



A Bayesian approach to analyse overdispersed longitudinal count data

Fernanda B. Rizzato, Roseli A. Leandro, Clarice G.B. Demétrio & Geert Molenberghs

To cite this article: Fernanda B. Rizzato, Roseli A. Leandro, Clarice G.B. Demétrio & Geert Molenberghs (2015): A Bayesian approach to analyse overdispersed longitudinal count data, Journal of Applied Statistics, DOI: [10.1080/02664763.2015.1126812](https://doi.org/10.1080/02664763.2015.1126812)

To link to this article: <http://dx.doi.org/10.1080/02664763.2015.1126812>

 View supplementary material [↗](#)

 Published online: 30 Dec 2015.

 Submit your article to this journal [↗](#)

 Article views: 15

 View related articles [↗](#)

 View Crossmark data [↗](#)

A Bayesian approach to analyse overdispersed longitudinal count data

Fernanda B. Rizzato^{a,b}, Roseli A. Leandro^b, Clarice G.B. Demétrio^b and Geert Molenberghs^c

^aStatistics, Universidade Federal do Paraná, Curitiba, Paraná, Brazil; ^bESALQ-USP, São Paulo, Brazil; ^cI-BioStat Universiteit Hasselt and Katholieke Universiteit Leuven, Diepenbeek, Belgium

ABSTRACT

In this paper, we consider a model for repeated count data, with within-subject correlation and/or overdispersion. It extends both the generalized linear mixed model and the negative-binomial model. This model, proposed in a likelihood context [17,18] is placed in a Bayesian inferential framework. An important contribution takes the form of Bayesian model assessment based on pivotal quantities, rather than the often less adequate DIC. By means of a real biological data set, we also discuss some Bayesian model selection aspects, using a pivotal quantity proposed by Johnson [12].

ARTICLE HISTORY

Received 7 December 2014
Accepted 29 November 2015

KEYWORDS

Bayesian analysis; Bayesian model assessment; count data; generalized linear mixed model; over dispersion

1. Introduction

Longitudinal count data are very common in practice, with counts collected for each unit or individually at various times. The Poisson model is convenient and mathematically elegant to analyze counts, but does come with restrictions. Two of its main limitations that are addressed in this contribution are: (i) the occurrence of overdispersion, that is, the phenomenon whereby variability in the data is not adequately captured by the model; and (ii) the accommodation of data hierarchies owing to, for example, repeatedly measuring the outcome on the same subject. One way of accommodating overdispersion in the model is by using the negative-binomial distribution, of which the genesis is the inclusion of a random, gamma-distributed effect into the Poisson model parameter.

To accommodate the correlation between measurements made on the same individual, one can include normal random effects in the linear predictor. Molenberghs *et al.* [17], in line with Booth *et al.* [2], proposed a model that simultaneously generalizes the Poisson–Normal generalized linear mixed model as well as conventional overdispersion models, in particular the negative-binomial model. See also Fotouhi [6], Gumedze and Chatora [10], Lee *et al.* [14], and Coelho-Barros *et al.* [3], for related work. A broader framework, encompassing other than count data types, was constructed as well [18].

CONTACT Fernanda B. Rizzato  fernandab@ufpr.br

 Supplemental data for this article can be accessed here. <http://dx.doi.org/10.1080/02664763.2015.1126812>

This work was done from a likelihood perspective, with implementations in the SAS procedure NLMIXED. One concern is that it relies on the validity of the asymptotics inherent in likelihood work.

Here, we take a Bayesian perspective, which is natural, given the random-effects formulation of the model. The difficulty of calculating the posterior marginal densities is overcome by obtaining a sample of the posterior joint distribution using stochastic simulation via Markov Chain (MCMC) algorithms. Thanks to the state of the art in Bayesian computation, and flexible implementations thereof in standard statistical software tools, such an endeavor has come within reach for routine use.

An important contribution of our work is that model assessment will be done using pivotal quantities. Methodology is illustrated on a set of data, not used before in this context. Of course, when referring to model assessment in a random-effects context, one needs to be aware that this refers only to the way the model fits the observed data, that is, the fit of the so-called marginal model. Indeed, as shown by Molenberghs and Verbeke [16], different hierarchical models can produce the same marginal model, illustrating that a good fit of the marginal model should not be seen as definitive evidence for any of the candidate hierarchical models.

An important advantage of this modeling approach is that it is based on full distributional assumptions, implying that analyses are valid under the assumption of missingness at random (MAR), which means that missingness can depend on covariates and observed outcome data, but given these not further on unobserved outcomes. Like when likelihood is used, the analyses reported here are valid under MAR, without the need to explicitly model the missing data mechanism. Of course, one cannot rule out a missing not at random mechanism to be operating, implying that, even given observed data, missing data further determine the missingness mechanism. Work needs to be done to allow for joint modeling of the longitudinal outcome and the missingness indicators. At the same time, a local influence-based method is being developed to assess the impact of incompleteness on the model conclusions. These developments fall outside of the scope of the current paper.

The structure of the paper is as follows. Section 2 describes the motivating data set. Section 3 presents a review on some basic concepts. In Section 4, we describe the hierarchical Bayesian analysis for the combined model and its three special cases (Poisson–Normal, negative-binomial and conventional Poisson models), using Markov Chain Monte Carlo methods and introduced a pivotal quantity for Bayesian model assessment. Section 5 gives results from the analysis performed for the epilepsy data. A brief discussion about a simulation study is presented in Section 6.

2. A clinical trial in epileptic patients

The data considered here are obtained from a randomized, double-blind, parallel group multi-center study for the comparison of placebo with a new anti-epileptic drug (AED), in combination with one or two other AEDs. The study is described in full detail in Faught *et al.* [5]. The randomization of epilepsy patients took place after a 12-week baseline period that served as a stabilization period for the use of AEDs, and during which the number of seizures were counted. After that period, 45 patients were assigned to the placebo group, 44 to the active (new) treatment group. This subgroup of 89 patients out of the original

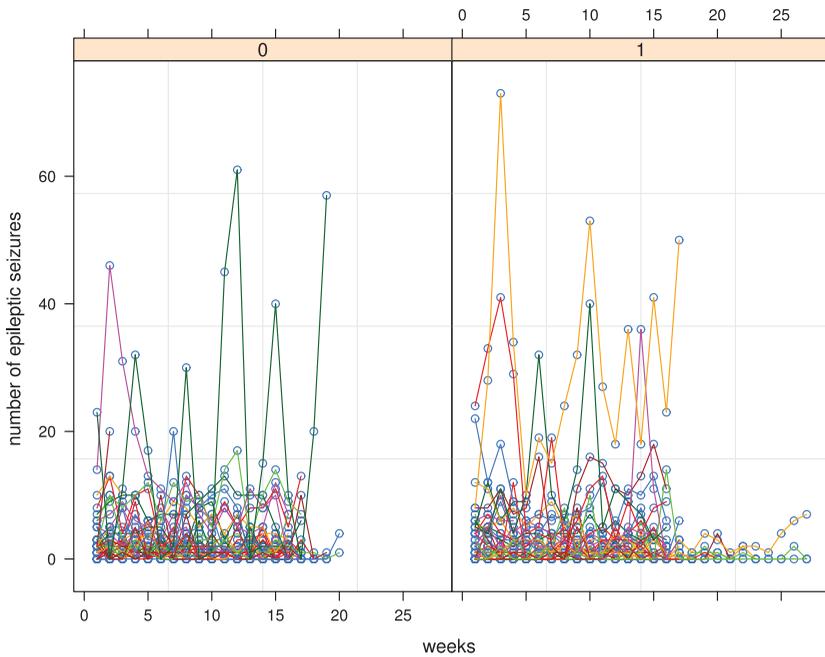


Figure 1. Graph of individual profiles for the number of epileptic seizures experienced during 27 weeks for patients assigned to the placebo group (0) and new treatment group (1) [colour online].

181 was used before in the statistical literature [16]. Patients were then measured weekly. They were followed (double-blind) during 16 weeks, after which they were entered into a long-term open-extension study. Some patients were followed for up to 27 weeks. The outcome of interest is the number of epileptic seizures experienced during the most recent week, that is, since the last time the outcome was measured. The key research question is whether or not the additional new treatment reduces the number of epileptic seizures. Individual profiles are shown in Figure 1. The unstable behavior can be explained by the presence of extreme values, but is also the result of the fact that very little observations are available at some of the time-points, especially past week 20.

3. Review of basic concepts

Background on generalized linear models is presented in Appendix A.

A straightforward and commonly used model for (i) is the Poisson–Normal model and for (ii) is the negative-binomial model. Combining ideas from (i) and (ii), Molenberghs *et al.* [17] specified, in line with Booth *et al.* [2], a model for repeated Poisson data with overdispersion, taking the form:

$$\begin{aligned}
 Y_{ij} &\sim \text{Poisson}(\lambda_{ij}), \\
 \lambda_{ij} &= \theta_{ij} \exp(\mathbf{x}'_{ij} \boldsymbol{\beta} + \mathbf{z}'_{ij} \mathbf{b}_i), \\
 \mathbf{b}_i &\sim N(\mathbf{0}, \mathbf{D}),
 \end{aligned}$$

$$\begin{aligned} E(\boldsymbol{\theta}_i) &= E[(\theta_{i1}, \dots, \theta_{in_i})'] = \boldsymbol{\Phi}_i, \\ \text{Var}(\boldsymbol{\theta}_i) &= \boldsymbol{\Sigma}_i, \end{aligned} \quad (1)$$

where Y_{ij} is the j th outcome measured for cluster (subject) i , $i = 1, \dots, N$, $j = 1, \dots, n_i$; \mathbf{Y}_i the n_i -dimensional vector of all measurements available for cluster i ; \mathbf{x}_{ij} and \mathbf{z}_{ij} , respectively, the p -dimensional and q -dimensional vectors of known covariate values, with $\boldsymbol{\beta}$ the p -dimensional vector of unknown fixed regression coefficients and \mathbf{b}_i the q -dimensional vector of random effects; $\boldsymbol{\theta}_i$ the n_i -dimensional vector of random effects for cluster i . The above model extends a GLM by (a) including normal random effects, like in a generalized linear mixed model (GLMM), and (b) including additional random effects $\boldsymbol{\theta}_i$ to flexibly accommodate overdispersion. Note that the components of $\boldsymbol{\theta}_i$ can be allowed to have constant as well as non-constant variance. At the same time, they can be correlated or independent of one another. While in the application the variance of the θ random effects will be assumed constant, also supported by the graphical exploration, it is important to see that the general framework is very flexible.

This model leads to several important special cases: the Poisson, Poisson–Normal, and negative binomial models, when $\theta_{ij} \sim \text{Gamma}(\alpha_1, \alpha_2)$ with $E(\theta_{ij}) = \alpha_1 \cdot \alpha_2$ and $\text{Var}(\theta_{ij}) = \alpha_1 \cdot \alpha_2^2$. Molenberghs *et al.* [17] derived explicit expressions for means, variance-covariance, and joint marginal probability of the outcome vector and discussed several estimation options. Vangeneugden *et al.* [27] derived explicitly correlation functions and showed the importance of having two different types of random effects. In the following section, the Bayesian version of this model and its special cases will be presented, against the background of the epilepsy data.

Another approach to incorporating heteroscedasticity and/or heterogeneity between clusters was proposed by Lee and Nelder [13] when they introduced a class of double hierarchical generalized linear mixed models (DHGLMs), in which random effects can be specified in both the mean and the dispersion components.

4. A Bayesian analysis of the epilepsy data

Let Y_{ij} represent the number of epileptic seizures patient i ($i = 1, \dots, N$) experiences during week j ($j = 1, \dots, n_i$) of the follow-up period. Also, let t_{ij} be the time-point at which Y_{ij} has been measured, $t_{ij} = 1, 2, \dots$, until at most 27; $\mathbf{y} = (y_{11}, \dots, y_{1n_1}, \dots, y_{N1}, \dots, y_{Nn_N})$ and $\mathbf{t} = (t_{11}, \dots, t_{1n_1}, \dots, t_{N1}, \dots, t_{Nn_N})$. The models used to analyse this data set are described below.

Model 1: Poisson model: We will first assume the model for inclusion of covariates, considering that count variable follows a Poisson distribution (1) with parameter

$$\lambda_{ij} = \begin{cases} \exp\{\beta_{00} + \beta_{01}t_{ij}\} & \text{if } i \text{ in placebo group,} \\ \exp\{\beta_{10} + \beta_{11}t_{ij}\} & \text{if } i \text{ in treatment group.} \end{cases}$$

We assume independence between the parameters $\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}$. The uncertainty related to them is incorporated into the model specifying the following prior distributions:

$$\beta_{kl} \sim N(a, b^2); \quad a, b \text{ known,} \quad (2)$$

with $k = 0, 1$ and $l = 0, 1$; the joint posterior distribution is proportional to

$$\begin{aligned} \pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11} | \mathbf{y}) &\propto \exp\left(-\sum_{i=1}^N \sum_{j=1}^{n_i} \lambda_{ij}\right) \prod_{i=1}^N \prod_{j=1}^{n_i} \lambda_{ij}^{y_{ij}} \\ &\times \prod_{k=0}^1 \prod_{l=0}^1 \exp\left\{-\frac{1}{2b^2}(\beta_{kl} - a)^2\right\}. \end{aligned} \tag{3}$$

Model 2: Poisson–Gamma model: We assume that the count variable follows the Poisson model (1) and to accommodate the variability among individuals, we consider a random effect θ_{ij} with gamma distribution, as follows:

$$\begin{aligned} Y_{ij} | \theta_{ij} &\sim \text{Poisson}(\lambda_{ij}) \\ \lambda_{ij} &= \begin{cases} \theta_{ij} \exp\{\beta_{00} + \beta_{01} t_{ij}\} & \text{if } i \text{ in placebo group,} \\ \theta_{ij} \exp\{\beta_{10} + \beta_{11} t_{ij}\} & \text{if } i \text{ in treatment group,} \end{cases} \\ \theta_{ij} &\sim \text{Gamma}(\alpha_1, \alpha_2), \end{aligned} \tag{4}$$

where $E(\theta_{ij}) = \alpha_1 \cdot \alpha_2$ and $\text{Var}(\theta_{ij}) = \alpha_1 \cdot \alpha_2^2$. To avoid identifiability problems, we assume $\alpha_1 = \alpha_2^{-1}$. For the hierarchical Bayesian analysis, first let us assume the prior distributions in Equation (2) for the parameter β_{kl} , $k = 0, 1$ and $l = 0, 1$. We also assume that θ_{ij} are independent and the following prior distribution for the parameter α_2

$$\alpha_2 \sim \text{Gamma}(c, d); \quad c, d \text{ known}, \tag{5}$$

which is chosen in a way that the values that α_2 assume are non-negative. In a frequentist setting, the above model is often called a negative-binomial model.

We further assume prior independence among the parameters. The joint posterior distribution is proportional to

$$\begin{aligned} \pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \alpha_2 | \mathbf{y}) &\propto \exp\left(-\sum_{i=1}^N \sum_{j=1}^{n_i} \lambda_{ij}\right) \prod_{i=1}^N \prod_{j=1}^{n_i} \lambda_{ij}^{y_{ij}} \\ &\times \prod_{k=0}^1 \prod_{l=0}^1 \exp\left\{-\frac{1}{2b^2}(\beta_{kl} - a)^2\right\} \\ &\times \prod_{i=1}^N \prod_{j=1}^{n_i} \frac{\alpha_2^{-\alpha_2^{-1}}}{\Gamma(\alpha_2^{-1})} \theta_{ij}^{\alpha_2^{-1}-1} e^{-\alpha_2^{-1} \theta_{ij}} \times \alpha_2^{c-1} e^{-d\alpha_2}, \end{aligned} \tag{6}$$

where $\boldsymbol{\theta} = (\theta_{11}, \dots, \theta_{1n_1}, \dots, \theta_{N1}, \dots, \theta_{Nn_N})$.

Model 3: Poisson–Normal model: We assume that the count variable follows the Poisson model (1) and that the random effects b_i to capture the correlation among longitudinal

measurements follows a normal distribution as follows:

$$\begin{aligned}
 Y_{ij}|b_i &\sim \text{Poisson}(\lambda_{ij}) \\
 \lambda_{ij} &= \begin{cases} \exp\{(\beta_{00} + b_i) + \beta_{01}t_{ij}\} & \text{if } i \text{ in placebo group,} \\ \exp\{(\beta_{10} + b_i) + \beta_{11}t_{ij}\} & \text{if } i \text{ in treatment group,} \end{cases} \\
 b_i &\sim N(0, \sigma^2).
 \end{aligned} \tag{7}$$

For the hierarchical Bayesian analysis, first let us assume the same prior distributions (2) for β_{kl} , $k = 0, 1$ and $l = 0, 1$. Under model 3, the b_i are assumed to be independent random variables with normal distribution $N(0, \sigma^2)$. Further, we assume the following prior distribution for the parameter σ^2

$$\sigma^2 \sim \text{Inverse Gamma}(g, h); \quad g, h \text{ known.} \tag{8}$$

We also assume prior independence among the parameters. The joint posterior distribution is given by

$$\begin{aligned}
 \pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \mathbf{b}, \sigma^2 | \mathbf{y}) &\propto \exp\left(-\sum_{i=1}^N \sum_{j=1}^{n_i} \lambda_{ij}\right) \prod_{i=1}^N \prod_{j=1}^{n_i} \lambda_{ij}^{y_{ij}} \\
 &\times \prod_{k=0}^1 \prod_{l=0}^1 \exp\left\{-\frac{1}{2b^2}(\beta_{kl} - a)^2\right\} \\
 &\times \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{b_i^2}{2\sigma^2}\right\} \\
 &\times \sigma^{2-(g+1)} \exp\left\{-\frac{h}{\sigma^2}\right\},
 \end{aligned} \tag{9}$$

where $\mathbf{b} = (b_1, \dots, b_N)$.

Model 4: Poisson–Normal–Gamma (combined) model: Finally, we combine models 2 and 3, that is,

$$\begin{aligned}
 Y_{ij}|b_i, \theta_{ij} &\sim \text{Poisson}(\lambda_{ij}), \\
 \lambda_{ij} &= \begin{cases} \theta_{ij} \exp\{(\beta_{00} + b_i) + \beta_{01}t_{ij}\} & \text{if } i \text{ in placebo group,} \\ \theta_{ij} \exp\{(\beta_{10} + b_i) + \beta_{11}t_{ij}\} & \text{if } i \text{ in treatment group.} \end{cases}
 \end{aligned}$$

θ_{ij} is as in Equation (4) and b_i is as in Equation (7). For the hierarchical Bayesian analysis, let us assume the prior distribution (2) for the parameters β_{kl} , $k = 0, 1$ and $l = 0, 1$. Considering that b_i and θ_{ij} are independent, let us assume the prior distribution (8) for the parameter σ^2 and the prior distribution (5) for the parameter α_2 . Note that this model goes beyond the negative-binomial model, because correlation among repeated measures is accounted for by the normal random effects. As such, it combines GLMM and negative-binomial concepts.

We further assume prior independence among the parameters. The joint posterior distribution is proportional to

$$\begin{aligned} \pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \mathbf{b}, \alpha_2, \sigma^2 | \mathbf{y}) &\propto \exp\left(-\sum_{i=1}^N \sum_{j=1}^{n_i} \lambda_{ij}\right) \prod_{i=1}^N \prod_{j=1}^{n_i} \lambda_{ij}^{y_{ij}} \\ &\times \prod_{k=0}^1 \prod_{l=0}^1 \exp\left\{-\frac{1}{2b^2}(\beta_{kl} - a)^2\right\} \\ &\times \prod_{i=1}^N \prod_{j=1}^{n_i} \frac{\alpha_2^{-\alpha_2^{-1}}}{\Gamma(\alpha_2^{-1})} \theta_{ij}^{\alpha_2^{-1}-1} e^{-\alpha_2^{-1}\theta_{ij}} \times \alpha_2^{c-1} e^{-d\alpha_2} \\ &\times \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{b_i^2}{2\sigma^2}\right\} \sigma^{2-(g+1)} \exp\left\{-\frac{h}{\sigma^2}\right\}. \end{aligned} \tag{10}$$

Model 5: Poisson–Gamma–Gamma (DHGLM) model: We assume that the count variable follows Poisson model (1); to accommodate the variability among individuals a random effect θ_{ij} with gamma distribution is incorporated, like in Model 2. Additionally, the random effects d_i , to capture the correlation among longitudinal measurements, follows a gamma distribution. The model formulation is as follows:

$$\begin{aligned} Y_{ij} | d_i, \theta_{ij} &\sim \text{Poisson}(\lambda_{ij}), \\ \lambda_{ij} &= \begin{cases} \theta_{ij} \exp\{(\beta_{00} + \log(d_i)) + \beta_{01} t_{ij}\} & \text{if } i \text{ in placebo group,} \\ \theta_{ij} \exp\{(\beta_{10} + \log(d_i)) + \beta_{11} t_{ij}\} & \text{if } i \text{ in treatment group.} \end{cases} \\ \theta_{ij} &\sim \text{Gamma}(\alpha_1, \alpha_2), \\ d_i &\sim \text{Gamma}(\alpha_3, \alpha_4). \end{aligned}$$

Like in Models 2 and 4, to avoid identifiability problems, we assume $\alpha_1 = \alpha_2^{-1}$ and $\alpha_3 = \alpha_4^{-1}$. For the hierarchical Bayesian analysis, first we assume prior distributions (2) for the parameter β_{kl} , $k = 0, 1$ and $l = 0, 1$. We also assume that the θ_{ij} 's are independent. Finally, we assume the following prior distribution for the parameter α_2 and α_4 :

$$\alpha_2 \quad \text{and} \quad \alpha_4 \sim \text{Gamma}(c, d); \quad c, d \text{ known}, \tag{11}$$

which are chosen in a way that α_2 and α_4 are guaranteed to be non-negative. We further assume prior independence among the parameters. The joint posterior distribution is proportional to

$$\begin{aligned} \pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \mathbf{d}, \alpha_2, \alpha_4, \sigma^2 | \mathbf{y}) &\propto \exp\left(-\sum_{i=1}^N \sum_{j=1}^{n_i} \lambda_{ij}\right) \prod_{i=1}^N \prod_{j=1}^{n_i} \lambda_{ij}^{y_{ij}} \\ &\times \prod_{k=0}^1 \prod_{l=0}^1 \exp\left\{-\frac{1}{2b^2}(\beta_{kl} - a)^2\right\} \end{aligned}$$

$$\begin{aligned}
& \times \prod_{i=1}^N \prod_{j=1}^{n_i} \frac{\alpha_2^{-\alpha_2^{-1}}}{\Gamma(\alpha_2^{-1})} \theta_{ij}^{\alpha_2^{-1}-1} e^{-\alpha_2^{-1}\theta_{ij}} \times \alpha_2^{c-1} e^{-d\alpha_2} \\
& \times \prod_{i=1}^N \prod_{j=1}^{n_i} \frac{\alpha_4^{-\alpha_4^{-1}}}{\Gamma(\alpha_4^{-1})} d_i^{\alpha_4^{-1}-1} e^{-\alpha_4^{-1}d_i} \times \alpha_4^{c-1} e^{-d\alpha_4}.
\end{aligned} \tag{12}$$

The joint posterior distributions (3), (6), (9), (10) and (12) are evidently important aspects of Bayesian inference, but come with the issue that they are not analytically tractable. So, the posterior summaries of interest are based on samples obtained from the joint posterior distribution using stochastic simulation via Markov Chains. Implementation was done in OpenBugs [26]. The tool requires the user to pass on the data, to specify the model, and to provide initial values for the parameters. Otherwise, for implementation it is necessary to obtain the full conditional posterior distributions. In this work, the history plot, showing subsequent draws, the autocorrelation function plot, and the Gelman–Rubin criterion were used as convergence diagnostics.

4.1. Bayesian model assessment using pivotal quantities

Model comparison is commonly based on the deviance information criterion (DIC) of Spiegelhalter *et al.* [25] because of its widespread use and ease of application. However, Millar [15] demonstrated that the use of DIC is inappropriate to assess the negative binomial and Poisson-lognormal models when gamma and lognormal latent parameters are included to incorporate extra-Poisson variability. The most familiar alternative to the DIC is the Bayes factor, but it comes with some drawbacks: the difficulty to obtain the marginal densities on the one hand and its sensitivity to prior specifications on the other. A second strategy for assessing model adequacy is to use posterior-predictive model checks. This approach was initially proposed by Guttman [11] and Rubin [23], and was extended to more general discrepancy functions by Gelman, Meng, and Stern [8]. Despite its relative ease of implementation, Bayarri and Berger [1] and Robins *et al.* [21] have noted that they are not (even asymptotically) uniformly distributed. This is a drawback because interpretation is then more complicated. Indeed, to properly gauge the extent of departures, the asymptotic reference distribution would have to be derived, which can be unwieldy in realistic settings.

As an alternative, Johnson [12] proposed to explore the relation between the distribution of pivotal quantities evaluated at the true parameter value and the distribution of the same pivotal quantities evaluated at parameter values sampled from the posterior distribution. In general, these two distributions are identical, which makes it straightforward to define a large number of Bayesian model diagnostics. For Johnson's approach, it is necessary to define a *pivotal quantity*, a function $S(\mathcal{Y}, \Theta) \mapsto \mathbb{R}^p$ for which the integral

$$G(\mathbf{s}) \equiv \int_{\mathcal{Y}} \mathcal{I}_{A(\omega, \mathbf{y})} f(\mathbf{y}|\omega) d\mathbf{y}, \quad A(\omega, \mathbf{y}) = \{(\mathbf{y}, \omega) : S(\mathbf{y}, \omega) \leq \mathbf{s}\}$$

depends only on the value of the vector $\mathbf{s} \in \mathbb{R}^p$ for all $\omega \in \Theta$. Johnson [12] showed that $S(\mathbf{Y}, \omega_Y)$ and $S(\mathbf{Y}, \omega_0)$ are identically distributed when ω_0 is a random vector drawn from

density π and ω_Y is a parameter vector drawn from the posterior distribution on ω given Y . But it is not the same for the joint distribution of pivotal quantities based on multiple draws of (Y, ω_Y^i) from the same posterior distribution, where $\{\omega_Y^i\}$ denotes posterior samples of ω based on the same data vector Y ; i indexes these samples. However, he considered strategies for combining information from multiple values of a pivotal quantity defined from the same posterior distribution.

The most direct is numerical evaluation of the joint sampling distribution of the pivotal quantity using prior-predictive-posterior (PPP) simulation [4], but PPP simulation comes with a disadvantage, that is, the high computational burden. To avoid PPP simulations, he suggested that model assessment be based on informal graphical comparisons of the posterior distribution of several pivotal quantities to their nominal sampling distributions. Assuming

$$\begin{aligned} \mathbf{y} &= (y_{1;1}, \dots, y_{1;n_1}, \dots, y_{89;1}, \dots, y_{89;n_{89}}), \\ \boldsymbol{\lambda} &= (\lambda_{1;1}, \dots, \lambda_{1;n_1}, \dots, \lambda_{89;1}, \dots, \lambda_{89;n_{89}}), \\ \boldsymbol{\theta} &= (\theta_{1;1}, \dots, \theta_{1;n_1}, \dots, \theta_{89;1}, \dots, \theta_{89;n_{89}}), \\ \mathbf{b} &= (b_{1;1}, \dots, b_{1;n_1}, \dots, b_{89;1}, \dots, b_{89;n_{89}}), \\ \mathbf{d} &= (d_{1;1}, \dots, d_{1;n_1}, \dots, d_{89;1}, \dots, d_{89;n_{89}}), \end{aligned}$$

where the first index refers to subject and the second one to repeated measurement within subject, we define the pivotal quantities

$$S(\mathbf{y}, \boldsymbol{\lambda}) = \sum_i \sum_j \frac{(y_{ij} - \lambda_{ij})^2}{\lambda_{ij}}, \tag{13}$$

$$S_b(\mathbf{b}, \sigma^2) = \sum_i \frac{(b_i - 0)^2}{\sigma^2}, \tag{14}$$

and considering the restrictions on the models with gamma random effect, $E(\theta_{ij}) = \alpha_1\alpha_2 = 1$ and $\text{Var}(\theta_{ij}) = \alpha_1\alpha_2^2 = \alpha_2$,

$$S_\theta(\boldsymbol{\theta}, \alpha_2) = \sum_i \sum_j \frac{(\theta_{ij} - 1)^2}{\alpha_2}, \tag{15}$$

$E(d_i) = \alpha_3\alpha_4 = 1$ and $\text{Var}(d_i) = \alpha_3\alpha_4^2 = \alpha_4$

$$S_d(\mathbf{d}, \alpha_4) = \sum_i \frac{(d_i - 1)^2}{\alpha_4}, \tag{16}$$

for each level in each proposed model that we will use. These measures are approximately pivotal quantities [12]. Intuitively, it follows from the fact that squared standardized residuals are added, leading to exact or asymptotic χ^2 distributions. The four pivotal quantities address the fixed-effects structure (Equation (13)), the normal random effects (Equation (14)), and the gamma random effects (Equations (15) and (16)), respectively. Combining the three into a single measure could be of interest, but is not pursued further here. Then, the analysis is based on the comparison of a draw from the posterior

distributions of these measures with the respective χ^2 distribution χ_T^2 , χ_{89}^2 , and χ_T^2 , with $T = \sum_{i=1}^{89} n_i = 1419$, for the epilepsy data. In practice, for the Poisson–Normal–Gamma model we plot the histogram-based estimates of $S(\mathbf{y}, \boldsymbol{\lambda})$, $S_\theta(\boldsymbol{\theta}, \alpha_2)$, and $S_b(\mathbf{b}, \sigma^2)$, evaluated at a posterior sample of $(\boldsymbol{\lambda}, \boldsymbol{\theta}, \alpha_2, \mathbf{b}, \sigma^2)$, and also here with the corresponding marginal χ^2 distribution for comparison. For the Poisson–Gamma–Gamma model, we plot the histogram-based estimates of $S(\mathbf{y}, \boldsymbol{\lambda})$, $S_\theta(\boldsymbol{\theta}, \alpha_2)$, and $S_d(\mathbf{d}, \alpha_4)$ evaluated at a posterior sample of $(\boldsymbol{\lambda}, \boldsymbol{\theta}, \alpha_2, \mathbf{d}, \alpha_4)$; likewise, the corresponding marginal χ^2 distribution is plotted, for the sake of comparison.

For the Poisson model, we plot the histogram estimate of Equation (13) for comparison with the χ_{1419}^2 distribution. For the Poisson–Gamma model we plot the histogram-based estimate of Equations (13) and (15) for comparison, respectively, both with χ_{1419}^2 . For the Poisson–Normal model, we plot the histogram-based estimate of Equations (13) and (14) for comparison, respectively, with χ_{1419}^2 and χ_{89}^2 . For the Poisson–Normal–Gamma model, we plot the histogram estimate of Equations (13)–(15) for comparison, respectively, with χ_{1419}^2 , χ_{89}^2 , and χ_{1419}^2 . Finally, for the Poisson–Gamma–Gamma model, we plot the histogram estimates of Equations (13), (16), and (15) for comparison, respectively, with χ_{1419}^2 , χ_{89}^2 , and χ_{1419}^2 .

5. Analysis of the epilepsy data

For a Bayesian analysis of the data set presented in Section 2, we assume for each model the following hyperparameter values for the prior distributions (2), (5), (8), and (11): $a = 1$, $b^2 = 1000$, and $c = d = 0.001$. These hyperparameters values were chosen to ensure approximately non-informative priors when Gibbs sampling for the joint posterior distribution of interest, using the OpenBugs software [26]. The code of the OpenBugs program is given in Appendix A. For each model, we simulated two chains, each one with 55,000 samples. The first 5000 samples (‘burn-in-samples’) were discarded to eliminate the effect of the initial values for the algorithm and to have approximately uncorrelated samples. We considered the samples 10^a , 20^a , 30^a , \dots , which result in final samples of size 5000 for each parameter. It takes around 15 minutes for the program to run. The chains were initialized at different points and the convergence was monitored by the conventional time series plots for simulated samples and also using the method proposed by Gelman and Rubin [9].

Although it does not provide reliable results for mixed models [15], in Table 1 we present the estimated DIC for each model; smaller values of DIC indicate better models, suggesting that the combined model exhibits a better fit to the data.

Figures 2–6 present the histogram-based estimates of the posterior distributions of measures S , S_θ , S_b , and S_d , described in Section 4, under Poisson, Poisson–Gamma,

Table 1. DIC criterion for each model obtained using the OpenBugs software.

| Model | λ_{ij} | DIC | p_D |
|----------------------------|--|------------|---------|
| Poissonmodel | $\exp\{\beta_0 + \beta_1 t_{ij}\}$ | 11,597.700 | 4.002 |
| Poisson–Gamma model | $\theta_{ij} \exp\{\beta_0 + \beta_1 t_{ij}\}$ | 4916.890 | 810.616 |
| Poisson–Normal model | $\exp\{(\beta_0 + b_i) + \beta_1 t_{ij}\}$ | 6047.630 | 86.845 |
| Poisson–Normal–Gamma model | $\theta_{ij} \exp\{(\beta_0 + b_i) + \beta_1 t_{ij}\}$ | 4838.770 | 558.346 |
| Poisson–Gamma–Gamma model | $\theta_{ij} \exp\{(\beta_0 + \log(d_i)) + \beta_1 t_{ij}\}$ | 4831.000 | 552.200 |

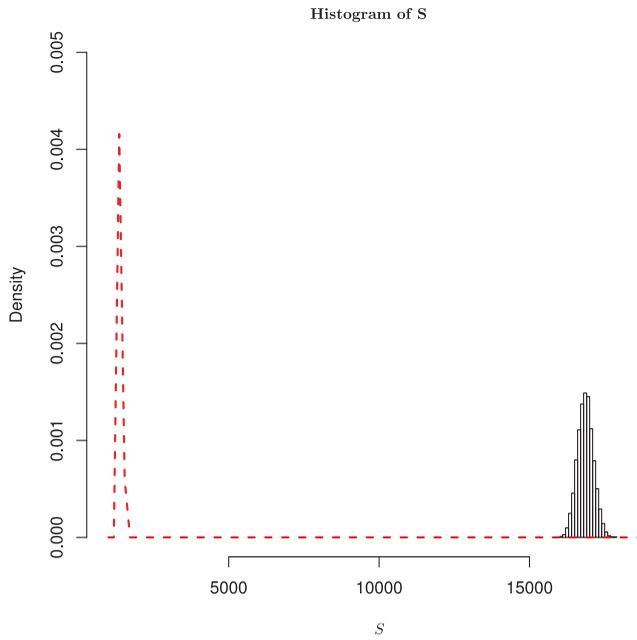


Figure 2. Histogram estimate of the posterior distribution of $S(\mathbf{y}, \lambda_y^i)$ under Poisson model. For comparison the graphs display in dashed line type the marginal χ_{1419}^2 distribution.

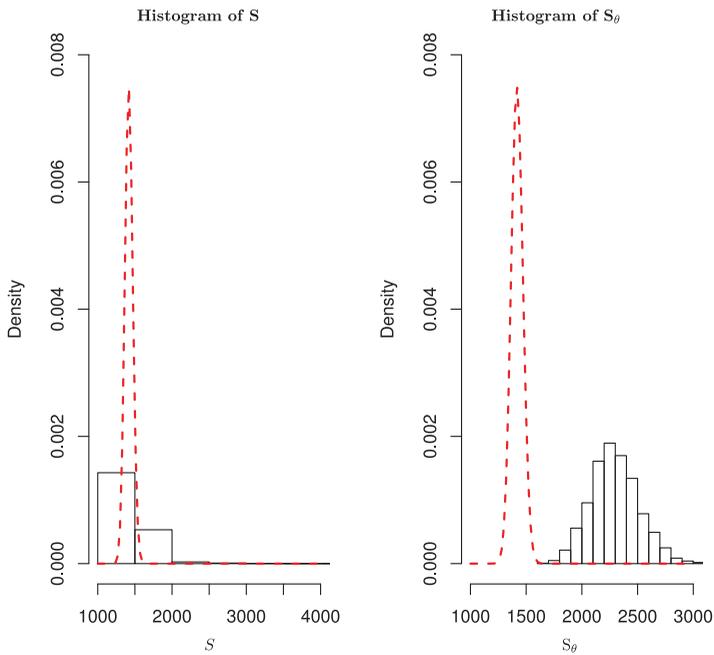


Figure 3. Histogram estimate of the posterior distribution of $S(\mathbf{y}, \lambda_y^i)$ and $S_\theta(\theta_y^i, \alpha_{1y}^i)$ under Poisson–Gamma model. For comparison the graphs display, in dashed line type the marginal χ_{1419}^2 or χ_{89}^2 distributions.

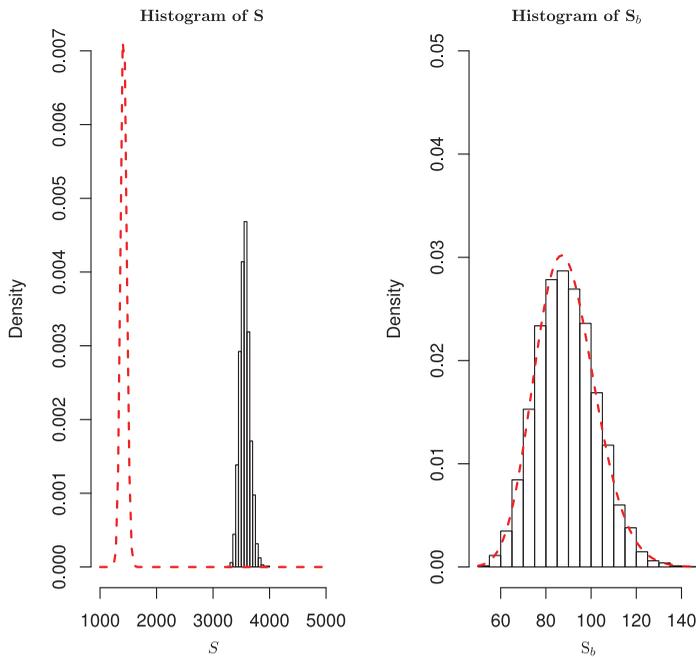


Figure 4. Histogram estimate of the posterior distribution of $S(\mathbf{y}, \lambda_{y^j})$ and $S_b(\theta_{y^j}, \sigma_{y^j}^2)$ under Poisson–Normal model. For comparison the graphs display, using dashed line type, the marginal χ_{1419}^2 or χ_{89}^2 distributions.

Poisson–Normal, Poisson–Normal–Gamma and Poisson–Gamma–Gamma models, respectively. Model assessment was based on informal graphical comparison of the posterior distribution of a pivotal quantity (or several pivotal quantities) to their nominal sampling distribution. From the graphical analysis, we concluded that the Poisson–Normal–Gamma model is the one that fits best to the data.

Further, in Figure 7(a)–(b), we compare the sample variances for the count data in each time \times treatment combination, and the variability captured by each model in each such combination. As in the analysis of the pivotal quantities, through the DIC and graphical analysis of the variability captured by each model, we conclude that the combined model fits best, reinforcing our choice.

In Table 2, we observe the posterior summaries for the parameters of the models 1–5. In order to compare the treatments, two measures were included $ds = \beta_{11} - \beta_{01}$ and $rs = \beta_{11}/\beta_{01}$, that is, the difference and ratio between parameters, respectively. If the values zero and one belong to the credible interval for these measures, respectively, then we could conclude that treatments do not differ. As in Molenberghs *et al.* [17], the Poisson model leads to significance for both the difference and the ratio, whereas for the Poisson–Normal this is not the case for the difference while some significance is maintained for the ratio. For the Poisson–Gamma, Poisson–Normal–Gamma and Poisson–Gamma–Gamma models, both the difference and the ratio are non-significant.

The posterior predictive model check is a standard fit assessment technique, consisting of drawing simulated values from the posterior predictive distribution of replicated data

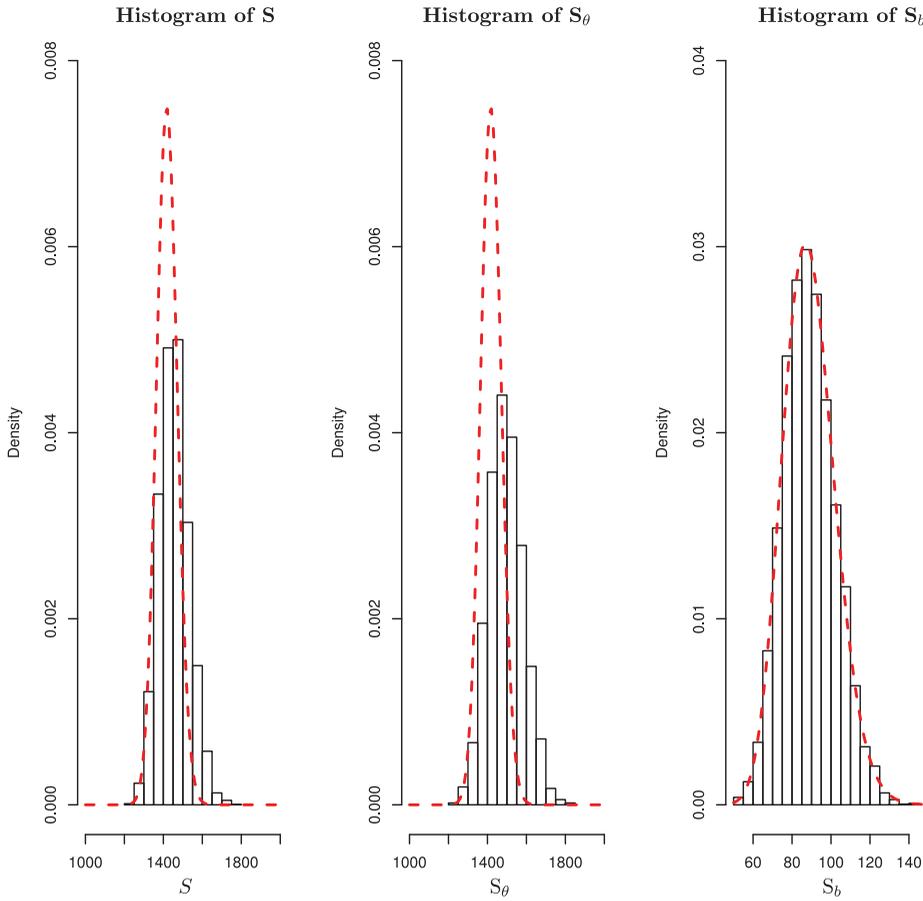


Figure 5. Histogram estimate of the posterior distribution of $S(\mathbf{y}, \lambda_y^i)$, $S_\theta(\theta_y^i, \alpha_{1y}^i)$, and $S_b(\theta_y^i, \sigma_y^{2i})$ under a Poisson–Normal–Gamma model. For comparison the graphs display, in dashed line type, the marginal χ^2_{1419} or χ^2_{89} distributions.

and comparing these samples to the observed data through the graphs of the observed data versus the mean of the predicted data. Figure 8 shows that the Poisson–Normal–Gamma model fits the data well.

5.1. Sensitivity analysis on the prior grid

There are several ways of investigating the sensitivity of the estimates and their posterior distributions to the choice of the hyperprior distributions. We investigated sensitivity of the posterior mean estimates for large perturbations of the gamma and normal hyperpriors by fitting the Poisson–Normal–Gamma model with different priors (informative and non-informative). For the gamma priors, the following 3×3 grid of prior values have been considered [7,22]: $G(a, b)$ where $a \in \{0.5, 1, 1.5\}$ and $b \in \{1/0.311, 1/0.622, 1/0.933\}$, giving nine pairs of values (a, b) while for the normal priors, $N(0, d)$, we considered three values $d \in \{1, 10^3, 10^6\}$, giving a total of 27 combinations. The results show that the values

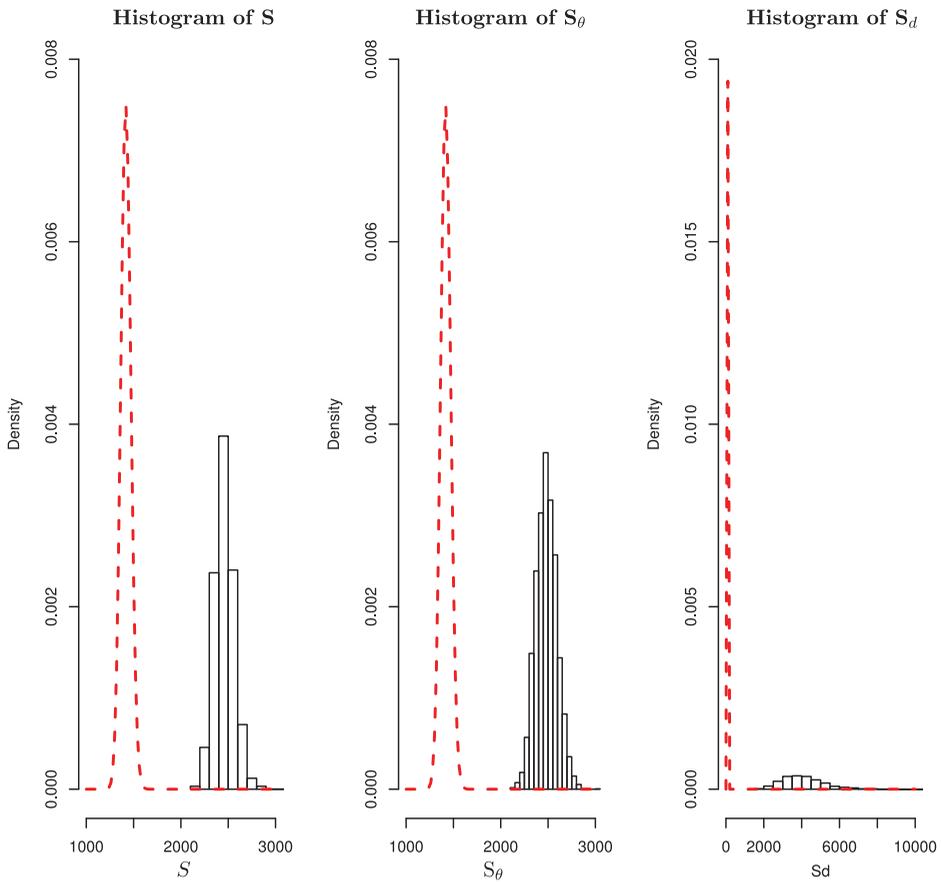


Figure 6. Histogram estimate of the posterior distribution of $S(\mathbf{y}, \lambda_y^i)$, $S_\theta(\theta_y^i, \alpha_{1y}^i)$, and $S_d(\theta_y^i, \sigma_y^{2i})$ under a Poisson–Gamma–Gamma model. For comparison the graphs display, using dashed line type, the marginal χ_{1419}^2 distribution.

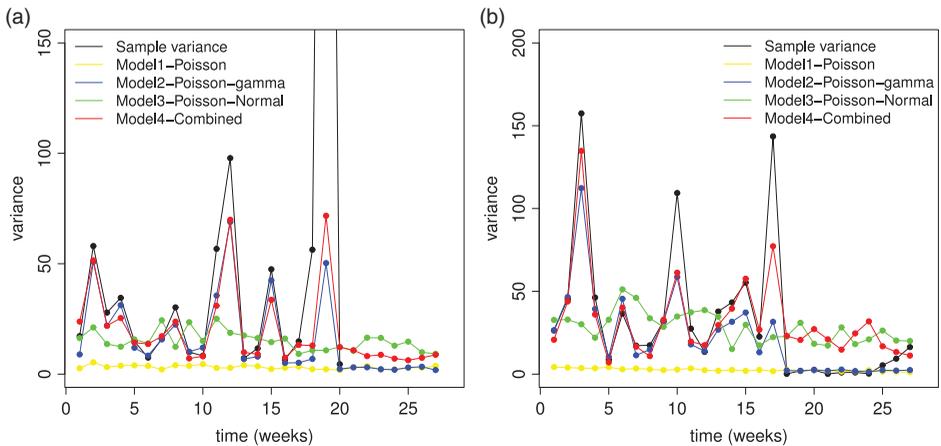


Figure 7. (a) Graphic representation of the variability among subjects who received placebo captured by different models for different times, with a variance of up to 150. (b) Display of the variability among subjects who received the new drug captured by the different models for different times [colour online].

Table 2. Posterior summaries considering Poisson, Poisson–Gamma, Poisson–Normal, Poisson–Normal–Gamma, and Poisson–Gamma–Gamma models.

| Parameter | Mean | sd | median | Credible interval | |
|--------------------------------------|--------|--------|--------|-------------------|--------|
| | | | | 2.5% | 97.5% |
| <i>Poisson model</i> | | | | | |
| β_{00} | 1.266 | 0.042 | 1.265 | 1.182 | 1.347 |
| β_{01} | -0.013 | 0.004 | -0.013 | -0.022 | -0.005 |
| β_{10} | 1.453 | 0.038 | 1.453 | 1.378 | 1.527 |
| β_{11} | -0.033 | 0.004 | -0.033 | -0.040 | -0.025 |
| $ds = \beta_{11} - \beta_{01}$ | -0.019 | 0.006 | -0.019 | -0.031 | -0.008 |
| $rs = \frac{\beta_{11}}{\beta_{01}}$ | 2.891 | 3.898 | 2.458 | 1.411 | 6.991 |
| <i>Poisson–Gamma model</i> | | | | | |
| β_{00} | 1.265 | 0.114 | 1.263 | 1.044 | 1.492 |
| β_{01} | -0.013 | 0.011 | -0.013 | -0.034 | 0.009 |
| β_{10} | 1.475 | 0.109 | 1.474 | 1.264 | 1.692 |
| β_{11} | -0.035 | 0.010 | -0.035 | -0.055 | -0.015 |
| α_1 | 0.526 | 0.025 | 0.525 | 0.477 | 0.578 |
| α_2 | 1.906 | 0.092 | 1.905 | 1.729 | 2.095 |
| $ds = \beta_{11} - \beta_{01}$ | -0.022 | 0.015 | -0.022 | -0.051 | 0.007 |
| $rs = \frac{\beta_{11}}{\beta_{01}}$ | 2.546 | 64.530 | 2.094 | -21.160 | 27.860 |
| <i>Poisson–Normal model</i> | | | | | |
| β_{00} | 0.808 | 0.170 | 0.810 | 0.465 | 1.137 |
| β_{01} | -0.014 | 0.004 | -0.014 | -0.023 | -0.006 |
| β_{10} | 0.641 | 0.170 | 0.643 | 0.310 | 0.973 |
| β_{11} | -0.012 | 0.004 | -0.012 | -0.020 | -0.003 |
| σ^2 | 1.211 | 0.201 | 1.189 | 0.877 | 1.663 |
| $ds = \beta_{11} - \beta_{01}$ | 0.002 | 0.006 | 0.002 | -0.010 | 0.015 |
| $rs = \frac{\beta_{11}}{\beta_{01}}$ | 0.957 | 1.556 | 0.841 | 0.221 | 2.429 |
| <i>Poisson–Normal–Gamma model</i> | | | | | |
| β_{00} | 0.895 | 0.183 | 0.894 | 0.536 | 1.253 |
| β_{01} | -0.025 | 0.008 | -0.025 | 0.040 | -0.009 |
| β_{10} | 0.657 | 0.184 | 0.659 | 0.293 | 1.014 |
| β_{11} | -0.012 | 0.008 | -0.012 | -0.027 | 0.003 |
| α_1 | 2.462 | 0.211 | 2.452 | 2.080 | 2.905 |
| α_2 | 0.409 | 0.035 | 0.408 | 0.344 | 0.481 |
| σ^2 | 1.191 | 0.203 | 1.168 | 0.855 | 1.649 |
| $ds = \beta_{11} - \beta_{01}$ | 0.013 | 0.011 | 0.013 | -0.008 | 0.034 |
| $rs = \frac{\beta_{11}}{\beta_{01}}$ | 0.564 | 1.730 | 0.480 | -0.128 | 1.633 |
| <i>Poisson–Gamma–Gamma model</i> | | | | | |
| β_{00} | 1.402 | 0.165 | 1.404 | 1.086 | 1.73 |
| β_{01} | -0.024 | 0.008 | -0.024 | -0.039 | -0.009 |
| β_{10} | 1.291 | 0.170 | 1.289 | 0.984 | 1.627 |
| β_{11} | -0.011 | 0.007 | -0.010 | -0.025 | 0.004 |
| α_1 | 2.456 | 0.211 | 2.446 | 2.080 | 2.902 |
| α_2 | 0.410 | 0.035 | 0.409 | 0.345 | 0.481 |
| α_3 | 0.984 | 0.141 | 0.976 | 0.733 | 1.286 |
| α_4 | 1.038 | 0.151 | 1.025 | 0.778 | 1.364 |
| $ds = \beta_{11} - \beta_{01}$ | 0.014 | 0.011 | 0.014 | -0.006 | 0.035 |
| $rs = \frac{\beta_{11}}{\beta_{01}}$ | 0.483 | 0.835 | 0.435 | -0.186 | 1.445 |

of the gamma and normal hyperpriors have almost no influence on the posterior mean estimates of all parameters as shown in Figure 9. It should be noted that some patterns, like in the bottom panel of Figure 9, should not be over-interpreted, given the compact vertical-axis scale.

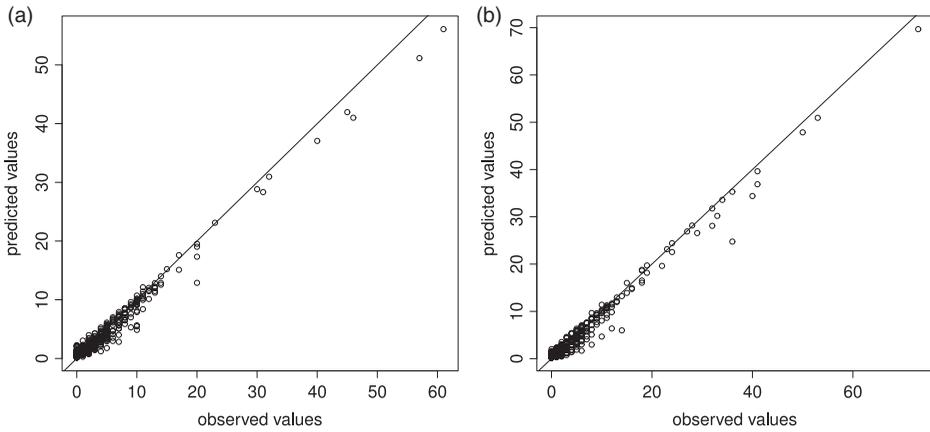


Figure 8. (a) Graphical display of observed values versus the average of the predicted values by the Poisson–Gamma–Normal model for individuals who were treated with placebo. (b) Representation of observed values versus the average of the predicted values by the Poisson–Gamma–Normal model for individuals under the new AED.

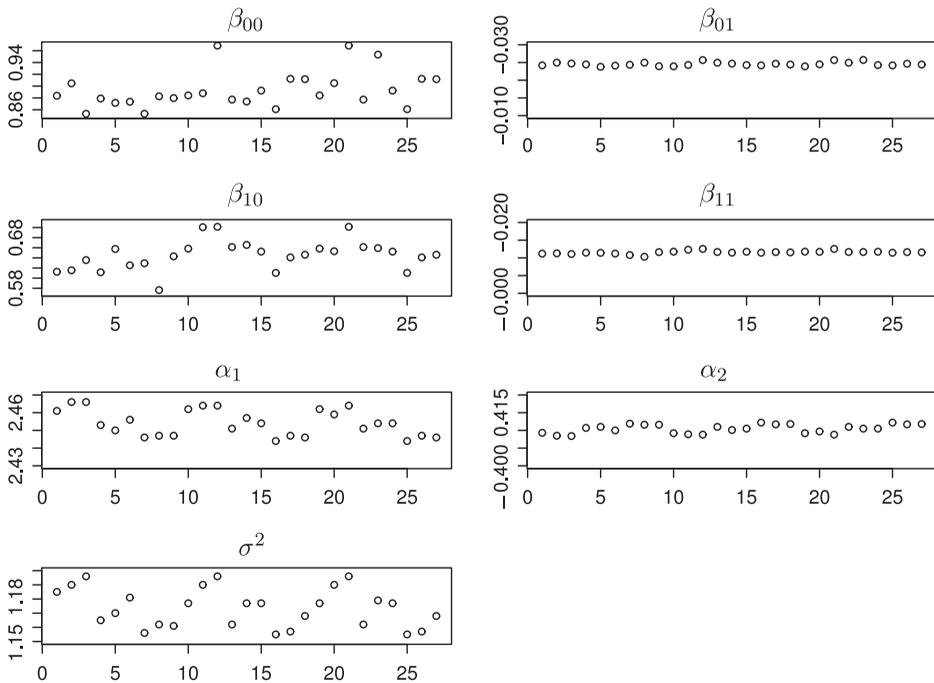


Figure 9. Range of posterior mean estimates of parameters in the Poisson–Normal–Gamma model for different values (a, b) of gamma and κ of the normal prior of the hyperparameters. The ordering of parameters from left to right on the x -axis is as follows: gamma(a, b) and $N(0, \kappa)$ 1: ($a = 1.5, b = 1/0.933, \kappa = 1$), 2: ($a = 1.5, b = 1/0.622, \kappa = 1$), 3: ($a = 1.5, b = 1/0.311, \kappa = 1$), 4: ($a = 1.5, b = 1/0.933, \kappa = 1,000,000$), 5: ($a = 1.5, b = 1/0.622, \kappa = 1,000,000$), 6: ($a = 1.5, b = 1/0.311, \kappa = 1,000,000$).

6. Simulation study

6.1. Data simulation

To investigate the efficiency of Bayesian model assessment using pivotal quantities we provide a simulation study. Based on the epilepsy data set we generated 100 data sets from

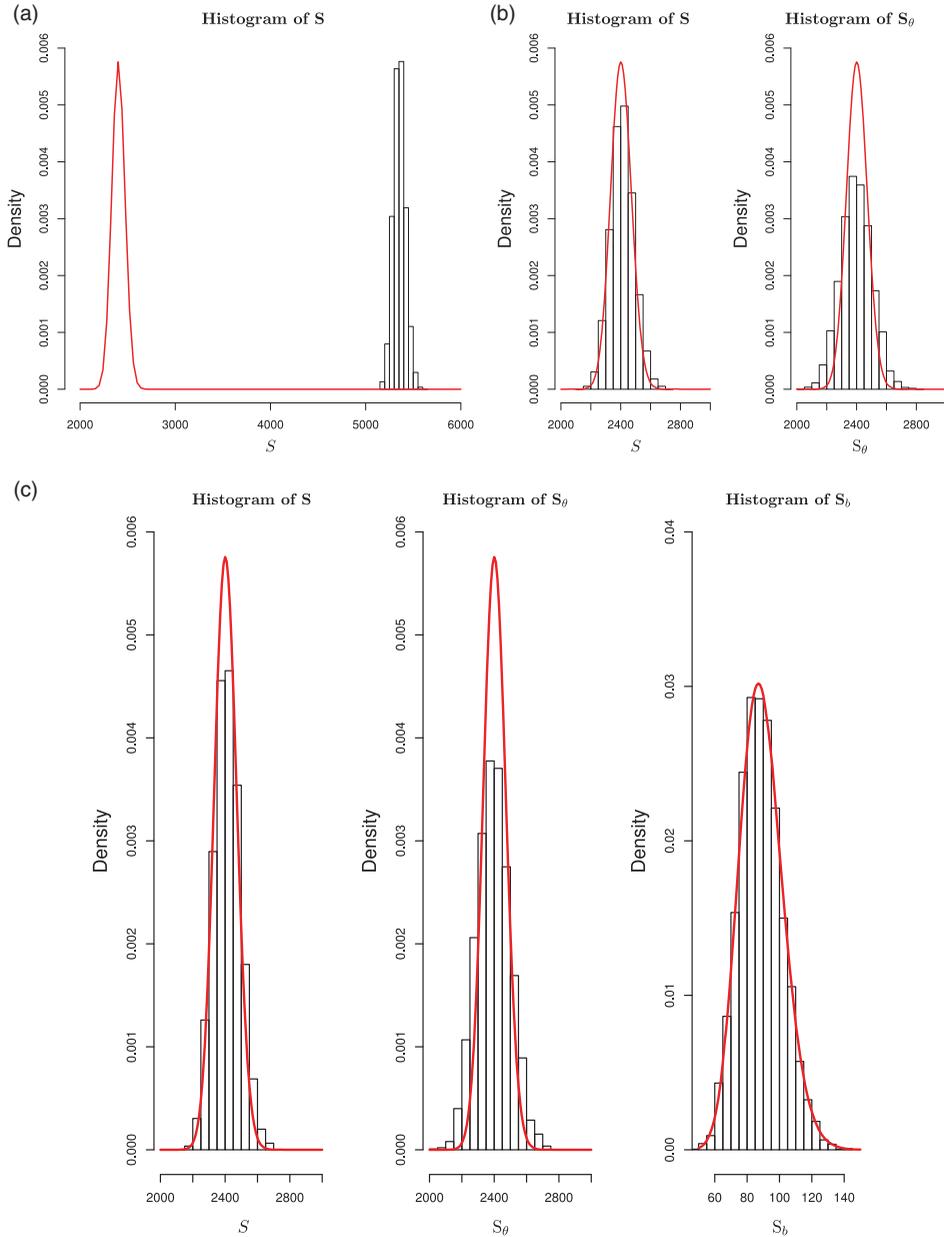


Figure 10. (a) Histogram estimate of the posterior distribution of measures S under the Poisson model and the corresponding marginal χ^2 distributions. (b) Histogram estimate of the posterior distribution of measures S and S_θ under the Poisson–Gamma model and the corresponding marginal χ^2 distributions. (c) Histogram estimate of the posterior distribution of measures S , S_θ and S_b under the combined model and the corresponding marginal χ^2 distributions.

three different models: Poisson–Gamma, Poisson–Normal, and Poisson–Normal–Gamma (combined), using the software R [24]. We considered 89 patients, 45 treated with placebo and 44 treated with a new AED, which will be observed for a period of 27 weeks. We employed the estimates resulting from analyzing the epilepsy data as input for the simulations.

Poisson–Gamma model: To generate data from the Poisson–Gamma model, the following steps were performed:

Step 1. We generated 2403 values for the gamma random effects, assuming $\alpha_1 = 2.4640$, $\alpha_2 = 1/\alpha_1$ and $\theta_{ij} \sim \text{Gamma}(\alpha_1, \alpha_2)$, $i = 1, \dots, 89$, $j = 1, \dots, 27$.

Step 2. The values for λ_{ij} , $i = 1, \dots, 89$, $j = 1, \dots, 27$ were calculated using $\lambda_{ij} = \theta_{ij} \exp(1.2662 - 0.0134t_j)$, $i = 1, \dots, 45$ and $t_j = 1, 2, \dots, 27$ for placebo and $\lambda_{ij} = \theta_{ij} \exp(1.4531 - 0.0328t_j)$, $i = 46, \dots, 89$ and $t_j = 1, 2, \dots, 27$ for the new AED.

Step 3. Finally, we generated 2403 values Y_{ij} from a Poisson distribution with mean λ_{ij} .

Such a data set was then analyzed using the Poisson, Poisson–Gamma, and Poisson–Normal–Gamma models to study the relative quality of the fits.

Poisson–Normal model: To generate data from the Poisson–Normal model, the following steps were performed:

Step 1. We generated 89 values for the normal random effects assuming $b_i \sim N(0, 1.1289)$, $i = 1, \dots, 89$. *Step 2.* The values for λ_{ij} , $i = 1, \dots, 89$, $j = 1, \dots, 27$ were calculated using $\lambda_{ij} = \exp(1.2662 + b_i - 0.0134t_j)$, $i = 1, \dots, 45$ and $t_j = 1, 2, \dots, 27$ for placebo and $\lambda_{ij} = \exp(1.4531 + b_i - 0.0328t_j)$, $i = 46, \dots, 89$ and $t_j = 1, 2, \dots, 27$ for the

Table 3. Posterior summaries considering Poisson, Poisson–Gamma, and Poisson–Normal–Gamma models.

| Parameter | Mean | sd | median | Credible interval | |
|-----------------------------------|---------|-----------------|---------|-------------------|---------|
| | | | | 2.5% | 97.5% |
| <i>Poisson model</i> | | | | | |
| β_{00} | 1.2180 | 0.0337 | 1.2180 | 1.1510 | 1.2840 |
| β_{01} | −0.0115 | 0.0022 | −0.0115 | −0.01573 | −0.0071 |
| β_{10} | 1.4470 | 0.0328 | 1.4470 | 1.3830 | 1.5129 |
| β_{11} | −0.0327 | 0.0023 | −0.0327 | −0.0373 | −0.0282 |
| DIC = 111962 | | $p_D = 4.018$ | | | |
| <i>Poisson–Gamma model</i> | | | | | |
| β_{00} | 1.2200 | 0.0519 | 12200 | 1.1180 | 1.3210 |
| β_{01} | −0.0116 | 0.0033 | −0.0116 | −0.0179 | −0.0051 |
| β_{10} | 1.4480 | 0.0518 | 1.4480 | 1.3460 | 1.5500 |
| β_{11} | −0.0328 | 0.0034 | −0.0327 | −0.0394 | −0.0263 |
| α_1 | 2.2900 | 0.1317 | 2.2870 | 2.0460 | 2.5610 |
| α_2 | 0.4381 | 0.0251 | 0.4373 | 0.3906 | 0.4889 |
| DIC = 9539.48 | | $p_D = 1105.13$ | | | |
| <i>Poisson–Normal–Gamma model</i> | | | | | |
| β_{00} | 1.2190 | 0.0521 | 1.2190 | 1.1170 | 1.3200 |
| β_{01} | −0.0116 | 0.0033 | −0.0116 | −0.0180 | −0.0050 |
| β_{10} | 1.4480 | 0.0518 | 1.4480 | 1.3460 | 1.5490 |
| β_{11} | −0.0328 | 0.0034 | −0.0328 | −0.0395 | −0.0263 |
| α_1 | 2.2970 | 0.1325 | 2.2910 | 2.0510 | 2.5740 |
| α_2 | 0.4368 | 0.0251 | 0.4364 | 0.3885 | 0.4877 |
| σ^2 | 0.0020 | 0.0025 | 0.0011 | 6.29E−5 | 0.0091 |
| DIC = 9549.39 | | $p_D = 1105.37$ | | | |

new AED. *Step 3.*] Finally, we generated 2403 values Y_{ij} from a Poisson distribution with mean λ_{ij} .

Such a data set was then analyzed using the same models as above.

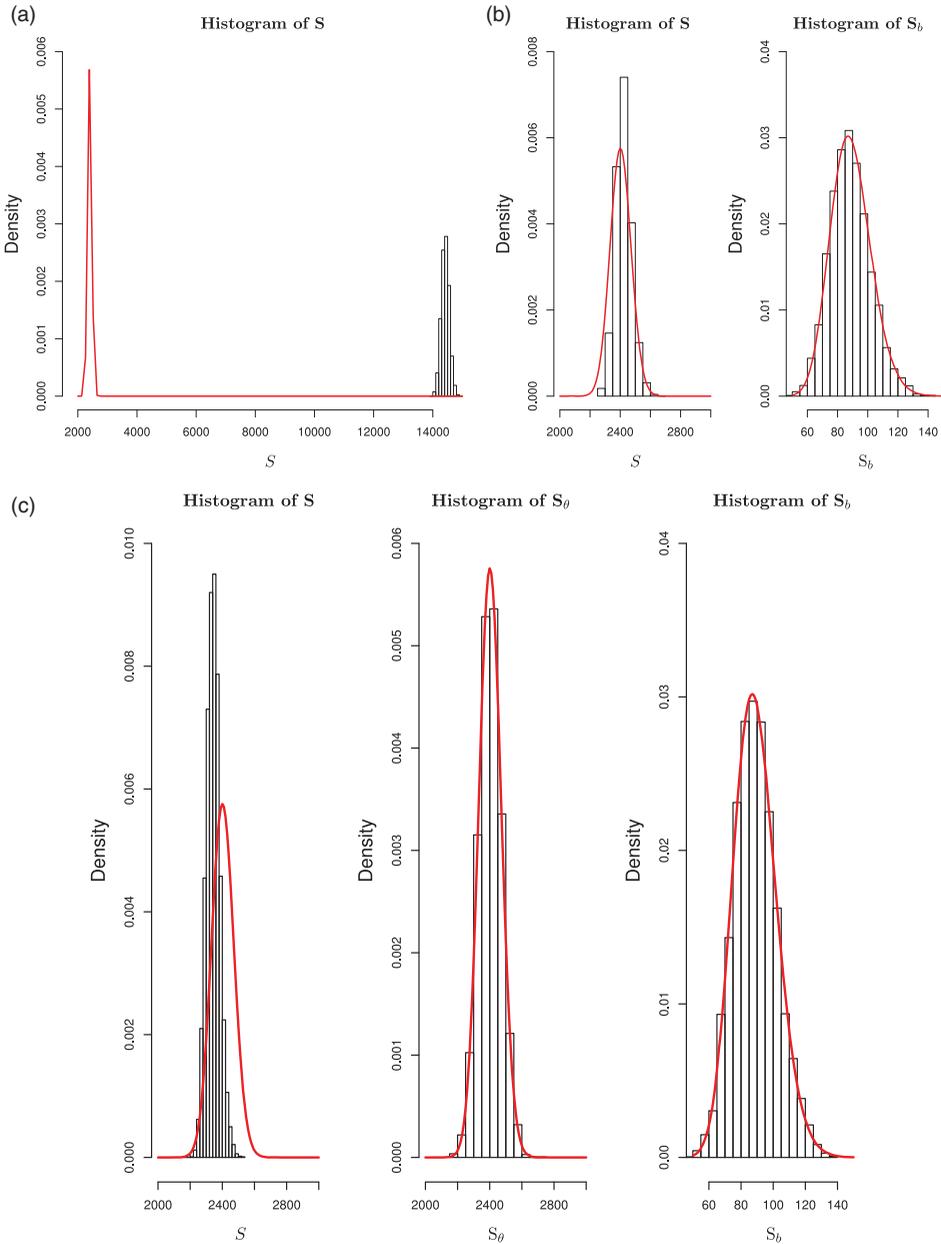


Figure 11. Poisson–Gamma simulated data: (a) Histogram estimate of the posterior distribution of measure S under the Poisson model and the corresponding marginal χ^2 distributions. (b) Histogram estimate of the posterior distribution of measures S and S_b under the Poisson–Normal model and the corresponding marginal χ^2 distributions. (c) Histograms estimate of the posterior distribution of measures S , S_θ and S_b under combined model and the corresponding marginal χ^2 distributions.

Poisson–Normal–Gamma model: To generate data from the combined model, the following steps were performed:

Step 1. We generated 2403 values for the gamma random effects assuming $\alpha_1 = 2.4640$, $\alpha_2 = 1/\alpha_1$ and $\theta_{ij} \sim \text{Gamma}(\alpha_1, \alpha_2)$, $i = 1, \dots, 89, j = 1, \dots, 27$.

Step 2. We generated 89 values for the normal random effects assuming $b_i \sim N(0, 1.1289)$, $i = 1, \dots, 89$. *Step 3.* The values for λ_{ij} , $i = 1, \dots, 89, j = 1, \dots, 27$ were calculated using $\lambda_{ij} = \theta_{ij} \exp(1.2662 + b_i - 0.0134t_j)$, $i = 1, \dots, 45$ and $t_j = 1, 2, \dots, 27$ for placebo and $\lambda_{ij} = \theta_{ij} \exp(1.4531 + b_i - 0.0328t_j)$, $i = 46, \dots, 89$ and $t_j = 1, 2, \dots, 27$ for the new AED.

Step 4. Finally, we generated 2403 values y_{ij} from a Poisson distribution with mean λ_{ij} .

Once more, the same models are then fitted to the data.

6.2. Results

For a Bayesian analysis of the simulated data set, we followed the procedures described in Section 5. For all models fitted, the chain convergence was monitored by the conventional time series plots and also using the method proposed by Gelman and Rubin [9]. We conclude that the chains are uncorrelated. Here, we restrict ourselves to one of the data sets only. In Appendix B, the graphics corresponding to Bayesian model assessment using pivots are shown for 20 of the simulated data sets.

(a) *Simulated data set from a Poisson–Gamma model:* From the histogram estimate of the joint posterior distribution of measures S, S_θ and S_b for the Poisson, Poisson–Gamma,

Table 4. Posterior summaries considering Poisson, Poisson–Normal and Poisson–Normal–Gamma models.

| Parameter | Mean | sd | median | Credible interval | |
|-----------------------------------|-------------------------|-------------------------|-------------------------|-------------------------|----------------------|
| | | | | 2.5% | 97.5% |
| <i>Poisson model</i> | | | | | |
| β_{00} | 1.7810 | 0.0253 | 1.7810 | 1.7310 | 1.8310 |
| β_{01} | −0.0139 | 0.0016 | −0.0139 | −0.0171 | −0.0106 |
| β_{10} | 1.8640 | 0.0259 | 1.8640 | 1.8120 | 1.9140 |
| β_{11} | −0.0319 | 0.0018 | −0.0319 | −0.0354 | −0.0283 |
| DIC = 18800.8 | | $p_D = 3.982$ | | | |
| <i>Poisson–Normal model</i> | | | | | |
| β_{00} | 1.2990 | 0.1599 | 1.3010 | 0.9710 | 1.6050 |
| β_{01} | −0.0138 | 0.0017 | −0.0138 | −0.0172 | −0.0105 |
| β_{10} | 1.3530 | 0.1573 | 1.3530 | 1.0460 | 1.6740 |
| β_{11} | −0.0319 | 0.0018 | −0.0319 | −0.0355 | −0.0283 |
| σ^2 | 1.0880 | 0.1735 | 1.0700 | 0.8027 | 1.4770 |
| DIC = 7843.95 | | $p_D = 88.177$ | | | |
| <i>Poisson–Normal–Gamma model</i> | | | | | |
| β_{00} | 1.2920 | 0.1595 | 1.2930 | 0.9770 | 1.5980 |
| β_{01} | −0.0139 | 0.0017 | −0.0138 | −0.0172 | −0.0106 |
| β_{10} | 1.3560 | 0.1613 | 1.3560 | 1.0390 | 1.6720 |
| β_{11} | −0.0319 | 0.0018 | −0.0319 | −0.0354 | −0.0284 |
| α_1 | 1.5150×10^6 | 3.9990×10^6 | 155900.0 | 731.6 | 1.3030×10^7 |
| α_2 | 1.6350×10^{-4} | 4.3940×10^{-4} | 6.4180×10^{-6} | 7.6940×10^{-8} | 0.0014 |
| σ^2 | 1.09 | 0.1742 | 1.073 | 0.7983 | 1.491 |
| DIC = 9040.06 | | $p_D = 93.574$ | | | |

and Poisson–Normal–Gamma (Figure 10 and Figures C.1–C20), we observe that Poisson–Gamma and Poisson–Normal–Gamma models fit much better than the Poisson model. From Table 3, we observe that the DIC value is smaller for the Poisson–Gamma model, the values for posterior summaries for the parameters of the Poisson–Gamma are similar to the assumed values, and that using the combined model nothing is added significantly, as expected. However the DIC values presented in Table C.1 are smaller for the Poisson–Gamma model. Also, Figure C.21) for the values of the posterior summaries indicate that the results for the parameters related to the mean do not vary much. Thus, it is clear that there is no need for an additional random effect b_i .

(b) *Simulated data set from a Poisson–Normal model:* From the histogram estimate of the joint posterior distribution of measures S , S_θ , and S_b for the Poisson, the Poisson–Normal,

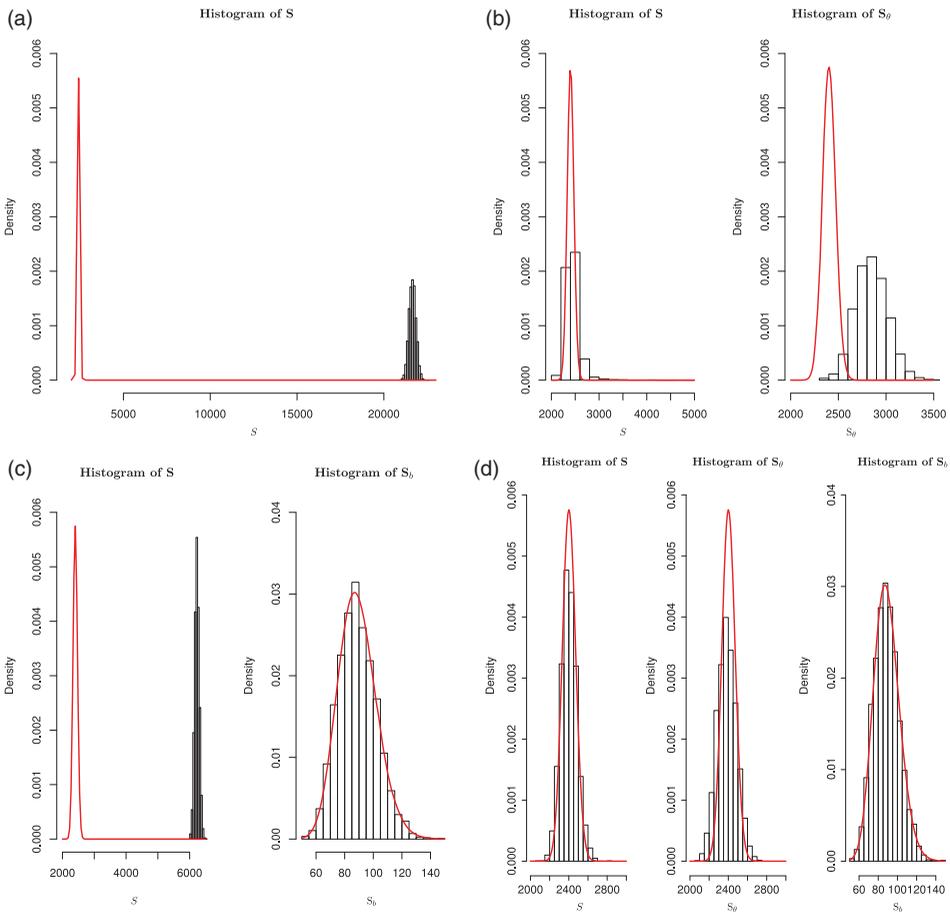


Figure 12. (a) Histogram estimates of the posterior distribution of measures S under the Poisson–Gamma model and the corresponding marginal χ^2 distributions. (b) Histogram estimates of the posterior distribution of measures S and S_θ under the Poisson–Gamma model and the corresponding marginal χ^2 distributions. (c) Histogram estimates of the posterior distribution of measures S and S_b under the Poisson–Normal model and the corresponding marginal χ^2 distributions. (d) Histogram estimates of the posterior distribution of measures S , S_θ , and S_b under the combined model and the corresponding marginal χ^2 distributions.

and the Poisson–Normal–Gamma (Figure 11), we observe that the Poisson–Normal model fits better to the data set. The graphics presented in Figures C.22–C.41 show that the Poisson–Normal and Poisson–Normal–Gamma models fit much better to the Poisson–Normal simulated data sets than the Poisson model itself. From Table 4, we observe that the DIC value is smaller for the Poisson–Normal model, the values for posterior summaries for the parameters of the Poisson–Normal model are similar to the assumed values and that by using the combined model nothing is added significantly, as expected. However, the DIC values presented in Table C.2 are the same for the Poisson–Gamma and Poisson–Normal–Gamma models. Figure C.42, presenting the values of the posterior summaries, indicates that the results for the parameters related to the random effect θ_{ij} do not vary much, underscoring that there is no need for the additional random effect θ_{ij} .

Simulated data set from a combined model: From the histogram estimate of the joint posterior distribution of measures S , S_θ , and S_b for the Poisson, Poisson–Gamma, Poisson–Normal, and Poisson–Normal–Gamma (Figure 12), we observe that the combined model fits better to the data, as expected. Figures C.43–C.62 indicate that the Poisson–Normal–Gamma model fits much better to the simulated data sets. From Table 5, we

Table 5. Posterior summaries considering Poisson, Poisson–Gamma, Poisson–Normal and Poisson–Normal–Gamma models.

| Parameter | Mean | sd | median | Credible interval | |
|-----------------------------------|---------|-----------------|---------|-------------------|---------|
| | | | | 2.5% | 97.5% |
| <i>Poisson model</i> | | | | | |
| β_{00} | 1.7810 | 0.0254 | 1.7810 | 1.7310 | 1.8310 |
| β_{01} | −0.0144 | 0.0017 | −0.0144 | −0.0178 | −0.0111 |
| β_{10} | 1.8700 | 0.0265 | 1.8700 | 1.8181 | 1.9220 |
| β_{11} | −0.0354 | 0.0019 | −0.0355 | −0.0391 | −0.0317 |
| DIC = 21576.8 | | $p_D = 4.007$ | | | |
| <i>Poisson–Gamma model</i> | | | | | |
| β_{00} | 1.7770 | 0.0764 | 1.7770 | 1.6290 | 1.9250 |
| β_{01} | −0.0142 | 0.0048 | −0.0142 | −0.0234 | −0.0049 |
| β_{10} | 1.8810 | 0.0775 | 1.8790 | 1.7360 | 2.0380 |
| β_{11} | −0.0361 | 0.0049 | −0.0361 | −0.0459 | −0.0267 |
| α_1 | 0.7045 | 0.0247 | 0.7040 | 0.6575 | 0.7541 |
| α_2 | 1.4210 | 0.0498 | 1.4210 | 1.3260 | 1.5210 |
| DIC = 9800.01 | | $p_D = 1546.28$ | | | |
| <i>Poisson–Normal model</i> | | | | | |
| β_{00} | 1.3000 | 0.1549 | 1.302 | 0.9966 | 1.6010 |
| β_{01} | −0.0144 | 0.0017 | −0.0144 | −0.0177 | −0.0112 |
| β_{10} | 1.3860 | 0.1611 | 1.3830 | 1.0730 | 1.6970 |
| β_{11} | −0.0354 | 0.0019 | −0.0352 | −0.0390 | −0.0317 |
| σ^2 | 1.0830 | 0.1730 | 1.0670 | 0.7961 | 1.4670 |
| DIC = 12245.9 | | $p_D = 89.175$ | | | |
| <i>Poisson–Normal–Gamma model</i> | | | | | |
| β_{00} | 1.2950 | 0.1655 | 1.2990 | 0.9588 | 1.6150 |
| β_{01} | −0.0140 | 0.0032 | −0.0140 | −0.0203 | −0.0078 |
| β_{10} | 1.3830 | 0.1577 | 1.3850 | 1.0780 | 1.6950 |
| β_{11} | −0.0355 | 0.0034 | −0.0355 | −0.0422 | −0.0289 |
| α_1 | 2.5720 | 0.1440 | 2.5700 | 2.3010 | 2.8670 |
| α_2 | 0.3900 | 0.0218 | 0.3890 | 0.3489 | 0.4346 |
| σ^2 | 1.0710 | 0.1736 | 1.055 | 0.7815 | 1.4610 |
| DIC = 9542.96 | | $p_D = 1110.56$ | | | |

observe that the DIC value is smaller for the Poisson–Normal–Gamma model, the values for posterior summaries for the parameters of the Poisson–Normal–Gamma and Poisson–Normal models are similar to the assumed values, a phenomenon not observed for the other models. Additionally, the DIC values presented in Table C.3 are smaller for the Poisson–Normal–Gamma model. Figure C.63, showing the values of the posterior summaries, indicates that the results for the parameters related to the random effect b_i and θ_{ij} are a necessary improvement of the model.

7. Concluding remarks

In this paper, we have developed Bayesian methods to analyze longitudinal count data, potentially subject to overdispersion. For this, the so-called Poisson–Normal–Gamma, or combined, model was used [27]. The original authors used maximum likelihood. The model was fitted to a set of data, introduced in Section 2. Apart from the Poisson–Gamma–Normal model, three special cases were considered: the Poisson–Normal, negative-binomial, and conventional Poisson models. Furthermore, we considered the Poisson–Gamma–Gamma model [13]. We found that a Bayesian approach is well feasible and hence a viable route, in particular thanks to recent software to simulate samples for the joint posterior distribution of interest. The OpenBugs software leads to great simplification to obtain the posterior summaries of interest.

A further novel contribution is the development of Bayesian model assessment. This is important, because until now no general model assessment tools for the combined model were developed. Our assessment is based on graphical analysis of pivotal quantities. We found that the Poisson–Normal–Gamma fits the data better than the simplified versions considered. The model reached the proposed objective of capturing the variability among individuals and the correlation between measurements across time. Finally, we concluded that, based on the Poisson–Normal–Gamma model, the treatments do not differ as in Molenberghs *et al.* [17] and Vangeneugden *et al.* [27].

A simulation study was conducted making use of the methodology proposed by Johnson [12], for which model comparison is based on the analysis of pivotal quantities. Precisely, our aim was to see whether this is a good strategy to select the best fitting models. The fit is usually best when the models used for simulation and analysis coincide. Evidently, this is in line with expectation, but is useful in confirming that the procedures work in the way they are supposed to. However, in some cases the Poisson–Gamma–Normal model fits better, computationally, to data generated from a Poisson–Normal model, than the Poisson–Normal model itself. This seems surprising but can be explained as follows: when the normal random effect generated is large, very large counts are generated. Fitting the Poisson–Normal model then becomes relatively unstable, because the normal random effect distribution needs to compromise between the large variance and the correlation. The Poisson–Gamma–Normal model is better equipped to describe the variance and correlation functions in their own terms. So, it can happen that the more elaborate model is, at least in some cases, easier to fit.

Relevant questions center around model fit, influential subjects, and outliers. The combined model, in itself, can be used to assess the fit of its sub-models, such as the GLMM. Importantly, we have proposed pivotal quantities to assess model fit from a Bayesian perspective. It is also of interest to assess the impact of outliers and influential subjects on the

conclusions. Work to this effect is being done, admittedly currently for a likelihood-based context [19,20].

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Fernanda Rizzato was partially supported by CNPq and CAPES (Brazilian Science Funding Agencies) and Clarice Demétrio by CNPq. Geert Molenberghs gratefully acknowledges support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy).

References

- [1] M.J. Bayarri and J.O. Berger, *P values for composite null models*, J. Amer. Statist. Assoc. 95 (2000), pp. 1127–1142.
- [2] J.G. Booth, G. Casella, H. Friedl, and J.P. Hobert, *Negative binomial loglinear mixed models*, Stat. Model 3 (2003), pp. 179–181.
- [3] E.A. Coelho-Barros, J.A. Achcar, and J. Mazucheli, *Longitudinal Poisson modeling: An application for CD4 counting in HIV-infected patients*, J. Appl. Statist. 5 (2010), pp. 865–880.
- [4] D.K. Dey, A.E. Gelfand, T.B. Swartz, and P.K. Vlachos, *A simulation-intensive approach for checking hierarchical models*, Test 7 (2003), pp. 325–346.
- [5] E. Faught, B.J. Wilder, R.E. Ramsay, R.A. Reife, L.D. Kramer, G.W. Pledger, and R.M. Karim, *Topiramate placebo-controlled dose-ranging trial in refractory partial epilepsy using 200-, 400-, and 600-mg daily dosages*, Neurology 46 (1996), pp. 1684–1690.
- [6] A. Fotouhi, *Modelling overdispersion in longitudinal count data in clinical trials with application to epileptic data*, Contemp. Clin. Trials 29 (2008), pp. 547–554.
- [7] S. Geisser, *Predictive Inference*, Chapman & Hall, New York, 1993.
- [8] A. Gelman, X.L. Meng, and H. Stern, *Posterior predictive assessment of model fitness via realized discrepancies*, Statist. Sinica 6 (1996), pp. 733–807.
- [9] A. Gelman and D.B. Rubin, *Inference from iterative simulation using multiple sequences*, Statist. Sci. 7(4) (1992), pp. 457–511.
- [10] F.N. Gumedze and T.D. Chatora, *Detection of outliers in longitudinal count data via overdispersion*, Comput. Statist. Data Anal. 79 (2014), pp. 192–202.
- [11] I. Guttman, *The use of the concept of a future observation in goodness-of-fit problems*, J. R. Stat. Soc. Ser. B 29 (1967), pp. 83–100.
- [12] V.E. Johnson, *Bayesian model assessment using pivotal quantities*, Bayesian Anal. 2 (2007), pp. 719–734.
- [13] Y. Lee and J.A. Nelder, *Two ways of modelling overdispersion in non-normal data*, Appl. Stat. 49 (2000), pp. 591–598.
- [14] Y. Lee, J.A. Nelder, and Y. Pawitan, *Generalized Linear Models with Random Effects: Unified Analysis via h-likelihood*, Chapman & Hall/CRC, Boca Raton, FL, 2006.
- [15] R.B. Millar, *Comparison of hierarchical Bayesian models for overdispersed count data using DIC and Bayes' factors*, Biometrics 65 (2009), pp. 962–969.
- [16] G. Molenberghs and G. Verbeke, *Models for Discrete Longitudinal Data*, Springer, New York, 2005.
- [17] G. Molenberghs, G. Verbeke, and C.G.B. Demétrio, *An extended random-effects approach to modeling repeated, overdispersion count data*, Lifetime Data Anal. 13 (2007), pp. 457–511.
- [18] G. Molenberghs, G. Verbeke, C.G.B. Demétrio, and A. Vieira, *A family of generalized linear models for repeated measures with normal and conjugate random effects*, Statist. Sci. 25 (2010), pp. 325–347.

- [19] T. Rakhmawati, G. Molenberghs, G. Verbeke, and C. Faes, *Local influence diagnostics for generalized linear mixed models with overdispersion*, submitted for publication, 2015.
- [20] T. Rakhmawati, G. Molenberghs, G. Verbeke, and C. Faes, *Local influence diagnostics for hierarchical count data models with overdispersion and excess zeros*, submitted for publication, 2015.
- [21] J.M. Robins, A. Van der Vaart, and V. Ventura, *Asymptotic distribution of p values in composite null models*, J. Amer. Statist. Assoc. 95 (2000), pp. 1143–1159.
- [22] M. Roos and L. Held, *Sensitivity analysis in Bayesian generalized linear mixed models for binary data*, Bayesian Anal. 6 (2011), pp. 259–278.
- [23] D.B. Rubin, *Bayesianly justifiable and relevant frequency calculations for the applied statistician*, Ann. Statist. 12 (1984), pp. 1151–1172.
- [24] R Development Core Team, *R: A Language and Environment for Statistical Computing*, Vienna: R Foundation for Statistical Computing, 2009. Available at <http://www.R-project.org>.
- [25] D.J. Spiegelhalter, N.G. Best, B.P. Carlin, and A. Linde, *Bayesian measures of model complexity and fit*, J. R. Stat. Soc. Ser. B 64 (2002), pp. 583–639.
- [26] D.J. Spiegelhalter, A. Thomas, N. Best, and W. Gilks, *Bugs: Bayesian inference Using Gibbs Sampling Manual*, 1996. Available at <http://www.mrc-bsu.cam.ac.uk/bugs>.
- [27] T. Vangeneugden, G. Molenberghs, G. Verbeke, and C.G.B. Demétrio, *Marginal correlation from an extended random-effects model for repeated, overdispersed counts*, J. Appl. Statist. 38 (2011), pp. 215–232.