Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
0(0) 1

DOI: 10.1177/0962280210393712
smm.sagepub.com
SAGE

Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: Methods for the absolute risk difference and relative risk

Haitao Chu,[1] Lei Nie,[2,*] Yong Chen,[3] Yi Huang[4] and Wei Sun[5]

Abstract

Multivariate meta-analysis is increasingly utilised in biomedical research to combine data of multiple comparative clinical studies for evaluating drug efficacy and safety profile. When the probability of the event of interest is rare, or when the individual study sample sizes are small, a substantial proportion of studies may not have any event of interest. Conventional meta-analysis methods either exclude such studies or include them through *ad hoc* continuality correction by adding an arbitrary positive value to each cell of the corresponding 2×2 tables, which may result in less accurate conclusions. Furthermore, different continuity corrections may result in inconsistent conclusions. In this article, we discuss a bivariate Beta-binomial model derived from Sarmanov family of bivariate distributions and a bivariate generalised linear mixed effects model for binary clustered data to make valid inferences. These bivariate random effects models use all available data without *ad hoc* continuity corrections, and accounts for the potential correlation between treatment (or exposure) and control groups within studies naturally. We then utilise the bivariate random effects models to reanalyse two recent meta-analysis data sets.

Keywords

beta-binomial distribution, bivariate generalised linear mixed models, bivariate random effects models, clustered binary data, meta-analysis

[1]Division of Biostatistics, School of Public Health, The Univerity of Minnesota, Minneapolis, MN 55455, USA.
[2]Division IV, Office of Biostatistics/OTS/CDER/FDA, Spring, MD 20993-0002, USA.
[3]Division of Biostatistics, School of Public Health, The University of Texas Health Science Center, Houston, TX 77030, USA.
[4]Departemt of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, MD 21250, USA.
[5]Department of Biostatistics and Genetics, The Univerity of North Carolina, Chapel Hill, NC 27599, USA.

*Views expressed in this paper are the author's professional opinions and do not necessarily represent the official positions of the U.S. Food and Drug Administration.

Corresponding author:
Haitao Chu, Division of Biostatistics, School of Public Health, The Univerity of Minnesota, Minneapolis, MN 55455 USA
Email: chux0051@umn.edu

1 Introduction

The growth of evidence-based medicine has led to an increase in attention to meta-analysis.[1] Meta-analysis, also known as systematic overview, is a statistical process commonly used in biomedical research of combining the information from several independent studies concerned with the same clinical question including the treatment or exposure effect, with the aim of being able to resolve contradictory issues that cannot be concluded from a single study alone.

In meta-analysis of a set of N clinical trials with a binary outcome comparing an experimental treatment with a placebo, data can be represented as a series of 2×2 tables. The standard fixed and random meta-analysis methods for providing an overall estimate of the treatment effect across all studies rely on some assumptions.[2] Specifically, the fixed effect model assumes homogeneous treatment effects across all studies. Let θ_i be the value of a chosen measure (e.g. risk difference, RD or log relative risk) of treatment effect in the ith study ($i = 1, 2, \ldots, N$), the homogeneity requires $\theta_i = \theta$ ($i = 1, 2, \ldots, N$). Let $\hat{\theta}_i$ be an estimate of θ_i, and w_i denote the weight, which is often taken to be the reciprocal of the estimated variance $\hat{v}_i$ of $\hat{\theta}_i$ (i.e. $\hat{w}_i = 1/\hat{v}_i$),[3] then the overall treatment effect based on the fixed effect model is estimated as a weighted average of the individual study estimated treatment effects, that is, $\hat{\theta}_w = \sum_i \hat{w}_i \hat{\theta}_i \big/ \sum_i \hat{w}_i$. Under the combined null hypothesis H_0: $\theta_i = 0 (i = 1, 2, \ldots, N)$, the test-statistic $U = (\sum_i \hat{w}_i \hat{\theta}_i)^2 \big/ \sum_i \hat{w}_i$ follows a χ^2 distribution with 1 degree of freedom. A formal test of homogeneity can be performed using the Cochran's Q-statistic, defined by $Q = \sum_i \hat{w}_i (\hat{\theta}_i - \hat{\theta}_w)^2$, which has approximately a χ^2_{N-1} distribution under the null hypothesis H_0: $\theta_i = \theta$ ($i = 1, 2, \ldots, N$).

Through a random-effects model, DerSimonian and Laird[4] provided a way of incorporating heterogeneity into the overall estimate by including a between-study variance component σ_b^2. It basically assumes that $\hat{\theta}_i \sim N(\theta_i, \hat{v}_i)$ and $\theta_i \sim N(\theta, \sigma_b^2)$.[2] The overall treatment effect is once again obtained as a weighted average with the weight being estimated as $\hat{w}_i^* = 1/(\hat{v}_i + \hat{\sigma}_b^2)$, i.e. $\hat{\theta}_w^* = \sum_i \hat{w}_i^* \hat{\theta}_i \big/ \sum_i \hat{w}_i^*$. Under the combined null hypothesis H_0: $\theta_i = 0 (i = 1, 2, \ldots, N)$, the test-statistic $U^* = (\sum_i \hat{w}_i^* \hat{\theta}_i)^2 \big/ \sum_i \hat{w}_i^*$ follows a χ^2 distribution with 1 degree of freedom. The method of moment's estimate of σ_b^2 is given by: $\hat{\sigma}_b^2 = \max\{[\sum_i \hat{w}_i - (\sum_i \hat{w}_i^2) \big/ \sum_i \hat{w}_i]^{-1} [Q - (N-1)], 0\}$.[4]

A concern on these conventional two-step meta-analysis methods is that, they require estimating study-specific treatment effect $\hat{\theta}_i$ (commonly expressed by log relative risk, log odds ratio; OR or RD) and its variance $\hat{v}_i$ based on the normal approximation. When the probability of the event of interest is rare or if the individual study sample sizes are small, this normality assumption might not hold. Furthermore, a substantial proportion of studies may not have any event of interest. To circumvent the issues of zero cells, the conventional meta-analysis methods either exclude such studies[5] or add an arbitrary positive value to each cell of the corresponding 2×2 tables in the analysis, which may lead to less accurate conclusions. For example, different continuity corrections may result in different conclusions.[6] An interesting yet challenging methodology question is how to use all available data without assigning an arbitrary number to the empty cells in meta-analysis.[7–10] Furthermore, it has been noted that these weighting-according-to-the-variance methods may introduce biases in meta-analyses of binary outcomes because this weighting scheme favours studies with certain frequencies of outcome events.[11] The relative weights for the individual studies in a meta-analysis may change considerably among different choices of effect measurements, which may lead to contradictory conclusions. This is particularly true for the sparse data scenario. Specifically, a study with zero event in both treatment and placebo groups, which would be excluded on a relative scale, would be included and even be given a large weight on a RD scale.[6]

Recently, multivariate random effects models for meta-analyses have become increasingly popular in biomedical research. For example, multivariate random effects models have been

proposed for meta-analyses of diagnostic test studies[12–18] and correlated multiple outcomes.[19,20] Given the potential issues of applying conventional meta-analysis methods based on a univariate outcome, we discuss bivariate random effects models to deal with those challenges for meta-analyses of comparative clinical trials with binary outcomes in this article. Although, the proposed methods were primarily presented for bivariate meta-analyses, they can be easily generalised to multivariate meta-analyses. Specifically, Section 2 shows the estimation of marginal treatment effects using the maximum likelihood methods under two models, i.e. a generalised linear mixed effects model and a bivariate Beta-binomial model. In Section 3, we reanalyse the data from two case studies: the study of type 2 diabetes mellitus after gestational diabetes[21] and the study of myocardial infarction (MI) with rosiglitazone.[5] Section 4 concludes this article with a brief discussion.

2 Bivariate random effects models for meta-analysis of comparative studies

Let n_{ki} be the number of subjects, and p_{ki} the probability of 'success' for the ith study ($i = 1, 2, \ldots, N$) in the kth treatment (or exposure) group with $k = 1$ denoting the placebo (or unexposed) group and $k = 2$ denoting the treated (or exposed) group. Let Y_{kij} denote a Bernoulli random variable with value 1 denoting a 'success' and value 0 denoting a 'failure' for the jth subject ($j = 1, 2, \ldots, n_{ki}$) of the ith study in the kth treatment group. Let $X_{ki} = \sum_{j=1}^{n_{ki}} Y_{kij}$ be the total number of 'success' in the kth treatment group in the ith study. In the first stage, conditional on the probability of events (i.e. p_{ki}) and the number of subjects (i.e. n_{ki}) of the kth treatment group in the ith study, the bivariate random effects model assumes that X_{ki} is independently binomially distributed as $\text{Bin}(n_{ki}, p_{ki})$ for $k = 1, 2$ and $i = 1, 2, \ldots, N$, that is,

$$P(X_{1i} = x_{1i}, \quad X_{2i} = x_{2i} | n_{1i}, n_{2i}, p_{1i}, p_{2i}) = \prod_{k=1}^{2} \binom{n_{ki}}{x_{ki}} (p_{ki})^{x_{ki}} (1 - p_{ki})^{n_{ki} - x_{ki}} \tag{1}$$

In the second stage, the joint distribution $f(p_{1i}, p_{2i})$, which is also denoted as $f(p_1, p_2)$ for ease of notation, is specified. Specifically, we first review the bivariate generalised linear mixed effects models (BGLMMs) and then propose the bivariate Beta-binomial models as an alternative, for the evaluation of marginal treatment or exposure effect. Note that bivariate models are commonly used when there are two outcomes (e.g. the response to a treatment and the appearance of a side effect), in this article, we use bivariate models to jointly model the study-specific response rates in the placebo group and the treatment group in a meta-analysis with multiple studies.

2.1 Bivariate generalised linear mixed effects models

In the second stage, the BGLMM assumes a bivariate normal distribution of (p_{1i}, p_{2i}) in a transformed scale, which implies a linear relationship between p_{1i} and p_{2i} on a transformed scale. It is generally specified as follows,

$$g(p_{1i}) = \nu_1 + \nu_{1i}, \quad g(p_{2i}) = \nu_2 + \nu_{2i} \text{ and } (\nu_{1i}, \nu_{2i})^T \sim N(\mathbf{0}, \mathbf{\Sigma}_\nu), \tag{2}$$

where $g()$ is the link function such as the commonly used logit, probit and complementary log–log transformation functions, (ν_1, ν_2) are the fixed effects and $\mathbf{\Sigma}_\nu = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ is the variance-covariance matrix. To implement the natural constraint of $-1 \le \rho \le 1$, one can use the Fisher's z transformation as $\rho = [\exp(2z) - 1]/[\exp(2z) + 1]$.

Based on the model in Equation (2), the median success probability in the kth treatment group for the population can be estimated as $M(p_k) = g^{-1}(\nu_k)$, $k = 1, 2$. And, its mean can be estimated as:

$$E(p_k) = \int_{-\infty}^{+\infty} g^{-1}(\nu_k + z)\sigma_k^{-1}\phi(z/\sigma_k)\mathrm{d}z, \quad k = 1, 2, \tag{3}$$

where $\phi()$ is the standard Gaussian density function. Based on the bivariate normality assumption of $(\nu_{1i}, \nu_{2i})^T$, the expected success probability in group k ($k = 1, 2$) at a given success probability in group l ($l = 1, 2$) in the transformed scale is given by:

$$\mathrm{E}[\mathrm{g}(p_k)|\mathrm{g}(p_l)] = \nu_k + \rho\sigma_k/\sigma_l[\mathrm{g}(p_l) - \nu_l] = (\nu_k - \rho\nu_l\sigma_k/\sigma_l) + \rho\sigma_k/\sigma_l\mathrm{g}(p_l), \quad k \neq l;\ k, l = 1, 2. \tag{4}$$

Thus, the BGLMM implies a linear relationship between p_1 and p_2 on a transformed scale.

2.2 Bivariate Beta-binomial models

As an alternative, Beta-binomial distributions can be used to model the success probabilities of the treatment and control groups to account for the within-study correlation. To allow for the possible correlation between the success probabilities in the treatment and control groups, Lee[22] introduced a class of bivariate Beta-binomial distributions using the framework introduced by Sarmanov.[23] Such bivariate Beta-binomial distributions can be used to model the success probabilities of (p_{1i}, p_{2i}) jointly as follows,

$$f(p_{1i}, p_{2i}) = f(p_1, p_2) = \left[1 + \omega\prod_{k=1}^{2}(p_k - \mu_k)\right]\prod_{k=1}^{2}\left[\frac{(p_k)^{\alpha_k - 1}(1 - p_k)^{\beta_k - 1}}{B(\alpha_k, \beta_k)}\right], \tag{5}$$

where $\alpha_k,\ \beta_k > 0$, $E(p_k) = \mu_k = \dfrac{\alpha_k}{\alpha_k + \beta_k}$ and $B(\alpha_k, \beta_k) = \int_0^1 x^{\alpha_k - 1}(1 - x)^{\beta_k - 1}\mathrm{d}x$. The bivariate Beta-binomial distribution specified by Equation (5) has several attractive features. First, the marginal distribution of p_k follows a Beta distribution $f(p_k) = Beta(\alpha_k,\ \beta_k)$. Second, a correlation between the success probabilities in the treatment and control groups is allowed and modelled by $\rho = \omega\delta_1\delta_2$, where $\delta_k^2 = \dfrac{\alpha_k\beta_k}{(\alpha_k + \beta_k)^2(\alpha_k + \beta_k + 1)}$ is the variance of p_k. When $\omega = 0$, Equation (5) collapses to the product of two univariate Beta densities, corresponding to independent Beta distributions for p_1 and p_2. To ensure a valid joint probability density function, ω must satisfy the condition

$$-[\max(\alpha_1\alpha_2, \beta_1\beta_2)]^{-1}\prod_{k=1}^{2}(\alpha_k + \beta_k) \leq \omega \leq [\max(\alpha_1\beta_2, \beta_1\alpha_2)]^{-1}\prod_{k=1}^{2}(\alpha_k + \beta_k). \tag{6}$$

To ensure a valid joint probability density function for (p_1, p_2) and avoid computational difficulties, we re-parameterise ω by the unconstrained parameter η as follows,

$$\omega = \prod_{k=1}^{2}(\alpha_k + \beta_k)\left\{\frac{\exp(\eta)}{1 + \exp(\eta)}[\max(\alpha_1\beta_2, \beta_1\alpha_2)]^{-1} - \frac{1}{1 + \exp(\eta)}[\max(\alpha_1\alpha_2, \beta_1\beta_2)]^{-1}\right\}. \tag{7}$$

The conditional distribution of p_k for a chosen p_l is $f(p_k|p_l)=\frac{f(p_1,p_2)}{f(p_l)}$ $=f(p_k)[1+\omega\prod_{k=1}^{2}(p_k-\mu_k)](k\neq l;\ k=1,\ 2;\ l=1,\ 2)$. The conditional mean of p_k for a chosen p_l is given by $E(p_k|p_l)=\mu_k+\rho\delta_k/\delta_l(p_l-\mu_l)$, and the conditional variance of p_k for a chosen p_l is given by:

$$\mathrm{Var}(p_k|p_l)=\delta_k^2\left\{1-\rho^2\frac{(p_l-\mu_l)^2}{\delta_l^2}\right\}+\rho\frac{p_l-\mu_l}{\delta_k\delta_l}\left\{\frac{2\beta_k\alpha_k(\beta_k-\alpha_k)}{(\alpha_k+\beta_k+2)(\alpha_k+\beta_k+1)(\alpha_k+\beta_k)^3}\right\}. \tag{8}$$

The bivariate Beta-binomial model implies a linear relationship between p_1 and p_2 on the original scale. The unconditional joint probability density function for $(X_1=x_{1i},\quad X_2=x_{2i})$ is,

$$\begin{aligned}&P(X_1=x_{1i},\quad X_2=x_{2i}|n_{1i},\,n_{2i},\,\alpha_1,\,\alpha_2,\,\beta_1,\,\beta_2,\,\omega) \qquad (9)\\ &=\int_0^1\int_0^1\prod_{k=1}^{2}\binom{n_{ki}}{x_{ki}}(p_{ki})^{x_{ki}}(1-p_{ki})^{n_{ki}-x_{ki}}\left[1+\omega\prod_{k=1}^{2}(p_{ki}-\mu_{ki})\right]\prod_{k=1}^{2}\left[\frac{(p_{ki})^{\alpha_k-1}(1-p_{ki})^{\beta_k-1}}{B(\alpha_k,\beta_k)}\right]\mathrm{d}p_{1i}\mathrm{d}p_{2i}\\ &=\prod_{k=1}^{2}\binom{n_{ki}}{x_{ki}}\frac{B(\alpha_k+x_{ki},\beta_k+n_{ki}-x_{ki})}{B(\alpha_k,\beta_k)}\times\left[1+\omega\prod_{k=1}^{2}\frac{x_{ki}-n_{ki}\mu_k}{\alpha_k+\beta_k+n_{ki}}\right],\end{aligned}$$

which leads to the following log-likelihood function for the observed 2×2 tables after ignoring some constant,

$$\begin{aligned}\log L(\omega,\alpha_k,\beta_k)=\sum_{i=1}^{N}\sum_{k=1}^{2}\Bigg\{&\log B(\alpha_k+x_{ki},\beta_k+n_{ki}-x_{ki})-\log B(\alpha_k,\beta_k)\\ &+\log\left[1+\omega\prod_{k=1}^{2}\frac{x_{ki}-n_{ki}\mu_k}{\alpha_k+\beta_k+n_{ki}}\right]\Bigg\},\end{aligned} \tag{10}$$

where ω must satisfy the condition in Equation (6) to ensure nonnegative probability.

2.3 Marginal treatment effects: risk difference and risk ratio

Although, the issue of deciding which effect measure to use in a particular application is non-trivial,[1] we focus on the estimation of risk ratio (RR) and RD here for two reasons: (1) the interpretation of OR as an estimate of RR often leads to exaggerated associations when the binary outcome of interest is common;[24–26] and (2) the well-known non-collapsibility issue related to OR makes it undesirable in interpretation and estimation.[27,28] For example, in the presence of effect modification, when an exposure increases risk but all risks are less than 0.5, it is possible for the relative risk and the RD to change in the same direction, but the OR to change in the opposite direction.[29] In this article, we focus on the overall marginal (or population averaged) treatment (or exposure) effect, as suggested by McCullagh,[30,31] which is defined as the RD $(\mathrm{RD})=E(p_1)-E(p_2)$ and the $\mathrm{RR}=E(p_1)/E(p_2)$, where $E(p_k)=\int_{-\infty}^{+\infty}g^{-1}(\nu_k+z)\sigma_k^{-1}\phi(z/\sigma_k)\mathrm{d}z$ for the BGLMM and $E(p_k)=\alpha_k/(\alpha_k+\beta_k)$ for the bivariate Beta-binomial model, $k=1,2$. Furthermore, although the computation of $E(p_k)$ $(k=1,2)$ from BGLMM involves integration, there is a closed-form formula of $E(p_k)=\Phi\left(\nu_k/\sqrt{1+\sigma_k^2}\right)$ $(k=1,2)$ for the bivariate probit random effects model, and

a well-established approximation formula of $E(p_k) \approx \text{expit}(\nu_i/\sqrt{1 + C^2\sigma_k^2})$ $(k = 1, 2)$ for the bivariate logit random effects models,[32] where $C = 16\sqrt{3}/(15\pi)$. For the complementary log–log random effects models, $E(p_k)$ can be easily computed by numerical integration, for example, by the trapezoidal rule with 1000 equal space subintervals as implemented in this article.

2.4 Model implementation

The bivariate Beta-binomial model and the bivariate generalised linear mixed model can be fitted using commonly used statistical software. We implement it through the SAS NLMIXED procedure (SAS Institute Inc., Cary, NC), which maximises the likelihood function by dual quasi-Newton optimisation techniques for the bivariate Beta-binomial model, and uses an adaptive Gaussian quadrature to approximate the likelihood integrated over the random effects for BGLMM.[33] Furthermore, the delta method built in SAS NLMIXED is used to compute the population averaged overall treatment effect estimates and their SEs based on the normal approximation. To select a model that can give a better goodness-of-fit, either the finite sample corrected Akaike's information criterion (AIC_C) or the Bayesian information criterion (BIC) can be used as the guideline.[34]

3 Two case studies

To illustrate and compare the performance of the bivariate Beta-binomial model and the BGLMMs discussed in this article, we apply them to two recently published meta-analyses.

3.1 Example 1: Meta-analysis of type 2 diabetes mellitus after gestational diabetes

Recently, Bellamy et al.[21] presented an interesting comprehensive systematic review and meta-analysis to assess the strength of association between gestational diabetes and type 2 diabetes mellitus. In summary, 20 cohort studies were included in the meta-analysis with 675 455 women and 10 859 type 2 diabetic events. Table 1 shows the frequencies of the diabetic events for these 20 studies.

We fitted the bivariate Beta-binomial and the BGLMMs as described in Section 2 to study the association between gestational diabetes and type 2 diabetes mellitus. Table 2 presents the parameter estimates and their SEs, including the population averaged risk of type 2 diabetes mellitus for those with and without gestational diabetes, population averaged RD and RR, and the goodness-of-fit measures including the finite sample corrected AIC_C and the BIC. As shown in Table 2, there is not enough evidence to support that the risks of type 2 diabetes mellitus for those with and without gestational diabetes are correlated within studies from both the bivariate Beta-binomial model and the bivariate generalised linear mixed models with three link functions, i.e. the models with $\rho = 0$ provide better goodness-of-fit for all four models considered. Note that the results from different models are very similar here. Based on AIC_C and BIC, the best fitted model is a bivariate logit generalised linear mixed effects model with $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$. Based on this model, the population averaged risk of type 2 diabetes mellitus for those with and without gestational diabetes are estimated to be 0.200 (SE = 0.031) and 0.025 (SE = 0.006), respectively. The population averaged RD is estimated to be 0.175 (SE = 0.031), where the population averaged RR is estimated to be 7.948 (SE = 2.167). It is interesting to note that the population averaged RR estimates from all models are slightly higher than what Bellamy et al.[21] reported (i.e. 7.43 with 95% confidence interval

Table 1. Example 1: Data from a meta-analysis of studies on type 2 diabetes mellitus after gestational diabetes[21].

	Type 2 diabetes mellitus with gestational diabetes		Type 2 diabetes mellitus without gestational diabetes	
Study	# events	# observations	# events	# observations
1	2874	21823	6628	637341
2	71	620	22	868
3	21	68	0	39
4	43	166	150	2242
5	53	295	1	111
6	405	5470	16	783
7	6	70	7	108
8	13	35	8	489
9	7	23	0	11
10	23	435	0	435
11	44	696	0	70
12	21	229	1	61
13	10	28	0	52
14	15	45	1	39
15	105	801	7	431
16	10	15	0	35
17	33	241	0	57
18	14	47	3	47
19	224	615	18	328
20	5	145	0	41

of 4.79 to 11.51) based on the random effects model by DerSimonian and Laird.[4] The reason might be the fact that an *ad hoc* continuity correction was implemented for the studies with zero diabetic events in the group without gestational diabetes in Bellamy et al.,[21] where our models do not.

Because one of the studies has almost all the cases (9502 out of 10859 cases), we did a sensitivity meta-analysis by excluding that study. The results are presented in Table A1. In summary, it suggests similar conclusions. Specifically, the best fitted model is a bivariate probit generalised linear mixed effects model with $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$, and the population averaged risk of type 2 diabetes mellitus for those with and without gestational diabetes are estimated to be 0.205 (SE = 0.031) and 0.025 (SE = 0.007), respectively. The population averaged RD is estimated to be 0.181 (SE = 0.032), where the population RR is estimated to be 8.371 (SE = 2.756).

3.2 Example 2: Meta-analysis of the risk of MI with rosiglitazone

To investigate whether rosiglitazone, a drug for treating type 2 diabetes mellitus, significantly increases the risk of MI or cardiovascular disease (CVD)-related death, Nissen and Wolski[5] performed a meta-analysis of 48 clinical trials that satisfied the inclusion criteria for their analysis. Among them, 10 studies have no MI events and 25 studies have no CVD-related deaths, which were simply excluded by Nissen and Wolski from their analysis. This meta-analysis data set has been reanalysed by Shuster et al.,[35] Tian et al.[8] and others,[36–38] and updated by Dahabreh.[39] For the illustration purpose, we will only focus on the association between rosiglitazone usage and the

Table 2. Point estimates (SEs) for meta-analysis of studies on type 2 diabetes mellitus after gestational diabetes[21].

Model description	Modification from the standard model	Probability $\hat{P}_1$	Probability $\hat{P}_2$	Risk difference $\widehat{RD}$	Relative risk $\widehat{RR}$	BIC	AIC_C
Bivariate probit GLMM							
	No change	0.2014 (0.0306)	0.0238 (0.0061)	0.1776 (0.0292)	8.4556 (2.1305)	306.1	302.8
	Random effects $v_{1i} = v_{2i}$	0.1934 (0.0262)	0.0254 (0.0064)	0.1680 (0.0202)	7.6029 (0.9612)	341.6	339.3
	Correlation $\rho = 0$	0.2036 (0.0313)	0.0226 (0.0059)	0.1810 (0.0319)	9.0077 (2.7280)	305.1	302.2
	Correlation $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$	0.2002 (0.0291)	0.0237 (0.0067)	0.1765 (0.0295)	8.4589 (2.6317)	302.5	300.1
Bivariate logit GLMM							
	No change	0.1973 (0.0307)	0.0266 (0.0068)	0.1707 (0.0293)	7.4154 (1.8853)	306.2	303.0
	Random effects $v_{1i} = v_{2i}$	0.2012 (0.0298)	0.0233 (0.0051)	0.1779 (0.0251)	8.6212 (0.8395)	356.3	354.0
	Correlation $\rho = 0$	0.1991 (0.0313)	0.0255 (0.0066)	0.1736 (0.0320)	7.8038 (2.3676)	305.2	302.4
	Correlation $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$	0.1999 (0.0312)	0.0252 (0.0059)	0.1748 (0.0314)	7.9476 (2.1669)	302.2	299.9
Bivariate complementary log–log GLMM							
	No change	0.2037 (0.0353)	0.0224 (0.0063)	0.1813 (0.0353)	9.0950 (2.8922)	306.2	303.0
	Random effects $v_{1i} = v_{2i}$	0.2113 (0.0361)	0.0197 (0.0055)	0.1916 (0.0355)	10.7288 (3.2038)	373.3	371.0
	Correlation $\rho = 0$	0.2051 (0.0359)	0.0209 (0.0059)	0.1842 (0.0375)	9.9088 (3.5280)	305.4	302.6
	Correlation $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$	0.2082 (0.0364)	0.0225 (0.0057)	0.1857 (0.0376)	9.2608 (3.0130)	302.6	300.3
Bivariate Beta-binomial model							
	No change	0.2039 (0.0308)	0.0222 (0.0056)	0.1817 (0.0313)	9.1890 (2.6929)	307.7	307.0
	Correlation $\rho = 0$	0.2041 (0.0308)	0.0222 (0.0054)	0.1819 (0.0313)	9.2114 (2.6571)	304.7	303.4

AIC_C = the finite sample corrected Akaike's information criterion and BIC = the Bayesian information criterion. $\hat{P}_1$ = the risk of type 2 Diabetes Mellitus with gestational diabetes, $\hat{P}_2$ = the risk of type 2 Diabetes Mellitus without gestational diabetes. $\widehat{RD} = \hat{P}_1 - \hat{P}_2$; $\widehat{RR} = \hat{P}_1 / \hat{P}_2$.

risk of MI. In summary, 86 out of 16 856 in the rosiglitazone group and 72 out of 12 962 in the control group had MI event in the 48 clinical trials.

Similar to Section 3.1, we fitted the bivariate Beta-binomial and the BGLMMs as described in Section 2 on those 48 clinical trials to study the association between rosiglitazone usage and the risk of MI in type 2 diabetes mellitus patients. Table 3 presents the parameter estimates and their SEs, including the population averaged risk of MI event for those with and without rosiglitazone usage, the population averaged RD and RR, and the goodness-of-fit measurements including the finite sample corrected AIC_C and the BIC. The BGLMM models assuming random effects $v_{1i} = v_{2i}$ with any of the three link functions provide better model fit than the bivariate Beta-binomial model with either $\rho \neq 0$ or $\rho = 0$. Based on AIC_C and BIC, the bivariate logit and complementary log–log generalised linear mixed effects models with random effects $v_{1i} = v_{2i}$ provide similar best fit.

It is worthy to mention that for the logit BGLMM model, it seems that the approximation of $E(p_k) \approx \text{expit}(\nu_k/\sqrt{1 + C^2\sigma_k^2})$, where $C = 16\sqrt{3}/(15\pi)$, slightly overestimate the population averaged probability of MI for each group. For example, for the logit BGLMM model assuming random effects $v_{1i} = v_{2i}$, the estimated population averaged probabilities of MI in the rosiglitazone treatment and control groups are 0.00627 (SE = 0.00133) and 0.00480 (SE = 0.00109) using the approximation of $E(p_k) \approx \text{expit}(\nu_k/\sqrt{1 + C^2\sigma_k^2})$. While, using the numerical integration by: $E(p_k) = \int_{-\infty}^{+\infty} 1/[1 + \exp(-\nu_k - z)]\sigma_k^{-1}\phi(z/\sigma_k)\mathrm{d}z$, the corresponding estimates are 0.00493 (SE = 0.00140) and 0.00366 (SE = 0.00114), which are consistent with the estimates from other models. However, we notice that the overestimation of the population averaged probability of MI using the approximation formula of the logit BGLMM does not seem to have any noticeable effects on the estimation of RD or RR.

4 Discussion

In this article, we discussed bivariate Beta-binomial models derived from Sarmanov family of bivariate distributions and BGLMMs using a general link function for meta-analysis of 2×2 tables in comparative clinical studies. Specifically, we have discussed logit, probit and complementary log–log link functions as special cases. These bivariate random effects models naturally account for the potential correlation between treatment (or exposure) and control groups within studies. Moreover, they can be used to make valid inferences using all available data without using *ad hoc* continuity corrections for the sparse data scenario. We illustrated the utilisation of the bivariate random effects models in two recent meta-analysis data sets, which emphasises the importance of model selection. In particular, based on AIC_C and BIC, in the example one, the best fitted model is a bivariate logit generalised linear mixed effects model with $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$, which suggests that the study-specific risks of type 2 diabetes mellitus (in logit scale) for those with and without gestational diabetes are independent and have similar variations. In the example two, both the bivariate logit and complementary log–log generalised linear mixed effects models with random effects $v_{1i} = v_{2i}$ provide similar best fit, which suggests that one can reasonably assume a fixed effect of rosiglitazone on the risk of MI (in logit or complementary log–log scale). Furthermore, we provided methods to estimate the population averaged RD and relative risk. It is worth to noting that the bivariate Beta-binomial model and the BGLMMs involves two different distributional assumptions, one would imagine that their performance would heavily depend on whether the distributional assumptions approximate the underline data generating mechanism. In particular, the BGLMMs implies a linear relationship between p_1 and p_2 on a

Table 3. Point estimates (SEs) for meta-analysis of the risk of MI with rosiglitazone[5].

Model description	Modification from the standard model	Probability $\hat{P}_1$	Probability $\hat{P}_2$	Risk difference $\widehat{RD}$	Relative risk $\widehat{RR}$	BIC	AIC_C
Bivariate probit GLMM							
	No change	0.00493 (0.00091)	0.00359 (0.00079)	0.00133 (0.00075)	1.3710 (0.2455)	262.4	253.7
	Random effects $v_{1i} = v_{2i}$	0.00494 (0.00093)	0.00364 (0.00077)	0.00130 (0.00074)	1.3577 (0.2325)	254.8	249.5
	Correlation $\rho = 0$	0.00486 (0.00095)	0.00379 (0.00914)	0.00107 (0.00132)	1.2811 (0.3973)	272.4	265.3
	Correlation $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$	0.00491 (0.00096)	0.00377 (0.00087)	0.00115 (0.00128)	1.3037 (0.3884)	269.2	263.8
Bivariate logit GLMM							
	No change	0.00624 (0.00131)	0.00484 (0.00115)	0.00141 (0.00086)	1.2929 (0.2023)	261.9	253.2
	Random effects $v_{1i} = v_{2i}$	0.00627 (0.00133)	0.00480 (0.00109)	0.00146 (0.00085)	1.3046 (0.1933)	254.2	248.9
	Correlation $\rho = 0$	0.00618 (0.00140)	0.00508 (0.00141)	0.00110 (0.00199)	1.2164 (0.4359)	272.7	265.7
	Correlation $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$	0.00627 (0.00134)	0.00496 (0.00113)	0.00131 (0.00153)	1.2643 (0.3441)	268.8	263.5
Bivariate complementary log–log GLMM							
	No change	0.00501 (0.00094)	0.00353 (0.00099)	0.00148 (0.00092)	1.4187 (0.3389)	261.9	253.2
	Random effects $v_{1i} = v_{2i}$	0.00499 (0.00094)	0.00373 (0.00088)	0.00126 (0.00075)	1.3390 (0.2399)	254.2	248.9
	Correlation $\rho = 0$	0.00494 (0.00097)	0.00367 (0.00106)	0.00126 (0.00133)	1.3423 (0.4361)	272.7	265.6
	Correlation $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$	0.00497 (0.00097)	0.00384 (0.00092)	0.00113 (0.00130)	1.2945 (0.3885)	268.8	263.5
Bivariate Beta-binomial model							
	No change	0.00491 (0.00090)	0.00385 (0.00138)	0.00111 (0.00126)	1.2915 (0.3823)	277.8	269.9
	Correlation $\rho = 0$	0.00491 (0.00090)	0.00380 (0.00088)	0.00111 (0.00126)	1.2932 (0.3826)	274.0	267.5

AIC_C = the finite sample corrected Akaike's information criterion and BIC = the Bayesian information criterion. $\hat{P}_1$ = the risk of myocardial infarction in the rosiglitazone group, $\hat{P}_2$ = the risk of myocardial infarction in the control group, $\widehat{RD} = \hat{P}P_1 - \hat{P}_2$; $\widehat{RR} = \hat{P}_1/\hat{P}_2$.

transformed scale, and after transforming back to the scale of probability, the relation between p_1 and p_2 is no longer linear. The Beta-binomial model implies a linear relationship between p_1 and p_2 on the original scale. So which model works better in a particular application depends on whether the relation between p_1 and p_2 is linear on the original scale or on the transformed scale. We suggest that fitting both models and comparing goodness-of-fit to select the best model to make inference in practice.

Alternative approaches using Bayesian methods can be fitted by free downloadable software such as WinBUGS, for example, by the Bayesian random effect models as in Warn, Thompson and Spiegelhalter.[40] However, Warn et al.[40] focused on the conditional treatment effects. Here, we focus on the overall marginal (or population averaged) treatment effects, as suggested by McCullagh.[30,31] Remark that our parameterisation of BGLMM is slightly different from the random effects models by Smith et al.[41] and Warn et al.[40] Specifically, Smith et al.[41] considered a Bayesian logit random effects model assuming $\text{logit}(p_{1i}) = \mu_i - \delta_i/2$, $\text{logit}(p_{2i}) = \mu_i + \delta_i/2$, $\delta_i \sim N(\delta, \sigma^2)$, and non-informative priors for the average event rates,μ_is, which are treated as the nuisance parameters. It implicitly restricts Var[logit(p_{1i})]= Var[logit(p_{2i})], i.e. restricting $\sigma_1^2 = \sigma_2^2$ as in the BGLMM model. Warn et al.[40] assumes that $g(p_{1i}) = \mu_i$, $g(p_{2i}) = \mu_i + \delta_i$ and $\delta_i \sim N(\delta, \sigma^2)$ where $g()$ is a link function, which implicitly restricts Var[$g(p_{1i})$] $\leq$ Var[g(p_{2i})], i.e. restricting $\sigma_1^2 \leq \sigma_2^2$ in our BGLMM parameterisation. It is worth pointing out that our purpose here is not to demonstrate the advantage of our approach over a Bayesian approach, because both BGLMM and bivariate Beta-binomial models can be fitted using a Bayesian approach. For the general model that we considered in Equation (2), we do not make any restrictions on the variances of σ_1^2 and σ_2^2. Furthermore, the bivariate Beta-binomial model and the bivariate generalised linear mixed models we proposed in this article do not include any study-level or individual level covariates. It is straightforward to include covariates when using the BGLMM through the SAS NLMIXED procedure.

Acknowledgements

Dr Haitao Chu was supported in part by 1P01CA142538-01 from the US National Cancer Institute. The authors are grateful to the Editor Brian Everitt and two anonymous reviewers for their helpful comments on an earlier version of this manuscript.

References

1. Egger M, Smith GD and Altman DG. *Systematic reviews in health care: meta-analysis in context*. London: BMJ Publishing Group, 2001.
2. Whitehead A and Whitehead J. A general parametric approach to the meta-analysis of randomized clinical-trials. *Stat Med* 1991; **10**(11): 1665–1677.
3. Woolf B. On estimating the relation between blood group and disease. *Ann Hum Genet* 1955; **19**(3): 251–253.
4. Dersimonian R and Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**(3): 177–188.
5. Nissen SE and Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med* 2007; **356**(24): 2457–2471.
6. Sweeting MJ, Sutton AJ and Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004; **23**: 1351–1375.
7. Bradburn MJ, Deeks JJ, Berlin JA and Localio AR. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med* 2007; **26**(1): 53–77.
8. Tian L, Cai T, Pfeffer MA, Piankov N, Cremieux PY and Wei LJ. Exact and efficient inference procedure for meta-analysis and its application to the analysis of independent 2 x 2 tables with all available data but without artificial continuity correction. *Biostatistics* 2009; **10**(2): 275–281.
9. Vandermeer B, Bialy L, Hooton N, et al. Meta-analyses of safety data: a comparison of exact versus asymptotic methods. *Stat Methods Med Res* 2009; **18**(4): 421–432.
10. Rucker G, Schwarzer G, Carpenter J and Olkin I. Why add anything to nothing? The arcsine difference as a measure of treatment effect in meta-analysis with zero cells. *Stat Med* 2009; **28**(5): 721–738.
11. Tang JL. Weighting bias in meta-analysis of binary outcomes. *J Clin Epidemiol* 2000; **53**(11): 1130–1136.
12. Chu H and Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol* 2006; **59**(12): 1331–1332.
13. Riley R, Abrams K, Sutton A, Lambert P and Thompson J. Bivariate random-effects meta-analysis and the

estimation of between-study correlation. *BMC Med Res Methodol* 2007; **7**(1): 3.
14. Harbord RM, Deeks JJ, Egger M, Whiting P and Sterne JAC. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007; **8**(2): 239–251.
15. Riley RD, Thompson JR and Abrams KR. An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. *Biostatistics* 2008; **9**(1): 172–186.
16. Chu H, Nie L, Cole SR and Poole C. Meta-analysis of diagnostic accuracy studies accounting for disease prevalence: alternative parameterizations and model selection. *Stat Med* 2009; **28**(18): 2384–2399.
17. Chu H, Chen S and Louis TA. Random effects models in a meta-analysis of the accuracy of two diagnostic tests without a gold standard. *J Am Stat Assoc* 2009; **104**: 512–523.
18. Chu H, Guo H and Zhou Y. Bivariate random effects meta-analysis of diagnostic studies using generalized linear mixed models. *Med Decis Making* 2010; **30**(4): 499–508.
19. Riley RD, Abrams KR, Lambert PC, Sutton AJ and Thompson JR. An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes. *Stat Med* 2007; **26**(1): 78–97.
20. Jackson D, White IR and Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Stat Med* 2010; **29**: 1282–1297.
21. Bellamy L, Casas JP, Hingorani AD and Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. *The Lancet* 2009; **373**(9677): 1773–1779.
22. Lee MLT. Properties and applications of the Sarmanov family of bivariate distributions. *Commun Stat Theory Methods* 1996; **25**(6): 1207–1222.
23. Sarmanov OV. Generalized normal correlation and two dimensional frechet classes. *Doklady Soviet Mathematics)* 1966; **168**: 596–599.
24. Schwartz LM, Woloshin S and Welch HG. Misunderstandings about the effects of race and sex on physicians' referrals for cardiac catheterization. *N Engl J Med* 1999; **341**(4): 279–283.
25. Cummings P. Early exposure to marijuana and risk of later drug use. *JAMA* 2003; **290**(3): 329–330.
26. Kalilani L and Atashili J. Measuring additive interaction using odds ratios. *Epidemiol Perspect Innovat* 2006; **3**(1): 5.
27. Greenland S. Interpretation and choice of effect measures in epidemiologic analyses. *Am J Epidemiol* 1987; **125**(5): 761–768.
28. Greenland S, Robins JM and Pearl J. Confounding and collapsibility in causal inference. *Stat Sci* 1999; **14**(1): 29–46.
29. Brumback B and Berg A. On effect-measure modification: relationships among changes in the relative risk, odds ratio, and risk difference. *Stat Med* 2008; **27**(18): 3453–3465.
30. McCullagh P. Sampling bias and logistic models. *J Roy Stat Soc B Stat Meth* 2008; **70**: 643–664.
31. Longford NT, Diggle PJ, Nelder J, et al. Sampling bias and logistic models - discussion. *J Roy Stat Soc B Stat Meth* 2008; **70**: 664–677.
32. Zeger SL, Liang KY and Albert PS. Models for longitudinal data - a generalized estimating equation approach. *Biometrics* 1988; **44**(4): 1049–1060.
33. Pinheiro JC and Bates DM. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *J Comput Graph Stat* 1995; **4**(1): 12–35.
34. Burnham KP and Anderson DR. *Model selection and inference: a practical information-theoretic approach*. New York: Springer-Verlag, 1998.
35. Shuster JJ, Jones LS and Salmon DA. Fixed vs random effects meta-analysis in rare event studies: the rosiglitazone link with myocardial infarction and cardiac death. *Stat Med* 2007; **26**(24): 4375–4385.
36. Shuster JJ, Jones LS and Salmon DA. Rebuttal to Carpenter et al. comments on 'fixed vs. random effects meta-analysis in rare event studies: the rosiglitazone link with myocardial infarction and cardiac death' - reply. *Stat Med* 2008; **27**(19): 3912–3914.
37. Carpenter J, Rucker G and Schwarzer G. Comments on 'fixed vs random effects meta-analysis in rare event studies: the rosiglitazone link with myocardial infarction and cardiac death'. *Stat Med* 2008; **27**(19): 3910–3912.
38. Waksman J and Kollar C. Comments on 'Rebuttal to Carpenter et al. 'comments on 'fixed vs random effects meta-analysis in rare event studies: the rosiglitazone link with myocardial infarction and cardiac death'. *Stat Med* 2009; **28**(3): 534–536.
39. Dahabreh IJ. Meta-analysis of rare events: an update and sensitivity analysis of cardiovascular events in randomized trials of rosiglitazone. *Clin Trials* 2008; **5**(2): 116–120.
40. Warn DE, Thompson SG and Spiegelhalter DJ. Bayesian random effects meta-analysis of trials with binary outcomes: methods for the absolute risk difference and relative risk scales. *Stat Med* 2002; **21**(11): 1601–1623.
41. Smith TC, Spiegelhalter DJ and Thomas A. Bayesian approaches to random-effects meta-analysis: a comparative study. *Stat Med* 1995; **14**(24): 2685–2699.

Appendix

Table A1. Point estimates (SEs) for meta-analysis of studies on type 2 diabetes mellitus after gestational diabetes: A sensitivity analysis with study one excluded[21].

Model description	Modification from the standard model	Probability $\hat{P}_1$	Probability $\hat{P}_2$	Risk difference $\widehat{RD}$	Relative risk $\widehat{RR}$	BIC	AIC_C
Bivariate probit GLMM							
	No change	0.2062 (0.0327)	0.0247 (0.0068)	0.1815 (0.0314)	8.3585 (2.2818)	266.1	268.9
	Random effects $v_{1i} = v_{2i}$	0.1821 (0.0254)	0.0351 (0.0084)	0.1471 (0.0183)	5.1959 (0.6773)	291.5	293.6
	Correlation $\rho = 0$	0.2083 (0.0333)	0.0236 (0.0066)	0.1847(0.0339)	8.8363 (2.8455)	265.0	267.6
	Correlation $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$	0.2053 (0.0313)	0.0245 (0.0074)	0.1808 (0.0318)	8.3707 (2.7556)	262.8	264.9
Bivariate logit GLMM							
	No change	0.2017 (0.0328)	0.0278 (0.0076)	0.1739 (0.0314)	7.2505 (1.9855)	266.5	269.3
	Random effects $v_{1i} = v_{2i}$	0.1849 (0.0276)	0.0365 (0.0076)	0.1484 (0.0213)	5.0613 (0.5335)	283.2	285.3
	Correlation $\rho = 0$	0.2034 (0.0333)	0.0268 (0.0074)	0.1765 (0.0342)	7.5806 (2.4442)	265.5	268.1
	Correlation $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$	0.2045 (0.0335)	0.0263 (0.0065)	0.1783 (0.0337)	7.7820 (2.2429)	263.0	265.2
Bivariate complementary log–log GLMM							
	No change	0.2084 (0.0379)	0.0230 (0.0072)	0.1854 (0.0381)	9.0751 (3.1809)	266.7	269.5
	Random effects $v_{1i} = v_{2i}$	0.1925 (0.0327)	0.0352 (0.0070)	0.1573 (0.0292)	5.4641 (0.9310)	282.6	284.7
	Correlation $\rho = 0$	0.2098 (0.0385)	0.0214 (0.0067)	0.1884 (0.0403)	9.8003 (3.8260)	265.8	268.4
	Correlation $\rho = 0$ and $\sigma_1^2 = \sigma_2^2$	0.2132 (0.0393)	0.0235 (0.0062)	0.1897 (0.0403)	9.0566 (3.0188)	263.6	265.7
Bivariate Beta-binomial model							
	No change	0.2085 (0.0327)	0.0229 (0.0061)	0.1856 (0.0332)	9.1019 (2.8148)	269.7	269.8
	Correlation $\rho = 0$	0.2087 (0.0328)	0.0229 (0.0061)	0.1859(0.0333)	9.1243 (2.8347)	266.0	266.9

AIC_C = the finite sample corrected Akaike's information criterion and BIC = the Bayesian information criterion. $\hat{P}_1$ = the risk of type 2 Diabetes Mellitus with gestational diabetes, $\hat{P}_2$ = the risk of type 2 Diabetes Mellitus without gestational diabetes. $\widehat{RD} = \hat{P}_1 - \hat{P}_2$; $\widehat{RR} = \hat{P}_1 / \hat{P}_2$.