

ISSUES WITH THE RANDOM EFFECTS APPROACH IN META-ANALYSIS OF SPARSE INCIDENCE DATA

**ELASMA MILANZI^{1,2}, MATTHEW SPITTAL¹
and LYLE GURRIN^{1,2}**

¹Melbourne School of Population and Global Health
University of Melbourne
Parkville, Victoria
Australia

²Victorian Centre for Biostatistics
Parkville, Victoria
Australia
e-mail: emilanzi@unimelb.edu.au

Abstract

The current interest in meta-analysis of count data in which some studies have zero events (sparse data) has led to re-assessment of commonly used meta-analysis methods to establish their validity in such scenarios. The general consensus is that methods which exclude studies with zero events should be avoided. In the family of parametric methods, random effects models come out highly recommended. While acknowledging the strength of this approach, one

2020 Mathematics Subject Classification: 62P10.

Keywords and phrases: meta-analysis, misspecification, random effects, sparse data, bias.

Received January 31, 2020; Revised March 06, 2020

© 2020 Scientific Advances Publishers

This work is licensed under the Creative Commons Attribution International License (CC BY 3.0).

http://creativecommons.org/licenses/by/3.0/deed.en_US

Open Access



of its aspects with potentially undesirable impact on the results, is often overlooked. The random effects approach accounts for the variation in the effect measure across studies by using models with random slopes. It has been shown that parameters associated with a random structure risk being estimated with biased unless the distribution of the random effects is correctly specified. In meta-analysis the parameter of interest, average effect measure, is associated with a random structure (random slope). Information on how the effect measure point and precision estimates are affected by misspecification of random effects distribution is still lacking. To fill in the information gap, we used a simulation study to investigate the impact of misspecification of distribution of random effects in this context and provide guidelines in using the random effects approach. Our results indicated that relative bias in the estimated effect measure could be as high as 30% and 95% confidence interval coverage as low as 0%. These results send a clear message that possible effects of misspecification of the distribution of random effects should not be ignored. In light of these findings, we have proposed a sensitivity analysis that also establishes whether a random slope model is necessary.

1. Introduction

There is an emerging interest in evaluation of methods for meta-analysis of incidence rates data in which some of the studies have no events. The inverse-variance method introduced by DerSimonian & Laird [1] is one of the widely used approaches to meta-analysis but it breaks down when some studies have zero events [2, 3]. Proposed solutions include omitting studies with zero events or adding a correction factor. However, previous work has shown that such methods tend to give a biased estimate of the average effect measure of interest, for example, incidence rate ratio for rate data [4-6].

Several approaches that avoid omitting studies with zero events or adding a correction factor have been proposed [2, 7-9] and out of these, the random effects (mixed models) approach comes highly recommended. Random effects are usually either used to acknowledge that event counts from the same study may be more similar than those from a different study or to account for the variation of intervention effect across different studies.

Conspicuously missing from the discussion of random effects approach in meta-analysis of sparse data is the need to correctly specify the distribution of the random effects which may result into twofold problems. Firstly, [10, 11] showed that fixed effects estimates for covariates associated with a random effects structure can be biased if the random effects distribution is misspecified. Further, [12] proved that if covariates are not associated with a random effects structure, and the true effect of such covariates is zero then the corresponding maximum likelihood estimators will estimate zero, but this is not generally true for covariates associated with a random effects structure. Based on these findings, a misspecified random effects distribution would either lead to biased estimates and/or concluding that there is an effect when it doesn't exist (inflation of type I error).

Recall that a random effects approach avoids omitting studies with zero events or adding an arbitrary correction factor, making it an important modelling tool for sparse data. Additionally, accounting for potential heterogeneity across studies is an important aspect of meta-analysis. In the random effects approach this is achieved through a mixed model with a random slope, implying that the covariate for treatment (intervention) effect is associated with a random structure. It follows that the a treatment fixed effect estimated from such a model is prone to bias and the corresponding inferences may be invalid due to inflation of type I error.

A lot of work has been done in assessing the impact of misspecification of random effects distribution [10-15]. However most of this work was done in the realm of longitudinal or clustered data where the parameter of interest was not associated with a random effects structure and cluster sizes larger than two. In these scenarios, fixed effects estimates for parameters of interest remain unbiased and the inferences are valid.

On the other hand, the extent to which misspecification of random effects distribution impacts the estimates of effect measure in meta-analysis for sparse data is not well documented. Our goal is therefore to assess the impact of misspecifying a random effects distribution on the estimate of an effect measure from meta-analysis of sparse data when intervention effects are suspected to vary across studies. Based on the findings recommendations on how to use the random effects approach will be provided. The suicide hot spots data, described in the next section motivated this work and will be used as an example.

The rest of the paper is organized as follows; after introducing the motivating case study in Section 2, we will give a brief recap of popular random effects approaches in Section 3. The design of the simulations studies will be presented in Subsection 4.1 followed by results in Subsection 4.2. The motivating study will be analysed and discussed in Section 5 and we will summarize with a brief discussion in Section 6.

2. Motivating Case Study

Suicides from known jumping hotspots

Our interest in this problem was initiated by the analysis of data in a study evaluating the effectiveness of installing barriers for reducing jumping deaths at known suicide hotspots [16]. Jumps from these sites (bridges, viaducts and cliffs) generally have high fatality rates, can cause significant distress or injury to bystanders and often receive prominent media coverage, increasing the risk of copycat acts [17]. A number of studies have investigated the effectiveness of structural interventions – such as barriers, fences or safety nets – on reducing suicide by jumping at these sites. This is based on the premise that restricting access to means is one of the few successful suicide prevention strategies [18]. Individual studies are therefore before-and-after designs, with the pre-intervention period considered the “control” group and the post-intervention the “intervention” group. A total of 8 studies had previously counted the number of suicide deaths at hotspots in the periods before and after the

installation of barriers and safety nets. The data from the eight studies that examined the number of suicides by jumping before and after the installation of barriers is shown in Table 1. Six of the studies had zero events after the introduction of barriers and exposure time ranged from approximately 5 months to 22 years. The data will be analysed in Section 5.

Table 1. Suicide counts and exposure time by study

Study no.	Pre-intervention		Post-intervention	
	No. of events	Time (years)	No. of events	Time (years)
1	19	6	0	4
2	41	5	20	5
3	221	14	0	0.4
4	25	7	1	5
5	14	22	0	22
6	7	3	0	3
7	96	10	0	5
8	13	10	0	2

3. Methods: Random Effects Models

The Poisson-normal mixed model is commonly used for meta-analysis of incidence rate data due to its easy implementation in many statistical software packages. If we denote Y_{ij} as the events count for the j -th group in the i -th study $i = 1 \dots, n$, $j = 1, \dots, J$, the model is specified as

$$Y_{ij} | \mathbf{b}_i \sim \text{Poi}(\lambda_{ij})$$

$$\ln[\lambda_{ij}] = \beta_0 + b_{i0} + (\beta_1 + b_{i1})X_{ij} + \ln[T_{ij}], \quad (1)$$

where T_{ij} defines the corresponding offset value, e.g., exposure time in our case study. X_{ij} is the intervention covariate which takes the value of 1 for the intervention group, and zero for the control group, hence β_1 is the log of effect measure of interest, log of incidence rate ratio. The study specific random effects with the joint density, $f(\mathbf{b}_i)$ are assumed to follow

bivariate normal distribution: $(b_{i0}, b_{i1}) \sim N(\mathbf{0}, \mathbf{D})$, with b_{i1} accounting for the variation in the incidence rate ratio estimates across studies. The marginal contribution of the i -th study to the likelihood is given by

$$f(Y_i) = \int \prod_{j=1}^J f(Y_{ij} | \mathbf{b}_i) f(\mathbf{b}_i) d\mathbf{b}_i, \quad (2)$$

and estimation proceeds by maximising the likelihood

$$L(\mathbf{Y} | \beta_0, \beta_1, \mathbf{D}) = \prod_{i=1}^n f(Y_i), \quad (3)$$

where n is the number of studies. Our interest is to assess the impact of assuming that $f(\mathbf{b}_i)$ follows a bivariate normal distribution when the true distribution $g(\mathbf{b}_i)$ is non-normal. A combination of Poisson and normal distributions in Model (1) precludes closed forms expressions in estimating the parameters thereby making analytical quantification of bias problematic. A simulation study will therefore be used to study the impact of misspecifying the random effects distribution by quantifying bias and inflation of type I error (through studying the confidence interval coverage).

4. Simulation Study

4.1. Design

We investigated the impact of misspecification of the random effects distribution on the estimation of the log incidence rate ratio in meta-analysis of sparse incidence data through two sets of simulations. First, we generated the data assuming

$$Y_{ij} | \mathbf{b}_i \sim \text{Poi}(\lambda_{ij}),$$

where

$$\ln[\lambda_{ij}] = \beta_0 + b_{i0} + (\beta_1 + b_{i1})X_{ij} + \ln[T_{ij}], \quad (4)$$

with (b_{i0}, b_{i1}) having a probability distribution density $g(\mathbf{b}_i)$. The density $g(\mathbf{b}_i)$ was varied to represent varying levels of skewness, mean and variance of b_{i1} , the random slope which accounts for heterogeneity effect measures across studies. Distributions considered for $g(\mathbf{b}_i)$ included; bivariate normal, bivariate mixtures of normal, beta, gamma, Poisson, and uniform distribution. These were chosen to represent a wide range of skewness and variability. The values of skewness ranged from -1.75 to 9 . Similarly, variance of b_{i1} ranged from 0.04 to 7 . Details of the distributions considered for $g(\mathbf{b}_i)$ are given in Table 2. Except for bivariate normal and mixture of normals, b_{i0} and b_{i1} , were sampled independently due to lack of a straightforward bivariate distribution equivalent for the other distributions. Parameters β_0 and β_1 were set to 1.6 and -4.5 , respectively. These were chosen based on the estimates from the study in Section 2, and ensured that a considerable number of studies have zero events for the intervention group. The average number of studies with zero events in the intervention group are given in Figure 1. The follow-up time was generated as $T_{i0} \sim U(2, 5)$ and $T_{i1} \sim U(2, 10)$, also to correspond to the time range in the case study.

Table 2. Parameter values for the studied distributions of the random effects

Parameter values			Properties of distribution for b_{i1}		
$g(\mathbf{b}_i)$	b_0	b_1	Mean(b_{i1})	Var(b_{i1})	Skewness
Beta	Beta(2, 3)	Beta(1.8, 0.3)	0.86	0.04	- 1.75
Uniform	U(- 3,0)	U(- 5,1)	- 2.00	3.00	0.00
Gamma	Gamma(0.01,1)	Gamma(0.05,1)	0.05	0.05	8.94
Gamma	Gamma(0.4,1)	Gamma(0.6,1)	0.60	0.60	2.58
Poisson	Poisson(1)	Poisson(1)	1.00	1.00	1.00
Poisson	Poisson(0.4)	Poisson(0.8)	0.80	0.8	1.12
Normal	$N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\right]$		0.00	2	0.00
Mixture normals	$0.5.N\left[\begin{pmatrix} -0.5 \\ -0.5 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\right] + 0.5.N\left[\begin{pmatrix} -5 \\ -5 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\right]$		- 2.75	7.06	0.00

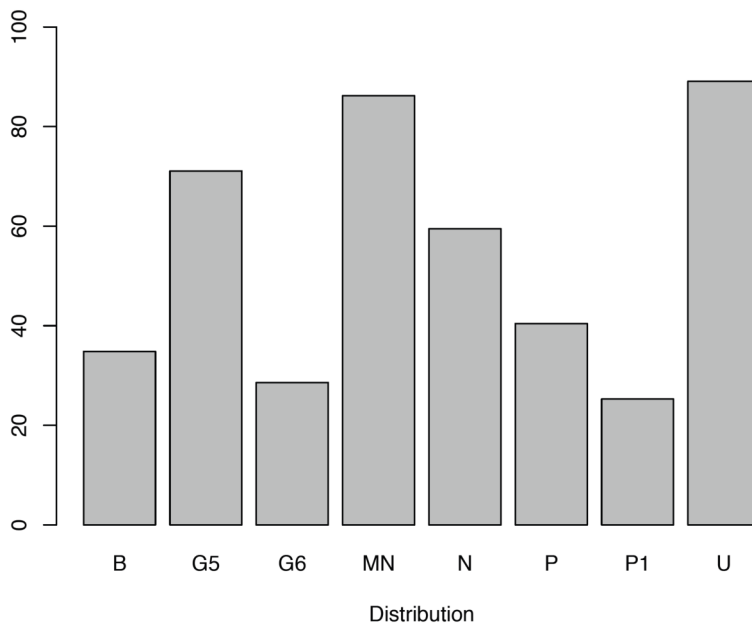


Figure 1. Average (over all simulation runs) percent of studies with zero events in the intervention group) in the different scenarios. B = Beta, G5 = Gamma(0.05,1), MN = Mixture of normals, N = Normal, P = Poisson (0.8), P1 = Poisson(1), and U = Uniform.

Further, we considered a meta-analysis with small number of studies (20), medium (50) and large (100) to assess the impact at different sample sizes.

The generated data were then fitted to Model (1), which will be referred to as the *Poisson-bivariate normal* (PBVN) model. We will assume that $f(\mathbf{b}_i)$ follows a bivariate normal distribution and thus misspecified when the true distribution for \mathbf{b}_i is non-normal. Since previous studies have shown that fixed effects of covariates not associated with a random effects structure are unbiased, there may be a temptation to just disassociate the effect measure (β_1) from random effects structure. This can be done by ignoring heterogeneity across studies, that is omitting b_{i1} from Model (1). To assess the impact of ignoring heterogeneity of the effect measure across the studies when it exists, the generated data was also fitted to random intercept Poisson-normal model which is specified as

$$Y_{ij}|b_i \sim \text{Poi}(\lambda_{ij})$$

$$\ln[\lambda_{ij}] = \beta_0 + b_{i0} + \beta_1 X_{ij} + \ln[T_{ij}], \quad (5)$$

where $b_0 \sim N(0, \sigma^2)$. Finally, we fitted the data using the DerSimonian and Laird random effects approach [1] which does not make explicit assumptions about the distribution of the random effects but omits all studies with zero events. These two approaches represent pragmatic solutions that can be employed to avoid misspecification of a random effects distribution.

The second set of simulations involved generating data by assuming

$$Y_{ij}|\mathbf{b}_i \sim \text{Poi}(\lambda_{ij}),$$

where

$$\ln[\lambda_{ij}] = \beta_0 + b_{i0} + \beta_1 X_{ij} + \ln[T_{ij}], \quad (6)$$

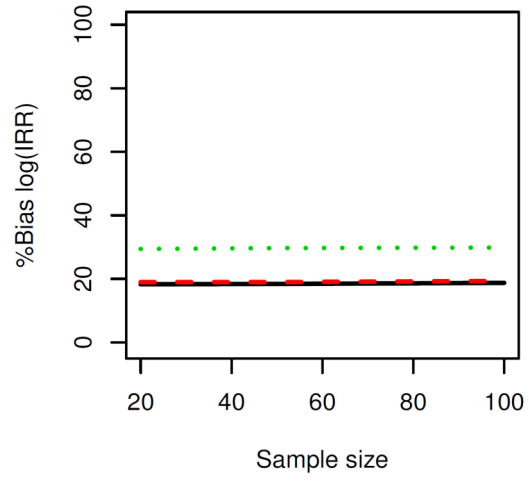
with b_{i0} having a probability distribution density $g(b_i)$, with the same distributions considered in the first set of simulations (the case where $g(b_i)$ is a gamma distribution is equivalent to a negative binomial model). The data was fitted to both Poisson-bivariate normal (PBVN) and random intercept Poisson-normal models. This is similar to simulations largely covered in literature but for the specific scenario of meta-analysis of sparse data. Results from this set of simulations will help identify patterns and or direction of bias, which will inform the sensitivity analysis developed in the later section. Each scenario was simulated a thousand times. The next section discusses results of the simulation study.

4.2. Results

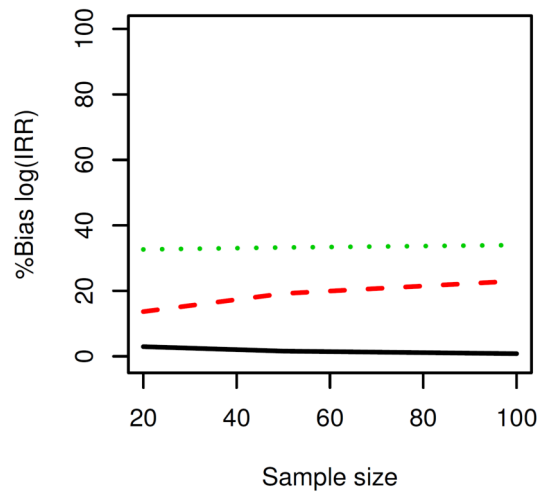
Bias

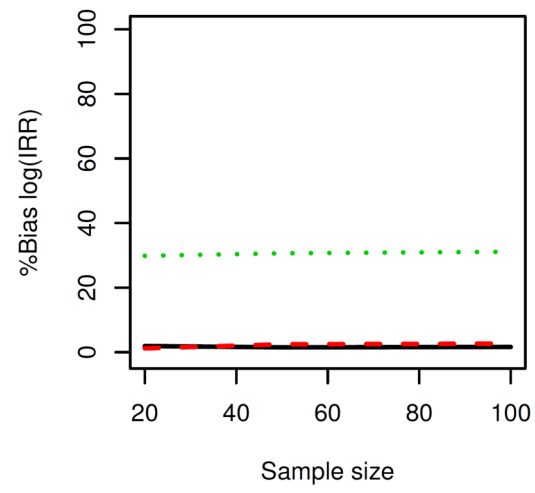
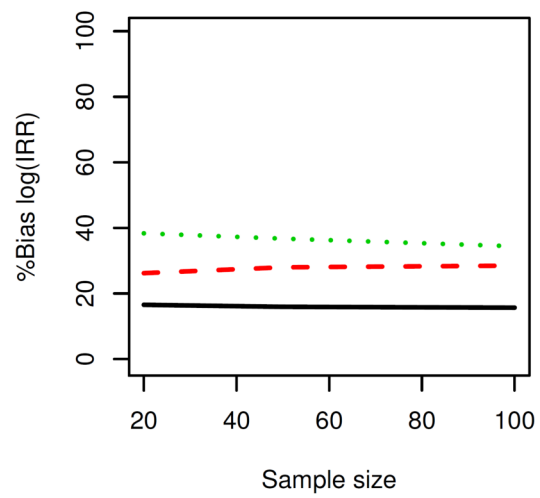
The final estimate for the log incidence rate ratio (IRR), $\hat{\beta}_1$ was obtained as $\frac{1}{S} \sum_{s=1}^S \hat{\beta}_{1s}$, where $S = 1000$ is the total number of simulations. Figure 2 summarises the percent bias obtained as $100 \times (|\hat{\beta}_1 - \beta_1| / \beta_1)$ for the different scenarios.

beta

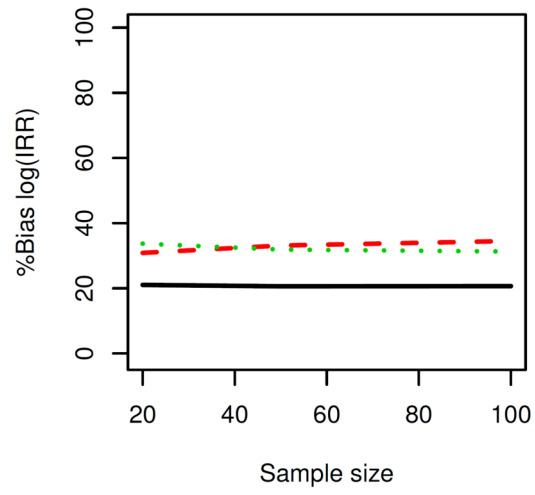


mixnormal

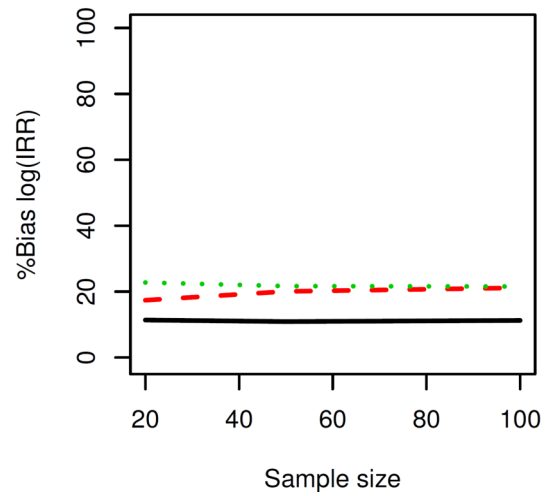


uniform**Poisson**

Poisson1



gamma06



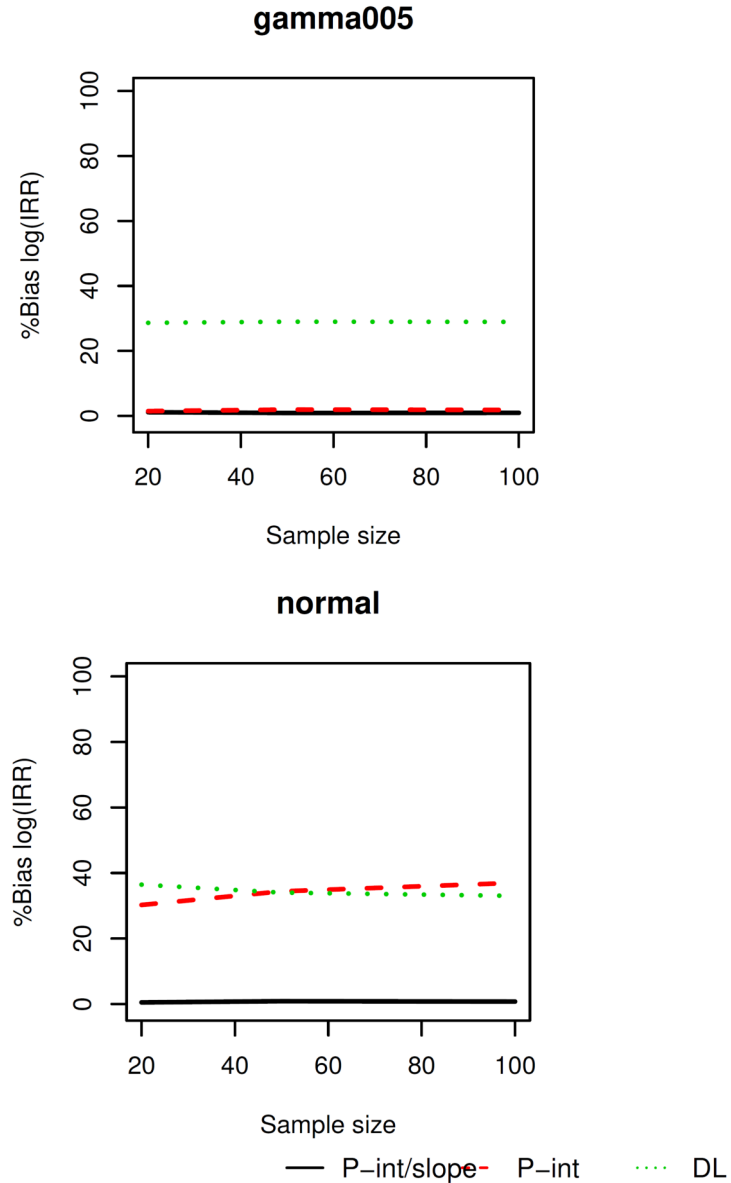


Figure 2. Relative bias (%) obtained as $100 \times (|\hat{\beta}_1 + 4.5| / 4.5)$. Poisson1 refers to the scenario where the b_{i1} random effects were generated from Poisson(1), similarly for gamma06, Gamma(0.6,1) and gamma005, Gamma(0.05,1). Black = Poisson random intercept and slope model, Red = Poisson random intercept model, Green = DL.

As expected, bias is non-existent for the Poisson-bivariate normal (PBVN) model when there is no misspecification, i.e., $g(\mathbf{b}_i) = f(\mathbf{b}_i)$, a bivariate normal distribution.

Little or no bias was observed for the PBVN where $g(\mathbf{b}_i)$ was the bivariate mixture of normals, and uniform probability density distributions. Both uniform and mixture of normals distributions are symmetrical just like the normal distribution. This seems to confirm results from [10] who observed minimal bias when both $g(\mathbf{b}_i)$ and $f(\mathbf{b}_i)$ were symmetrical. However, for data generated with uniform random effects distribution, we correctly specified the correlation structure (independence) between the intercept and slope random effects. When random intercept and slope were allowed to be correlated, bias in the estimated log of IRR was as high as 26% (results not shown), suggesting that symmetry of a distribution is beneficial when the covariance matrix of the random effects is also correctly specified.

For the scenarios where $g(b_{i1})$ was Gamma(0.6, 1), Beta(1.8, 0.3) and Poisson(0.8), relative bias ranged from 15-20% for the PBVN model. All these distributions have skewness in excess of 1 and relatively small variance for b_{i1} . These results suggest that misspecification of the distribution of random effects result into considerable bias for the effect measure.

Results for methods that try to elude misspecification of the random effects by either ignoring the heterogeneity of measures of effect across studies (Poisson-normal intercept (PNI) model), or omitting studies with zero events (inverse variance methods), reveal that such methods should be avoided. Relative bias of around 30% was observed when PNI model was fitted to data generated from PBVN model. These results agree with [19] who noted that ignoring the correlation between random effects and covariates can result into serious bias for the fixed covariates estimates. Average estimates of log IRR and relative bias estimated from PBVN and PNI models are given in Table 3.

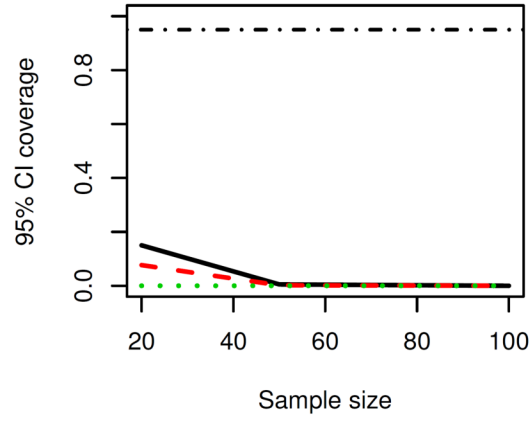
Similarly, the DerSimonian and Laird models which omits studies with zero events had bias of greater than 30% in all scenarios which is also in agreement with previous findings [5], [9], [3].

Confidence interval coverage

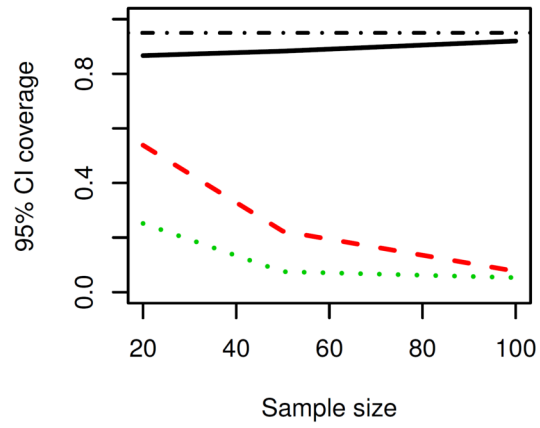
We further looked at the 95% confidence interval coverage for scenarios that exhibited bias, that is where $g(\mathbf{b}_i)$ was Beta, Poisson and Gamma(0.6, 1) distribution. Figure 3 shows that the coverage was close to 0% even for PBVN model. The highest coverage observed in these scenarios was 50%. These results indicate that use of a random effects approach in the meta-analysis of sparse data is delicate and underscores the need to conduct a sensitivity analysis for the distribution.

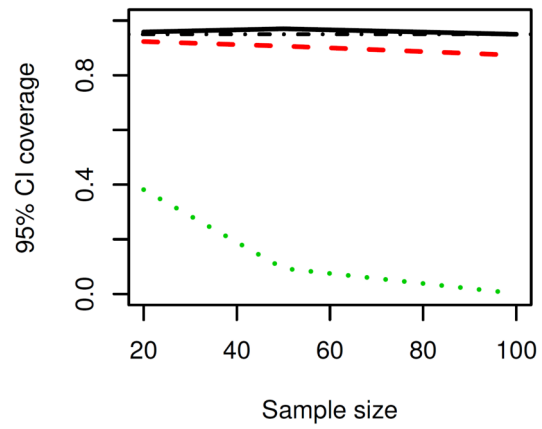
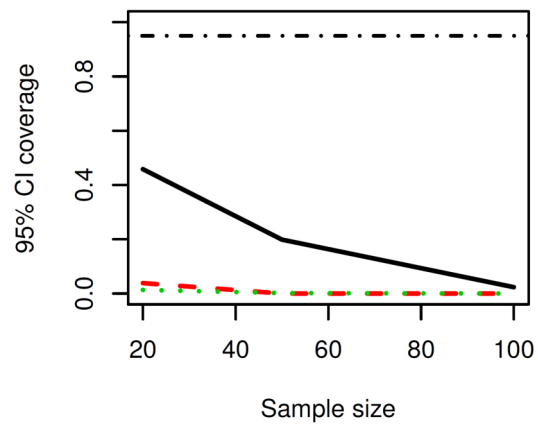
Results from the second set of simulations also in Table 3, were in line with the available literature where no or little bias was observed in the effect measure estimated both from PNI and PBVN models.

beta

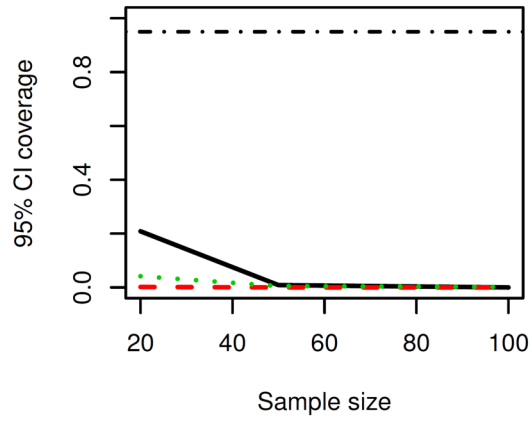


mixnormal

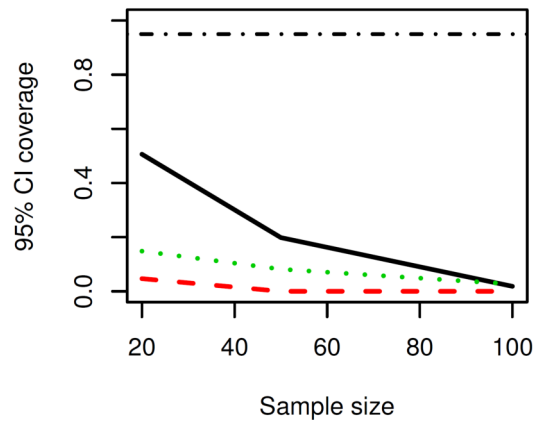


uniform**Poisson**

Poisson1



gamma06



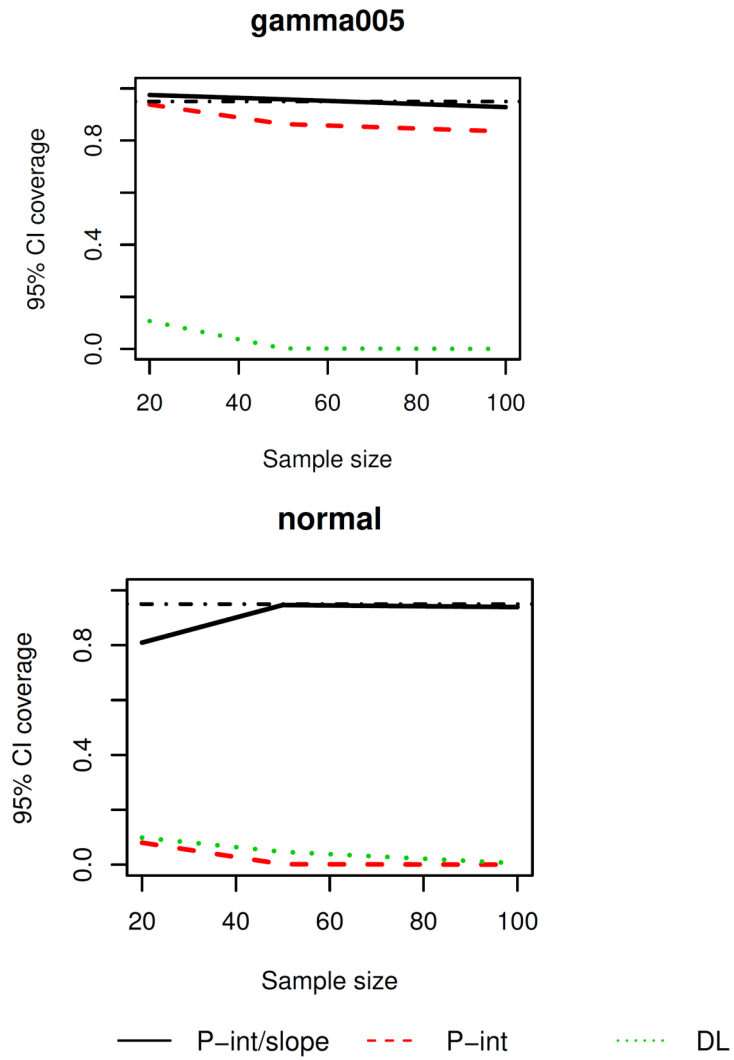


Figure 3. 95% Confidence interval coverage for β_1 . Poisson1 refers to the scenario where the b_{i1} random effects were generated from Poisson(1), similarly for gamma06, Gamma(0.6,1) and gamma005, Gamma(0.05,1). Black = Poisson random intercept and slope model, Red = Poisson random intercept model, Green= DL.

Table 3. The average estimates $\hat{\beta}_1$ (std.err) as obtained from the simulation study. $g(\mathbf{b}_i)$ is the true distribution of random effects. The column “Model” is the model that was fitted generated from the model in “Generating model” column. PBNV is Poisson-bivariate normal, and PNI is Poisson normal intercept model

$g(\mathbf{b}_i)$	Model	Generating model			
		PBNV		PNI	
		$\hat{\beta}_1$ (std.err)	%bias	$\hat{\beta}_1$ (std.err)	%bias
Poisson(0.8)	PBVN	- 3.805 (0.171)	15.4	- 4.537 (0.145)	0.83
	PNI	- 3.185 (0.068)	29.2	- 4.512 (0.127)	0.28
Poisson(1)	PBVN	- 3.575 (0.146)	20.6	- 4.508 (0.083)	0.18
	PNI	- 2.892 (0.037)	35.7	- 4.503 (0.078)	0.06
Beta	PBVN	- 3.659 (0.107)	18.7	- 4.551 (0.171)	1.14
	PNI	- 3.633 (0.094)	19.3	- 4.510 (0.144)	0.22
Gamma(0.05,1)	PBVN	- 4.571 (0.230)	1.68	- 4.584 (0.222)	1.86
	PNI	- 4.415 (0.168)	1.89	- 4.519 (0.176)	0.42
Gamma(0.6,1)	PBVN	- 4.000 (0.138)	11.1	- 4.505 (0.069)	0.11
	PNI	- 3.435 (0.038)	23.7	- 4.501 (0.063)	0.02
Mixture of normals	PBVN	- 4.526 (0.420)	0.58	- 4.549 (0.235)	1.09
	PNI	- 3.401 (0.125)	24.4	- 4.519 (0.204)	0.42
Normal	PBVN	- 4.500 (0.298)	0.00	- 4.515 (0.120)	0.34
	PNI	- 2.775 (0.050)	38.3	- 4.501 (0.109)	0.03
Uniform	PBVN	- 4.640 (0.421)	0.31	- 4.641 (0.390)	3.13
	PNI	- 4.425 (0.273)	1.67	- 4.523 (0.285)	0.51

5. Case Study Analysis

5.1. Suicide spots

The case study introduced in Section 2 was analysed in [3] and [16]. The latter [16] reported a pooled incidence rate ratio of 0.14 with 95% CI 0.09 to 0.21 obtained by using the inverse variance method. The former used the random effects approach whose results will be replicated in this section and we will further propose an analysis routine that should be followed when meta-analysing sparse data.

Given the impact of misspecification of the random effects distribution discussed in Subsection 4.2, the first step should be to establish if it is necessary to associate the effect measure with a random effect. To do this we propose to utilise the knowledge that if there is little or no heterogeneity in the effect measure across studies, random intercept only model will produce an unbiased estimate for the effect measure even under misspecification. In addition, the model with both random intercept and slope also give unbiased estimates when there is no heterogeneity across studies, see Table 3. We therefore propose to perform a sensitivity analysis by fitting the data to both random intercept only and random intercept and slope models, and do this assuming at least two different sampling distributions for the random effects. If there is no heterogeneity across studies, we should obtain similar estimates for the effect measure from all models. In such a case a random intercept model will suffice. This form of sensitivity analysis was performed for the hot suicide spots data and the random intercept only and random intercept and slope models were defined as

$$\ln[\lambda_{ij}] = \beta_0 + b_{i0} + \beta_1 X_{ij} + \ln[T_{ij}], \quad (7)$$

$$\ln[\lambda_{ij}] = \beta_0 + b_{i0} + (\beta_1 + b_{i1})X_{ij} + \ln[T_{ij}], \quad (8)$$

where $\mathbf{b}_i \sim g(\mathbf{b}_i)$, and X_{ij} is defined as 1 for the intervention group and 0 for the control group. Distributions considered for $g(\mathbf{b}_i)$ were normal,

gamma, beta, and uniform distributions. The probability integral transform method proposed by [20] and implemented through SAS NLMIXED procedure was used to fit the non-normal random effects. SAS codes are provided in the Appendix.

The first four rows of Table 4 give results for Model (7) assuming the above mentioned distributions for the random effects and the remaining rows are for Model (8). It is evident that the estimates for the log of incidence rate ratio (β_1) are different for the two models. This suggest the need to take into account heterogeneity across studies and make conclusions based on Model (8).

Table 4. Results for fitting the models to the case study on suicides from known jumping hotspots

Model		Parameter estimates			
	$g(\mathbf{b}_i)$	β_0 (std.err)	β_1 (std.err)	$\text{Var}(b_{0i})$ (std.err)	$\text{Var}(b_{1i})$ (std.err)
P-Intercept	Normal	1.263 (0.390)	- 1.968 (0.229)	1.163 (0.603)	-
	Gamma	2.206 (0.057)	- 1.992 (0.226)	1.000 (0.141)	-
	Beta	2.863 (0.187)	- 1.979 (0.229)	0.417 (0.312)	-
	Uniform	1.956 (21.67)	- 2.030 (0.227)	0.799 (140.5)	-
P-Bivariate	Normal	1.290 (0.370)	- 4.993 (1.942)	1.039 (0.542)	6.388 (7.585)
	Gamma	1.797 (0.057)	- 2.551 (0.781)	1.318 (0.074)	0.150 (0.007)
	Beta	2.876 (0.219)	- 0.440 (0.531)	0.067 (0.027)	0.042(0.028)
	Uniform	2.457 (8.218)	- 1.639 (58.140)	0.162 (2.662)	0.261(30.298)

The estimate for log of incidence rate ratio ranged from -4.5 to -0.4 representing a 99% to 32% reduction in suicide rates by the intervention. This poses the challenge of determining which model to use. It is not straightforward to decipher this from sensitivity results hence we recommend reporting point and precision estimates from all models. Expert opinion should also play a role in deciding on reasonable models. For example, results from a model with Uniform random effects seem less reliable due to relatively large precision estimates. Alternatively, one can report the more conservative result, for example, in this case we would report the model with the least reduction in suicide rate, the model with Beta random effects.

6. Discussion

One of the hurdles in meta-analysis of sparse event data is how to handle studies with zero events. A random effects approach is often used to avoid bias arising from omitting such studies or use of arbitrary correction terms. Our aim was to highlight the drawback of this approach in the specific yet common setting where there is suspected heterogeneity of the effect measure across studies. For the random effects approach such heterogeneity is accounted for by associating the parameter for the effect measure with a random effects structure. This requires correct specification of the distribution of random effects, otherwise an estimate of effect measure would be biased and its corresponding type I error inflated. Results from simulation studies in the context of incidence rate data show that the relative bias for the estimate of log of incidence rate ratio (IRR) could be as high as 20% when the correct model is used but the random effects distribution is misspecified. This could go up to 30% if both the model and random effects distribution are misspecified.

We further investigated coverage for the 95% confidence intervals under misspecification of the random effects distribution to assess inflation of type I error. The percent coverage was as low as 0% for some of the studied scenario.

To this end, we have proposed a sensitivity analysis routine in meta-analysis of sparse data where the first step is to establish whether there is substantial heterogeneity of the effect measure. This is achieved by comparing estimates of the effect measure from PNI and PBVN models, which based on previous results and confirmed with our simulations, should be similar if there is no heterogeneity of the effect measure across studies.

Our recommendation is that meta-analysis for sparse data using the random effects approach should always be accompanied by such a sensitivity analysis.

Acknowledgements

This work was supported by the Victorian Centre for Applied Biostatistics (“ViCBiostat”), an Australian National Health and Medical Research Council (NHMRC) Centre for Research Excellence, Application No. 1035261.

References

- [1] R. DerSimonian and N. Laird, Meta-analysis in clinical trials, *Controlled Clinical Trials* 7(3) (1986), 177-188.
DOI: [https://doi.org/10.1016/0197-2456\(86\)90046-2](https://doi.org/10.1016/0197-2456(86)90046-2)
- [2] T. Cai, L. Parast and L. Ryan, Meta-analysis for rare events, *Statistics in Medicine* 29(20) (2010), 2078-2089.
DOI: <https://doi.org/10.1002/sim.3964>
- [3] M. J. Spittal, J. Pirkis and L. C. Gurrin, Meta-analysis of incidence rate data in the presence of zero events, *BMC Medical Research Methodology* 15 (2015); Article 42.
DOI: <https://doi.org/10.1186/s12874-015-0031-0>

- [4] M. J. Sweeting, A. J. Sutton and P. C. Lambert, What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data, *Statistics in Medicine* 23(9) (2004), 1351-1375.
DOI: <https://doi.org/10.1002/sim.1761>
- [5] O. Kuss, Statistical methods for meta-analyses including information from studies without any events—add nothing to nothing and succeed nevertheless, *Statistics in Medicine* 34(7) (2015), 1097-1116.
DOI: <https://doi.org/10.1002/sim.6383>
- [6] J. Cheng, E. Pullenayegum, J. K. Marshall, A. Iorio and L. Thabane, Impact of including or excluding both-armed zero-event studies on using standard meta-analysis methods for rare event outcome: A simulation study, *BMJ Open* 6(8) (2016), 1-10.
DOI: <http://dx.doi.org/10.1136/bmjopen-2015-010983>
- [7] J. J. Shuster and M. A. Walker, Low-event-rate meta-analyses of clinical trials: Implementing good practices, *Statistics in Medicine* 35(14) (2016), 2467 -2478.
DOI: <https://doi.org/10.1002/sim.6844>
- [8] D. Bohning, K. Mylona and A. Kimber, Meta-analysis of clinical trials with rare events, *Biometrical Journal* 57(4) (2015), 633-648.
DOI: <https://doi.org/10.1002/bimj.201400184>
- [9] T. Stijnen, T. H. Hamza and P. Ozdemir, Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data, *Statistics in Medicine* 29(29) (2010), 3046-3067.
DOI: <https://doi.org/10.1002/sim.4040>
- [10] C. E. McCulloch and J. M. Neuhaus, Misspecifying the shape of a random effects distribution: Why getting it wrong may not matter, *Statistical Science* 26(3) (2011), 388-402.
DOI: <https://doi.org/10.1214/11-STS361>
- [11] J. M. Neuhaus, C. E. McCulloch and R. Boylan, Estimation of covariate effects in generalized linear mixed models with a misspecified distribution of random intercepts and slopes, *Statistics in Medicine* 3142 (2013), 2419-2429.
DOI: <https://doi.org/10.1002/sim.5682>
- [12] S. Litiere, A. Alonso and G. Molenberghs, The impact of a misspecified random-effects distribution on the estimation and the performance of inferential procedures in generalized linear mixed models, *Statistics in Medicine* 27(16) (2008), 3125-3144.
DOI: <https://doi.org/10.1002/sim.3157>

- [13] E. Milanzi, A. Alonso and G. Molenberghs, Ignoring overdispersion in hierarchical loglinear models: Possible problems and solutions, *Statistics in Medicine* 31(14) (2012), 1475-1482.
DOI: <https://doi.org/10.1002/sim.4482>
- [14] S. Litiere, A. Alonso and G. Molenberghs, Type I and type II error under random-effects misspecification in generalized linear mixed models, *Biometrics* 63(4) (2007), 1038-1044.
DOI: <https://doi.org/10.1111/j.1541-0420.2007.00782.x>
- [15] G. Verbeke and E. Lesaffre, The effect of misspecifying the random-effects distribution in linear mixed models for longitudinal data, *Computational Statistics & Data Analysis* 23(4) (1997), 541-556.
DOI: [https://doi.org/10.1016/S0167-9473\(96\)00047-3](https://doi.org/10.1016/S0167-9473(96)00047-3)
- [16] J. Pirkis, M. J. Spittal, G. Cox, J. Robinson, Y. T. D. Cheung and D. Studdert, The effectiveness of structural interventions at suicide hotspots: A meta-analysis, *International Journal of Epidemiology* 42(2) (2013), 541-548.
DOI: <https://doi.org/10.1093/ije/dyt021>
- [17] Y. Y. Chen, K. C. C. Yu and P. S. F. Yip, Suicide Prevention Through Restricting Access to Suicide Means and Hotspots In: *International Handbook for Suicide Prevention: Research, Policy and Practice*, Wiley-Blackwell, Chichester, 2011.
DOI: <https://doi.org/10.1002/9781119998556.ch31>
- [18] J. J. Mann, A. Apter, J. Bertolote, A. Beautrais, D. Currier et al., Suicide prevention strategies: A systematic review, *JAMA* 294(16) (2005), 2064-2074.
DOI: <https://doi.org/10.1001/jama.294.16.2064>
- [19] C. E. McCulloch, S. R. Searle and J. M. Neuhaus, *Generalized Linear, and Mixed Models*, Wiley & Sons, 2008.
- [20] K. P. Nelson, S. R. Lipsitz, G. M. Fitzmaurice, J. Ibrahim, M. Parzen and R. Strawderman, Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects, *Journal of Computational and Graphical Statistics* 15(1) (2006), 39-57.
DOI: <https://doi.org/10.1198/106186006X96854>



Appendix**SAS code for the case study**

```
/*Modelling: Sensitivity Analysis*/
```

```
/*Obtain starting values*/
```

```
proc nlmixed data=m.main maxiter=5000 qpoints=40;
eta = beta0 + (beta1)*group + log(time);
lambda = exp(eta);
model deaths ~ poisson(lambda);
estimate 'rr' exp(beta1);
run;
```

```
/*Normal distribution*/
```

```
/*Bivariate*/
```

```
proc nlmixed data=m.main maxiter=5000 qpoints=40;
parms beta0=1.7 beta1=-2.5;
eta = beta0 + b0 + (beta1+b1)*group + log(time);
lambda = exp(eta);
model deaths ~ poisson(lambda);
random b0 b1 ~ normal([0,0],[d11**2, d12, d22**2]) subject=id;
estimate 'var1' d11**2;
estimate 'var2' d22**2;
estimate 'rr' exp(beta1);
run;
```

```
/*Random intercept*/
```

```
proc nlmixed data=m.main maxiter=5000 qpoints=40;
parms beta0=1.7 beta1=-2.5;
eta = beta0 + b0 + (beta1)*group + log(time);
lambda = exp(eta);
model deaths ~ poisson(lambda);
random b0 ~ normal(0,d11**2) subject=id;
estimate 'var1' d11**2;
estimate 'rr' exp(beta1);
run;
```

```
/*Gamma Distribution*/
```

```
/*Bivariate*/
```

```
proc nlmixed data=main maxiter=5000 noad qpoints=80;
parms beta0=1.7 beta1=-1.9;
p_1 = CDF( 'NORMAL' , b0) ;
p_2 = CDF( 'NORMAL' , b1) ;
if (p_1 > 0.999999) then p_1 = 0.999999 ;
if (p_2 > 0.999999) then p_2 = 0.999999 ;
g_i1 = quantile( 'GAMMA' , p_1, 1/bb1) ;
g_i2 = quantile( 'GAMMA' , p_2, 1/bb2) ;
g_ia = bb1*g_i1;
g_ib = bb2*g_i2;
```

```

eta=beta0 + log(g_ia) + betal*group + log(g_ib)*group + log(time);
lambda=exp(eta);
model deaths ~ poisson(lambda);
random b0 b1 ~ normal([0,0],[1,0,1]) subject=id;
estimate 'rr' exp(betal);
estimate 'var1' 1/bb1;
estimate 'var2' 1/bb2;
run;

```

```

/*Random intercept model*/

```

```

proc nlmixed data=main maxiter=5000 noad qpoints=80;
parms beta0=1.7 betal=-2.5;
p_1 = CDF( 'NORMAL' ,b0) ;
if (p_1 > 0.999999) then p_1 = 0.999999 ;
g_il = quantile( 'GAMMA' ,p_1,1/bb1) ;
g_ia = bb1*g_il;
eta=beta0 + log(g_ia) + betal*group + log(time);
lambda=exp(eta);
model deaths ~ poisson(lambda);
random b0 ~ normal(0,1) subject=id;
estimate 'rr' exp(betal);
estimate 'var1' 1/bb1;
run;

```

```
/*Beta Distribution*/
```

```
/*Bivariate*/
```

```
proc nlmixed data=main maxiter=5000 noad qpoints=80;
parms beta0=1.7 betal=-2.5;
p_1 = CDF( 'NORMAL' , b0) ;
p_2 = CDF( 'NORMAL' , b1) ;
if (p_1 > 0.999999) then p_1 = 0.999999 ;
if (p_2 > 0.999999) then p_2 = 0.999999 ;
g_i1 = quantile( 'Beta' , p_1, a1, bb1) ;
g_i2 = quantile( 'Beta' , p_2, a2, bb2) ;
eta=beta0 + log(g_i1) + betal*group + log(g_i2)*group + log(time);
lambda=exp(eta);
model deaths ~ poisson(lambda);
random b0 b1 ~ normal([0,0],[1,0,1]) subject=id;
estimate 'rr' exp(betal);
run;
```

```
/*Random intercept*/
```

```
proc nlmixed data=main maxiter=5000 noad qpoints=80;
parms beta0=1.7 betal=-2.5;
p_1 = CDF( 'NORMAL' , b0) ;
if (p_1 > 0.999999) then p_1 = 0.999999 ;
g_i1 = quantile( 'Beta' , p_1, a1, bb1) ;
```



```

eta = beta0 + log(g_il) + betal*group + log(time);
lambda = exp(eta);
model deaths ~ poisson(lambda);
random b0 ~ normal(0,1) subject=id;
estimate 'rr' exp(betal);
estimate 'var1' (a1/bb1)/(a1 + bb1)**2 * (a1 + bb1 + 1);
run;

/*Uniform Distribution*/

/*Bivariate*/
proc nlmixed data=main maxiter=5000 noad qpoints=80;
parms beta0=1.7 betal=-2.5;
p_1 = CDF( 'NORMAL' , b0) ;
p_2 = CDF( 'NORMAL' , b1) ;
if (p_1 > 0.999999) then p_1 = 0.999999 ;
if (p_2 > 0.999999) then p_2 = 0.999999 ;
g_ilq = quantile( 'uniform' , p_1) ;
g_i2q = quantile( 'uniform' , p_2) ;
g_il = a1 + exp(log(bb1))*g_ilq; g_i2 = a2 + exp(log(bb2))*g_i2q;
eta = beta0 + log(g_il) + betal*group + log(g_i2)*group + log(time);
lambda = exp(eta);
model deaths ~ poisson(lambda);
random b0 b1 ~ normal([0,0],[1,0,1]) subject=id;

```

```
estimate 'rr' exp(beta1);
estimate 'var1' (bb1)**2/12;
estimate 'var2' (bb2)**2/12;
run;

/*Random intercept*/
proc nlmixed data=main maxiter=5000 noad qpoints=80;
parms beta0=1.7 beta1=-2.5;
p_1 = CDF( 'NORMAL' , b0) ;
if (p_1 > 0.999999) then p_1 = 0.999999 ;
g_ilq = quantile( 'uniform' , p_1) ;
g_il = a1 + exp(log(bb1))*g_ilq;
eta = beta0 + log(g_il) + beta1*group + log(time);
lambda = exp(eta);
model deaths ~ poisson(lambda);
random b0 ~ normal(0,1) subject=id;
estimate 'rr' exp(beta1);
estimate 'var1' (bb1)**2/12;
run;
```