DOI: 10.1111/biom.128

Power and Sample Size Calculation Incorporating Misspecifications of the Variance Function in Comparative Clinical Trials with Over-Dispersed Count Data

Masataka Igeta[1,2,*] **Kunihiko Takahashi**[1,**] **and Shigeyuki Matsui**[1,***]

[1]Department of Biostatistics, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya, Aichi 466-8550, Japan
[2]Department of Biostatistics, Hyogo College of Medicine, 1-1 Mukogawa-cho, Nishinomiya, Hyogo, 663-8501, Japan
email: masataka.igeta@gmail.com
**email:* kunihiko@med.nagoya-u.ac.jp
***email:* smatsui@med.nagoya-u.ac.jp

Summary. Over-dispersed count data are frequently observed in clinical trials where the primary endpoint is occurrence of clinical events. Sample sizes of comparative clinical trials with these data are typically calculated under negative binomial models or quasi–Poisson models with specified variance functions, or under the assumption that the specified "working" variance functions are correctly specified. In this article, we propose a sample size formula anticipating misspecifications of the working variance function. We derived a method based on the asymptotic distribution of a Wald test statistic with a sandwich-type robust variance estimator under quasi–Poisson models. Under misspecifications of the working variance function, the asymptotic variance of the estimator of the treatment effect is expressed as a form involving both true and working variance functions. Our sample size formula includes several existing formulas as special cases when the working variance function is correctly specified as the true variance function. We also consider a sensitivity analysis for possible misspecifications of the "true" variance function when estimating sample sizes using our formula. A simulation study demonstrated the adequacy of our formulas in finite sample size settings. An application to a clinical trial to evaluate the treatment effect on prevention of COPD exacerbation is provided.

Key words: Misspecification of the variance function; Over-dispersed count data; Power and sample size; Sandwich-type robust variance estimator.

1. Introduction

Count data have been frequently obtained in clinical trials where the primary endpoint is occurrence of clinical events. Examples of such events include exacerbations of chronic obstructive pulmonary disease (COPD) (Calverley et al., 2009) and seizures in asthma patients (Keene et al., 2007) during certain periods. The purpose of clinical trials with count data is to evaluate the effect of new treatments on reducing the rate of event occurrence. In the analysis of count data, Poisson or related models are commonly employed. In order to incorporate overdispersion in count data, quasi–Poisson models with a linear variance function or negative binomial models with a quadratic variance function are often applied (Keene et al., 2007; Friede and Schmidli, 2010; Li et al., 2015; Tang, 2015). To accommodate possible misspecifications of the variance function, a sandwich-type robust variance estimator can be utilized to obtain a consistent estimator of the standard error for the estimated treatment effects (e.g., Cameron and Trivedi, 2013, pp. 73–75). This is particularly important for rigorous control of Type I error rates in confirmatory clinical trials with relatively large sample sizes, even if the working variance function is misspecified at the analysis stage, although a sandwich-type robust test is not necessarily able to control the Type I error rate in small sample size cases (Mancl and DeRouen, 2001).

There have been various proposals on sample size estimation in randomized clinical trials with count data. Self and Mauritsen (1988), Signorini (1991), and Yee (1998) proposed sample size calculation formulas when count data follows simple Poisson models. Mahnken (2009) and Friede et al. (2010) proposed modified sample size formulas to incorporate overdispersion using quasi–Poisson models. Keene et al. (2007) and Zhu and Lakkis (2014) developed another formula based on negative binomial models. Li, Wang, and Wei (2015) and Tang (2015) considered an extension to incorporate variations in the follow-up time among subjects in quasi–Poisson and negative binomial models, respectively. It should be noted that these methods can achieve the nominal power only when the specified or induced variance function is correct. However, since reliable estimates for the variance function are rarely available from earlier trial data or previous studies on similar treatments, it is difficult to specify the variance function exactly during the planning stage. Besides, the influence of the misspecification of the variance function on the power in testing treatment efficacy is generally unclear at the planning stage. We, therefore, argue that it is necessary to invoke the

concept of the "working" variance function at the planning stage, as well as at the analysis stage.

In this article, we propose a sample size formula that anticipates misspecifications of the working variance function. Specifically, we derived a method based on the asymptotic distribution of a Wald test statistic with a sandwich-type robust variance estimator under quasi–Poisson models. Under misspecifications of the working variance function, the asymptotic variance of the estimator of the treatment effect is expressed as a form involving both true and working variance functions. We also consider a sensitivity analysis for possible misspecifications of the "true" variance function when estimating sample sizes using our formula. We derive an equation to assess the influence of misspecifications of the specified true variance function on the power in testing treatment efficacy. With a catalog of anticipated true variance functions, including those which differ from the specified true variance function, we can identify a sample size estimate that maintains the power at a certain level (under the anticipated misspecifications of the true variance function). Of note, since our method is semi-parametric and is based on a quasi–likelihood model, the assessment of power and sample size can be made without full distributional assumptions for overdispersed count data.

Our sample size formula includes some of the existing sample size formulas as special cases when the working variance function is set as an assumed true variance function (and the assumed true variance function is specified correctly). That is, our method for sample size estimation can be seen as a generalization of the previously developed methods in an attempt to incorporate misspecifications of the working and true variance functions.

The rest of the article is organized as follows. In Section 2, we provide the Wald-type test using the sandwich-type robust variance estimator under the null hypothesis. In Section 3, we propose the sample size formula under misspecifications of the working variance function and the sensitivity analysis for possible misspecifications of the true variance function. In Section 4, we evaluate the validity of our method under finite samples via a simulation study. In Section 5, we provide an application of our method to a clinical trial to evaluate the treatment effect on prevention of COPD exacerbation. Lastly, we provide some discussion in Section 6.

2. Robust Two-Sample Test and Its Asymptotic Distributions

We consider a randomized clinical trial where n subjects are randomly assigned to the control group or the treatment group. Let n_i be the sample size of the ith group ($i = 1, 2$) and let $q_i = n_i/n$ be the proportion of subjects allocated to the ith group, where $n = n_1 + n_2$. Let π_i be the fixed allocation ratio to the ith group, where $\lim_{n\to\infty} q_i = \pi_i$ (> 0). Let X_{ij} be an indicator of the treatment assignment for the jth subject ($j = 1, \ldots, n_i$) in the ith group, such that $X_{1j} = 0$ and $X_{2j} = 1$ indicate assignment to the control group and treatment group, respectively. Let Y_{ij} be the number of events that occur during the period $[0, T_{ij}]$, from the point of treatment assignment to the end of follow-up for each patient.

We consider a statistical test for group comparison based on a quasi–likelihood using a sandwich-type robust variance estimator. The expectation of Y_{ij} is expressed as $E(Y_{ij}) = \mu_{ij} = T_{ij}\lambda_i$ as in the Poisson models. Here the incidence rate, λ_i, is specified as $\lambda_i = \exp(\beta_0 + \beta_1 X_{ij})$, where β_0 represents the (baseline) incidence rate in the control group, and β_1 represents the effect of the treatment relative to the control, $\beta_1 = \log(\lambda_2/\lambda_1)$. The true variance function for Y_{ij} is a function of μ_{ij}, denoted by $V(\mu_{ij};\phi)$, where ϕ is a dispersion parameter, but the form of the function is generally unknown. At the analysis stage, we employ a certain variance function $\tilde{V}(\mu_{ij};\tilde{\phi})$, which is called the working variance function (Cameron and Trivedi, 2013). Commonly used forms of $\tilde{V}(\mu_{ij};\tilde{\phi})$ are $\tilde{V}(\mu_{ij};\tilde{\phi}) = \tilde{\phi}\mu_{ij}$ and $\tilde{V}(\mu_{ij};\tilde{\phi}) = \mu_{ij} + \tilde{\phi}\mu_{ij}^2$. In what follows, we denote $V(\mu_{ij};\phi)$ and $\tilde{V}(\mu_{ij};\tilde{\phi})$ as V_{ij} and $\tilde{V}_{ij}$, respectively, for brevity. We also drop subscripts i and j when expressing models for a particular subject.

The maximum quasi–likelihood estimate of $\boldsymbol{\beta} = (\beta_0,\ \beta_1)'$, denoted by $\hat{\boldsymbol{\beta}}$, is obtained by solving the following estimation equation (Liang and Zeger, 1986).

$$\sum_{i=1}^{2}\sum_{j=1}^{n_i} \boldsymbol{D}_{ij}'\tilde{V}_{ij}^{-1}(Y_{ij} - \mu_{ij}) = \mathbf{0},$$

where $\boldsymbol{D}_{ij} = \partial\mu_{ij}/\partial\boldsymbol{\beta}'$. The asymptotic distribution of $\hat{\boldsymbol{\beta}}$ is expressed as follows.

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}),$$

where

$$\boldsymbol{A} = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\tilde{V}_{ij}^{-1}\boldsymbol{D}_{ij}'\boldsymbol{D}_{ij},$$

$$\boldsymbol{B} = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\tilde{V}_{ij}^{-2}V_{ij}\boldsymbol{D}_{ij}'\boldsymbol{D}_{ij}.$$

In particular, the asymptotic variance of $\hat{\beta}_1$, denoted by $n^{-1}W_1$, is the [2,2] element of $n^{-1}\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}$, where

$$W_1 = \frac{v_1}{\pi_1 w_1^2} + \frac{v_2}{\pi_2 w_2^2},$$

$$w_i = \lim_{n_i\to\infty}\frac{1}{n_i}\sum_{j=1}^{n_i}\tilde{V}_{ij}^{-1}\mu_{ij}^2, \quad v_i = \lim_{n_i\to\infty}\frac{1}{n_i}\sum_{j=1}^{n_i}\tilde{V}_{ij}^{-2}V_{ij}\mu_{ij}^2, \quad i = 1, 2.$$

Following Shieh (2001), we consider the asymptotic variance under the null hypothesis $H_0 : \beta_1 = 0$ by replacing μ_{ij} in W_1 with $\mu_{ij}^* = T_{ij}\exp(\beta_0^*)$, where β_0^* is the parameter regarding a common incidence rate under H_0. The dispersion parameters, $\tilde{\phi}$ and ϕ, in $\tilde{V}_{ij}$ and V_{ij} are also replaced with those under H_0, which are denoted by $\tilde{\phi}^*$ and ϕ^*, respectively. Likewise, the true and working variance functions under H_0, $V(\mu_{ij};\phi^*)$ and $\tilde{V}(\mu_{ij};\tilde{\phi}^*)$, are denoted by V_{ij}^* and $\tilde{V}_{ij}^*$, respectively. Thus, the

asymptotic variance of $\hat{\beta}_1$ under H_0 is $n^{-1}W_0$, where

$$W_0 = \frac{v_1^*}{\pi_1 w_1^{*2}} + \frac{v_2^*}{\pi_2 w_2^{*2}},$$

$$w_i^* = \lim_{n_i\to\infty} \frac{1}{n_i}\sum_{j=1}^{n_i} \tilde{V}_{ij}^{*-1}\mu_{ij}^{*2},\ v_i^* = \lim_{n_i\to\infty} \frac{1}{n_i}\sum_{j=1}^{n_i} \tilde{V}_{ij}^{*-2}V_{ij}^*\mu_{ij}^{*2},\ i=1,2.$$

We employ a Wald-type test statistic using the sandwich-type robust variance estimator under the null hypothesis,

$$Z = \frac{\hat{\beta}_1}{\sqrt{n^{-1}\widehat{W}_0}}, \tag{1}$$

where $\widehat{W}_0$ is the [2,2] element of $\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{B}}\hat{\boldsymbol{A}}^{-1}$. Here,

$$\hat{\boldsymbol{A}} = \frac{1}{n}\sum_{i=1}^{2}\sum_{j=1}^{n_i} \hat{\tilde{V}}_{ij}^{*-1}\hat{\boldsymbol{D}}_{ij}^{*'}\hat{\boldsymbol{D}}_{ij}^*,$$

$$\hat{\boldsymbol{B}} = \frac{1}{n}\sum_{i=1}^{2}\sum_{j=1}^{n_i} \hat{\tilde{V}}_{ij}^{*-2}\hat{V}_{ij}^*\hat{\boldsymbol{D}}_{ij}^{*'}\hat{\boldsymbol{D}}_{ij}^*,$$

$\hat{\boldsymbol{D}}_{ij}^* = [\hat{\mu}_{ij}^*,\ X_{ij}\hat{\mu}_{ij}^*]$, $\hat{V}_{ij}^* = (Y_{ij}-\hat{\mu}_{ij}^*)^2$, $\hat{\tilde{V}}_{ij}^* = \tilde{V}(\hat{\mu}_{ij}^*;\hat{\tilde{\phi}}^*)$, and $\hat{\mu}_{ij}^* = T_{ij}\exp(\hat{\beta}_0^*)$. The estimate $\hat{\beta}_0^*$ is the maximum quasi–likelihood estimate of β_0^*, and $\hat{\tilde{\phi}}^*$ is a consistent estimate of the dispersion parameter of $\tilde{V}$ under H_0. For large sample sizes, the distribution of $\hat{\beta}_1$ can be approximated by $N(0, n^{-1}W_0)$ under H_0 and by $N(\beta_1, n^{-1}W_1)$ under H_1. The Wald test statistic using the variance estimate $\widehat{W}_0$ under H_0 in (1) would be preferable because the size of such a test would more closely approximate the desired Type I error rate under large sample size cases (Lasker and King, 1997; Lachin, 2010, p. 563). We therefore expect that the use of the sandwich-type robust variance estimator of $\hat{\beta}_1$ under H_0 in the test statistic (1) would provide a more accurate p-value under H_0 even when the true variance function is misspecified.

3. Proposed Methods

Based on the asymptotic distribution of the robust two-sample test (1) derived from the quasi–Poisson model, we derive a quasi–likelihood method for power and sample size calculations without full distributional assumptions for over-dispersed count data.

3.1. *Sample Size Formula Incorporating Misspecifications of the Variance Function*

The asymptotic power of the test using the test statistic Z with two-sided significance level α is expressed as the following equation.

$$\Pr\left(Z > z_{1-\alpha/2}\right) = 1 - \Phi\left(z_{1-\alpha/2}\sqrt{\frac{W_0}{W_1}} - \sqrt{n}\cdot\frac{\beta_1}{\sqrt{W_1}}\right), \tag{2}$$

where z_a represents the lower a point of the standard normal distribution. Therefore, the sample size that provides a power greater than or equal to $1-\beta$ is expressed as

$$n \geq \frac{(z_{1-\alpha/2}\sqrt{W_0} + z_{1-\beta}\sqrt{W_1})^2}{\left\{\log(\lambda_2/\lambda_1)\right\}^2}. \tag{3}$$

The derivation of (2) and (3) are given in Web Appendix A. When evaluating W_0 and W_1 in (3), we should determine how to handle the follow-up period T_{ij} depending on the study design. One simple setting is to use a constant $T_{ij} = \tau$, say, representing a common follow-up duration. In this case, W_0 and W_1 may reduce to $W_0 = V^*/\mu^{*2}(\pi_1^{-1}+\pi_2^{-1})$ and $W_1 = V_1/(\mu_1^2\pi_1) + V_2/(\mu_2^2\pi_2)$, where $V^* = V(\mu^*;\phi^*)$, $\mu^* = \tau\exp(\beta_0^*)$, $V_i = V(\mu_i;\phi^*)$, $\mu_i = \tau\exp(\beta_0+\beta_1 X_{ij})$. Of note, when the follow-up time is constant, the working variance is canceled out, so that W_0 and W_1 depend on the true variance function only. We can also incorporate varying T_{ij}s among subjects, possibly related to subjects' early withdrawal, by introducing a distribution of follow-up periods, for example, an exponential distribution. Integration of the component quantities, v_i^* and w_i^* in W_0 and v_i and w_i in W_1, with regard to a specified distribution of follow-up periods might be attempted, although this is generally intractable. One simple, but practical method to evaluate these quantities is to simulate T_{ij} for a sufficiently large number of n_i, such as $n_i = 1{,}000{,}000$. The dispersion parameter in $\tilde{V}$ of (3) should be determined corresponding to V in the sample size calculation (through determining W_0 and W_1 under the null and alternative hypotheses, respectively). Specifically, we consider using the working variance function, $\tilde{V}_{ij} = \mu_{ij} + \tilde{\phi}_p\mu_{ij}^p$, which is a generalization of the negative binomial model (Cameron and Trivedi, 1986). Under the alternative hypothesis, for a given V_{ij} and μ_{ij}, the dispersion parameter, $\tilde{\phi}_p$, can be determined by

$$\begin{aligned}\tilde{\phi}_p &= \lim_{n\to\infty}\arg\min_{\tilde{\phi}_p}\left\{\sum_{i=1}^{2}\sum_{j=1}^{n_i}(V_{ij}-\tilde{V}_{ij})^2\right\}\\ &= \lim_{n\to\infty}\frac{\sum_{i=1}^{2}\sum_{j=1}^{n_i}(V_{ij}-\mu_{ij})\mu_{ij}^p}{\sum_{i=1}^{2}\sum_{j=1}^{n_i}\mu_{ij}^{2p}}.\end{aligned} \tag{4}$$

Under the null hypothesis, the dispersion parameter is determined by replacing V_{ij} with V_{ij}^* and μ_{ij} with μ_{ij}^* in (4). We can employ other working variance functions such as $\phi\mu^p$ and $\phi_1\mu+\phi_2\mu^2$ (Wang and Lin, 2005; Wang and Zhao, 2007).

Since the true variance function V is generally unknown, we specify V in (3) as $V^{(S)}$ to approximate V. By considering settings with $V^{(S)} \neq \tilde{V}$, we can estimate the required sample sizes to achieve the nominal power under misspecifications of the working variance function for $V^{(S)}$. For example, by setting $\tilde{V} = \tilde{\phi}\mu$ and $V^{(S)} = \mu + \phi\mu^p$ for constants $\tilde{\phi}$, ϕ, and p, we can obtain the required sample size when applying the quasi–Poisson models with the linear working variance function at the analysis stage, while the true variance function is the p-th degree variance function. It is noted that by setting $V^{(S)} = \tilde{V}$, sample size formula (3) reduces to some existing formulas for sample size calculation based on the model-based variance function. Table 1 summarizes the relationship between our sample size formula and existing formulas.

Table 1
Relationship between proposed method and existing methods

		Settings in (3)	
Previous research	Model	Variance function	Follow-up time
Signorini (1991)	Poisson model	$V^{(S)} = \tilde{V} = \mu$	T
Friede et al. (2009)	Quasi–likelihood model	$V^{(S)} = \tilde{V} = \phi\mu$	T
Mahnken (2009)	Quasi–likelihood model	$V^{(S)} = \tilde{V}$	T
Zhu and Lakkis (2014)	Negative binomial model	$V^{(S)} = \tilde{V} = \mu + \phi\mu^2$	T
Tang (2015)	Negative binomial model	$V^{(S)} = \tilde{V} = \mu + \phi\mu^2$	T_{ij}
Li et al. (2015)	Quasi–likelihood model	$V^{(S)} = \tilde{V}$	T_{ij}

Note: $V^{(S)} = \tilde{V}$ in the "Variance function" column indicates that the sample size formula was derived without assuming a specific variance function. T in the "Follow-up time" column indicates that the sample size formula was derived assuming a constant follow-up time for all subjects. T_{ij} in the "Follow-up time" column indicates that the sample size formula was derived assuming a different follow-up time with a distribution for each subject.

3.2. *Sensitivity Analysis for Unknown True Variance Function*

Since the specified true variance function $V^{(S)}$ is just an approximation of the true variance function V, it is warranted to consider a sensitivity analysis that assesses whether the power is maintained for a range of plausible variance functions for V, which may include $V^{(S)}$. When the sample size is determined based on $V^{(S)}$ using (3), the power under the true variance function V is expressed as the following equation,

$$\text{Power} = 1 - \Phi(z_\beta \cdot \nu),$$
$$\nu = \frac{z_{1-\alpha/2}}{z_{1-\beta}} \sqrt{\frac{W_0}{W_1}} \left(\sqrt{\frac{W_0^{(S)}}{W_0}} - 1 \right) + \sqrt{\frac{W_1^{(S)}}{W_1}}, \quad (5)$$

where $W_1^{(S)}$ is the one based on $V^{(S)}$, rather than V, in W_1; the dispersion parameter of $\tilde{V}$ in $W_1^{(S)}$ is determined (using (4)) based on $V^{(S)}$, rather than V. $W_0^{(S)}$ is obtained by replacing the dispersion parameters of $V^{(S)}$ and $\tilde{V}$ in $W_1^{(S)}$ with those under the null hypothesis. The derivation of (5) is given in Web Appendix B. It should be noted that if $V^{(S)}$ is equivalent to V, then ν becomes 1 and the power reduces to $1 - \beta$.

3.3. *Proposed Sample Size Determination*

We propose a procedure of sample size determination anticipating the misspecification of the variance function using (3) and (5) below.

(i) Consider a catalog of plausible true variance functions, $V^{(k)}$, $k = 1, \ldots, K$, to approximate the true variance function, V. It is recommended to consider a general form for $V^{(k)}$ to cover V, involving popular variance functions such as quasi–Poisson and negative binomial models, for example $V^{(k)} = \mu + \phi_k \mu^{p_k}$ for some constants ϕ_k and p_k, $k = 1, \ldots, K$. Among $V^{(k)}$, $k = 1, \ldots, K$, select one which is deemed the most plausible for V, denoted as $V^{(S)}$, which could be used for determining the sample size of the clinical trial.

(ii) Specify the working variance function $\tilde{V}$. Typically, this may be specified as a model-based variance function in the analysis model to well approximate $V^{(S)}$. For example, $\tilde{V}(\mu; \tilde{\phi}^{(S)}) = \tilde{\phi}^{(S)}\mu$ is used for the quasi–Poisson models (which can be implemented using the "scale = p" option in the GENMOD procedure in SAS), while $\tilde{V}(\mu; \tilde{\phi}^{(S)}) = \mu + \tilde{\phi}^{(S)}\mu^2$ is used for the negative binomial models. The dispersion parameter in $\tilde{V}$ should be determined using (4) for $V^{(S)}$. Then, we calculate the required sample size, $n^{(S)}$, to achieve the nominal power $1–\beta$ when the true variance function is $V^{(S)}$ using (3).

(iii) With the possible true variance functions, $V^{(k)}$, $k = 1, \ldots, K$, for the true variance function, V, we implement a sensitivity analysis for the sample size $n^{(S)}$ calculated based on $V^{(S)}$. As done for $V^{(S)}$, we evaluate the working variance $\tilde{V}$ for each $V^{(k)}$ ($k = 1, \ldots, K$); the dispersion parameter in $\tilde{V}$ is determined using (4) for $V^{(k)}$ ($k = 1, \ldots, K$). We then assess the power for each of $V^{(1)}, \ldots, V^{(K)}$, using (5) under the sample size of $n^{(S)}$,

$$\text{Power}^{(k)} = 1 - \Phi(z_\beta \cdot \nu^{(k)}),$$
$$\nu^{(k)} = \frac{z_{1-\alpha/2}}{z_{1-\beta}} \sqrt{\frac{W_0^{(k)}}{W_1^{(k)}}} \left(\sqrt{\frac{W_0^{(S)}}{W_0^{(k)}}} - 1 \right) + \sqrt{\frac{W_1^{(S)}}{W_1^{(k)}}},$$

where $W_1^{(k)}$ and $W_0^{(k)}$ are those based on $V^{(k)}$ for V in W_1 and W_0, respectively. If the $\text{Power}^{(k)}$, $k = 1, \ldots, K$ are acceptable, the sample size is determined as $n^{(S)}$. Otherwise, we adjust the sample size (e.g., consider a larger sample size) and assess the power again.

In specifying the catalog of true variance functions in Step (i), it is generally advisable to incorporate data available from previous clinical trials or observational studies on the same or a similar treatment. If raw data is available, point estimates and confidence intervals for ϕ and p in $V(\mu) = \mu + \phi\mu^p$ can be obtained by fitting the variance function to the squared residuals (Wang and Lin, 2005). Referring to confidence intervals, we can take into consideration uncertainty in the estimation of the variance function, and thus determine a plausible range of the true variance function. When only summary data is available, estimation becomes difficult (see Web Appendix C for an estimation of the dispersion parameter using summary data from previous studies.)

The impact of misspecification of the true variance function on the power can be captured by a contour plot of the power over grids of $(\phi; p)$ in $V = \mu + \phi\mu^p$. This is a useful tool for ensuring stability of the power for a range of the true variance function before finalizing sample size determination.

We provide an application of this procedure in Section 5.

4. Simulation Studies

In this section, we describe simulation studies performed to evaluate the adequacy of the sample size estimation formula (3) and the power equation (5) under finite samples. We assumed equal allocation ratios for respective treatment groups, $q_1 = q_2 = 0.5$. We set the rate of event incidence in the control group, λ_1, to be 1.25 (/year), and the relative risk, $\exp(\beta_1) = \lambda_2/\lambda_1$ representing the effect of the treatment, $\exp(\beta_1) = \lambda_2/\lambda_1$ to be 0.4, 0.6, or 0.8. We set the same follow-up period, $\tau = 1$ (year), for all subjects. We assumed random early withdrawal or dropout during the follow-up period and the time to dropout followed an exponential distribution with the rate λ_d (/year). We set $\lambda_d = 0.356$, such that the expected cumulative proportion of dropout during the follow-up period was 30%. As such, the observed follow-up time, T, for a particular subject was set as a minimum of the time to dropout and τ.

In sample size calculations, we considered the catalog of plausible true variance functions to be $V^{(k)}(\mu) = \mu + \phi_k\mu^{p_k}$, $(p_1, p_2, p_3, p_4) = (0.5, 1, 2, 3)$ (see the column $V^{(k)}$ in Tables 2 and 3). Since the variance functions $V^{(k)}$s should approximate the same true variance function, we link $V^{(k)}$s to have the same variance value at the global mean under the null hypothesis, $\bar{\mu}^* = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{2} \sum_{j=1}^{n_i} T_{ij} \exp(\beta_0^*)$, where $\exp(\beta_0^*)$ is the parameter regarding a common incidence rate under H_0. Specifically, we set $V^{(k)}(\bar{\mu}^*) = V^{(2)}(\bar{\mu}^*)$ (i.e., the variance when $p_k = 1$), resulting in ϕ_k, satisfying $\phi_k = \phi_2/\bar{\mu}^{*p_k-1}$, where ϕ_2 was set as 2 or 5. The value of $\phi_2 = 2$ represents a value close to the point estimate of the dispersion parameter based on summary data given by Calverley et al. (2009) (see Web Appendix C for the estimation method). As confidence intervals on the dispersion parameter were not obtainable (based on summary data), we arbitrarily set $\phi_2 = 5$ to represent more dispersed data. The value of ϕ_k was fixed when evaluating the variance function under H_1. Regarding the dispersion parameter for the working variance $\tilde{V}$, as described in Section 3.1, the sample size formula (3) requires determination of the dispersion parameters for $\tilde{V}$ under the null and alternative hypotheses based on (4).

We assumed quasi–Poisson models with $\tilde{V} = \tilde{\phi}_1\mu$ or negative binomial models with $\tilde{V} = \mu + \tilde{\phi}_2\mu^2$. The empirical power was evaluated for the robust two-sample test using the Wald-type test statistic (1) with a one-sided significance level of 2.5% in testing treatment efficacy. Following Li et al. (2015), to simulate count data, we employed a parametric, negative binomial model modulated to have the specified variance function, $V_{ij} = \mu_{ij} + \phi_p\mu_{ij}^p$, where the variance function of the negative binomial model, $V_{ij} = \mu_{ij} + \psi_{ij}\mu_{ij}^2$, was modulated at the subject level, such that $\psi_{ij} = \phi_p\mu_{ij}^{p-2}$. This is to assume subject-specific gamma distributions for subject-level event rates to obtain over-dispersed count data with the variance function $V_{ij} = \mu_{ij} + \phi_p\mu_{ij}^p$, although other distributions, such as log-normal distributions, can be used. We conducted 10,000 simulations for each scenario.

4.1. *Adequacy of Sample Size Estimation Formula (3)*

Table 2 summarizes the required sample size for an asymptotic power of 90%, as well as the empirical power under the sample sizes for each scenario. Depending on the effect size, $\exp(\beta_1) = \lambda_2/\lambda_1$, or the relative risk, a wide range of required sample sizes was obtained. The empirical power was generally close to the nominal level, 90%, for all scenarios. In cases where the sample size was less than 100 per group (for a large effect size, $\exp(\beta_1) = 0.4$), the empirical power was slightly greater than the nominal level, indicating the conservativeness of our formula in such situations.

4.2. *Adequacy of Power Equation (5)*

The sample size per treatment group, n_1, was determined for the asymptotic power of 90% under $V^{(S)} = \tilde{V}$. For various variance functions for the true one V, Table 3 summarizes the asymptotic power based on (5) and the empirical power under n_1 based on 10,000 simulations. The empirical power was almost the same with asymptotic powers for small or moderate effect sizes, while slightly greater than the asymptotic power for large effect size (i.e., $\exp(\beta_1) = 0.4$), which again indicates the conservativeness of our method.

5. Application

We applied the method for sample size determination to a confirmatory clinical trial for developing a novel phosphodiesterase 4 (PDE4) inhibitor for COPD treatment. Calverley et al. (2009) provided the results of two phase III confirmatory clinical trials, M2-124 and M2-125, of a PDE4 inhibitor, roflumilast. These were placebo-controlled trials in subjects followed up for one year, with co-primary endpoints of the change in the forced expiratory volume in one second (FEV 1) and the incidence rate of moderate or severe COPD exacerbation. Regarding the analysis of the COPD exacerbation rate, a quasi–Poisson model with a linear variance function, $V = \phi\mu$, was used in the primary analysis and a negative binomial model was used in a sensitivity analysis.

Here, we present the design of a placebo-controlled clinical trial for a novel PDE4 inhibitor based on the previous studies for roflumilast by Calverley et al. (2009). In comparing the COPD exacerbation rate between the two treatment groups, we accommodated the quasi–Poisson model used in Calverley et al. (2009) and specified the working variance function as $\tilde{V} = \tilde{\phi}\mu$. However, since we had limited knowledge about the true variance function for the frequency of the COPD exacerbation, we planned to perform the robust two-sample test using the test statistic (1).

Regarding the parameters for sample size calculation, the COPD exacerbation rate of the placebo group, λ_1, was assumed to be 1.35 (/year), referring to Calverley et al. (2009). We expected a 20% risk reduction of the COPD exacerbation rate by the treatment using the novel PDE4 inhibitor compared to placebo (i.e., $\exp(\beta_1) = 0.8$), such that the COPD exacerbation rate in the novel PDE4 inhibitor group was set as $\lambda_2 = 1.08$ (/year). The proportion of dropouts after one-year follow-up was assumed to be 30%. Assuming an exponential distribution for the time to dropout or censoring,

Table 2
Required sample size per treatment group (n_1) and empirical power (EP) based on simulations

			$\tilde{V} = \tilde{\phi}_1\mu$		$\tilde{V} = \mu + \tilde{\phi}_2\mu^2$	
$\exp(\beta_1)$	k	$V^{(k)}$	n_1	EP (%)	n_1	EP (%)
0.8	1	$\mu + 0.97\mu^{0.5}$	889	90.13	926	89.96
	2	$\mu + 1.00\mu^{1.0}$	901	89.70	913	89.74
	3	$\mu + 1.06\mu^{2.0}$	956	90.00	944	90.00
	4	$\mu + 1.12\mu^{3.0}$	1042	90.34	1012	90.21
	1	$\mu + 3.39\mu^{0.5}$	1978	90.08	2393	89.11
	2	$\mu + 4.00\mu^{1.0}$	2245	90.06	2427	89.66
	3	$\mu + 4.23\mu^{2.0}$	2464	89.88	2361	90.13
	4	$\mu + 4.48\mu^{3.0}$	2810	90.54	2555	90.45
0.6	1	$\mu + 0.92\mu^{0.5}$	200	91.11	208	90.44
	2	$\mu + 1.00\mu^{1.0}$	200	90.53	203	91.30
	3	$\mu + 1.19\mu^{2.0}$	213	91.23	211	91.13
	4	$\mu + 1.41\mu^{3.0}$	239	91.10	231	90.80
	1	$\mu + 3.67\mu^{0.5}$	492	90.45	602	89.52
	2	$\mu + 4.00\mu^{1.0}$	495	90.40	532	90.02
	3	$\mu + 4.74\mu^{2.0}$	545	90.53	522	90.72
	4	$\mu + 5.63\mu^{3.0}$	649	90.95	587	90.92
0.4	1	$\mu + 0.86\mu^{0.5}$	79	91.59	82	90.89*
	2	$\mu + 1.00\mu^{1.0}$	78	92.04	78	91.36
	3	$\mu + 1.36\mu^{2.0}$	83	92.43	82	91.51*
	4	$\mu + 1.85\mu^{3.0}$	99	91.88	95	91.76
	1	$\mu + 3.43\mu^{0.5}$	193	91.04	230	91.34
	2	$\mu + 4.00\mu^{1.0}$	188	91.37	201	91.24
	3	$\mu + 5.43\mu^{2.0}$	209	91.26	200	91.79*
	4	$\mu + 7.37\mu^{3.0}$	272	91.67	243	91.21

Note: *For one or two out of 10,000 simulated datasets, the GENMOD procedure in Web Appendix D-4 did not return estimation results with an error message on floating point overflow. The estimation results for the datasets were obtained successfully using the GLIMMIX procedure.

we set the dropout rate $\lambda_d = 0.357$ (/year). We also assumed a uniform distribution for the time to dropout, but almost the same results were obtained (results not shown). We set the same allocation ratio for the two treatments, that is, $n_1 = n_2$. We used a two-sided significance level of 5%.

We then applied the procedure proposed in Section 3.3.

(i) We first determined the catalog of the plausible true variance functions as $V^{(k)} = \mu + \phi_k\mu^{p_k}$, where $p_k \in \{0.5, 1, 2, 3\}$ (see the column $V^{(k)}$ in Table 4). We hoped that the power was maintained at a sufficiently high value under possible misspecifications of p_k within the range (0.5, 1, 2, 3), which represents a wide range for the true variance function, covering quasi–Poisson ($p_k = 1$) and negative binomial ($p_k = 2$) models. As estimates for the dispersion parameters were not reported in Calverley et al. (2009), we employed a method proposed by Zhu and Lakkis (2014) that allows estimating the dispersion parameters from the limited published information; assuming $V = \phi\mu$ (as in Calverley et al. (2009)), the dispersion parameter, ϕ, was estimated between 2 and 3 based on the results of Calverley et al. (2009) (see Web Appendix C for the details of the calculation). Taking an intermediate value, 2.5, we determined one of the plausible variance functions as $V^{(2)} = \mu + 1.5\mu^{1.0} = 2.5\mu$ (i.e.,$\phi_2 = 1.5$). Furthermore, the dispersion parameters of the other variance functions were derived in the same way as in Section 4, as $\phi_k = \phi_2/\bar{\mu}^{*p_k-1}$, $1 \le k \le 4$, where $\bar{\mu}^* = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{2}\sum_{j=1}^{n_i} T_{ij}\exp(\beta_0^*)$ and $\exp(\beta_0^*) = 1.215$ (/year). To examine the impact of an underestimation of the true variance on power reduction, we also considered a case with an upper limit of the estimated variance, 3.0μ, that is, $V^{(6)} = \mu + 2.0\mu^{1.0} = 3.0\mu$ (i.e.,$\phi_6 = 2.0$) and $\phi_k = \phi_6/\bar{\mu}^{*p_k-1}$, $5 \le k \le 8$. We tentatively selected $V^{(S)} = 2.5\mu$ as the true variance function to calculate the required sample size.

(ii) We obtained convergence values for the dispersion parameter, $\tilde{\phi}$, of the working variance function, $\tilde{V} = \tilde{\phi}\mu$, using (4) for the selected true variance function, $V^{(S)}$, under the null and alternative hypotheses. We then calculated the required sample size using (3) to ensure a power of 90% when the true variance function was $V^{(S)} = 2.5\mu$. We obtained a required sample size of $n_1 = n_2 = 1042$.

Table 3
Comparison of the asymptotic power (AP) and empirical power (EP) under misspecifications of the true variance function

			$\tilde{V} = \tilde{\phi}_1\mu$			$\tilde{V} = \mu + \tilde{\phi}_2\mu^2$		
$\exp(\beta_1)$	k	$V^{(k)}$	n_1	AP (%)	EP (%)	n_1	AP (%)	EP (%)
0.8	1	$\mu + 0.97\mu^{0.5}$	901	90.39	90.68	944	90.57	90.46
	2	$\mu + 1.00\mu^{1.0}$	901	90.03	89.70	944	90.93	90.50
	3	$\mu + 1.06\mu^{2.0}$	901	88.24	88.18	944	90.00	90.00
	4	$\mu + 1.12\mu^{3.0}$	901	85.40	86.07	944	87.92	87.68
	1	$\mu + 3.39\mu^{0.5}$	2245	93.22	93.58	2361	89.62	89.02
	2	$\mu + 4.00\mu^{1.0}$	2245	90.00	90.06	2361	89.21	88.76
	3	$\mu + 4.23\mu^{2.0}$	2245	87.16	87.39	2361	90.00	90.13
	4	$\mu + 4.48\mu^{3.0}$	2245	82.49	82.38	2361	87.60	87.68
0.6	1	$\mu + 0.92\mu^{0.5}$	200	90.08	91.11	211	90.53	90.56
	2	$\mu + 1.00\mu^{1.0}$	200	90.01	90.53	211	91.12	91.78
	3	$\mu + 1.19\mu^{2.0}$	200	88.12	88.73	211	90.12	91.13
	4	$\mu + 1.41\mu^{3.0}$	200	84.12	85.31	211	87.17	88.39
	1	$\mu + 3.67\mu^{0.5}$	495	90.20	90.39	522	85.63	86.03
	2	$\mu + 4.00\mu^{1.0}$	495	90.01	90.40	522	89.46	89.56
	3	$\mu + 4.74\mu^{2.0}$	495	87.03	87.59	522	90.01	90.72
	4	$\mu + 5.63\mu^{3.0}$	495	80.35	81.27	522	86.24	86.87
0.4	1	$\mu + 0.86\mu^{0.5}$	78	89.82	91.57	82	90.29	90.89*
	2	$\mu + 1.00\mu^{1.0}$	78	90.31	92.04	82	91.33	92.71
	3	$\mu + 1.36\mu^{2.0}$	78	88.33	90.70	82	90.27	91.51*
	4	$\mu + 1.85\mu^{3.0}$	78	81.59	84.21	82	85.19	86.85*
	1	$\mu + 3.43\mu^{0.5}$	188	89.32	91.03	200	86.00	89.08*
	2	$\mu + 4.00\mu^{1.0}$	188	90.01	91.37	200	89.94	91.25
	3	$\mu + 5.43\mu^{2.0}$	188	86.63	88.71	200	90.03	91.79*
	4	$\mu + 7.37\mu^{3.0}$	188	74.95	78.30	200	82.85	84.81*

Note: *For one or two out of 10,000 simulated datasets, the GENMOD procedure in Web Appendix D-4 did not return estimation results with an error message on floating point overflow. The estimation results for the datasets were obtained successfully using the GLIMMIX procedure.

(iii) For the sensitivity analysis for possible misspecifications of the true variance functions, we calculated the asymptotic power using (5) when each $V^{(k)}$, $k = 1, \ldots, 8$ was the true variance function. The results are shown in Table 4. Again, we obtained the empirical power via 10,000 simulations based on the modulated negative binomial model as in Section 4. The asymptotic power well approximated the empirical power in all the scenarios. The following provides some examples of sample size adjustment to ensure the power at a certain level even under misspecifications of the true variance function.

- When $V = \mu + 1.44\mu^{3.0}$, the asymptotic power was 84.45%. The misspecification of the true variance function caused a 5.56% reduction in power. To ensure a power of 90%, we should increase the sample size to $n_1 = 1238$.
- When $V = \mu + 1.96\mu^{2.0}$, the asymptotic power was 81.22%, with a power reduction of 8.78% by the misspecification of the variance function. We should increase the sample size to $n_1 = 1155$ to ensure a power of 85% in this case.
- When $V = \mu + 1.91\mu^{3.0}$, the asymptotic power was 76.84%, a power reduction of 13.17% by the misspecification of the variance function. We should increase the sample size to $n_1 = 1127$ to ensure a power of 80% in this case.

As described in Section 3.3, it is helpful to investigate the stability of the power for a range of the true variance functions. Figure 1 shows a contour plot of the power over grids of (ϕ, p) of $V(\mu) = \mu + \phi\mu^p$ under $n_1 = n_2 = 1042$. The range of (ϕ, p) in Figure 1 was determined based on a simulation under settings similar to Calverley et al. (2009) (see Web Appendix E for the details of the simulation). To take account of uncertainty in estimating the variance function, we set the range of $p = 0.5$–4.5 and $\phi = 0.5$–2.5 based on the width of the confidence intervals of p and ϕ that were estimated via the simulation. If there is some subregion in the plausible range of (ϕ, p) for which the power is unacceptably low, the sample size should be modified to improve the power over this subregion.

Table 4
Asymptotic power (AP) and empirical power (EP) for the required sample size (n_1) in evaluating the efficacy of a new PDE4 inhibitor for COPD treatment.

			$\tilde{V} = \tilde{\phi}_1 \mu$		
$\exp(\beta_1)$	k	$V^{(k)}$	n_1	AP (%)	EP (%)
0.8	1	$\mu + 1.52\mu^{0.5}$	1042	90.37	90.59
	2	$\mu + 1.50\mu^{1.0}$	1042	90.02	89.88
	3	$\mu + 1.47\mu^{2.0}$	1042	87.89	87.83
	4	$\mu + 1.44\mu^{3.0}$	1042	84.45	84.73
	5	$\mu + 2.02\mu^{0.5}$	1042	84.76	85.01
	6	$\mu + 2.00\mu^{1.0}$	1042	84.18	84.09
	7	$\mu + 1.96\mu^{2.0}$	1042	81.22	81.74
	8	$\mu + 1.91\mu^{3.0}$	1042	76.84	77.26

6. Discussion

We have developed new methods for sample size determination in comparative clinical trials with over-dispersed count data, incorporating possible misspecification of both working and true variance functions. Our methods involve many of the existing sample size methods as special cases (see Table 1) under the specification that the working variance function, $\tilde{V}$, is equivalent to the selected true variance function, $V^{(S)}$, and also under the assumption that $V^{(S)}$ accords with the true variance function V. The proposed methods would, therefore, allow a robust evaluation of sample size under many situations in which there is limited information about the true variance function at the design of the clinical trials.

In Sections 4 and 5, we ascertained a close agreement of the asymptotic power based on the proposed formula, with the empirical power based on a parametric, negative binomial model with a modulation of the dispersion parameter at the subject level. We expect to obtain similar results for other parametric models (such as log-normal distributions, rather than gamma distributions, to capture over-dispersion in the count data). This property derives from the fact that our method is based on quasi–likelihood models. Our method is distribution free, requiring specification of only the mean and variance functions for over-dispersed count data.

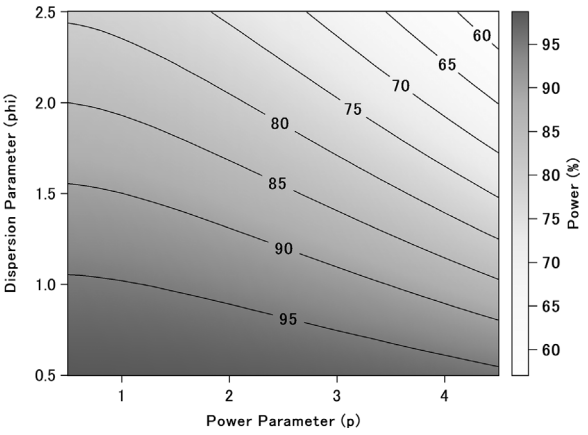

Figure 1. A contour plot of the power over grids of the dispersion and power parameters (ϕ, p) of the true variance function $V(\mu) = \mu + \phi\mu^p$ under $n_1 = n_2 = 1042$. Darker color corresponds to higher power.

The sensitivity analysis based on (5) is particularly helpful for identifying a set of possible true variance functions that may yield underpowered trials, as illustrated in Section 5. Final selection of the true variance function for determining the sample size should be made on a case-by-case basis, taking into account various factors such as the characteristics of the disease, availability of other treatments, and so forth, as well as the cost and time that can be invested and available recourses for clinical trials.

When severe uncertainty is expected in specifying the true variance function at the planning stage, sample size re-estimation, such as that conducted in the internal pilot study design, can be a good solution to avoid an underpowered study. Point estimates and confidence intervals for the parameters of the variance function $V = \mu + \phi\mu^p$ at an interim analysis would provide more reliable insight into a plausible range for the true variance function. The contour plot described in Section 3.3 and 5 is still useful for evaluating the stability of the power for a modified sample size.

Although we have presented our sample size formulas for fixed subject follow-up periods, T_{ij}s, we can derive formulas that incorporate the distribution of the follow-up time T_{ij}. As an approximation, one approach is to replace the numerator, v, and denominator, w^2, of W in (2) or (5) with $E(v)$ and $E(w^2)$, respectively, under a distribution for T_{ij}. However, as explicit formulas are unavailable, except for the variance function $V_{ij} = \phi\mu_{ij}$ (Li et al., 2015), numerical calculations are generally needed. Another approach is to consider an upper bound for a variance component (such as W) and derive an explicit formula for conservative sample size calculation (Tang, 2015). Alternatively, as a practical approach to incorporating the distribution of T_{ij}, as was done in Sections 4 and 5, we can simulate T_{ij} based on a specified underlying distribution (such as an exponential distribution) and employ our formula to calculate the required sample size.

Our approach to incorporating misspecifications of the working and true variance functions for over-dispersed count data could be extended to more complex situations, including comparative trials with clustered data, such as repeated

measurement data. For clustered data, misspecifications of the correlation structures, as well as those of the variance function, are frequently an important concern, although almost all of the sample size formulas assume that the working correlation structure is equivalent to the true one (Liu and Liang, 1997; Rochon, 1998). Incorporation of misspecified correlation structures in sample size determination is an important subject for future research.

7. Supplementary Materials

Web Appendices A, B, C, D, and E referenced in Sections 3.1, 3.2, 4.1, and 5, and the SAS code are available with this article at the *Biometrics* website on Wiley Online Library.

Acknowledgements

This research was supported by CREST-JST (JPMJCR1412) and a Grant-in-Aid for Scientific Research (16H06299) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Calverley, P. M., Rabe, K. F., Goehring, U. M., Kristiansen, S., Fabbri, L. M., and Martinez, F. J. (2009). Roflumilast in symptomatic chronic obstructive pulmonary disease: Two randomised clinical trials. *The Lancet* **374**, 685–694.

Cameron A. C. and Trivedi P. K. (1986). Econometric models based on count data: Comparisons and applications of some estimators and tests. *Journal of Applied Econometrics* **1**, 29–53.

Cameron, A. C. and Trivedi, P. K. (2013). *Regression Analysis of Count Data*, 2nd edition. New York: Cambridge University Press.

Friede, T. and Schmidli, H. (2010). Blinded sample size reestimation with count data: Methods and applications in multiple sclerosis. *Statistics in Medicine* **29**, 1145–1156.

Keene, O. N., Jones, M. R., Lane, P. W., and Anderson, J. (2007). Analysis of exacerbation rates in asthma and chronic obstructive pulmonary disease: Example from the TRISTAN study. *Pharmaceutical Statistics* **6**, 89–97.

Lachin, J. M. (2010). *Biostatistical Methods: The Assessment of Relative Risks*, 2nd Edition. Hoboken, New Jersey: John Wiley & Sons, Inc.

Laskar M. R. and King M. L. (1997). Modified Wald test for regression disturbances. *Economics Letters* **56**, 5–11.

Li, H., Wang, L., Wei, L., and Quan, H. (2015). Sample size calculation for count data in comparative clinical trials with nonuniform patient accrual and early dropout. *Journal of Biopharmaceutical Statistics* **25**, 1–15.

Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Liu, G. and Liang, K. Y. (1997). Sample size calculations for studies with correlated observations. *Biometrics* **53**, 937–947.

Mahnken, J. D. (2009). Power and sample size calculations for models from unknown distributions. *Statistics in Biopharmaceutical Research* **1**, 328–336.

Mancl, L. A. and DeRouen, T. A. (2001). A covariance estimator for GEE with improved small-sample properties. *Biometrics* **57**, 126–134.

Rochon, J. (1998). Application of GEE procedures for sample size calculations in repeated measures experiments. *Statistics in Medicine* **17**, 1643–1658.

Self, S. G. and Mauritsen, R. H. (1988). Power/sample size calculations for generalized linear models. *Biometrics* **44**, 79–86.

Shieh, G. (2001). Sample size calculations for logistic and Poisson regression models. *Biometrika* **88**, 1193–1199.

Signorini, D. F. (1991). Sample size for Poisson regression. *Biometrika* **78**, 446–450.

Tang, Y. (2015). Sample size estimation for negative binomial regression comparing rates of recurrent events with unequal follow-up time. *Journal of Biopharmaceutical Statistics* **25**, 1100–1113.

Wang, Y. G. and Lin, X. (2005). Effects of variance-function misspecification in analysis of longitudinal data. *Biometrics* **61**, 413–421.

Wang, Y. G. and Zhao, Y. (2007). A modified pseudolikelihood approach for analysis of longitudinal data. *Biometrics* **63**, 681–689.

Yee, K. F. (1998). Poisson regression sample size estimation in clinical trials. In *Proceedings of the Biopharmaceutical Section, Annual Meeting of the American Statistical Association*, 114–118. Dallas, Texas.

Zhu, H. and Lakkis, H. (2014). Sample size calculation for comparing two negative binomial rates. *Statistics in Medicine* **33**, 376–387.

Received June 2017. *Revised February* 2018. *Accepted February* 2018.