Ignoring overdispersion in hierarchical loglinear models: Possible problems and solutions

Ignoring overdispersion in hierarchical loglinear models: Possible
problems and solutions
Peer-reviewed author version

Made available by Hasselt University Library in

Ignoring overdispersion in hierarchical loglinear models: Possible problems and solutions

Elasma Milanzi[1], Ariel Alonso[2] and Geert Molenberghs[1,3]

[1] I-BioStat, Universiteit Hasselt, B-3590 Diepenbeek, Belgium

[2] Department of Methodology and Statistics. Maastricht University. The Netherlands

[3] I-BioStat, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium

Abstract

Poisson data frequently exhibit overdispersion and, for univariate models, many options exist to go around this problem. Nonetheless, in complex scenarios like, for example, in longitudinal studies, accounting for overdispersion is a more challenging task. Recently, [1] presented a model that accounts for overdispersion by combining two sets of random effects. However, introducing a new set of random effects implies additional distributional assumptions for intrinsically unobservable variables. Using the combined model as a framework, we explored the impact of ignoring overdispersion in complex longitudinal settings via simulations. Furthermore, we also evaluated the effect of misspecifying the random effects distribution on both the combined model and the classical Poisson hierarchical model. Our results indicate that even though inferences may be affected by ignored overdispersion, the combined model is a promising tool in this scenario.

Keywords: Poisson-normal model; Overdispersion; Hierachical; Combined model; Type I error.

1 Introduction

The introduction of random effects to model correlated responses coming from the same subject, was a milestone contribution to the analysis of complex data [2]. Over the last decades, these hierarchical models have been applied in a multitude of areas like, item response theory [3], toxicology [4], survival analysis [5] and non-linear mixed models [6]. Many of the models used in these fields fall under the umbrella of generalized linear mixed models (GLMM) [7]. Basically, GLMMs are used to account for the heterogeneity that arises from correlated measurements. However, in several applications, there may be additional sources of heterogeneity that can affect our inferences if ignored. Poisson longitudinal data are an archetypical example where heterogeneity may arise, not only from the repeated measurements, but also from additional overdispersion[8].

1

Several studies have explored the impact of misspecifying different aspects of GLMMs on the inferential procedures emanating from them. For instance, [7] addressed the impact of omitting important confounding factors, [9][10] investigated the effect of misspecifying the random effects distribution and [11] assessed the impact of fitting an incomplete multilevel structure. All the previous research has clearly shown that misspecification may seriously affect our conclusions. Along these lines, in the present work we study the effect of ignoring overdispersion in hierarchical loglinear models.

[1] presented a model that deals with overdispersion by introducing an additional set of random effects in the classical Poisson loglinear mixed model. In the following, we shall refer to this model as the combined model. We will use this combined model as a framework to assess the impact of ignoring overdispersion via simulations. Essentially, we will study the impact of the misspecification on the consistency of the maximum likelihood estimators (MLE) and the Type I error rates.

Another important concern that arises when using this type of models, is the distributional assumptions one needs to make for the random effects. Indeed,[12] showed that misspecfying the random effect distribution in a logistic model may result in estimates that are asymptotically biased, though the bias is typically small. In a similar setting [13] found that misspecification of the random effects distribution may produce a loss of efficiency. Through extensive simulations [9] [10] investigated the impact of this misspecification on the consistency of the MLE, the power and Type I error rate of commonly used inferential procedures in GLMM. They observed that, although in most scenarios the estimates of many fixed effects were little affected, the estimates of variance components were severely biased and the power and Type I error rates were also gravely impacted.

Introducing an additional set of random effects obviously implies additional distributional assumptions for intrinsically unobservable latent variables. Therefore, in the present work, we will also explore the impact of misspecifying the random effects distribution on both the classical Poisson loglinear mixed model and the combined model introduced by [1].

The paper is organized as follows, Section 2 will describe the case study. The combined model is presented in Section 3. Section 4 and 5 explains the design and results of simulation studies respectively. A re-analysis of the case studied is carried out in Section 6 and some final conclusions are given in Section 7.

2 Case study

The data come from a randomized, double-blind, parallel group and multi-center clinical trial for the comparison of placebo with a new anti-epileptic drug (AED), in combination with one or two other AED's. The study is described in full detail in [14]. Randomization took place after a 12-week baseline period that served as a stabilization time for the use of AED's, and during which the number of seizures were counted. After that period, 45 patients were assigned to the placebo group and 44 to the active (new) treatment group. Patients were then measured weekly and after a followed up of 16 weeks (double-blind) they were entered into a long-term open-extension study. Consequently, some patients were followed for up to 27 weeks. The outcome of interest was the number of epileptic seizures experienced during the most recent week. The research question was whether or not the new treatment could reduce the number of epileptic seizures.

Let Y_{ij} represent the number of epileptic seizures patient i experienced during week j of the follow-up period. Further, let t_{ij} be the time-point at which Y_{ij} is measured,

$t_{ij} = 1, 2, \ldots$ until at most 27. Following [4] the next model was used to analyze the data

$$Y_{ij} \sim \text{Poisson}(\kappa_{ij}),$$

$$\ln(\kappa_{ij}) = \begin{cases} (\xi_0 + b_i) + \xi_1 t_{ij} & \text{if placebo} \\ (\xi_2 + b_i) + \xi_3 t_{ij} & \text{if treated} \end{cases}$$

where b_i is assumed to follow $N(0, \sigma_b^2)$. The first part of Table 1 summarizes the main findings. The results indicate that the expected number of seizures significantly decreases over time in both the placebo and experimental group with p-values 0.0017 and 0.0067 respectively . Importantly, the rate of decrease was the same for both groups, i.e., no significant difference between the placebo and the new treatment was detected with a p-value of 0.7115. Obviously, the preceding results are conditional on the validity of the model used for the analysis. In that line, one relevant question is if the previous model suffices to account for all the variability present in the data and the impact of ignoring extra sources of variability on the inferences previously described. We will address this important issue at the end of the manuscript.

3   Combining conjugate and normal random effects

In this section we will briefly introduce the model proposed by [1]. To that effect, let us denote by Y_{ij} the jth outcome in cluster $i = 1, \ldots, N$ with $j = 1, \ldots, n_i$. Furthermore, it will be assumed that, conditionally upon two q-n_i-dimensional vectors of random effects $\boldsymbol{b_i}$ and $\boldsymbol{\theta}_i$, the outcomes Y_{ij} are independent with density function of the form

$$f_i(y_{ij}|\boldsymbol{b_i}, \boldsymbol{\xi}, \theta_{ij}, \phi) = \exp\left\{\phi^{-1}[y_{ij}\lambda_{ij} - \psi(\lambda_{ij})] + c(y_{ij}, \phi)\right\}, \tag{1}$$

where the conditional mean μ_{ij}^c is further modeled as

$$E(Y_{ij}|\boldsymbol{b_i}, \boldsymbol{\xi}, \theta_{ij}) = \mu_{ij}^c = \theta_{ij}\kappa_{ij}. \tag{2}$$

In the preceding expression the random variable $\theta_{ij} \sim \mathcal{G}_{ij}(\vartheta_{ij}, \sigma_{ij}^2)$ with ϑ_{ij} and σ_{ij}^2 denoting the mean and the variance of θ_{ij} respectively and $\kappa_{ij} = g(\boldsymbol{x}'_{ij}\boldsymbol{\xi} + \boldsymbol{z}'_{ij}\boldsymbol{b_i})$. Moreover, it will be typically assumed that $\boldsymbol{b_i} \sim N(\boldsymbol{0}, \boldsymbol{D})$.

It is convenient, but not strictly necessary, to consider that the two sets of random effects $\boldsymbol{\theta}_i$ and $\boldsymbol{b_i}$ are independent of each other. Regarding the components θ_{ij} of $\boldsymbol{\theta}_i$, three useful special cases result from assuming that: (1) they are independent; (2) they are correlated, implying that the collection of univariate distributions $\mathcal{G}_{ij}(\vartheta_{ij}, \sigma_{ij}^2)$ needs to be replaced with a multivariate one; and (3) they are equal to each other, useful in applications with exchangeable outcomes Y_{ij}.

Obviously, parameterization (2) allows for random effects θ_{ij} capturing overdispersion, and formulated directly at the mean scale, whereas κ_{ij} could be considered the GLMM component. The relationship between the mean and natural parameter is

$$\lambda_{ij} = h(\mu_{ij}^c) = h(\theta_{ij}\kappa_{ij}). \tag{3}$$

We can still apply standard GLM ideas, in particular, to derive the mean and variance, combined with iterated-expectation-based calculations. For the mean, it follows that

$$E(Y_{ij}) = E(\theta_{ij})E(\kappa_{ij}) = E[h^{-1}(\lambda_{ij})]. \tag{4}$$

3.1 Combined poisson model for count data

From the general developments above, the Poisson model with gamma and normal random effects naturally follows. By way of overview, let us assemble all model elements

$$Y_{ij} \sim \text{Poisson}(\theta_{ij}\kappa_{ij}),$$
$$\kappa_{ij} = \exp\left(\boldsymbol{x}'_{ij}\boldsymbol{\xi} + \boldsymbol{z}'_{ij}\boldsymbol{b_i}\right),$$
$$\boldsymbol{b_i} \sim N(\boldsymbol{0}, \boldsymbol{D}),$$
$$E(\boldsymbol{\theta}_i) = E[(\theta_{i1}, \ldots, \theta_{in_i})'] = \boldsymbol{\vartheta}_i,$$
$$\text{Var}(\boldsymbol{\theta}_i) = \boldsymbol{\Sigma}_i.$$

Essentially, this model has the same structure of the one by [15] The θ_{ij} can be assumed independent and following a gamma distribution, producing, what we could term, a Poisson-gamma-normal model or, equivalently, a negative-binomial-normal model. This is natural in many cases in the sense that the $\boldsymbol{b_i}$ will induce association between repeated measurements, with then the θ_{ij} taking care of additional dispersion. In this case, $\boldsymbol{\Sigma}_i$ reduces to a diagonal matrix. Nevertheless, it is perfectly possible to allow for general covariance structures. When a fully distributional specification would be desired, then one could choose, for example, multivariate extensions of the gamma distribution .

The Poisson-gamma-normal model can be fitted following a two step procedure. In fact, integrating the previous conditional model over the gamma random effects, leaving the normal random effects untouched, leads to

$$f(y_{ij}|\boldsymbol{b_i}) = \begin{pmatrix} \alpha_j + y_{ij} - 1 \\ \alpha_j - 1 \end{pmatrix} \cdot \left(\frac{\beta_j}{1 + \kappa_{ij}\beta_j}\right)^{y_{ij}} \cdot \left(\frac{1}{1 + \kappa_{ij}\beta_j}\right)^{\alpha_j} \kappa_{ij}^{y_{ij}}, \qquad (5)$$

where $\kappa_{ij} = \exp(\boldsymbol{x}'_{ij}\boldsymbol{\xi} + \boldsymbol{z}'_{ij}\boldsymbol{b_i})$. It is important to point out that in this approach the gamma random effects are assumed to be independent. Model (5) can now be easily fitted using maximization routines like the one implemented in PROC NLMIXED in SAS.

4 Simulation studies

4.1 Impact of ignoring overdispersion

In this set of simulations, and using the Poisson-gamma-normal model as suitable framework to generate the data, we explore the impact of ignoring overdispersion on the parameter estimates and their standard errors. Mimicking the case study, longitudinal poisson responses Y_{ij} were generated with mean $\theta_{ij}\kappa_{ij}$, where θ_{ij} was randomly sampled from $\Gamma(\alpha, \beta)$ and

$$\kappa_{ij} = \exp\left(\xi_0 + b_i + \xi_1 t_j + \xi_2 z_i + \xi_3 t_j z_i.\right) \qquad (6)$$

In the previous expression $i = 1, ..., 500$ and $t_j = 1, 2, 3, 4, 5, 6$ denote the subject and the time of measurement respectively. Moreover, $b_i \sim N(0, \sigma_b^2)$ and z_i is a treatment group indicator variable taking values $0/1$.

The data were generated using three sets of parameters, i) $\xi_0 = -2$, $\xi_1 = -0.5$, $\xi_2 = -3$, $\xi_3 = 1$, $\alpha = 2$, ii) $\xi_0 = 0.1$, $\xi_1 = 0.2$, $\xi_2 = 0.3$, $\xi_3 = 0.5$, $\alpha = 4$ and iii) like (ii) but with

$\alpha = 0.5$. In order to vary the amount of overdispersion in the data, the θ_{ij}s were sampled from three different gamma distributions: $\Gamma(4, 0.25)$, $\Gamma(2, 0.5)$ and $\Gamma(0.5, 1)$ and in all the cases $\sigma_b^2 = 4$. In total 500 data sets were generated in each setting and analyzed using the correct model

$$E(Y_{ij}|\theta_{ij}b_i) = \theta_{ij}\exp(\xi_0 + b_i + \xi_1 t_j + \xi_2 z_i + \xi_3 t_j z_i) \qquad (7)$$

and a model that ignores the overdispersion, i.e., the misspecified model

$$E(Y_{ij}|b_i) = \exp(\xi_0 + b_i + \xi_1 t_j + \xi_2 z_i + \xi_3 t_j z_i) \qquad (8)$$

4.2 Impact on wrongly assuming overdispersion

In these simulations we studied the performance of model (7) when it is used to analyze data with no additional overdispersion. Basically, and using $\xi_0 = 0.1$, $\xi_1 = 0.2$, $\xi_2 = 0.3$ and $\xi_3 = 0.5$, $\sigma_b^2 = 4$ as values for the parameters, the data were generated using model (8) and latter analyzed using (7) and (8).

4.3 Impact of misspecification of random effects

In order to study the impact of misspecifying the random effects distribution a new set of simulations was designed. Essentially, data were generated following the scheme presented in Section 4.1 using the following parameter values; $\xi_0 = 0.1$, $\xi_1 = 0.2$, $\xi_2 = 0.3$ and $\xi_3 = 0.5$, but three main variations were introduced: Firstly, the b_is were generated using five different distributions: $N(0, 2)$, $\exp(\sqrt{0.5})$, t_4, $\Gamma(2, 0.5)$ and $\chi^2(1)$. Note that the parameters for the distributions were chosen such that $\text{var}(b_i) = 2$. The overdispersion random effects θ_{ij}s were always sampled from a $\Gamma(4, 0.25)$. Secondly, the θ_{ij}s were generated from $\Gamma(4, 0.25)$ and $\chi^2(2)$ and the b_is were always sampled from $N(0, 2)$. Finally in the third scenario, the b_is were sampled from the distributions aforementioned in the first setting and the θ_{ij}s were generated from $\chi^2(2)$.

Eventually, Models 7 and 8 were fitted to the generated data with the distributional assumptions described in Section 3.1. The goal is to explore the impact of misspecifying every set of random effects separately or simultaneously on the inferences obtained from these models.

4.4 Type I error

With these simulations we seek to establish whether the Type I error is preserved in the combined model and also when we have ignored the overdispersion. The data generation is as in section 4.1 but using $\xi_0 = -2$, $\xi_1 = -0.5$, $\xi_2 = 0$ and $\xi_3 = 1$, $\sigma_b^2 = 4$ $\alpha = 2$ as parameter values. The data were also fitted to both Models 7 and 8. For all situations, 500 datasets were generated and the sample size was 500.

5   Simulation results

5.1   Ignoring overdispersion

Table 2 summarizes the main findings of this study. Even though ignoring overdispersion may have a negative impact on the parameter estimates, particularly when the overdispersion distribution is highly skewed (when $\alpha = 0.5$), this impact was in general very mild. Indeed, for the covariates effect the relative bias never surpassed 7% and was frequently much smaller. Nevertheless, the intercept was severely affected by the misspecification in some scenarios and the variance of the random effect was seriously biased in some settings as well.

Importantly, the standard errors for the Poisson-normal model were always underestimated, especially for the interaction parameter which, in many cases, provides the answer to the main research question.

Much as it is harmful to ignore overdispersion, it was found (results not shown) that there is no harm in fitting the combined model when there is no overdispersion. Actually, in this scenario, the parameter estimates and standard errors for the two models (combined and Poisson-normal) were very close. When fitting the combined model, the estimate of α was very large in all cases, implying a very flat overdispersion distribution. Basically, the combined model converges to the Poisson-normal model when it is fitted to data that is not overdispersed.

5.2   Misspecification of random effects distribution

The intangible nature of random effects makes the selection of their distributional assumptions quite arbitrary. In Table 3 we show the impact of misspecifying the distribution of b_i on inferences based on both, the combine and the Poisson-normal model. In the combined model the covariate effects were rather robust with respect to the misspecification. Indeed, the relative bias was always smaller than 3% and the associated standard errors were close to those obtained when the random effect distribution was not misspecified. Nevertheless, like before, the intercept and the estimates of σ_b^2 were seriously affected in some scenarios.

We also studied the impact of misspecifying the distribution of θ_{ij} on inferences emanating from the combined model. The main results are presented in Table 4. Clearly, the impact of this misspecification is almost negligible and the estimates and the standard errors of all covariate effects are very close to the those obtained under the correctly specified model.

Turning to the Poisson-normal model, it is important to point out that in this setting this model misses two important features of the data: the presence of overdispersion and the real distribution of the random effect b_i. When b_i was sampled from a gamma and chisquare distribution, the Poisson-normal model never converged. For the t-distribution with 4 degrees of freedom (t_4) the parameter estimates for ξ_2 had a relative bias as large as 20%. This is considerable larger than the bias found when the model only missed the overdispersion in the data. Here again severe underestimation of the standard errors was observed as well.

Another important issue that emerged in this study was the low rate of convergence observed for both models. In fact, the rates of convergence for the rows of Table 3 from top to bottom were: 100%, 55.8%, 36.6%, 75.6% and 31.6% respectively.

Perhaps the most relevant situation for practical purposes is when both distributions are

misspecified. Table 5 illustrates our findings in this scenario. In this situation, estimates associated to the variance components of both distributions are largely biased. Nonetheless, apart from the intercept, the other effects estimates are generally close to the true values. Therefore, even when the distributions associated with both sets of random effects are misspecified, the covariate effects can be reliably estimated.

5.3 Type I error

In section 5.1 we discussed the impact of ignoring overdispersion on parameter estimates and standard errors, in which we saw that there is a large impact on standard errors which can possibly lead to erroneous conclusions. In this section we studied the impact of ignoring overdispersion on the Type I error by simulating data with no treatment effect and fitting both the combined model and the Poisson-normal model with treatment effect. The pre-specified Type I error was 5%. For the combined mode, out of the 500 datasets, the treatment effect was found to be significant in 27, which translates into 5.4% Type I error. On the other hand, out of the 331 datasets which converged for the Poisson-normal model, 61 found a significant treatment effect which represents 18.4% Type I error. In the best case scenario, if we assume that the models that failed to converge would not have detected a treatment effect, 61 out of 500 would translate into 12.2% which is still highly inflated. This finding is inline with the underestimation of standard errors by the Poisson-normal model as discussed in section 5.1.

6 Re-analyzing the case study

We also fitted the combined model to the data introduced in Section 2. The main findings are presented in the second part of Table 1. A number of remarks come into place here. Note first that, informally assessing the estimates of the parameters for the overdispersion random effects distribution, one can conclude that the overdispersion in these data should not be ignored. Complementing our findings is the observation that, the standard errors of the Poisson-normal model, especially the slope parameters are generally underestimated in the Poisson-normal. However, like before, the difference in the expected number of seizures between the experimental and placebo groups was not significant (p-value=0.2260) when the data were analyzed with the combined model. Note finally that, comparing the likelihood values, the combined model achieves a much better fitting of the observable data than the Poisson-normal model.

7 Concluding Remarks

Blindly assuming that unobserved heterogeneity in repeated measurements data only comes from the correlation in the responses can be too restrictive and sometimes can lead to invalid conclusions. We have shown through simulations that ignoring overdispersion in count data can have dire consequences on estimation of some covariate effects and their standard errors, as well as, on the variance components and the Type I error rates. Importantly, we found that the Type I error rates were considerably inflated when overdispersion was ignored, implying that the probability of detecting a spurious effect increases. Remarkably, our findings are strikingly similar to those reported by [9] [10] when studying the impact of misspecifying the random effect distribution in a logistic model with a random intercept. It is interesting to see that two related but different types of misspec-

ification, i.e., ignoring overdispersions and misspecifying the random effect distribution, may have very similar consequences.

Our simulations also indicate that the combined model may be a reasonable alternative in this situation. In fact, this model can be easily fitted using standard software packages like PROC NLMIXED in SAS. As discussed in section 4.1 when the combined model is fitted to data that has no overdispersion, it converges to the Poisson-normal model so no numerical issues emerge in this situation. Furthermore, the model is rather robust to misspecification of the random effects distributions. All the previous characteristics seem to indicate that the combined model is a useful tool for the analysis of Poisson data with overdispersion.

Acknowledgments

References

[1] Molenberghs, G., Verbeke, G., and Demétrio, C., Vieira,A A Family of Generalized Linear Models for Repeated Measures With Normal and Conjugate Random Effects. *Statistical Science* 2010; **25**, 325–347.

[2] Fisher, R.A The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* 1918; **52**, 399–433.

[3] De Boeck, P. and wison, M. *Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach.* New York: Springer, 2004.

[4] Molenberghs, G. and Verbeke, G. *Models for Discrete Longitudinal Data.* New York: Springer, 2005.

[5] Duchateau, L. and Janssen, P. *The Frailty Model.* New York: Springer, 2007.

[6] Davidian, M. and Giltinan, D.M *Nonlinear Models for Repeated Measurement Data* London, New York: Chapman & Hall, 1995.

[7] Agresti, A. *Categorical Data Analysis* (2nd ed.). New York: John Wiley & Sons, 2002.

[8] Hinde, J. and Demétrio, C.G.B. *Overdispersion: Models and Estimation.* São Paulo: XIII Sinape, 1998.

[9] Litieère,S., Alonso, A. and Molenberghs, G. Type I and Type II error under random-effects misspecification in generalized linear mixed models. *Biometrics* 2007 ; **63**, 1038–1044

[10] Litieère,S., Alonso, A. and Molenberghs, G. The impact of a misspecified random-effects distribution on the estimation and the perfomance of inferential procedures in generalized linear mixed models. *Statistics In Medicine* 2008; **27**, 3125–3144

[11] Ten Have, T. R., and Tran, L. A comparison of mixed effects logistic regression models for binary response data with two nested levels of clustering. *Statistics In Medicine* 1999; **18**, 947–960.

[12] Neuhaus, J. M., Hauck, W. W. and Kalbfleisch, J. D. The effects of mixture distribution misspecification when fitting mixed-effects logistic models. *Biometrika* 1992;**79** 755-762

[13] Agresti, A., Caffo, B., and Ohman-Strickland, P. Examples in which misspecifcation of a random effects distribution reduces efficiency, and possible remedies. *Computational Statistics & Data Analysis* 2004; **47**, 639–653

[14] Faught, E., Wilder, B.J., Ramsay, R.E., Reife, R.A., Kramer, L.D.,Pledger, G.W., and Karim, R.M. Topiramate placebo-controlled dose-ranging trial in refractory partialepilepsy using 200-, 400-, and 600-mg daily dosages. *Neurology* 1996; **46**, 1684–1690.

[15] Booth, J.G., Casella, G., Friedl, H., and Hobert, J.P. Negative binomial loglinear mixed models. *Statistical Modelling* 2003; **3**, 179–181.

[16] Molenberghs, G., Verbeke, G., and Demétrio, C. An extended random-effects approach to modeling repeated, overdispersed count data. *Lifetime Data Analysis* 2007; **13**, 513–531.

Table 1: Epilepsy Study. Parameter estimates and standard errors for the regression coefficients in the Poisson-normal model, and the combined model. Estimation was done by maximum likelihood using numerical integration over the normal random effect, if present.

Effect	Parameter	Poisson-normal Estimate (s.e.)	Combined Estimate (s.e.)
Intercept placebo	ξ_0	0.8179 (0.1677)	0.9112 (0.1755)
Slope placebo	ξ_1	-0.0143 (0.0044)	-0.0248 (0.0077)
Intercept treatment	ξ_2	0.6475 (0.1701)	0.6555 (0.1782)
Slope treatment	ξ_3	-0.0120 (0.0043)	-0.0118 (0.0074)
overdispersion parameter	α	—	2.4640 (0.2113)
overdispersion parameter	$\beta = 1/\alpha_1$	—	0.4059 (0.0348)
Variance of random intercepts	σ_b^2	1.1568 (0.1844)	1.1289 (0.1850)
-2log-likelihood		-6810	-7664

Table 2: Median of Parameter Estimates, {relative bias in %} and (standard errors) for simulations studying the impact of ignoring overdispersion. Data were generated from combined model with normal (b_i) and gamma (θ_{ij}) random effects using different levels of skewness (α). The data were analyzed with Comb=combined model, PN=Poisson-normal. TV denotes the true values.

	Parameter Estimates					
	ξ_0	ξ_1	ξ_2	ξ_3	σ_b^2	α
TV	0.1	0.2	0.3	0.5	4	4
Comb	0.102{2}	0.200{0}	0.288{4}	0.500{0}	3.990{0.25}	4.020{0.5}
	(0.139)	(0.012)	(0.1926)	(0.0155)	(0.0694)	(0.176)
PN	0.112{12}	0.192{4}	0.315{5.1}	0.488{2.4}	3.974{0.65}	
	(0.1315)	(0.0040)	(0.1828)	(0.0044)	(0.0684)	
TV	-2	-0.5	-3	1	4	2
Comb	-2.004{0.2}	-0.499{0.2}	-2.982{0.6}	1.003{0.3}	4.149{3.7}	2.002{0.1}
	(0.2450)	(0.0520)	(0.3404)	(0.0679)	(0.1365)	(0.4438)
PN	-2.074{3.7}	-0.496{0.8}	-2.964{1.2}	0.995{0.6}	4.140{3.5}	
	(0.2312)	(0.0395)	(0.2989)	(0.0498)	(0.1343)	
TV	0.1	0.2	0.3	0.5	4	0.5
Comb	0.106{6}	0.202{1}	0.271{9.7}	0.502{0.4}	3.938{1.6}	0.500{0}
	(0.1837)	(0.0300)	(0.2433)	(0.0403)	(0.0792)	(0.0190)
PN	-0.706{805}	0.186{7}	0.287{4.3}	0.476{4.8}	4.485{12.1}	
	(0.1441)	(0.0057)	(0.1976)	(0.0061)	(0.0763)	

11

Table 3: Median parameter estimate, {relative bias in %} and (standard errors) for simulations studying the impact of misspecfying the distribution of b_i. Data were generated using the combined model in which b_i was sample from RE-dist. Further the Comb=combined and PN=Poisson-normal models were fitted assuming normality.

RE-dist.		Parameter Estimates					
		ξ_0	ξ_1	ξ_2	ξ_3	σ_b^2	α
		0.1	0.2	0.3	0.5	2	2
normal	comb	0.105{5}	0.199{0.5}	0.291{3}	0.501{0.2}	1.989{0.55}	4.027{0.68}
		(0.1048)	(0.0123)	(0.1435)	(0.0156)	(0.0491)	(0.1803)
gamma	Comb	2.078{1978}	0.201{0.5}	0.300{0}	0.501{0.2}	1.955{2.4}	4.016{0.4}
		(0.0956)	(0.0090)	(0.1343)	(0.0122)	(0.0453)	(0.1319)
chisquare	Comb	1.038{938}	0.200{0}	0.307{2.3}	0.500{0}	1.797{10.5}	3.992{0.2}
		(0.0951)	(0.0105)	(0.1325)	(0.0136)	(0.0434)	(0.1534)
t_4	Comb	0.107{7}	0.200{0}	0.301{0.3}	0.500{0}	1.732{13.4}	4.011{0.3}
		(0.0996)	(0.0124)	(0.1360)	(0.0157)	(0.0460)	(0.1824)
t_4	PN	0.093{7}	0.196{2}	0.360{20}	0.481{3.8}	1.724{13.8}	
		(0.0896)	(0.0060)	(0.1228)	(0.0064)	(0.0447)	

Table 4: Median parameter estimates, {relative bias in %} and (standard errors) for simulations studying the impact of misspecfying the distribution of θ_{ij}. The θ_{ij}s were sampled from Re-dist. The combined model was fitted assuming gamma distributed overdispersion random effects.

RE-dist.	Parameter Estimates					
	ξ_0	ξ_1	ξ_2	ξ_3	σ_b^2	α
	0.1	0.2	0.3	0.5	2	4
gamma	0.105{5}	0.199{0.5}	0.291{3}	0.501{0.2}	1.989{0.55}	4.027{0.68}
	(0.1048)	(0.0123)	(0.1435)	(0.0156)	(0.0491)	(0.1803)
chisquare	0.784{684}	0.201{0.5}	0.290{0.3}	0.501{0.2}	1.992{2}	1.002{75}
	(0.1483)	(0.0191)	(0.2069)	(0.0258)	(0.0700)	(0.0324)

Table 5: Median parameter estimates, {relative bias in %} and (standard errors) for simulations studying the impact of misspecfying both θ_{ij} and b_i distributions. Data were generated from a combined model with b_i sampled from Re-dist. and θ_{ij} sampled from a chisquare distribution. The data were analyzed with the combined model assuming normal and gamma random effects.

RE dist.	Parameter Estimates					
	ξ_0	ξ_1	ξ_2	ξ_3	σ_b^2	α
	0.1	0.2	0.3	0.5	2	4
normal	0.105{5}	0.199{0.5}	0.291{3}	0.501{0.2}	1.989{0.55}	4.027{0.68}
	(0.1048)	(0.0123)	(0.1435)	(0.0156)	(0.0491)	(0.1803)
t_4	0.804{704}	0.200{0}	0.276{8}	0.501{0.2}	1.714{14.3}	1.001{75}
	(0.1112)	(0.0184)	(0.1546)	(0.0251)	(0.0485)	(0.0320)
chisquare	1.751{1651}	0.192{4}	0.262{12.7}	0.504{8}	1.620{19}	1.002{75}
	(0.009)	(0.0171)	(0.1438)	(0.0238)	(0.0418)	(0.0285)

14