meta-analysis allows for the combination of resulties from different independent studies. Meta-analysis is very useful in medical research, especially in the comparison of different groups of studies, usually involving many risk factors.

With the advance of computational methods, the use of Bayesian methods is becoming more popular in meta-analysis. In special, the simulation methods based on Markov Chain Monte Carlo (MCMC) has been a practical alternative of great interest in meta-analysis, where the proposed models can have too many parameters, and the usual classical inference approach could be difficult or, sometimes, impossible to be obtained.

It is important to point out that the use of the new proposed models considered in this paper could be applied to any other situation of combination of studies from different research centers. The use of models considering the Bayesian approach gives simple interpretations and could be a powerful control mechanism to be used by medical researchers.

Acknowledgements

The authors are thankful to the editor and to a referee for some useful suggestions.

(Received September, 1999. Revised April, 2001.)

References

Adami, H. O. (1992). Long-term consequences of estrogen and estrogen-progestin replacement. *Cancer Causes and Control*, **3**, 83–90.

Brinton L. A.; Schairer, C. (1993). Estrogen replacement therapy and breast cancer risk. *Epidemiologic Reviews*, **15**, 66–79.

Bulpitt, C. J. (1988). Meta-analysis. *Lancet*, ii. 93–94.

Chib, S.; Greenberg, E. (1995). Understand the Metropolis-Hastings algorithm. *The American Statistician*, **49**, 327–335.

DuMouchel, W. H. (1996). Predictive Cross-validation in Bayesian Meta-analysis (J. M. Bernardo et al., eds). *Bayesian Statistics 5*, Oxford University Press, 107–127 (with discussion).

DuMouchel, W. H.; Harris, J.E. (1983). Bayes methods for combining the results of cancer studies in humans and other species. *Journal of the American Statistical Association*, **78**, 293–315.

Fleiss, J. L. (1973). *Statistical Methods for Rates & Proportions.* New York: John Wiley & Sons.

Gelfand, A. E.; Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**, 398–409.

Gelman, A. E.; Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**, 457–472.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, **5**, 3–8.

Glass, G. V.; McGraw, B.; Smith, M. L. (1981). *Meta-analysis in Social Research.* Beverly Hills, Calif.: Sage Publications.

Hedges, L. U. (1992). Meta-analysis. *Journal of Educational Statistics*, **17**, 229–296.

Hedges, L. U. (1983). Combining independent estimators in research synthesis. *Br. J. Math. Stat. Psychol*, **35**, 123–131.

Hedges, L. U. (1982). Estimation of effect size from a series of independent experiments. *Psychol Bull.*, **92**, 490–499.

Hunter, J. E.; Schimidt, F. L.; Jackson, B. G. (1982). *Meta-analysis: Cumulating Research Finding Across Studies.* Beverly Hills, Calif: Sage Publications.

Jeffreys, H. (1939). *Theory of Probability.* Oxford: University Press.

Kaufman, D. W. et al (1991). Estrogen replacement therapy and the risk of breast cancer: results from the case control surveillance study. *American Journal of Epidemiology*, **134**, 1375–1385.

L'Abbe, K. A.; Detsky, A. S.; O'Rourke, K. (1987). Meta-analysis in clinical research. *Ann. Intern. Med.*, **107**, 224–233.

Larose, D. T.; Dey, D. K. (1997). Grouped random effects models for Bayesian meta- analysis. *Statistics in Medicine*, **16**, 1817–1829.

Larose, D. T.; Dey, D. K. (1995). Is estrogen linked to cancers of the breast and endometrium? New meta-analysis using Bayesian grouped random effects model. *Technical Reporter,* 93-95, Department of Statistics, University of Connecticut, Storrs.

Light, R. J.; Pillemer, D. B. (1984). *Summing up: The Science of Reviewing Research.* Cambridge, Mass: Harvard University Press.

Roberts, G. O.; Smith, A. F. M. (1993). Bayesian methods via the Gibbs Sampling and related Markov Chain Monte Carlo methods. *Journal of the Royal Statistical Society B*, **55**, 3–23.

Rosenthal, R. (1978). Combining results of independent studies. *Psychol Bull.*, **85**, 185–193.

Sacks, H. S. et al (1988). Meta-analysis of randomized controlled trials. *N. Engl. J. Med.*, **316**, 450–455.

Van Belle, G. (1992). A meta-analysis of the effectiveness of progabide a an antiepileptic drug. Unpublished manuscript.