A generalized Bayesian nonlinear mixed effects regression model for zero inflated longitudinal count data in tuberculosis trials

Divan Aristo Burger[1] | **Robert Schall**[2,3] | **Rianne Jacobs**[4] | **Ding-Geng Chen**[1]

[1] *Department of Statistics, University of Pretoria, Pretoria, South Africa*

[2] *Department of Mathematical Statistics and Actuarial Science, University of the Free State, Bloemfontein, South Africa*

[3] *IQVIA, Biostatistics, Bloemfontein, South Africa*

[4] *Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University Groningen, Groningen, The Netherlands*

KEY WORDS: Bayesian, bactericidal activity, longitudinal, mixed effects, zero inflated

---

Address correspondence to Divan Aristo Burger, Department of Statistics, University of Pretoria, Pretoria 0028, South Africa; E-mail: divanaburger@gmail.com

ABSTRACT

In this paper we investigate Bayesian generalized nonlinear mixed effects (NLME) regression models for zero inflated longitudinal count data. The methodology is motivated by and applied to colony forming unit (CFU) counts in extended bactericidal activity tuberculosis (TB) trials. Furthermore, for model comparisons we present a generalized method for calculating the marginal likelihoods required to determine Bayes factors. A simulation study shows that the proposed zero inflated negative binomial regression model has good accuracy, precision and credibility interval coverage. In contrast, conventional normal NLME regression models applied to log-transformed count data, which handle zero counts as left censored values, may yield credibility intervals that undercover the true bactericidal activity of anti-TB drugs. We therefore recommend that zero inflated NLME regression models should be fitted to CFU count on the original scale, as an alternative to conventional normal NLME regression models on the logarithmic scale.

1. INTRODUCTION

Generalized linear models specifying either the Poisson or negative binomial distributions are basic models for count data. These models can be extended to relatively complex situations, including regression models, mixed effects models, and models that accommodate overdispersion, underdispersion, and excess zeros.[1–3] However, in particular when very large counts are to be modeled, such as counts of viral or bacterial load in medical applications,[4–7] or large counts in biology and ecology,[8,9] researchers often log-transform the data. After doing so, the data might be expected to be normally distributed, at least approximately, and standard statistical procedures, such as normal linear models, including mixed effects linear models, are applied. However, this approach has been criticized as a remnant of traditional statistical practice from an era when practical techniques were not yet available to model count data on the original scale.[10] Nevertheless, in relatively complex cases such as nonlinear models for hierarchically structured data,[11] the practice of log-transformation of count data is still common, despite the fact that generalized mixed effects models[12] may serve as alternatives. Moreover, imputation of zero counts on the logarithmic scale (as a workaround) can lead to biased parameter estimates when the variability of the counts is large and the mean counts are small.[10,13] Thus an important advantage of generalized mixed effects models for count data on the original scale is that zero counts do not have to be imputed, or specified as left censored values,[14,15] on the logarithmic scale.

Generalized linear mixed effects models and generalized nonlinear mixed effects (NLME) models for count data should adequately accommodate both overdispersion and excess amounts of zeros through, for example, zero inflated Poisson or zero inflated negative binomial models.[16,17] Zero inflated count models fit a mixture distribution, with a point mass at zero, and a Poisson or negative binomial distribution for the “count” part.[3,18] Ignoring zero inflation in the data may yield biased estimates of fixed and random effects, and thus may lead to inaccurate

inferences for model parameters of interest.[3]

Although a wide range of literature is available on Bayesian implementations of zero inflated generalized linear mixed effects models,[19–21] relatively little work has been done on zero inflated generalized NLME models. For example, Sae-Lim et al.[22] fitted Bayesian zero inflated Poisson and negative binomial NLME models to tick counts from animals on tick infested fields, to measure within-breed genetic variation and heritability.

For the purpose of model comparisons of zero inflated generalized NLME models using Bayes factors, the marginal likelihoods need to be approximated. Approximation techniques for marginal likelihoods include the Laplace-Metropolis approximation[23,24] (the so-called compound Laplace-Metropolis marginal likelihood (CLMML)), and the widely used harmonic mean sampling estimator.[25] Here it may be noted that in some cases the harmonic mean sampling estimator has infinite variance (which does not adhere to the central limit theorem)[26] and preferably should be abandoned in favor of CLMMLs.

The objectives of the present paper are to propose and evaluate zero inflated generalized NLME regression models for longitudinal count data that belong to the exponential family of distributions. The methodology is motivated by and applied to colony forming unit (CFU) counts in extended bactericidal activity tuberculosis (TB) trials. Furthermore, for model comparisons we present a generalized method for calculating the marginal likelihoods required to determine CLMMLs and ultimately Bayes factors, and we provide methods for calculating the deviance information criterion (DIC) statistic and Pearson goodness of fit measure. In addition, the NLME regression model is formulated as a zero inflated generalized Poisson (ZIGP) regression model to serve as an extension to the proposed regression models that fall within the exponential family of distributions.[27]

The paper is organized as follows: Section 2 provides an overview of analysis methods for CFU count data in extended bactericidal activity TB trials. Section 3 introduces a general zero

inflated generalized NLME regression model for CFU count over time. Special cases of the general zero inflated generalized NLME regression model are outlined in Section 4. The ZIGP regression model (as an extended version of the regression models presented in Section 4) is outlined in the online appendix of this paper. DIC statistics and Bayes factors for discriminating between the various NLME regression models, and the Pearson goodness of fit measure, are introduced in Section 5. Section 6 provides applications of the proposed methodology using a CFU count dataset of a recently published extended bactericidal activity TB trial. Section 7 presents a simulation study to compare the performance of the recommended regression models. Section 8 discusses the results and key findings of the paper.

2. MOTIVATING DATA

The methods in this paper are motivated by CFU count data in extended bactericidal activity TB trials. In such trials the bactericidal and sterilization activity of TB drugs are characterized by the rate of change in CFU count. Decline in CFU count during a particular treatment period typically is bilinear or biphasic over time, and has conventionally been modeled through linear and nonlinear regression, after log-transformation of the count data.

2.1. Early and extended bactericidal activity of tuberculosis drugs

The assessment of the early bactericidal activity (EBA) of anti-TB drugs, namely bactericidal activity during the first 14 days of treatment, is a key phase of the development of new anti-TB drugs and treatments.[28] Thus EBA trials are conducted to characterize the short-term efficacy of anti-TB drugs in development Phase 2a trials.[6] In such trials, overnight sputum samples are collected, usually daily, from Day 0 up to Day 14 after the start of treatment.[6] In these sputum samples, so-called colony forming units (CFUs) are counted. Based on these counts, the EBA of anti-TB drugs is characterized by the rate of decline in CFU count over time, on

the logarithmic scale to the base of 10 (or $\log_{10}(\text{CFU})$ count).[6] The EBA over a certain time interval, calculated from a $\log_{10}(\text{CFU})$ versus time profile, is expressed as follows[29]:

$$\text{EBA}(t_1\text{–}t_2) = -\frac{\hat{f}(t_2) - \hat{f}(t_1)}{t_2 - t_1} \tag{1}$$

where $f(t)$ is an appropriate regression function for $\log_{10}(\text{CFU})$ count versus time, and $\hat{f}(t_1)$ and $\hat{f}(t_2)$ are the corresponding fitted values at Day t_1 and Day t_2, respectively. From Equation (1) it can be seen that the anti-TB activity of a given drug or treatment becomes larger as $\text{EBA}(t_1\text{–}t_2)$ increases. Most anti-TB drugs, such as isoniazid,[30,31] cause a relatively fast decline in $\log_{10}(\text{CFU})$ count during the initial phase of treatment, thereby eradicating most of the TB bacteria during this time.

In contrast to the concept of EBA, the sterilization property of TB drugs refers to the rate of decline in $\log_{10}(\text{CFU})$ count after the initial phase of treatment (i.e. the rate of decline once the majority of TB bacteria have been eradicated).[32] That is, the sterilization activity of anti-TB drugs refers to their activity against persistent TB microorganisms surviving the first few days of treatment.[32] Extended bactericidal activity trials have longer treatment duration than EBA trials and typically span 56 days. For example, in the trial reported by Rustomjee et al.[33] overnight sputum samples were collected from TB infected patients on Days 0, 2, 7, 14, 21, 28, 35, 42, 49 and 56 following the start of treatment. Extended bactericidal activity trials might be conducted during later stages of development of anti-TB drugs, such as in late Phase 2 (Phase 2b) trials, but before the drug enters Phase 3 trials.[34] Rustomjee et al.[33] clearly distinguish between the EBA and sterilizing activity of each of the treatment regimens studied. In fact, for each treatment the $\log_{10}(\text{CFU})$ counts over time suggest that the initial slope (rate of decline of CFU counts) is substantially larger than the terminal slope.

Regression slopes of $\log_{10}(\text{CFU})$ count over 56 days of treatment, time to sputum culture conversion in liquid media, and the proportion of patients with negative sputum culture after 56

days of treatment have been used as surrogate markers of the long term efficacy of anti-TB drugs (Phase 3; favorable versus unfavorable outcome).[35,36] Figure 1, adapted from Mitchison and Davies[35], illustrates the relationship between the aforementioned standard efficacy endpoints of 56-day extended bactericidal activity trials.

CFU counts in anti-TB trials are conventionally calculated as:

$$\text{CFU} = \frac{1}{n}\sum_{x=1}^{n} \text{CFU}_x \times \text{factor} \times 10^{\text{dilution}} \tag{2}$$

where CFU_x is the count of culture plate x (from n replicate plates in total), and "factor $\times$ 10^{dilution}" compensates for the dilution used in the counting process.

Inspection of CFU count data from a wide range of previous trials (see Burger and Schall[14]) showed that the variance of CFU counts on the original scale increases with their mean. That is, if one assumes that the variance of the data Y is of the form $\text{Var}\,(Y) = \sigma^2\left[\text{E}\,(Y)\right]^2 = \sigma^2\mu^2$ for some constant σ^2, then $\text{CV}\,(Y) = \frac{\sqrt{\text{Var}(Y)}}{\text{E}(Y)} = \frac{\sigma\mu}{\mu} = \sigma$. Thus, the coefficient of variation (CV) of Y is constant over all μ. Approximately, constant variance of CFU count on the logarithmic scale was confirmed by the extensive empirical study reported by Burger and Schall[14]. The observation of constant variance on the logarithmic scale suggests that the mean-variance relationship of the distribution of CFU count on the original scale is characterized by a constant CV. As a consequence, one promising approach to modeling CFU count over time was to regress the data against time on the logarithmic scale.

CFU counts of zero are considered valid data, and in analyses based on log-transformed counts, zero counts have been accommodated as censored observations. The censoring limit for zero counts on the logarithmic scale was chosen as follows: In Equation (2), suppose that $n = 2$, factor $= 20$ & dilution $= 0$[6,15], then the smallest possible count above zero is 1 for one plate, and zero for the other plate, and therefore the smallest value of $\log_{10}(\text{CFU})$ from a non-zero count is 1. Thus, on the logarithmic scale to the base of 10, a CFU count

of zero is specified as a left censored value of 1.[33] Combination therapy with effective anti-TB drugs such as moxifloxacin, pretomanid (or PA-824) and pyrazinamide causes substantial eradication of TB bacteria (hence, a steep decline in $\log_{10}(\text{CFU})$ count over time), which yields CFU counts of zero relatively early in the course of treatment.[37]

2.2. Conventional mixed effects models for colony forming unit count

As pointed out in Section 2.1, the EBA of an anti-TB drug or treatment regimen is conventionally calculated based on a regression fit of $\log_{10}(\text{CFU})$ counts over time. When jointly modeling the data of all patients in a given treatment group, or when modeling jointly the data of all patients from several treatment groups in a clinical trial, hierarchical mixed effects models of increasing complexity need to be fitted, resulting in random coefficients models. Most mixed effects regression models used in practice are linear and assume that the residual terms are normally distributed.[38] However, $\log_{10}(\text{CFU})$ count typically follows a bilinear or nonlinear pattern over time (e.g. as constituted by the sterilization activity of anti-TB drugs; see Figure 1). In order to account for the biphasic nature of $\log_{10}(\text{CFU})$ versus time curves, two types of nonlinear regression models have been described in the literature, namely bilinear and bi-exponential regression.

The use of NLME regression models for $\log_{10}(\text{CFU})$ count from TB trials was first proposed by Davies et al.[39] in the form of bi-exponential mixed effects regression models. However, the bi-exponential regression model is not appropriate for $\log_{10}(\text{CFU})$ versus time profiles that are decreasing slowly during the early phase of treatment, followed by a faster decline. The NLME regression model of Burger and Schall[14,15,40] has recently been introduced for the modeling of $\log_{10}(\text{CFU})$ count versus time profiles, and was shown to be more flexible than bi-exponential regression models in the sense that they allow for terminal rates of decline to be greater than initial rates of decline and *vice versa*. Burger and Schall[14,15] proposed a Bayesian method

for fitting the required hierarchical nonlinear regression models, for estimation of the model parameters and for inference about relevant parameter contrasts.

The Bayesian NLME regression of Burger and Schall[14,15] was fitted to CFU counts from several clinical trials that assessed the sterilization activity of anti-TB drugs.[37,41,42] The regression model fitted to these datasets assumed normally and skew Student t distributed residuals (on the logarithmic scale), and zero counts were specified as left censored values of 1, to avoid underestimation of the variances in random effects.[33]

As an alternative to performing “normal-theory" nonlinear regression of $\log_{10}(\text{CFU})$ count against time (where zero counts are treated as left censored values), in the following section we propose zero inflated generalized NLME regression models fitted to CFU count on the original scale (therefore treating zero counts as observed data instead of left censored observations).

3. ZERO INFLATED GENERALIZED NONLINEAR MIXED EFFECTS REGRESSION MODEL

We investigate four distributions for CFU count on the original scale: (i) Poisson, (ii) zero inflated Poisson (ZIP), (iii) negative binomial (NEGBIN) and (iv) zero inflated negative binomial (ZINB). The ZINB model, in particular, might be most appropriate when dealing with cases where the data exhibit overdispersion, together with an excess amount of zeros.[43] Furthermore, the magnitude of overdispersion, and the excess amount of zeros in CFU count may be time dependent (thus, may either increase or decrease over time). The dispersion and zero inflation parameters should therefore be allowed to vary freely over time.

Consider the following likelihood function, belonging to the exponential family of distributions,

of the zero inflated generalized NLME regression model for CFU count over time:

$$L\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk} | y_{ijk}\right) = \pi_{jk} I\left(y_{ijk} = 0\right) + \left(1 - \pi_{jk}\right) h\left(y_{ijk}, \boldsymbol{\tau_{jk}}\right) \times \tag{3}$$
$$\exp\left[\boldsymbol{\eta}\left(f\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, t_{ijk}, o_{ijk}\right), \boldsymbol{\tau_{jk}}\right) \cdot \boldsymbol{T}\left(y_{ijk}\right) - A\left(f\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, t_{ijk}, o_{ijk}\right), \boldsymbol{\tau_{jk}}\right)\right]$$

where $y_{ijk} = \sum_{l=1}^{n_{ijk}} \text{CFU}_{ijkl}$ is the total of n_{ijk} bacterial plate counts for patient $i = 1, \ldots, N_j$ in treatment group $j = 1, \ldots, J$ at timepoint $k = 1, \ldots, K_{ij}$ (or K), $t_{ijk} \geq 0$ is the corresponding measurement time, and $I\left(a\right)$ denotes an indicator function taking the value 1 if a is true, and 0 otherwise. Furthermore, the function f describes the nonlinear relationship between t_{ijk} and vectors of model parameters $\boldsymbol{\phi_{ij}}$ & $\boldsymbol{\theta_j}$, and $\boldsymbol{\tau_{jk}}$ are vectors of additional model parameters that describe the underlying distribution of the data per timepoint (e.g. characteristics such as dispersion). π_{jk} are the proportions of excess zeros, and o_{ijk} are offset constants, comprising the total number of replicate plates & dilution factor of the sputum sample in question, to warrant the modeling of CFU counts as per Equation (2). Furthermore, $\boldsymbol{\eta}$ and $\boldsymbol{T}$ are respectively vectors of canonical link functions and sufficient statistics of the data, whereas h and A are respectively the carrier measure and normalization factor.[1] Here, $\boldsymbol{\eta} \cdot \boldsymbol{T}$ signifies the sum of the dot product of vectors $\boldsymbol{\eta}$ and $\boldsymbol{T}$. The regression model is fitted jointly to the data of all patients from all treatment groups from a given trial.

The vectors of parameters $\boldsymbol{\phi_{ij}}$ are assumed to vary between patients as follows:

$$\boldsymbol{\phi_{ij}} = \boldsymbol{\phi_j} + \boldsymbol{\varphi_{ij}} \tag{4}$$

where $\boldsymbol{\varphi_{ij}}$ (or $\boldsymbol{\phi_{ij}}$) represent vectors of random effects, whereas $\boldsymbol{\phi_j}$ and $\boldsymbol{\theta_j}$ represent vectors of fixed effects. The fixed effects represent the average effect for each treatment group. Eventually, the fixed effects allow one to make inferences on the average bactericidal activity or sterilization activities of the treatment groups, and about the differences between treatment groups. The random effects, in the other hand, allow for separate regression curves for each

patient.

Assume $\boldsymbol{\phi_{ij}}$ follow multivariate distributions g with mean $\boldsymbol{\phi_j}$ and unstructured covariance matrices Ψ_j, so that:

$$\boldsymbol{\phi_{ij}} \sim g\left(\boldsymbol{\phi_j}, \Psi_j\right) \tag{5}$$

Accordingly, the likelihood of $\boldsymbol{\phi_{ij}}$, $\boldsymbol{\phi_j}$, Ψ_j, $\boldsymbol{\theta_j}$, $\boldsymbol{\tau_{jk}}$ and π_{jk} for patient i assigned to treatment group j can be written as follows:

$$L\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\phi_j}, \Psi_j, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk} | \boldsymbol{y_{ij}}\right) = \left(\prod_{k=1}^{K_{ij}} L\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk} | y_{ijk}\right)\right) P\left(\boldsymbol{\phi_{ij}} | \boldsymbol{\phi_j}, \Psi_j\right) \tag{6}$$

where $P\left(\boldsymbol{\phi_{ij}} | \boldsymbol{\phi_j}, \Psi_j\right)$ denotes the probability density function of $\boldsymbol{\phi_{ij}} | \boldsymbol{\phi_j}, \Psi_j$, and $\boldsymbol{y_{ij}}$ denote $K_{ij} \times 1$ vectors containing $\left(y_{ij1}, \ldots, y_{ijk}, \ldots, y_{ijK_{ij}}\right)'$.

The resulting joint posterior distribution of $\boldsymbol{\phi_{ij}}$, $\boldsymbol{\phi_j}$, Ψ_j, $\boldsymbol{\theta_j}$, $\boldsymbol{\tau_{jk}}$ and π_{jk} (for all $j = 1, \ldots, J$, $i = 1, \ldots, N_j$ and $k = 1, \ldots, K$) can be written as follows:

$$\begin{aligned} &P\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\phi_j}, \Psi_j, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk}, j = 1, \ldots, J, i = 1, \ldots, N_j, k = 1, \ldots, K | \boldsymbol{y}\right) \\ &= \left(\prod_{j=1}^{J} \prod_{i=1}^{N_j} L\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\phi_j}, \Psi_j, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk} | \boldsymbol{y_{ij}}\right)\right) \prod_{j=1}^{J} \left(P\left(\boldsymbol{\phi_j}\right) P\left(\Psi_j\right) P\left(\boldsymbol{\theta_j}\right) \prod_{k=1}^{K} \left[P\left(\boldsymbol{\tau_{jk}}\right) P\left(\pi_{jk}\right)\right]\right) \end{aligned} \tag{7}$$

where $\boldsymbol{y}$ denotes the $\sum_{j=1}^{J} \sum_{i=1}^{N_j} K_{ij} \times 1$ vector containing $\boldsymbol{y_{ij}}$ for all $j = 1, \ldots, J$ and $i = 1, \ldots, N_j$. $P\left(\boldsymbol{\phi_j}\right)$, $P\left(\Psi_j\right)$, $P\left(\boldsymbol{\theta_j}\right)$, $P\left(\boldsymbol{\tau_{jk}}\right)$ and $P\left(\pi_{jk}\right)$ respectively denote the prior density functions of $\boldsymbol{\phi_j}$, Ψ_j, $\boldsymbol{\theta_j}$, $\boldsymbol{\tau_{jk}}$ and π_{jk}.

For typical extended bactericidal activity trials lasting 56 days, in order to avoid numerical overflow, the regression models should be fitted with the times t_{ijk} expressed in weeks rather than days.

4. MODELS FOR COLONY FORMING UNIT COUNT ON ORIGINAL SCALE

The Bayesian NLME regression model of Burger and Schall[14,15] can be implemented as Poisson, ZIP, NEGBIN, ZINB and lognormal NLME regression models by setting the quantities in Equation (3) to those presented in Appendix Table 1 (see Appendix A). These models can thus be specified as special cases of the zero inflated generalized NLME regression model in Section 3. The likelihood functions of these distributions and the corresponding expected values and variances are presented in Table 1. The same quantities for the ZIGP regression model are presented in Section 1 of the online appendix of this paper.

The models in Table 1 include the following coefficients for patient i in treatment group j (see Section 3 of Burger and Schall[14]): Intercepts (α_{ij}); two slopes characterizing the rate of change over time (β_{1ij} and β_{2ij}). Furthermore, each treatment is characterized by the node (or inflection point) at which transition from one slope to another occurs (κ_j); and by a “smoothness" parameter governing the “speed" of transition (γ_j).

As a special case of Equation (4), the terms α_{ij}, β_{1ij} and β_{2ij} are the sums of fixed effects and associated random coefficients, namely:

$$\boldsymbol{\mu_{ij}} = \begin{bmatrix} \alpha_{ij} \\ \beta_{1ij} \\ \beta_{2ij} \end{bmatrix} = \begin{bmatrix} \alpha_j \\ \beta_{1j} \\ \beta_{2j} \end{bmatrix} + \begin{bmatrix} u_{0ij} \\ u_{1ij} \\ u_{2ij} \end{bmatrix} = \boldsymbol{\mu_j} + \begin{bmatrix} u_{0ij} \\ u_{1ij} \\ u_{2ij} \end{bmatrix} \tag{8}$$

where $\boldsymbol{\phi_{ij}} = \boldsymbol{\mu_{ij}} = (\alpha_{ij}, \beta_{1ij}, \beta_{2ij})'$ (or $\boldsymbol{\varphi_j} = [u_{0ij}, u_{1ij}, u_{2ij}]'$) and $\boldsymbol{\phi_j} = \boldsymbol{\mu_j} = (\alpha_j, \beta_{1j}, \beta_{2j})'$ are respectively the vectors of random and mean intercepts and slopes. Detail on the specification of the random effects and prior distributions is provided in Appendix B.

The mean bactericidal activity of treatment group j is expressed as the daily rate of change in

mean $\log_{10}(\text{CFU})$ count over timepoints k_1 and k_2, namely:

$$\text{BA}_j\left(t_{k_1}\text{–}t_{k_2}\right) = -\frac{\log\left[M_j\left(t_{k_2}\right)\right] - \log\left[M_j\left(t_{k_1}\right)\right]}{\log\left(10\right)\left(t_{k_2} - t_{k_1}\right)} \tag{9}$$

where $M_j\left(t_k\right) = \delta_{jk} = \exp\left(\alpha_j - \beta_{1j}t_k - \beta_{2j}\gamma_j \log\left[\frac{e^{\frac{t_k-\kappa_j}{\gamma_j}} + e^{-\frac{t_k-\kappa_j}{\gamma_j}}}{e^{\frac{\kappa_j}{\gamma_j}} + e^{-\frac{\kappa_j}{\gamma_j}}}\right]\right)$ for the Poisson, NEGBIN, and lognormal regression models, and $M_j\left(t_k\right) = \left(1 - \pi_{jk}\right)\delta_{jk}$ for the ZIP, ZINB and ZIGP regression models.

It should be noted that the lognormal regression model is equivalent to the regression model of Burger and Schall[14] when alternatively regressed to $z_{ijk} = \log\left(y_{ijk}\right) + o_{ijk}$ (see Table 1). Hence the z_{ijk} are the logarithm of the CFU counts as per Equation (2) and follow normal distributions, namely:

$$z_{ijk} \sim \text{Normal}\left(\log\left[\delta_{jk}\right], \sigma_{jk}^2\right) \tag{10}$$

The regression models were fitted using OpenBUGS.[44] Posterior samples were monitored and convergence was confirmed using iteration and autocorrelation plots, and Brooks-Gelman-Rubin statistics of parallel chains.[14,45,46]

5. MODEL DISCRIMINATION AND GOODNESS OF FIT

In order to compare the models investigated here, model discrimination statistics, namely DIC statistics[47] conditional on the random coefficients of the regression models and Bayes factors (or CLMMLs),[24] were calculated.

5.1. Deviance information criterion statistic

The DIC is defined under Model M as follows (see Table 1 and Appendix Table 1):

$$\text{DIC}(M) = 2\overline{\text{D}(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk})} - \text{D}(\hat{\boldsymbol{\phi}}_{\boldsymbol{ij}}, \hat{\boldsymbol{\theta}}_{\boldsymbol{j}}, \hat{\boldsymbol{\tau}}_{\boldsymbol{jk}}, \hat{\pi}_{jk}) \tag{11}$$

where $\text{D}(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk}) = -2\log\left(\prod_{j=1}^{J}\prod_{i=1}^{N_j}\prod_{k=1}^{K_{ij}} L\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk}|y_{ijk}\right)\right)$ is the deviance measure. Here, $\hat{\boldsymbol{\phi}}_{\boldsymbol{ij}}$, $\hat{\boldsymbol{\theta}}_{\boldsymbol{j}}$, $\hat{\boldsymbol{\tau}}_{\boldsymbol{jk}}$ and $\hat{\pi}_{jk}$ are respectively the mean of the posterior distribution of $\boldsymbol{\phi_{ij}}$, $\boldsymbol{\theta_j}$, $\boldsymbol{\tau_{jk}}$ and π_{jk}, and $\overline{\text{D}(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk})}$ is the mean of the posterior distribution of $\text{D}(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk})$.

Models with small DIC are favored. The DICs were calculated outside OpenBUGS using libraries available in the R project.

5.2. Bayes factors

When comparing two of any of the count models defined in Section 4, say Model M_0 and Model M_1, based on the posterior probability of each of the models conditional on the data, the Bayes factor in favor of M_0 is defined as[48]:

$$B_{01} = \frac{f\left(\boldsymbol{y}|M_0\right)}{f\left(\boldsymbol{y}|M_1\right)} \tag{12}$$

where $f(\boldsymbol{y}|M_0)$ and $f(\boldsymbol{y}|M_1)$ are the marginal likelihoods of $\boldsymbol{y}$ under Model M_0 and Model M_1, respectively. Equivalently:

$$\log(B_{01}) = \log(f\left[\boldsymbol{y}|M_0\right]) - \log(f\left[\boldsymbol{y}|M_1\right]) \tag{13}$$

The model with the larger log-marginal likelihood is favored. Detail on the Laplace-Metropolis approximation of Bayes factors is provided in Appendix C.

5.3. Pearson goodness of fit measure

Model checks were performed by evaluating the predictive performance of the candidate models using the following Pearson goodness of fit measure[49,50] (see Table 1 and Appendix Table 1):

$$\chi^2\left(y_{ijk}, \boldsymbol{\phi_{ij}^{(z)}}, \boldsymbol{\theta_j^{(z)}}, \boldsymbol{\tau_{jk}^{(z)}}, \pi_{jk}^{(z)}\right) = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\sum_{k=1}^{K_{ij}}\left(\frac{\left[y_{ijk} - E\left(y_{ijk}|\boldsymbol{\phi_{ij}^{(z)}}, \boldsymbol{\theta_j^{(z)}}, \boldsymbol{\tau_{jk}^{(z)}}, \pi_{jk}^{(z)}\right)\right]^2}{\text{Var}\left(y_{ijk}|\boldsymbol{\phi_{ij}^{(z)}}, \boldsymbol{\theta_j^{(z)}}, \boldsymbol{\tau_{jk}^{(z)}}, \pi_{jk}^{(z)}\right)}\right)$$

where $\boldsymbol{\phi_{ij}^{(z)}}$, $\boldsymbol{\theta_j^{(z)}}$, $\boldsymbol{\tau_{jk}^{(z)}}$ and $\pi_{jk}^{(z)}$ are respectively the posterior samples from $\boldsymbol{\phi_{ij}}$, $\boldsymbol{\theta_j}$, $\boldsymbol{\tau_{jk}}$ and π_{jk} at iteration z ($z = 1, \ldots, Z$). In addition, suppose $y_{ijk}^{*,(z)}$ is a random copy drawn from the joint posterior distribution of $y_{ijk}|\boldsymbol{\phi_{ij}^{(z)}}, \boldsymbol{\theta_j^{(z)}}, \boldsymbol{\tau_{jk}^{(z)}}, \pi_{jk}^{(z)}$ at iteration z. The corresponding posterior p-value under Model M is then defined as:

$$\text{p}\,(M) = \frac{1}{Z}\sum_{z=1}^{Z}\left(I\left[\chi^2\left(y_{ijk}^{*,(z)}, \boldsymbol{\phi_{ij}^{(z)}}, \boldsymbol{\theta_j^{(z)}}, \boldsymbol{\tau_{jk}^{(z)}}, \pi_{jk}^{(z)}\right) > \chi^2\left(y_{ijk}, \boldsymbol{\phi_{ij}^{(z)}}, \boldsymbol{\theta_j^{(z)}}, \boldsymbol{\tau_{jk}^{(z)}}, \pi_{jk}^{(z)}\right)\right]\right) \quad (14)$$

Values of p (M) close to 0.5 would suggest that the candidate model fits the data adequately, whereas values close to 0 or 1 indicate poor model fit.[45]

6. APPLICATION

We apply the models proposed in this paper in a reanalysis of the CFU count data of Dawson et al.[37]. In this trial, drug-sensitive TB patients were randomized to receive 8-week combination therapy of either moxifloxacin, PA-824 (100 mg) and pyrazinamide (M-PA100-Z; 60 pa-

tients); or moxifloxacin, PA-824 (200 mg) and pyrazinamide (M-PA200-Z; 62 patients); or Rifafour (59 patients; control treatment). Two 16-hour overnight sputum samples were collected pre-treatment and were used for the calculation of CFU count at Day 0. Thereafter, overnight sputum samples were collected on Days 3, 7, 14, 21, 28, 35, 42, 49 and 56.

Summary statistics of the observed CFU count over time are presented in Appendix Table 2 (see Appendix D) by treatment group. The preliminary investigation of the data suggests that the CFU counts are overdispersed. Furthermore, it is observed that the percentage of zero counts significantly increases over time. These findings therefore motivate the fit of zero inflated regression models to the data on the original scale as an alternative to fitting regression models on logarithmic scale.

The lower and upper bounds of κ_j and γ_j were set to $L_\kappa = 3$, $U_\kappa = 11$, $L_\gamma = 0.05$ and $U_\gamma = 2$. In order to avoid model overparameterization and overfit, the time-dependent model parameters (π_{jk}, ρ_{jk}, ω_{jk} and σ^2_{jk}) were pooled by (i) Days 0, 3 and 7, (ii) Days 14 and 21, (iii) Days 28 and 35, and (iv) Days 42, 49, 56.

The SAS®, OpenBUGS and R code for the implementation of the ZINB regression model, including the calculation of the DIC statistic, CLMML and Pearson goodness of fit measure, are included in Section 2 of the online appendix of this paper. In the code, OpenBUGS and R are called remotely from SAS®, and accordingly, posterior samples are exported back to SAS® for further computation.

For the ZINB regression model, 80000 samples were simulated from the joint posterior distribution for two parallel chains. Among those 80000 samples (per chain), the initial 30000 samples were discarded (burn-in). The thinning factor was set to 50 to reduce autocorrelation among the samples.

Regression fits were checked using Pearson goodness of fit measures, and compared us-

ing DIC statistics and CLMMLs, for cases where CFU counts of zero were treated either as (i) left censored observations, and (ii) observed data. The posterior p-values (of the Pearson goodness of fit measure) and model comparison statistics for the various Bayesian NLME regression models are provided in Table 2. The posterior p-values suggest that the NEGBIN and ZINB regression models fit the data very well, whereas the Poisson and ZIP regression models fit the data very poorly (due to overdispersion in the data). Both model comparison statistics, namely DICs and CLMMLs:

- Significantly favor the regression models with negative binomial distributions (NEGBIN and ZINB) over the Poisson distributions (Poisson and ZIP).
- Favor the ZINB regression model over the ZIGP regression model.
- Favor the NEGBIN regression model over the lognormal regression model.
- Favor the zero inflated regression models (ZIP and ZINB) over their non-zero inflation counterparts (Poisson and NEGBIN).

In order to investigate the effect of treating zero counts as left censored values, inferential statistics of the lognormal regression model (which fits zero counts as left censored observations) were compared to those of the NEGBIN, ZINB and ZIGP regression models (which fit zero counts as observed data). The Poisson and ZIP regression models were not further investigated due to poor model performance.

Figure 2 presents the posterior estimates and 95% Bayesian credibility intervals (BCIs) of mean $\text{BA}_j\,(0\text{–}56)$ for the NEGBIN, ZINB, ZIGP and lognormal regression models by treatment group: The posterior estimates for the NEGBIN, ZINB and ZIGP regression models are smaller than those of the lognormal regression model, whereas the 95% BCIs for the NEGBIN, ZINB and ZIGP regression models are generally somewhat wider than those of the lognormal regression model. Similarly, Table 3 presents the posterior estimates and 95% BCIs

of mean BA$_j$ $(0$–$56)$, and the corresponding differences versus Rifafour (control), for the ZINB and lognormal regression models. The difference between M-PA200-Z versus Rifafour in mean BA$_j$ $(0$–$56)$, based on both the ZINB and lognormal regression model, is statistically significant.

Table 4 presents the posterior estimates and 95% BCIs of π_{jk} for the ZINB regression model by treatment group. The analysis suggests that, as one would expect, the excess amount of zeros (in CFU count) increases over the course of treatment.

In a sensitivity analysis we investigated the influence of the prior specification of ρ_{jk} and π_{jk} on the inference about the bactericidal activity of treatments. The ZINB regression model specified noninformative gamma and beta distributions for ρ_{jk} and π_{jk} (see Appendix B). Alternatively, the ρ_{jk} and π_{jk} were assigned weakly informative prior distributions such as truncated Cauchy (or $t(1)$) and uniform distributions, namely $\rho_{jk} \sim t(1)I\,(0.01, \infty)$ and $\pi_{ij} \sim$ Uniform$(0, 1)$. The inferences about the bactericidal activity of the treatment groups are very similar under both prior specifications of ρ_{jk} and π_{jk} (data not shown). We therefore conclude that the bactericidal activity is not sensitive to the two types of prior specifications of ρ_{jk} and π_{jk}.

The ZINB regression model is the most general regression model among the exponential family of count models proposed in Section 4, and has the largest flexibility with respect to zero inflation and overdispersion. The findings of this reanalysis of the CFU count of Dawson et al.[37] point to the ZINB regression model as the most suitable count model for this data (compared to the Poisson, ZIP, NEGBIN and ZIGP regression models).

7. SIMULATION STUDY

We assessed the performance of the ZINB and lognormal regression models in a simulation study. Furthermore, we assessed the effect of fitting zero counts as left censored values in the lognormal regression model. Datasets were simulated from the ZINB regression model where

model parameters were chosen to mimic log(CFU) count versus time profiles of a moderately and highly efficacious anti-TB drugs, each with and without zero inflation in CFU count over time.

The slope parameters for the two treatments were chosen as $\beta_{11} = 0.461$, $\beta_{21} = -0.268$, $\beta_{12} = 0.443$ and $\beta_{22} = -0.203$, while the following parameter values were chosen for both treatments ($j = 1, 2$): $\alpha_j = 15$, $\kappa_j = 5$, $\rho_{j1} = 1.2$ (Days 0, 3 and 7), $\rho_{j2} = 1.3$ (Days 14 and 21), $\rho_{j3} = 1.4$ (Days 28 and 35), $\rho_{j4} = 1.5$ (Days 42, 49 and 56), $o_{ijk} = -5.0633$, and

$$\Omega_{\boldsymbol{\mu}j} = \begin{bmatrix} 1.000 & 0.001 & -0.005 \\ 0.001 & 0.020 & -0.005 \\ -0.005 & -0.005 & 0.015 \end{bmatrix}$$

Both parameter scenarios were investigated for the following two sets of zero inflation probabilities (hence, a total of four parameter scenarios): (i) $\pi_{j1} = \pi_{j2} = \pi_{j3} = \pi_{j4} = 0$ (without zero inflation), and (ii) $\pi_{j1} = 0$, $\pi_{j2} = 0.01$, $\pi_{j3} = 0.05$ and $\pi_{j4} = 0.1$ (with zero inflation). The accuracy and precision characteristics bias, relative bias, standard error (SE), and root mean square error (RMSE) of the BA$_j$ $(0$–$56)$ estimates for the four parameter scenarios were calculated, as was the empirical coverage probability of the associated 95% BCIs. The two candidate regression models (ZINB and lognormal) were fitted to 1000 simulated datasets, each dataset consisting of 35 profiles per treatment.

From Appendix Table 3 (see Appendix E) we observe that the bias of estimates of BA$_j$ $(0$–$56)$ from the ZINB regression model is noticeably smaller than that of estimates from the lognormal regression model. The SE and RMSE suggest that the ZINB regression model performs substantially better than the lognormal regression model. For larger values of BA$_j$ $(0$–$56)$ and π_{jk}, the coverage probability of the 95% BCI of the lognormal regression model is considerably smaller than the nominal value (therefore suggesting that the credibility intervals are anti-conservative). Overall, the ZINB regression model yields credibility interval coverage prob-

abilities that are quite close to the nominal value.

8. DISCUSSION

This paper investigates alternatives to published nonlinear regression models for CFU count versus time data which exhibit zero inflation. Our results suggest that practitioners should carefully consider whether transformation of count data (for example, for the purpose of variance stabilization, and adherence to the normality assumption) and subsequent "normal theory" analysis, rather than fitting models to count data on the original scale, are appropriate. In particular, this paper proposes a general approach for the Bayesian implementation of zero inflated generalized nonlinear mixed effects regression models for longitudinal count data.

In Bayesian statistics the suitability of candidate regression models is often judged by calculating their posterior marginal likelihoods (hence, Bayes factors). In the application of mixed effects regression modeling, marginal likelihoods are usually approximated by computing the harmonic mean of the likelihood with respect to the posterior distribution, based on samples drawn from the posterior distribution. Although the harmonic mean may serve as an unbiased estimator for the marginal likelihood, it may have infinite variance. Alternative methods for approximating the marginal likelihood should therefore be employed. Here, we propose the so-called Laplace approximation of the marginal distribution. For hierarchical models such as those considered here, the Laplace method involves marginalization (or integration) at two different levels: At the second level, the random effects need to be integrated out (therefore referred to as CLMMLs). The Laplace method seems to be greatly underutilized due to its complexity (because of the required second layer of integration). We propose the use of the "R2Cuba" library of the R project to reduce the computational burden of integrating out (multi-dimensional) random coefficients that describe the nonlinear relationship between longitudinal count data and time.

The model comparison statistics (DICs and CLMMLs) and Pearson goodness of fit measures obtained from a reanalysis of the dataset of Dawson et al.[37] suggest that the NEGBIN regression model is more suitable than the lognormal (or conventional normal) regression model for CFU count data with excess zeros. Since for effective anti-TB drugs zero inflation in CFU count greatly increases towards the end of CFU count versus time profile, ZINB regression models should be fitted to such data, rather than log-normal regression models that treat zero counts as "left censored" data. Furthermore, our simulation study shows that the ZINB regression model has good properties in terms of accuracy, precision and credibility interval coverage, and that the lognormal regression model yields biased estimates and credibility intervals that may severely undercover the true bactericidal activity of anti-TB drugs with zero inflation in CFU count. We therefore recommend that zero inflated NLME regression models should be fitted to CFU count on the original scale, as an alternative to conventional normal (or lognormal) NLME regression models on the logarithmic scale.

We also extended the proposed NLME regression model by assuming ZIGP distributions for CFU count. The methods introduced in this paper can also be extended to accommodate outliers in CFU count (e.g. considering heavy tailed distributions such as the generalized Sichel distribution[51]).

The application of the proposed zero inflated generalized NLME regression models is not limited to CFU count data: Thus, the methods introduced here (in particular, the programming code supporting this paper as part of the supplementary material) can be used by practitioners for other applications to hierarchically structured zero inflated longitudinal count data.

The application of the suggested NLME regression models within the framework of extended bactericidal activity anti-TB trials may allow accurate decision-making for advancing anti-TB drugs into Phase 3 trials, and motivates further extensions of theoretical models for longitudinal count data in clinical trials. Overall, we recommend that researchers should continuously

illustrate developments in mixed effects regression models via their added value in clinical research.

Data Availability Statement

The programming code supporting this paper has been included as part of the supplementary material.

ACKNOWLEDGEMENTS

We thank Emmanuel Lesaffre, Johannes T Ferreira and Janet van Niekerk for comments that improved the quality of the manuscript.

REFERENCES

[1] McCullagh P, Nelder JA. *Generalized Linear Models*. 2^{nd} ed.; Washington, DC: Chapman & Hall/CRC; 1989.

[2] Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. *J. Am. Stat. Assoc*. 1993; 88(421): 9–25. DOI: 10.2307/2290687.

[3] Kassahun W, Neyens T, Faes C, Molenberghs G, Verbeke G. A zero-inflated overdispersed hierarchical Poisson model. *Stat. Model*. 2014; 14(5): 439–456. DOI: 10.1177/1471082X14524676.

[4] Bandyopadhyay D, Lachos VH, Castro LM, Dey DK. Skew-normal/independent linear mixed models for censored responses with applications to HIV viral loads. *Biom. J*. 2012; 54(3): 405–425. DOI: 10.1002/bimj.201000173.

[5] Chen R, Huang Y. Mixed-effects models with skewed distributions for time-varying decay rate in HIV dynamics. *Commun. Stat. Simul. Comput*. 2016; 45(2): 737–757. DOI: 10.1080/03610918.2013.873129.

[6] Diacon AH, Dawson R, Von Groote-Bidlingmaier F, Symons G, Venter A, Donald PR, Van Niekerk C, Everitt D, Winter H, Becker P, *et al.* 14-day bactericidal activity of PA-824, bedaquiline, pyrazinamide, and moxifloxacin combinations: A randomized trial. *The Lancet* 2012; 380(9846): 986–993. DOI: 10.1016/S0140–6736(12)61080–0.

[7] Fujikawa H, Sakha MZ. Prediction of competitive microbial growth in mixed culture at dynamic temperature patterns. *Biocontrol Sci*. 2014; 19(3): 121–127. DOI: 10.4265/bio.19.121.

[8] St-Pierre AP, Shikon V, Schneider DC. Count data in biology – Data transformation or model reformation? *Ecol. Evol*. 2018; 8(6): 3077–3085. DOI: 10.1002/ece3.3807.

[9] Gebeyehu S, Samways MJ. Grasshopper assemblage response to a restored national park (Mountain Zebra National Park, South Africa). *Biodivers. & Conserv*. 2002; 11(2): 283–304. DOI: 10.1023/A:1014514007744.

[10] O'Hara RB, Kotze DJ. Do not log-transform count data. *Methods Ecol. Evol*. 2010; 1: 118–122. DOI: 10.1111/j.2041–210X.2010.00021.x.

[11] Lu X, Huang Y, Chen J, Zhou R, Yu S, Yin P. Bayesian joint analysis of heterogeneous- and skewed-longitudinal data and a binary outcome, with application to AIDS clinical studies. *Stat. Methods Med. Res*. 2017; 0(0): 0–0. DOI: 10.1177/0962280217689852.

[12] Engel B, Keen A. A simple approach for the analysis of generalized linear mixed models. *Stat. Neerlandica* 1994; 48: 1–22. DOI: 10.1111/j.1467–9574.1994.tb01428.x.

[13] Vock DM, Davidian M, Tsiatis AA, Muir AJ. Mixed model analysis of censored longitudi-

nal data with flexible random-effects density. *Biostat*. 2012; 13: 61–73. DOI: 10.1093/biostatistics/kxr026.

[14] Burger DA, Schall R. A Bayesian nonlinear mixed-effects regression model for the characterization of early bactericidal activity of tuberculosis drugs. *J. Biopharm. Stat*. 2015; 25(6): 1247–1271. DOI: 10.1080/10543406.2014.971170.

[15] Burger DA, Schall R. Robust fit of Bayesian mixed effects regression models with application to colony forming unit count in tuberculosis research. *Stat. Medicine* 2018; 37(4): 544–556. DOI: 10.1002/sim.7529.

[16] Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 1992; 34: 1–14. DOI: 10.2307/1269547.

[17] Ridout M, Hinde J, Demétrio CG. A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biom*. 2001; 57(1): 219–223. DOI: 10.1111/j.0006–341X.2001.00219.x.

[18] Zhu H, Luo S, DeSantis SM. Zero-inflated count models for longitudinal measurements with heterogeneous random effects. *Stat. Methods Med. Res*. 2017; 26(4): 1774–1786. DOI: 10.1177/0962280215588224.

[19] Neelon BH, Malley AJ, Normand SLT. A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use. *Stat. Model*. 2010; 10(4): 421–439. DOI: 10.1177/1471082X0901000404.

[20] Hall DB. Zero-inflated Poisson and binomial regression with random effects: a case study. *Biom*. 2000; 56(4): 1030–1039. DOI: 10.2307/2677034.

[21] Dagne GA. Bayesian semiparametric zero-inflated Poisson model for longitudinal count data. *Math. Biosci*. 2010; 224(2): 126–130. DOI: 10.1016/j.mbs.2010.01.004.

[22] Sae-Lim P, Grøva L, Olesen I, Varona L. A comparison of nonlinear mixed models and response to selection of tick-infestation on lambs. *PLoS ONE* 2017; 12(3): e0172711. DOI: 10.1371/journal.pone.0172711.

[23] Raftery AE. Hypothesis testing and model selection. In: Gilks WR, Richardson S, Spiegelhalter DJ, eds. Markov Chain Monte Carlo in Practice. London, UK: Chapman and Hall; 1996:163–188.

[24] Lewis SM, Raftery AE. Estimating Bayes factors via posterior simulation with the Laplace-Metropolis estimator. *J. Am. Stat. Assoc*. 1997; 92(438): 648–655. DOI: 10.2307/2965712.

[25] Newton MA, Raftery AE. Approximate Bayesian inference with the weighted likelihood bootstrap. *J. Royal Stat. Soc. Ser. B (Statistical Methodol*. 1994; 56: 3–48. URL: http://www.jstor.org/stable/2346025.

[26] Wolpert RL, Schmidler SC. α-Stable limit laws for harmonic mean estimators of marginal likelihoods. *Stat. Sinica* 2012; 22(3): 1233–1251. DOI: 10.5705/ss.2010.221.

[27] Zhu H, DeSantis SM, Luo S. Joint modeling of longitudinal zero-inflated count and time-to-event data: A Bayesian perspective. *Stat. Methods Med. Res*. 2018; 27(4): 1258–1270. DOI: 10.1177/0962280216659312.

[28] O'Brien RJ. Scientific blueprint for tuberculosis drug development. *Tuberc*. 2001; 81: 19–45.

[29] Jindani A, Doré CJ, Mitchison DA. Bactericidal and sterilizing activities of antituberculosis drugs during the first 14 days. *Am. J. Respir. Critical Care Medicine* 2003; 167: 1348–1354. DOI: 10.1164/rccm.200210–1125OC.

[30] Jindani A, Aber VR, Edwards EA, Mitchison DA. The early bactericidal activity of drugs

in patients with pulmonary tuberculosis. *Am. Rev. Respir. Dis*. 1980; 121: 939–949. DOI: 10.1164/arrd.1980.121.6.939.

[31] Mitchison DA, Strum AW. The measurement of early bactericidal activity. *Baillieres Clin. Infect. Dis*. 1997; 4(2): 185–206.

[32] Brindle R, Odhiambo JA, Mitchison DA. Serial counts for Mycobacterium tuberculosis in sputum as surrogate markers for the sterilising activity of rifampicin and pyrazinamide in treating pulmonary tuberculosis. *BMC Pulm. Medicine* 2001; 1: 2. DOI: 10.1186/1471–2466–1–2.

[33] Rustomjee R, Lienhardt C, Kanyok T, Davies GR, Levin J, Mthiyane T, Reddy C, Sturm AW, Sirgel FA, Allen J, *et al.* A Phase II study of the sterilising activities of ofloxacin, gatifloxacin and moxifloxacin in pulmonary tuberculosis. *The Int. J. Tuberc. Lung Dis*. 2008; 12(2): 128–138.

[34] Weiner M, Prihoda TJ, Burman W, Johnson JL, Goldberg S, Padayatchi N, Duran P, Engle M, Muzanye G, Mugerwa RD, *et al.* Evaluation of time to detection of Mycobacterium tuberculosis in broth culture as a determinant for end points in treatment trials. *J. Clin. Microbiol*. 2010; 48(12): 4370–4376. DOI: 10.1128/JCM.00757–10.

[35] Mitchison DA, Davies GR. Assessment of the efficacy of new anti-tuberculosis drugs. *The Open Infect. Dis. J*. 2008; 2: 59–76. DOI: 10.2174/1874279300802010059.

[36] Svensson EM, Karlsson MO. Modelling of mycobacterial load reveals bedaquiline's exposure-response relationship in patients with drug-resistant TB. *J. Antimicrob. Chemother*. 2017; 72(12): 3398–3405. DOI: 10.1093/jac/dkx317.

[37] Dawson R, Diacon AH, Everitt D, Van Niekerk C, Donald PR, Burger DA, Schall R, Spigelman M, Conradie A, Eisenach K, *et al.* Efficiency and safety of the combination of moxifloxacin, pretomanid (PA-824), and pyrazinamide during the first 8 weeks of antituberculosis treatment: A phase 2b, open-label, partly randomised trial in patients with

drug-susceptible or drug-resistant pulmonary tuberculosis. *The Lancet* 2015; 385(9979): 1738–1747. DOI: 10.1016/S0140–6736(14)62002–X.

[38] Brown H, Prescott R. *Applied Mixed Models in Medicine*. 2^{nd} ed.; Hoboken, New Jersey: John Wiley & Sons, Inc.; 2006.

[39] Davies GR, Brindle R, Khoo SH, Aarons LJ. Use of nonlinear mixed-effects analysis for improved precision of early pharmacodynamic measures in tuberculosis treatment. *Antimicrob. Agents Chemother*. 2006a; 50: 3154–3156. DOI: 10.1128/AAC.00774–05.

[40] Burger DA, Schall R, Chen DG. Robust Bayesian nonlinear mixed-effects modeling of time to positivity in tuberculosis trials. *Pharm. Stat*. 2018; 17(5): 615–628. DOI: 10.1002/pst.1877.

[41] Diacon AH, Dawson R, Von Groote-Bidlingmaier F, Symons G, Venter A, Donald PR, Van Niekerk C, Everitt D, Hutchings J, Burger DA, *et al.* Bactericidal activity of pyrazinamide and clofazimine alone and in combinations with pretomanid and bedaquiline. *Am. J. Respir. Critical Care Medicine* 2015; 191(8): 943–953. DOI: 10.1164/rccm.201410–1801OC.

[42] Phillips PPJ, Mendel CM, Burger DA, Crook AM, Nunn AJ, Dawson R, Diacon AH, Gillespie SH. Limited role of culture conversion for decision-making in individual patient care and for advancing novel regimens to confirmatory clinical trials. *BMC Medicine* 2016; 14: 19. DOI: 10.1186/s12916–016–0565–y.

[43] Yau KKW, Wang K, Lee AH. Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biom. J*. 2003; 45(4): 437–452. DOI: 10.1002/bimj.200390024.

[44] Lunn DJ, Spiegelhalter DJ, Thomas A, Best NG. The BUGS project: Evolution, critique and future directions. *Stat. Medicine* 2009; 28(25): 3049–3067. DOI: 10.1002/sim.3680.

[45] Ntzoufras I. *Bayesian Modeling Using WinBUGS*. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2009.

[46] Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat*. 1998; 7(4): 434–455. DOI: 10.1080/10618600.1998.10474787.

[47] Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A. Bayesian measures of model complexity and fit. *J. Royal Stat. Soc. Ser. B (Statistical Methodol*. 2002; 64(4): 583–639. DOI: 10.1111/1467–9868.00353.

[48] Kass RE, Raftery AE. Bayes factors. *J. Am. Stat. Assoc*. 1995; 90: 773–795. DOI: 10.1080/01621459.1995.10476572.

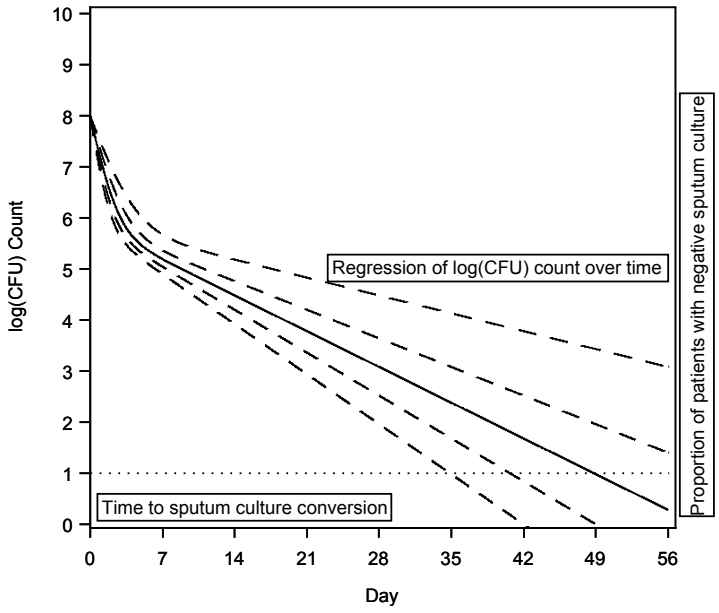
[49] Johnson VE. A Bayesian χ^2 test for goodness-of-fit. *The Annals Stat*. 2004; 32(6): 2361–2384. DOI: 10.1214/009053604000000616.

[50] Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis, Texts in Statistical Science*. 3^{rd} ed.; London: Chapman & Hall/CRC; 2013.

[51] Low YC, Ong SH, Gupta RC. Generalized Sichel distribution and associated inference. *J. Stat. Theory Appl*. 2017; 16: 322–336. DOI: 10.2991/jsta.2017.16.3.4.

[52] Hahn H, Bouvier A, Kiêu K. `R2Cuba`: Multidimensional numerical integration; 2013. R package Version 1.0-11; URL `http://CRAN.R-project.org/package=R2Cuba`.

Figure 1: Efficacy endpoints in extended bactericidal activity TB trials



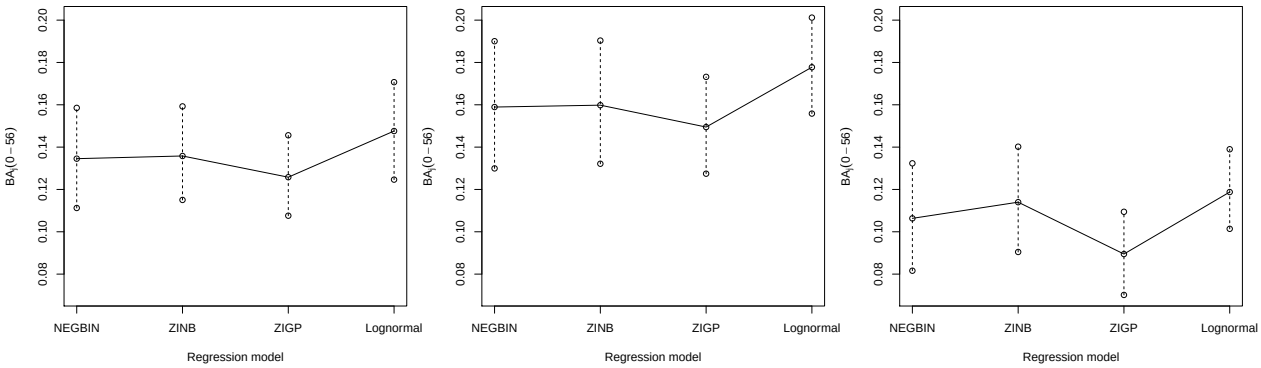
Dashed lines denote decline in $\log_{10}(\text{CFU})$ count in individual patients. Solid line denotes mean decline in $\log_{10}(\text{CFU})$ count of all patients. Dotted horizontal line denotes the censoring limit of $\log_{10}(\text{CFU})$ count.[35]

Table 1: Likelihood functions of zero inflated generalized NLME regression model for CFU count over time

y_{ijk}	$L\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, \boldsymbol{\tau_{jk}}, \pi_{jk} \middle\vert y_{ijk}\right)$	$E\left(y_{ijk}\right)$	Var $\left(y_{ijk}\right)$
Poisson $\left(\lambda_{ijk}\right)$	$\frac{\lambda_{ijk}^{y_{ijk}} e^{-\lambda_{ijk}}}{y_{ijk}!}$	λ_{ijk}	λ_{ijk}
ZIP $\left(\lambda_{ijk}, \pi_{jk}\right)$	$\pi_{jk} I\left(y_{ijk}=0\right)+\left(1-\pi_{jk}\right) \frac{\lambda_{ijk}^{y_{ijk}} e^{-\lambda_{ijk}}}{y_{ijk}!}$	$\left(1-\pi_{jk}\right) \lambda_{ijk}$	$\left(1-\pi_{jk}\right)\left(\lambda_{ijk}+\pi_{jk} \lambda_{ijk}^2\right)$
NEGBIN $\left(\lambda_{ijk}, \rho_{jk}\right)$	$\binom{y_{ijk}+\rho_{jk}-1}{y_{ijk}}\left(\frac{\rho_{jk}}{\lambda_{ijk}+\rho_{jk}}\right)^{\rho_{jk}}\left(\frac{\lambda_{ijk}}{\lambda_{ijk}+\rho_{jk}}\right)^{y_{ijk}}$	λ_{ijk}	$\frac{\lambda_{ijk}\left(\lambda_{ijk}+\rho_{jk}\right)}{\rho_{jk}}$
ZINB $\left(\lambda_{ijk}, \rho_{jk}, \pi_{jk}\right)$	$\pi_{jk} I\left(y_{ijk}=0\right)+\left(1-\pi_{jk}\right) \times$ $\binom{y_{ijk}+\rho_{jk}-1}{y_{ijk}}\left(\frac{\rho_{jk}}{\lambda_{ijk}+\rho_{jk}}\right)^{\rho_{jk}}\left(\frac{\lambda_{ijk}}{\lambda_{ijk}+\rho_{jk}}\right)^{y_{ijk}}$	$\left(1-\pi_{jk}\right) \lambda_{ijk}$	$\left(1-\pi_{jk}\right) \times$ $\left(\frac{\lambda_{ijk}\left(\lambda_{ijk}+\rho_{jk}\right)}{\rho_{jk}}+\pi_{jk} \lambda_{ijk}^2\right)$
LN $\left(\lambda_{ijk}, \sigma_{jk}^2\right)$	$\frac{1}{y_{ijk} \sigma_{jk} \sqrt{2 \pi}} e^{\left(-\frac{1}{2 \sigma_{jk}^2}\left[\log\left(y_{ijk}\right)-\log\left(\lambda_{ijk}\right)\right]^2\right)}$	$e^{\left(\log\left(\lambda_{ijk}\right)+\frac{\sigma_{jk}^2}{2}\right)}$	$e^{\left(\sigma_{jk}^2-1\right)} e^{\left(2 \log\left(\lambda_{ijk}\right)+\sigma_{jk}^2\right)}$

Expressions for $E\left(y_{ijk}\right)$ and Var $\left(y_{ijk}\right)$ are conditional on $\boldsymbol{\phi_{ij}}$, $\boldsymbol{\theta_j}$, $\boldsymbol{\tau_{jk}}$ and π_{jk}. CFU: Colony forming unit. NEGBIN: Negative binomial. NLME: Nonlinear mixed effects. ZINB: Zero inflated negative binomial. ZIP: Zero inflated Poisson. $\lambda_{ijk}=\exp\left(\alpha_{ij}-\beta_{1ij} t_{ijk}-\beta_{2ij} \gamma_j \log\left[\frac{e^{\frac{t_{ijk}-\kappa_j}{\gamma_j}}+e^{-\frac{t_{ijk}-\kappa_j}{\gamma_j}}}{e^{\frac{\kappa_j}{\gamma_j}}+e^{-\frac{\kappa_j}{\gamma_j}}}\right]-o_{ijk}\right)$. $o_{ijk}=\log\left(c_{ijk} 10^{d_{ijk}} / n_{ijk}\right)$. n_{ijk}, c_{ijk} & d_{ijk} are respectively n, "factor" and "dilution" as per Equation (2).

Figure 2: Posterior estimates and 95% BCIs of mean $\text{BA}_j\ (0\text{–}56)$ by regression model



(a) M-PA100-Z **(b)** M-PA200-Z **(c)** Rifafour

$\text{BA}_j\ (0\text{–}56)$: Daily rate of decline in $\log_{10}(\text{CFU})$ count from Day 0 to Day 56 of treatment group j. BCI: Bayesian credibility interval. CFU: Colony forming unit. NEGBIN, ZINB & ZIGP: CFU counts of zero treated as observed data. Lognormal: CFU counts of zero treated as left censored observations.

Table 2: Model discrimination statistics and Pearson goodness of fit measure by regression model

	Censored		Uncensored		
Regression model	DIC	CLMML	DIC	CLMML	p (M)
Poisson	343350.59$^{\{3\}}$	-52528.42$^{\{3\}}$	347293.00$^{\{5\}}$	-53071.68$^{\{5\}}$	0.0000
ZIP			297587.01$^{\{4\}}$	-45483.67$^{\{4\}}$	0.0000
NEGBIN	13601.56$^{\{1\}}$	-7099.78$^{\{1\}}$	13747.80$^{\{3\}}$	-7180.99$^{\{2\}}$	0.6260
ZINB			13500.67$^{\{1\}}$	-7124.05$^{\{1\}}$	0.5445
ZIGP			13515.15$^{\{2\}}$	-7303.18$^{\{3\}}$	NC
Lognormal	13923.44$^{\{2\}}$	-7227.44$^{\{2\}}$			

DIC: Deviance information criterion. CFU: Colony forming unit. CLMML: Compound Laplace-Metropolis marginal likelihood on the logarithmic scale. NC: Not calculated. p (M): Posterior p-value of Pearson goodness of fit measure. Censored: CFU counts of zero treated as left censored observations. Uncensored: CFU counts of zero treated as observed data. Superscripts indicate the ranking of model comparison statistics from most favored to least favored.

Table 3: Posterior estimates and 95% BCIs of mean $\text{BA}_j\ (0\text{–}56)$: ZINB & lognormal regression model

		ZINB		Lognormal	
			Difference vs.		Difference vs.
Treatment	Statistic	Mean	Rifafour	Mean	Rifafour
M-PA100-Z	PE	0.136	0.022	0.148	0.029
	95% BCI	[0.115; 0.159]	[-0.012; 0.056]	[0.125; 0.171]	[-0.001; 0.057]
M-PA200-Z	PE	0.160	0.046	0.178	0.059
	95% BCI	[0.132; 0.190]	[0.007; 0.083]	[0.156; 0.201]	[0.030; 0.089]
Rifafour	PE	0.114		0.119	
	95% BCI	[0.090; 0.140]		[0.101; 0.139]	

$\text{BA}_j\ (0\text{–}56)$: Daily rate of decline in $\log_{10}(\text{CFU})$ count from Day 0 to Day 56 of treatment group j. BCI: Bayesian credibility interval. CFU: Colony forming unit. PE: Posterior estimate. ZINB: CFU counts of zero treated as observed data. Lognormal: CFU counts of zero treated as left censored observations.

Table 4: Posterior estimates and 95% BCIs of π_{jk}: ZINB regression model

		Timepoint (days)			
Treatment	Statistic	0, 3, 7	14, 21	28, 35	42, 49, 56
M-PA100-Z	PE	0.045	0.035	0.239	0.507
	95% BCI	[0.021; 0.076]	[0.000; 0.160]	[0.000; 0.596]	[0.321; 0.682]
M-PA200-Z	PE	0.027	0.019	0.168	0.236
	95% BCI	[0.009; 0.053]	[0.000; 0.135]	[0.000; 0.488]	[0.000; 0.785]
Rifafour	PE	0.037	0.144	0.113	0.378
	95% BCI	[0.013; 0.068]	[0.060; 0.240]	[0.000; 0.302]	[0.018; 0.576]

BCI: Bayesian credibility interval. CFU: Colony forming unit. PE: Posterior estimate. ZINB: CFU counts of zero treated as observed data.

APPENDICES

Appendix A: Exponential family of distributions

Table 1: Exponential family of distributions of zero inflated generalized NLME regression model for CFU count over time

Distribution	π_{ij}	$\boldsymbol{\tau_{jk}}$	$\boldsymbol{\eta}\left(f\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, t_{ijk}, o_{ijk}\right), \boldsymbol{\tau_{jk}}\right)$	$h\left(y_{ijk}, \boldsymbol{\tau_{jk}}\right)$	$\boldsymbol{T}\left(y_{ijk}\right)$	$A\left(f\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, t_{ijk}, o_{ijk}\right), \boldsymbol{\tau_{jk}}\right)$
Poisson	0		$\log\left(\lambda_{ijk}\right)$	$\frac{1}{y_{ijk}!}$	y_{ijk}	λ_{ijk}
ZIP	π_{ij}		$\log\left(\lambda_{ijk}\right)$	$\frac{1}{y_{ijk}!}$	y_{ijk}	λ_{ijk}
NEGBIN	0	ρ_{jk}	$\log\left(\frac{\lambda_{ijk}}{\lambda_{ijk}+\rho_{jk}}\right)$	$\binom{y_{ijk}+\rho_{jk}-1}{y_{ijk}}$	y_{ijk}	$-\rho_{jk}\log\left(\frac{\rho_{jk}}{\lambda_{ijk}+\rho_{jk}}\right)$
ZINB	π_{ij}	ρ_{jk}	$\log\left(\frac{\lambda_{ijk}}{\lambda_{ijk}+\rho_{jk}}\right)$	$\binom{y_{ijk}+\rho_{jk}-1}{y_{ijk}}$	y_{ijk}	$-\rho_{jk}\log\left(\frac{\rho_{jk}}{\lambda_{ijk}+\rho_{jk}}\right)$
Lognormal	0	σ^2_{jk}	$\begin{bmatrix}\frac{\log\left(\lambda_{ijk}\right)}{\sigma^2_{jk}} \\ -\frac{1}{2\sigma^2_{jk}}\end{bmatrix}$	$\frac{1}{\sqrt{2\pi}y_{ijk}}$	$\begin{bmatrix}\log\left(y_{ijk}\right) \\ \left[\log\left(y_{ijk}\right)\right]^2\end{bmatrix}$	$\frac{\log\left(\lambda_{ijk}\right)^2}{2\sigma^2_{jk}}+\log\left(\sigma_{jk}\right)$

CFU: Colony forming unit. NEGBIN: Negative binomial. NLME: Nonlinear mixed effects. ZINB: Zero inflated negative binomial. ZIP: Zero inflated Poisson. $\lambda_{ijk} = f\left(\boldsymbol{\phi_{ij}}, \boldsymbol{\theta_j}, t_{ijk}, o_{ijk}\right) = \exp\left(\alpha_{ij} - \beta_{1ij}t_{ijk} - \beta_{2ij}\gamma_j \log\left[\frac{e^{\frac{t_{ijk}-\kappa_j}{\gamma_j}}+e^{-\frac{t_{ijk}-\kappa_j}{\gamma_j}}}{e^{\frac{\kappa_j}{\gamma_j}}+e^{-\frac{\kappa_j}{\gamma_j}}}\right] - o_{ijk}\right)$. $\boldsymbol{\phi_{ij}} = \left(\alpha_{ij}, \beta_{1ij}, \beta_{2ij}\right)'$. $\boldsymbol{\theta_j} = \left(\kappa_j, \gamma_j\right)'$. $o_{ijk} = \log\left(c_{ijk}10^{d_{ijk}}/n_{ijk}\right)$. n_{ijk}, c_{ijk} & d_{ijk} are respectively n, "factor" and "dilution" as per Equation (2).

Appendix B: Specification of random effects and prior distributions

In Equation (5), the random coefficients $\boldsymbol{\phi_{ij}} = \boldsymbol{\mu_{ij}}$ are assumed to follow tri-variate normal distributions as follows:

$$\boldsymbol{\mu_{ij}} \sim \text{Normal}\left(\boldsymbol{\mu_j}, \Omega_{\boldsymbol{\mu}j}\right) \tag{15}$$

where $\Psi_j = \Omega_{\boldsymbol{\mu}j}$ are the covariance matrices of $\boldsymbol{\phi_{ij}} = \boldsymbol{\mu_{ij}}$.

Multivariate normal and Wishart prior distributions are specified, respectively, for $\boldsymbol{\mu_j}$ and $\Omega_{\boldsymbol{\mu}j}^{-1}$, namely:

$$\boldsymbol{\mu_j} \sim \text{Normal}(\mathbf{0}, 10^4 \times I_3) \tag{16}$$

$$\Omega_{\boldsymbol{\mu}j}^{-1} \sim \text{Wishart}(3, 3 \times R_j) \tag{17}$$

where $\mathbf{0} = (0, 0, 0)'$ and I_3 denotes the 3×3 identity matrix. R_j represent 3×3 inverse scale matrices.

The parameters κ_j and γ_j are assumed to follow uniform prior distributions, namely:

$$\kappa_j \sim \text{Uniform}(L_\kappa, U_\kappa) \tag{18}$$

$$\gamma_j \sim \text{Uniform}(L_\gamma, U_\gamma) \tag{19}$$

where L_κ, U_κ, L_γ and U_γ are the pre-specified lower and upper bounds for parameters κ_j and γ_j, respectively.

For the choice of R_j, we fitted the model as a generalized linear mixed effects regression model under the assumption that the node and smoothness parameters (κ_j and γ_j) are fixed at $(U_\kappa + L_\kappa)/2$ and $(U_\gamma + L_\gamma)/2$, respectively. We calculated the “frequentist" estimates for $\Omega_{\boldsymbol{\mu}j}$ via maximum likelihood estimation using the SAS® procedure PROC GLIMMIX, to serve

as R_j.

The dispersion parameters ρ_{jk}, scale parameters σ_{jk}^{-2}, and zero inflation probabilities π_{jk} are assumed to follow vague gamma and beta prior distributions, namely:

$$\rho_{jk} \sim \text{Gamma}\,(0.1, 0.1) \tag{20}$$

$$\sigma_{jk}^{-2} \sim \text{Gamma}\,(0.0001, 0.0001) \tag{21}$$

$$\pi_{jk} \sim \text{Beta}\,(0.1, 0.1) \tag{22}$$

The prior distributions for the ZIGP regression model are presented in Section 1 of the online appendix of this paper.

The ZIP distribution can alternatively be specified as a mixture of Poisson and Bernoulli distributions as follows:

$$y_{ijk} \sim \text{Poisson}\,(\lambda_{ijk}\,[1 - u_{ijk}]) \tag{23}$$

$$u_{ijk} \sim \text{Bernoulli}\,(\pi_{jk}) \tag{24}$$

Similarly, the ZINB distribution can alternatively be specified as a mixture of NEGBIN and Bernoulli distributions as follows:

$$y_{ijk} \sim \text{NEGBIN}\,(\lambda_{ijk}\,[1 - u_{ijk}]\,, \rho_{jk}) \tag{25}$$

$$u_{ijk} \sim \text{Bernoulli}\,(\pi_{jk}) \tag{26}$$

From the law of total probability, the distributions marginalized over u_{ijk} result in the ZIP and ZINB distributions.

Appendix C: Laplace-Metropolis approximation of Bayes factors

The Laplace-Metropolis approximation of $\log(f\,[\boldsymbol{y}|M])$ (that is, CLMML) under Model M can be written as[15,23,24] (see Table 1 and Appendix Table 1):

$$\log(f\,[\boldsymbol{y}|M]) = \frac{1}{2}\log(2\pi)\,pJ + \frac{1}{2}\log\left|R_{\left(\boldsymbol{\phi_j},\boldsymbol{\theta_j},\boldsymbol{\tau_{jk}},\pi_{jk},j=1,\ldots,J\right)}\right| + s_{\left(\boldsymbol{\phi_j},\boldsymbol{\theta_j},\boldsymbol{\tau_{jk}},\pi_{jk},j=1,\ldots,J\right)} + \quad (27)$$
$$\sum_{j=1}^{J}\sum_{i=1}^{N_j}\left(\log\left[P\left(\boldsymbol{y_{ij}}|\hat{\boldsymbol{\phi}}_{\boldsymbol{j}},\hat{\Psi}_j,\hat{\boldsymbol{\theta}}_{\boldsymbol{j}},\hat{\boldsymbol{\tau}}_{\boldsymbol{jk}},\hat{\pi}_{jk}\right)\right]\right) + \sum_{j=1}^{J}\left(\log\left[P\left(\hat{\boldsymbol{\phi}}_{\boldsymbol{j}},\hat{\Psi}_j,\hat{\boldsymbol{\theta}}_{\boldsymbol{j}},\hat{\boldsymbol{\tau}}_{\boldsymbol{jk}},\hat{\pi}_{jk}\right)\right]\right)$$

where p is the number of parameters among $\boldsymbol{\phi_j}$, Ψ_j, $\boldsymbol{\theta_j}$, $\boldsymbol{\tau_{jk}}$ & π_{jk} of treatment group j, and $P\left(\boldsymbol{y_{ij}}|\hat{\boldsymbol{\phi}}_{\boldsymbol{j}},\hat{\Psi}_j,\hat{\boldsymbol{\theta}}_{\boldsymbol{j}},\hat{\boldsymbol{\tau}}_{\boldsymbol{jk}},\hat{\pi}_{jk}\right) = \int P\left(\boldsymbol{y_{ij}}|\boldsymbol{\phi_{ij}},\hat{\boldsymbol{\phi}}_{\boldsymbol{j}},\hat{\Psi}_j,\hat{\boldsymbol{\theta}}_{\boldsymbol{j}},\hat{\boldsymbol{\tau}}_{\boldsymbol{jk}},\hat{\pi}_{jk}\right)\mathrm{d}\boldsymbol{\phi_{ij}}$ (see Equation (6)). Here, $\hat{\boldsymbol{\phi}}_{\boldsymbol{j}}$, $\hat{\Psi}_j$, $\hat{\boldsymbol{\theta}}_{\boldsymbol{j}}$, $\hat{\boldsymbol{\tau}}_{\boldsymbol{jk}}$ and $\hat{\pi}_{jk}$ are respectively the mean of the posterior distribution of $\boldsymbol{\phi_j}$, Ψ_j, $\boldsymbol{\theta_j}$, $\boldsymbol{\tau_{jk}}$ and π_{jk}. $\left|R_{\left(\boldsymbol{\phi_j},\boldsymbol{\theta_j},\boldsymbol{\tau_{jk}},\pi_{jk},j=1,\ldots,J\right)}\right|$ and $s_{\left(\boldsymbol{\phi_j},\boldsymbol{\theta_j},\boldsymbol{\tau_{jk}},\pi_{jk},j=1,\ldots,J\right)}$ respectively denote the determinant of the correlation matrix and the sum of the logarithm of the standard deviations of the posterior distributions of $\boldsymbol{\phi_j}$, $\boldsymbol{\theta_j}$, $\boldsymbol{\tau_{jk}}$ and π_{jk}.

The multidimensional integration library “R2Cuba” of the R project was used to approximate the Laplace integrals.[52]

Appendix D: Summary statistics

Table 2: Summary statistics of observed CFU count over time

Treatment	Day	n	Mean	SD	CV	Minimum	Median	Maximum	Zeros (%)
M-PA100-Z	Day 0	110	4022762	8592024	214	0	509500	55500000	1.8
	Day 3	54	5252617	36022027	686	0	62150	265000000	5.6
	Day 7	51	43318	91253	211	0	7800	570000	9.8
	Day 14	49	13708	26824	196	0	2350	121000	18.4
	Day 21	47	17594	71960	409	0	245	460000	34.0
	Day 28	42	1272	4744	373	0	0	28600	66.7
	Day 35	39	1890	7482	396	0	0	45000	74.4
	Day 42	36	468	2017	431	0	0	12000	72.2
	Day 49	34	190	840	441	0	0	4800	85.3
	Day 56	35	47	205	437	0	0	1200	85.7
M-PA200-Z	Day 0	107	4472336	9638695	216	0	380000	51333333	0.9
	Day 3	53	432749	1043900	241	0	48000	4700000	3.8
	Day 7	47	94440	211236	224	0	15900	1090000	6.4
	Day 14	47	32462	173048	533	0	1830	1190000	17.0
	Day 21	49	1718	5884	342	0	160	38000	34.7
	Day 28	40	3741	22128	592	0	0	140000	57.5
	Day 35	40	160	492	307	0	0	2200	77.5
	Day 42	37	19131	115050	601	0	0	700000	86.5
	Day 49	32	313	1768	565	0	0	10000	93.8
	Day 56	35	1830	10818	591	0	0	64000	94.3

n = Number of CFU counts. CFU: Colony forming unit. CV: Coefficient of variation. SD: Standard deviation.

Table 2: Summary statistics of observed CFU count over time

Treatment	Day	n	Mean	SD	CV	Minimum	Median	Maximum	Zeros (%)
Rifafour	Day 0	104	3035046	7754022	255	50	330000	57000000	0.0
	Day 3	51	391262	816412	209	0	46150	4000000	7.8
	Day 7	47	90493	169618	187	0	29100	1000000	10.6
	Day 14	44	98332	387679	394	0	6088	2500000	15.9
	Day 21	42	51665	246875	478	0	845	1600000	23.8
	Day 28	37	1969	3556	181	0	370	15990	35.1
	Day 35	38	23117	128034	554	0	0	790000	55.3
	Day 42	37	5976	34337	575	0	0	209000	70.3
	Day 49	37	4087	17002	416	0	0	97000	73.0
	Day 56	32	20	78	383	0	0	410	87.5

n = Number of CFU counts. CFU: Colony forming unit. CV: Coefficient of variation. SD: Standard deviation.

Appendix E: Simulation results

Table 3: Accuracy & precision of BA $(0\text{–}56)$ estimates, and BCI coverage: ZINB & lognormal regression model

	Timepoint (days): π_{jk}					Regression model	
BA$_j$ $(0\text{–}56)$	0, 3, 7	14, 21	28, 35	42, 49, 56	Statistic	ZINB	Lognormal
0.1046	0	0	0	0	Bias	-0.0039	0.0113
					Absolute bias	0.0092	0.0147
					SE	0.0111	0.0145
					RMSE	0.0118	0.0183
					95% BCI coverage	94.6	90.7
0.1054	0	0.01	0.05	0.1	Bias	-0.0042	0.0175
					Absolute bias	0.0097	0.0190
					SE	0.0114	0.0147
					RMSE	0.0121	0.0228
					95% BCI coverage	94.3	82.3
0.1200	0	0	0	0	Bias	-0.0033	0.0147
					Absolute bias	0.0099	0.0171
					SE	0.0119	0.0149
					RMSE	0.0123	0.0209
					95% BCI coverage	95.4	86.4

BA$_j$ $(0\text{–}56)$: Daily rate of decline in $\log_{10}(\text{CFU})$ count from Day 0 to Day 56 of treatment group j. BCI: Bayesian credibility interval. CFU: Colony forming unit. RMSE: Root mean square error. SE: Standard error. ZINB: CFU counts of zero treated as observed data. Lognormal: CFU counts of zero treated as left censored observations.

Table 3: Accuracy & precision of BA $(0–56)$ estimates, and BCI coverage: ZINB & lognormal regression model

	Timepoint (days): π_{jk}					Regression model	
BA$_j$ $(0–56)$	0, 3, 7	14, 21	28, 35	42, 49, 56	Statistic	ZINB	Lognormal
0.1208	0	0.01	0.05	0.1	Bias	-0.0045	0.0215
					Absolute bias	0.0099	0.0225
					SE	0.0114	0.0153
					RMSE	0.0123	0.0264
					95% BCI coverage	95.9	75.0

BA$_j$ $(0–56)$: Daily rate of decline in $\log_{10}(\text{CFU})$ count from Day 0 to Day 56 of treatment group j. BCI: Bayesian credibility interval. CFU: Colony forming unit. RMSE: Root mean square error. SE: Standard error. ZINB: CFU counts of zero treated as observed data. Lognormal: CFU counts of zero treated as left censored observations.