Johannes Forkman*

Generalized Confidence Intervals for Intra- and Inter-subject Coefficients of Variation in Linear Mixed-effects Models

DOI 10.1515/ijb-2016-0093

Abstract: Linear mixed-effects models are linear models with several variance components. Models with a single random-effects factor have two variance components: the random-effects variance, i. e., the inter-subject variance, and the residual error variance, i. e., the intra-subject variance. In many applications, it is practice to report variance components as coefficients of variation. The intra- and inter-subject coefficients of variation are the square roots of the corresponding variances divided by the mean. This article proposes methods for computing confidence intervals for intra- and inter-subject coefficients of variation using generalized pivotal quantities. The methods are illustrated through two examples. In the first example, precision is assessed within and between runs in a bioanalytical method validation. In the second example, variation is estimated within and between main plots in an agricultural split-plot experiment. Coverage of generalized confidence intervals is investigated through simulation and shown to be close to the nominal value.

Keywords: bioanalytical method validation, generalized pivotal quantity, linear mixed model, semiparametric mixed-effects model, split-plot experiment

1 Introduction

The coefficient of variation is the sample standard deviation divided by the sample mean. This measure of dispersion is used in a wide range of applications where the standard deviation tends to increase linearly with the level of the observations. In these applications, observations cannot take negative values. Specifically, measurements are made on a ratio scale. The coefficient of variation can be used as an estimator of the population coefficient of variation, defined as the square root of the population variance, i. e., the square root of the second central moment, divided by the expected value.

Much research has been made on inference for a single coefficient of variation. When the sample consists of independent observations from a normal distribution, an exact confidence interval for the coefficient of variation can be computed using the noncentral t-distribution [1]. Approximate confidence intervals have been developed using asymptotic normality [2] and Taylor series expansion [3]. Vangel [4] and Forkman [5] advocated the use of McKay's approximation [6, 7] for the coefficient of variation. The approximate confidence interval method proposed by Forkman [5] can be used when there are samples from populations with different expected values but with a common coefficient of variation, which is a situation that was also considered by Verrill and Johnson [8]. Hayter [9] constructed a bounded confidence interval for a coefficient of variation when observations are normally distributed. Wong and Wu [10] proposed a confidence interval using a modified signed log likelihood ratio method, which can also be used for gamma distributed observations. They showed that for small samples of normally distributed observations, this method performs as well as the Vangel [4] method.

The present article considers linear random-effects and mixed-effects models for balanced and unbalanced sets of normally distributed observations. In balanced datasets, each sub-class contains the same number of observations (the exact mathematical definition of balance is complex but has been formalized by a number of authors; for further references, see [11, p.4]). Linear random-effects and mixed-effects

*Corresponding author: Johannes Forkman, Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Box 7043, SE-750 07 Uppsala, Sweden, E-mail: johannes.forkman@slu.se
<http://orcid.org/0000-0002-5796-0710>

models include several variance components. It is often practical to express the estimates of the variances as coefficients of variation, i. e., as standard deviations divided by the mean.

The intra-subject coefficient of variation assesses the reproducibility or reliability of a measurement [12]. In bioanalytical method validation, coefficients of variation specify within- and between-run precision [13]. In crop variety testing, expressing variability as coefficients of variation facilitates comparison of variability between crops [14].

Shoukri et al. [15] derived confidence intervals, based on a variance-stabilizing transformation, for the intra-subject coefficient of variation in the balanced one-way random-effects model. The present article derives generalized confidence intervals for intra- and inter-subject coefficients of variation in balanced and unbalanced random-effects and mixed-effects models with two variance components. Generalized confidence intervals are constructed using generalized pivotal quantities [16]. This type of confidence intervals has already been suggested for some other inferential problems on the coefficient of variation: Tian [17], Behboodian and Jafari [18] and Jafari [19] proposed generalized confidence intervals for a single coefficient of variation shared by many populations, and Krishnamoorthy and Lee [20] introduced generalized confidence intervals for differences and ratios of two coefficients of variation. The generalized pivotal quantities approach may also be used for the problem of testing equality of coefficients of variation [18, 19].

In the balanced one-way random-effects model, no exact bounds for the inter-subject variance are available that are functions of complete sufficient statistics [21]. A method for construction of generalized pivotal quantities and generalized confidence intervals for variance components was developed by Chiang [22], Iyer and Mathew [23], Iyer and Patterson [24] and Hannig et al. [25]. Their method is here extended to the problem of computing generalized confidence intervals for coefficients of variation. Burdick et al. [26] presented generalized confidence intervals for functions of variance components in mixed-effects models with balanced datasets and random-effects models with balanced or unbalanced datasets. For mixed-effects models and unbalanced datasets, the approach of the present article is similar to the method that was proposed for ‘a more general model’ by Gamage et al. [27], who considered the problem of computing confidence intervals for best linear unbiased predictions. However, their method uses transformation of data, but the method proposed in the present article does not.

The objective of this article is to propose methods for computing generalized confidence intervals for intra- and inter-subject coefficients of variation in mixed-effects models with two variance components. Section 2 provides theory and proposes methods. Section 3 studies, through simulation, the performance of the proposed methods and compares with the method proposed by Shoukri et al. [15]. Section 4 gives two examples: a bioanalytical method validation exemplifies the use of a one-way random-effects model, and an agricultural field experiment exemplifies a mixed-effects split-plot model. R scripts for these examples are published online as supplementary material on the Journal’s web page.

2 Theory and methods

2.1 Generalized confidence intervals

Let $\mathbf{Y}$ denote a random vector with a distribution that is dependent on a vector $\boldsymbol{\xi}$ of parameters, and let $\mathbf{y}$ be a realization, i. e. an observation, of $\mathbf{Y}$. Assume that we are interested in a function θ of the parameters: $\theta = f(\boldsymbol{\xi})$. A pivotal quantity for θ is a function g of $\mathbf{Y}$ and θ if the distribution of $g(\mathbf{Y}, \theta)$ does not depend on $\boldsymbol{\xi}$ [28]. A generalized pivotal quantity for θ is a function h of $\mathbf{y}$, $\mathbf{Y}$ and $\boldsymbol{\xi}$ if i) the distribution of $h(\mathbf{y}, \mathbf{Y}, \boldsymbol{\xi})$, where $\mathbf{y}$ is fixed, does not depend on $\boldsymbol{\xi}$, and ii) $h(\mathbf{y}, \mathbf{Y}, \boldsymbol{\xi} | \mathbf{Y} = \mathbf{y})$, i. e., $h(\mathbf{y}, \mathbf{y}, \boldsymbol{\xi})$, is a function of $\mathbf{y}$ and θ only [16]. If specifically $h(\mathbf{y}, \mathbf{y}, \boldsymbol{\xi}) = \theta$, then the function h is a fiducial generalized pivotal quantity for θ [25]. A $(1 - \alpha)$ equal-tailed generalized confidence interval for θ can be obtained as the interval between the $\alpha/2$ th and the $(1 - \alpha/2)$ th quantiles of the distribution of $h(\mathbf{y}, \mathbf{Y}, \boldsymbol{\xi})$, which can be estimated through simulation.

2.2 Statistical model

Consider the linear mixed-effects model

$$\mathbf{y} = \mathbf{B}\boldsymbol{\beta} + \mathbf{C}\mathbf{u} + \mathbf{e} \quad (1)$$

where $\mathbf{y}$ is an N -vector of observations, $\mathbf{B}$ is a design matrix for fixed effects, $\boldsymbol{\beta}$ is a vector of fixed effects parameters, $\mathbf{C}$ is an $N \times a$ design matrix for random effects, $\mathbf{u}$ is an a -vector of realized random effects, and $\mathbf{e}$ is an N -vector of residual error effects. The vectors of random effects $\mathbf{u}$ and $\mathbf{e}$ are realizations of the random vectors $\mathbf{U} \sim N(0, \sigma_A^2 \mathbf{I}_a)$ and $\mathbf{E} \sim N(0, \sigma_E^2 \mathbf{I}_N)$, which are independent. Thus, the vector of observations $\mathbf{y}$ is a realization of the random vector $\mathbf{Y} = \mathbf{B}\boldsymbol{\beta} + \mathbf{C}\mathbf{U} + \mathbf{E}$, and $\mathbf{Y} \sim N(\mathbf{B}\boldsymbol{\beta}, \mathbf{V})$, where $\mathbf{V} = \mathbf{C}\mathbf{C}^T \sigma_A^2 + \mathbf{I}_N \sigma_E^2$. Let μ denote the expected value of the mean of the elements of $E(\mathbf{Y})$ when the dataset is balanced. The intra- and inter-subject coefficients of variation are defined as $\gamma_E = \sigma_E/\mu$ and $\gamma_A = \sigma_A/\mu$, respectively.

In a more general setting, the random variables of $\mathbf{U}$ need not be normally distributed. When these variables are independent and identically distributed, but not necessarily normal, model (1) is the semiparametric mixed-effects model for clustered data proposed by Tao et al. [29] When the distributions of the random effects and errors are not known, or when the linear form of the conditional expectation might be questioned, model (1), including the assumptions of normality, may be considered as a working model. Generally, σ_A^2 and σ_E^2 can be defined as the variance of the distribution of the random variables of $\mathbf{U}$ and $\mathbf{E}$.

2.3 Balanced datasets

Let n be the number of observations at each level of the random-effects factor, let X_A and X_E denote the random sum of squares associated with the random-effects factor and the residual errors, respectively, and let $\bar{Y}$ denote the random overall mean, i. e., $\bar{Y} = \mathbf{1}_N^T \mathbf{Y}/N$. Then

$$U_A = \frac{X_A}{n\sigma_A^2 + \sigma_E^2}, U_E = \frac{X_E}{\sigma_E^2} \quad (2)$$

are central chi-square distributed with numbers of degrees of freedom depending on the design, and

$$Z = \frac{(\bar{Y} - \mu)\sqrt{N}}{\sqrt{n\sigma_A^2 + \sigma_E^2}}, \quad (3)$$

where $\mu = E(\bar{Y})$, is standard normal distributed. Let x_A and x_E denote the observed sum of squares associated with the random-effects factor and the residual errors, respectively, and let $\bar{y}$ denote the observed overall mean, i. e., $\bar{y} = \mathbf{1}_N^T \mathbf{y}/N$. If, in (2) and (3), x_A , x_E and $\bar{y}$ are substituted for X_A , X_E and $\bar{Y}$, respectively, the following quantities are obtained:

$$\hat{\sigma}_E^2 = \frac{x_E}{U_E}, \hat{\sigma}_A^2 = \frac{x_A}{nU_A} - \frac{\hat{\sigma}_E^2}{n}, \quad (4)$$

$$\hat{\mu} = \bar{y} - Z \sqrt{\frac{n\hat{\sigma}_A^2 + \hat{\sigma}_E^2}{N}}. \quad (5)$$

The expressions in (4) and (5) are fiducial generalized pivotal quantities for σ_E^2 , σ_A^2 and μ . Indeed, since conditioned on $\mathbf{y}$, the statistics x_E , x_A and $\bar{y}$ are constants, the quantities in (4) and (5) do not depend on $\boldsymbol{\xi} = (\sigma_A^2, \sigma_E^2, \mu)$, and when $\mathbf{Y} = \mathbf{y}$, the random variables X_A , X_E and $\bar{Y}$ equals x_E , x_A and $\bar{y}$, respectively, which by (2) and (3) implies $\hat{\sigma}_E^2|(\mathbf{Y} = \mathbf{y}) = \sigma_E^2$, $\hat{\sigma}_A^2|(\mathbf{Y} = \mathbf{y}) = \sigma_A^2$ and $\hat{\mu}|(\mathbf{Y} = \mathbf{y}) = \mu$. Fiducial generalized pivotal quantities $\hat{\gamma}_E$ and $\hat{\gamma}_A$, for the intra- and inter-subject coefficients of variation, respectively, are easily obtained as

$$\hat{\gamma}_E = \frac{\sqrt{\hat{\sigma}_E^2}}{\hat{\mu}}, \hat{\gamma}_A = \frac{\sqrt{\max(0, \hat{\sigma}_A^2)}}{\hat{\mu}}. \quad (6)$$

Generalized $(1-\alpha)$ confidence intervals are obtained through simulation of generalized pivotal quantities. The confidence interval limits are the $\alpha/2$ th and the $(1-\alpha/2)$ th quantiles of the simulated values. Technically, negative confidence interval limits can be obtained for μ , σ_A^2 , γ_E and γ_A , although these parameters are known to be positive. Negative confidence interval limits can be replaced with zeros.

2.4 Unbalanced datasets

Let $\mathbf{A}^-$ denote the Moore-Penrose generalized inverse of $\mathbf{A}$, where $\mathbf{A}$ is any matrix. Let $\mathbf{D} = (\mathbf{B}\mathbf{C})$ and $X_E = \mathbf{Y}^T(\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T\mathbf{D})^- \mathbf{D}^T)\mathbf{Y}$. Since $(\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T\mathbf{D})^- \mathbf{D}^T)\mathbf{V}/\sigma_E^2$ is idempotent (see Appendix) and $(\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T\mathbf{D})^- \mathbf{D}^T)\mathbf{B} = \mathbf{0}$, the quadratic form $U_E = X_E/\sigma_E^2$ is central chi-square distributed with $d = \text{rank}(\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T\mathbf{D})^- \mathbf{D}^T)$ degrees of freedom [11, p. 467]. If the observed error sum of squares $x_E = \mathbf{y}^T(\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T\mathbf{D})^- \mathbf{D}^T)\mathbf{y}$ is substituted for X_E in the expression for U_E , a fiducial generalized pivotal quantity $\hat{\sigma}_E^2$ for σ_E^2 is obtained as $\hat{\sigma}_E^2 = x_E/U_E$.

Similarly, let $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{V}^{-1}\mathbf{B})^- \mathbf{B}^T\mathbf{V}^{-1}$ and $X_T = \mathbf{Y}^T\mathbf{P}\mathbf{Y}$. Since $\mathbf{P}\mathbf{V}$ is idempotent (see Appendix) and $\mathbf{P}\mathbf{B} = \mathbf{0}$, the quadratic form X_T is central chi-square distributed with $t = \text{rank}(\mathbf{P})$ degrees of freedom. Let $\hat{\sigma}_A^2$ be the v such that $\mathbf{y}^T\mathbf{P}_v\mathbf{y} - \mathbf{Y}^T\mathbf{P}\mathbf{Y} = 0$, where $\mathbf{P}_v = \mathbf{V}_v^{-1} - \mathbf{V}_v^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{V}_v^{-1}\mathbf{B})^- \mathbf{B}^T\mathbf{V}_v^{-1}$ with $\mathbf{V}_v = \mathbf{C}\mathbf{C}^T v + \mathbf{I}_N \hat{\sigma}_E^2$. Then $\hat{\sigma}_A^2$ is a fiducial generalized pivotal quantity for σ_A^2 . This is so, because $\hat{\sigma}_A^2$ is a function of $\mathbf{y}$, $\mathbf{Y}$ and $\boldsymbol{\xi} = (\sigma_A^2, \sigma_E^2, \mu)$, which equals σ_A^2 when $\mathbf{Y} = \mathbf{y}$, and the distribution of $\hat{\sigma}_A^2$, conditioned on $\mathbf{y}$, does not depend on $\boldsymbol{\xi}$. Indeed, the distribution of $\mathbf{y}^T\mathbf{P}_v\mathbf{y} - \mathbf{Y}^T\mathbf{P}\mathbf{Y}$ is the same as the distribution of $g(v) = \mathbf{y}^T\mathbf{P}_v\mathbf{y} - x_E/\hat{\sigma}_E^2 - U_A$, where $U_A \sim \chi_{t-d}^2$. Since the distribution of $g(v)$, conditional on $\mathbf{y}$, does not depend on $\boldsymbol{\xi}$, neither does the distribution of $\hat{\sigma}_A^2$, conditional on $\mathbf{y}$. Given random realizations of U_E and U_A , the fiducial generalized pivotal quantity $\hat{\sigma}_A^2$ can be computed through solving the equation $g(v) = 0$. This equation can be solved using the Newton-Raphson procedure, utilizing that the derivative of $g(v)$ is $g'(v) = -\mathbf{y}^T\mathbf{P}_v\mathbf{C}\mathbf{C}^T\mathbf{P}_v\mathbf{y}$ [30, p. 325].

The unbalanced dataset can be regarded as a subset of a corresponding balanced dataset with M observations, where $M > N$. Let $\mathbf{B}_0$ denote the fixed-effects design matrix for the balanced dataset. The population mean can then be defined as $\mu = \mathbf{1}_M^T \mathbf{B}_0 \boldsymbol{\beta} / M$. Consider the random variable $\bar{Y} = \mathbf{1}_M^T \mathbf{B}_0 \hat{\boldsymbol{\beta}} / M$, where $\hat{\boldsymbol{\beta}} = (\mathbf{B}^T \mathbf{V}^{-1} \mathbf{B})^- \mathbf{B}^T \mathbf{V}^{-1} \mathbf{Y}$ and $\mathbf{V}$ is known. Since $\bar{Y}$ is $N(\mu, \mathbf{1}_M^T \mathbf{B}_0 (\mathbf{B}^T \mathbf{V}^{-1} \mathbf{B})^- \mathbf{B}_0^T \mathbf{1}_M / M^2)$, we can write

$$\begin{aligned} \hat{\mu} &= \bar{y} - Z \sqrt{\mathbf{1}_M^T \mathbf{B}_0 (\mathbf{B}^T \hat{\mathbf{V}}^{-1} \mathbf{B})^- \mathbf{B}_0^T \mathbf{1}_M / M^2} \\ &\stackrel{d}{=} \bar{y} - \frac{(\bar{Y} - \mu) \sqrt{\mathbf{1}_M^T \mathbf{B}_0 (\mathbf{B}^T \hat{\mathbf{V}}^{-1} \mathbf{B})^- \mathbf{B}_0^T \mathbf{1}_M / M^2}}{\sqrt{\mathbf{1}_M^T \mathbf{B}_0 (\mathbf{B}^T \mathbf{V}^{-1} \mathbf{B})^- \mathbf{B}_0^T \mathbf{1}_M / M^2}}, \end{aligned}$$

where $\hat{\mathbf{V}} = \mathbf{C}\mathbf{C}^T \hat{\sigma}_A^2 + \mathbf{I}_N \hat{\sigma}_E^2$, $Z \sim N(0, 1)$. Since $\hat{\mathbf{V}} = \mathbf{V}$ when $\mathbf{y} = \mathbf{Y}$, the quantity $\hat{\mu}$ takes the value μ when $\mathbf{y} = \mathbf{Y}$. The distribution of $\hat{\mu}$, conditioned on $\bar{y}$, is not dependent on $\boldsymbol{\xi} = (\sigma_A^2, \sigma_E^2, \mu)$. Thus, $\hat{\mu}$ is a fiducial generalized pivotal quantity for μ . Given $\hat{\sigma}_E^2$, $\hat{\sigma}_A^2$ and $\hat{\mu}$, as specified for unbalanced datasets, fiducial generalized pivotal quantities for γ_E and γ_A can be obtained through the same eq. (6) as used for balanced datasets. In repeated sampling, $\mathbf{V}_v$ and $\hat{\mathbf{V}}$ can become computationally singular. For this reason, $\mathbf{V}_v^-$ and $\hat{\mathbf{V}}^-$ may be used instead of $\mathbf{V}_v^{-1}$ and $\hat{\mathbf{V}}^{-1}$.

3 Simulation studies

3.1 The one-way random-effects model

Consider the one-way random-effects model $y_{ij} = \mu + a_i + e_{ij}$, where y_{ij} is the j th observation at the i th level, μ is an intercept, a_i is an effect of the i th level, and e_{ij} is a residual error of the j th observation at the i th level, $i = 1, 2, \dots, a$; $j = 1, 2, \dots, n$. It is assumed that y_{ij} , a_i and e_{ij} are realized values of the random variables Y_{ij} , A_i and E_{ij} , respectively, where $A_i \sim N(0, \sigma_A^2)$, $E_{ij} \sim N(0, \sigma_E^2)$, and $Y_{ij} = \mu + A_i + E_{ij}$.

Following Shoukri et al. [15], observations were simulated with intercept $\mu = 10$ and $\gamma_E \in \{0.01, 0.02, 0.04\}$, which corresponds to intra-subject variances $\sigma_E^2 \in \{0.01, 0.04, 0.16\}$. Shoukri et al. [15] studied $n \in \{2, 3, 5\}$ and $\rho \in \{0.3, 0.4, 0.6, 0.7, 0.8\}$, where ρ is the intra-class correlation coefficient: $\rho = \sigma_A^2 / (\sigma_A^2 + \sigma_E^2)$. The present evaluation was confined to $n \in \{2, 5\}$ and $\rho \in \{0.3, 0.6, 0.8\}$. Finally, Shoukri et al. [15] investigated numbers of subjects $a \in \{12, 25, 50, 75\}$, but the present study was extended to include cases $a \in \{4, 8, 12, 25, 50, 75\}$. In order to evaluate the performance of the method for unbalanced data, the same cases were explored, but in each simulated dataset $\lfloor 0.125N \rfloor$ randomly selected observations were removed before analysis. The symbol $\lfloor \cdot \rfloor$ denotes the floor function, so that $\lfloor 0.125N \rfloor$ is the largest integer not greater than $0.125N$, where $N = an$. For each of the 54 cases, 10,000 balanced and 1000 unbalanced datasets were randomly generated. Since analyses of unbalanced datasets were more time consuming, less unbalanced than balanced datasets were simulated. For the 10,000 balanced datasets, 95 % confidence intervals for γ_E were computed using the method proposed by Shoukri et al. [15]. In addition, generalized 95 % confidence intervals were computed for γ_A and γ_E , each using 1000 random samples of the fiducial generalized pivotal quantities $\hat{\gamma}_E$ and $\hat{\gamma}_A$. Generalized 95 % confidence intervals for γ_E and γ_A were also computed for the 1000 unbalanced datasets, each based on 1000 random samples of the fiducial generalized pivotal quantities $\hat{\gamma}_E$ and $\hat{\gamma}_A$.

The Newton-Raphson procedure employed in the method for unbalanced datasets used $v = 0$ as starting value. The procedure was stopped when $|g(v)| < 10^{-10}$, $|v| > 10^{10}$, or the maximum number of iterations, 100, was reached. Results obtained at the maximum number of iterations were discarded.

Tables 1 and 2 present observed frequencies of confidence intervals covering the true values of γ_E and γ_A , for cases with $n = 2$ and $n = 5$, respectively. In the simulations of balanced datasets, the generalized 95 % confidence

Table 1: Estimated coverage of 95 % confidence intervals (C.I.) for the intra- and inter-subject coefficients of variation γ_E and γ_A in a one-way random-effects model when the number of observations per subject is $n = 2$. Random effects are normally distributed. Coverage depends on the number of subjects, a , the intra-subject coefficient of variation, γ_E , the intra-class correlation coefficient ρ , and whether the dataset is balanced or not.

a	γ_E	ρ	Shoukri 0.95 C.I.	Generalized 0.95 C.I.			
			Balanced	Balanced		Unbalanced	
			γ_E	γ_E	γ_A	γ_E	γ_A
4	0.01	0.3	0.906	0.948	0.961	0.949	0.947
		0.6	0.907	0.946	0.954	0.934	0.959
		0.8	0.909	0.954	0.957	0.940	0.950
	0.02	0.3	0.910	0.949	0.953	0.946	0.950
		0.6	0.906	0.949	0.955	0.940	0.958
		0.8	0.904	0.948	0.955	0.938	0.948
	0.04	0.3	0.908	0.948	0.955	0.944	0.950
		0.6	0.905	0.950	0.957	0.940	0.958
		0.8	0.906	0.949	0.952	0.936	0.954
8	0.01	0.3	0.933	0.949	0.953	0.949	0.945
		0.6	0.933	0.952	0.952	0.945	0.953

(continued)

Table 1: (continued)

a	γ_E	ρ	Shoukri 0.95 C.I.	Generalized 0.95 C.I.			
			Balanced	Balanced		Unbalanced	
			γ_E	γ_E	γ_A	γ_E	γ_A
12	0.02	0.8	0.928	0.953	0.949	0.956	0.944
		0.3	0.926	0.946	0.954	0.952	0.949
		0.6	0.927	0.946	0.951	0.945	0.964
		0.8	0.932	0.951	0.949	0.946	0.953
	0.04	0.3	0.929	0.949	0.952	0.946	0.943
		0.6	0.931	0.950	0.951	0.959	0.954
		0.8	0.929	0.949	0.951	0.940	0.957
	0.01	0.3	0.937	0.948	0.949	0.938	0.936
		0.6	0.938	0.952	0.948	0.947	0.943
		0.8	0.934	0.949	0.947	0.928	0.938
	0.02	0.3	0.932	0.947	0.949	0.939	0.953
		0.6	0.934	0.948	0.952	0.934	0.939
		0.8	0.935	0.948	0.948	0.935	0.948
	0.04	0.3	0.935	0.948	0.948	0.944	0.951
		0.6	0.933	0.948	0.951	0.946	0.947
		0.8	0.937	0.946	0.952	0.943	0.944
	Mean		0.924	0.949	0.952	0.943	0.949

Table 2: Estimated coverage of 95 % confidence intervals (C.I.) for the intra- and inter-subject coefficients of variation γ_E and γ_A in a one-way random-effects model when the number of observations per subject is $n = 5$. Random effects are normally distributed. Coverage depends on the number of subjects, a , the intra-subject coefficient of variation, γ_E , the intra-class correlation coefficient ρ , and whether the dataset is balanced or not.

a	γ_E	ρ	Shoukri 0.95 C.I.	Generalized 0.95 C.I.			
			Balanced	Balanced		Unbalanced	
			γ_E	γ_E	γ_A	γ_E	γ_A
4	0.01	0.3	0.940	0.949	0.947	0.954	0.952
		0.6	0.942	0.953	0.948	0.940	0.945
		0.8	0.936	0.945	0.952	0.952	0.953
	0.02	0.3	0.941	0.949	0.951	0.956	0.945
		0.6	0.943	0.952	0.945	0.945	0.950
		0.8	0.940	0.948	0.950	0.940	0.941
	0.04	0.3	0.944	0.957	0.950	0.957	0.938
		0.6	0.934	0.947	0.951	0.941	0.955
		0.8	0.940	0.958	0.953	0.954	0.946
	0.01	0.3	0.944	0.948	0.951	0.965	0.951
		0.6	0.947	0.949	0.952	0.949	0.943
		0.8	0.944	0.950	0.950	0.938	0.933
8	0.02	0.3	0.944	0.945	0.951	0.938	0.947
		0.6	0.944	0.947	0.948	0.945	0.952
		0.8	0.944	0.947	0.950	0.931	0.949
	0.04	0.3	0.940	0.946	0.950	0.948	0.950
		0.6	0.946	0.953	0.948	0.944	0.940
		0.8	0.947	0.952	0.947	0.944	0.950
	0.01	0.3	0.946	0.947	0.949	0.948	0.937
		0.6	0.945	0.947	0.950	0.942	0.939

(continued)

Table 2: (continued)

			Shoukri 0.95 C.I.		Generalized 0.95 C.I.		
			Balanced	Balanced		Unbalanced	
a	γ_E	ρ	γ_E	γ_E	γ_A	γ_E	γ_A
	0.02	0.8	0.946	0.948	0.949	0.949	0.945
		0.3	0.950	0.953	0.946	0.941	0.962
		0.6	0.946	0.948	0.951	0.952	0.956
		0.8	0.947	0.947	0.947	0.957	0.964
	0.04	0.3	0.951	0.951	0.949	0.942	0.951
		0.6	0.945	0.948	0.949	0.946	0.946
		0.8	0.945	0.946	0.947	0.947	0.955
		Mean		0.944	0.949	0.949	0.947

intervals showed coverage close to the nominal level 0.95, although it was noted that coverage of generalized 95 % confidence intervals for γ_A was slightly too large for small values of a when $n=2$ (Table 1). Also in simulations of unbalanced datasets, the generalized 95 % confidence intervals covered the true values in approximately 95 % of the cases. However, confidence intervals tended to be somewhat too narrow, especially for γ_E when $n=2$. Shoukri et al. [15] showed that their 95 % confidence intervals cover the true γ_E with approximately probability 0.95 when $a=12$ and $n=5$. When $a=12$ and $n=2$, the probability is lower than 0.95. The present study confirms their results, but reveals that coverage of the Shoukri et al. [15] 95 % confidence interval is too small when $a=4$ and $n=2$ (Table 1). In this situation, the generalized 95 % confidence interval introduced here performs better. Tables A1 and A2, published online as supplementary material on the Journal's web page, present the results for $a \in \{25, 50, 75\}$. In these cases, observed coverage was close to 0.95.

The semiparametric mixed-effects model for clustered data (Section 2.2) allows for other random-effects distributions than the normal. In a study of robustness, three other distributions were investigated: a logistic distribution, a gamma distribution with shape parameter 4, and a gamma distribution with shape parameter 16. These distributions differ with respect to skewness and kurtosis. For a random variable Y with expected value μ_Y and standard deviation σ_Y , skewness and kurtosis may be defined as $E(Y - \mu_Y)^3 / \sigma_Y^3$ and $E(Y - \mu_Y)^4 / \sigma_Y^4 - 3$, respectively. Using these definitions, skewness and kurtosis are 0 for the normal distribution. The shape of the density function of the logistic distribution is similar to the normal, because skewness is 0, but kurtosis is 1.2. Skewness and kurtosis of a gamma distribution with shape parameter α is $2/\sqrt{\alpha}$ and $6/\alpha$, respectively. A gamma distribution with shape parameter 4 is more skewed and has heavier tails than a gamma distribution with shape parameter 16.

Tables 3 and 4 shows estimated coverage when the random-effects distribution is non-normal, for $n=2$ and $n=5$, respectively. Each case was simulated 10,000 times. Computations of confidence intervals were carried out using the same methods as previously specified, based on the now incorrect assumption of normally distributed random effects of subjects. Generalized confidence intervals were computed using 1000 random samples of fiducial generalized pivotal quantities.

Comparisons of Tables 3 and 4 with Tables 1 and 2, respectively, shows that generalized confidence intervals for γ_E were robust against non-normally distributed effects of subjects. Also the Shoukri et al. [15] 95 % confidence intervals for γ_E were robust. The generalized 95 % confidence intervals for γ_E still outperformed the Shoukri et al. [15] 95 % confidence intervals for γ_E .

Coverage of 95 % generalized confidence intervals for γ_A decreased as a consequence of non-normality. However, estimated coverage was never below 0.90 in these cases. Specifically, coverage of 95 % generalized confidence intervals for γ_A increased with γ_E and decreased with the intra-class correlation coefficient ρ . Moreover, coverage of 95 % generalized confidence intervals for γ_A decreased with the number of subjects. Cases with larger number of subjects ($a=25, 50, 75$) were also investigated through simulation. Tables A3 and A4, published online as supplementary material on the Journal's web page, presents the results. Coverage did not continue to decrease dramatically with larger numbers of subjects, but stabilized at levels lower than 0.95.

Table 3: Estimated coverage of 95 % confidence intervals (C.I.) for the intra- and inter-subject coefficients of variation γ_E and γ_A in a one-way random-effects model when the number of observations per subject is $n = 2$. Between-subjects effects are logistic distributed, gamma distributed with shape parameter 16, and gamma distributed with shape parameter 4. Error effects are normally distributed. Coverage depends on the number of subjects, a , the intra-subject coefficient of variation, γ_E , and the intra-class correlation coefficient ρ .

			Shoukri 0.95 C.I.			Generalized 0.95 C.I.						
			Logistic	$\Gamma(16)$	$\Gamma(4)$	Logistic		$\Gamma(16)$		$\Gamma(4)$		
a	γ_E	ρ	γ_E	γ_E	γ_E	γ_E	A	γ_E	γ_A	γ_E	γ_A	
4	0.01	0.3	0.910	0.909	0.903	0.949	0.954	0.950	0.952	0.944	0.955	
		0.6	0.907	0.908	0.911	0.949	0.949	0.949	0.954	0.951	0.946	
		0.8	0.907	0.909	0.907	0.944	0.941	0.943	0.952	0.947	0.934	
	0.02	0.3	0.906	0.905	0.903	0.949	0.950	0.948	0.948	0.948	0.949	
		0.6	0.909	0.910	0.912	0.947	0.947	0.953	0.955	0.950	0.946	
		0.8	0.911	0.903	0.907	0.949	0.937	0.949	0.950	0.953	0.936	
	0.04	0.3	0.911	0.906	0.909	0.952	0.954	0.949	0.954	0.956	0.951	
		0.6	0.906	0.907	0.908	0.948	0.944	0.949	0.957	0.950	0.950	
		0.8	0.910	0.908	0.907	0.956	0.940	0.950	0.952	0.951	0.945	
	8	0.01	0.3	0.926	0.925	0.926	0.949	0.948	0.946	0.946	0.947	0.939
			0.6	0.927	0.930	0.930	0.947	0.937	0.948	0.950	0.948	0.930
			0.8	0.930	0.921	0.929	0.948	0.923	0.945	0.939	0.947	0.921
0.02		0.3	0.930	0.930	0.925	0.950	0.948	0.950	0.952	0.948	0.945	
		0.6	0.927	0.929	0.928	0.949	0.931	0.949	0.951	0.951	0.934	
		0.8	0.926	0.928	0.927	0.948	0.922	0.947	0.943	0.947	0.924	
0.04		0.3	0.925	0.930	0.927	0.948	0.943	0.950	0.951	0.946	0.949	
		0.6	0.928	0.931	0.926	0.947	0.933	0.949	0.947	0.950	0.933	
		0.8	0.929	0.931	0.929	0.948	0.921	0.949	0.945	0.949	0.927	
12	0.01	0.3	0.934	0.936	0.936	0.948	0.946	0.947	0.950	0.949	0.936	
		0.6	0.932	0.934	0.936	0.946	0.922	0.945	0.948	0.952	0.926	
		0.8	0.934	0.935	0.940	0.947	0.913	0.951	0.940	0.952	0.907	
	0.02	0.3	0.934	0.935	0.940	0.953	0.946	0.949	0.944	0.954	0.939	
		0.6	0.933	0.935	0.934	0.948	0.928	0.950	0.946	0.947	0.927	
		0.8	0.938	0.934	0.934	0.949	0.913	0.949	0.947	0.948	0.916	
	0.04	0.3	0.933	0.937	0.936	0.946	0.943	0.950	0.948	0.947	0.941	
		0.6	0.935	0.932	0.936	0.950	0.931	0.947	0.945	0.948	0.926	
		0.8	0.936	0.937	0.938	0.949	0.912	0.951	0.940	0.951	0.919	
	Mean			0.923	0.923	0.943	0.949	0.936	0.949	0.948	0.949	0.932

In the semiparametric mixed-effects model for clustered data proposed by Tao et al. [29], between-subjects effects might not be normally distributed, but error effects are. Robustness was also investigated for the reverse situation, i. e., with normally distributed between-subject effects and non-normally distributed error effects. In the supplementary material, Tables A5 and A6 show the results of this investigation, which was carried out in the same way as the other simulation studies. Coverage of 95 % generalized confidence intervals for γ_A was close to 0.95, but coverage of 95 % generalized confidence intervals for γ_E become too small, often approximately 0.90. The Shoukri et al. [15] 95 % confidence intervals for γ_E performed in a similar way. Generally, the 95 % generalized confidence intervals for γ_E performed slightly better than the Shoukri et al. [15] 95 % confidence intervals.

3.2 The mixed-effects split-plot model

A balanced split-plot experiment with main-plot factor A and subplot factor B can be analyzed [31] using the model $y_{ijk} = \phi + \kappa_i + \alpha_j + \beta_k + (\alpha\beta)_{jk} + a_{ij} + e_{ijk}$, where y_{ijk} is the observation from the subplot of the i th replicate

Table 4: Estimated coverage of 95 % confidence intervals (C.I.) for the intra- and inter-subject coefficients of variation γ_E and γ_A in a one-way random-effects model when the number of observations per subject is $n=5$. Between-subjects effects are logistic distributed, gamma distributed with shape parameter 16, and gamma distributed with shape parameter 4. Error effects are normally distributed. Coverage depends on the number of subjects, a , the intra-subject coefficient of variation, γ_E , and the intra-class correlation coefficient ρ .

a	γ_E	ρ	Shoukri 0.95 C.I.			Generalized 0.95 C.I.					
			Logistic	$\Gamma(16)$	$\Gamma(4)$	Logistic		$\Gamma(16)$		$\Gamma(4)$	
			γ_E	γ_E	γ_E	γ_E	γ_A	γ_E	γ_A	γ_E	γ_A
4	0.01	0.3	0.940	0.936	0.934	0.947	0.940	0.949	0.950	0.944	0.937
		0.6	0.942	0.937	0.939	0.946	0.932	0.950	0.948	0.947	0.927
		0.8	0.940	0.944	0.941	0.949	0.926	0.951	0.941	0.949	0.930
	0.02	0.3	0.940	0.943	0.938	0.951	0.940	0.949	0.946	0.947	0.935
		0.6	0.946	0.938	0.941	0.953	0.931	0.950	0.941	0.949	0.935
		0.8	0.934	0.941	0.937	0.950	0.926	0.951	0.942	0.949	0.931
	0.04	0.3	0.942	0.938	0.939	0.953	0.937	0.949	0.949	0.949	0.939
		0.6	0.938	0.940	0.940	0.950	0.930	0.952	0.947	0.954	0.934
		0.8	0.938	0.941	0.943	0.958	0.931	0.956	0.946	0.956	0.936
8	0.01	0.3	0.945	0.943	0.943	0.950	0.929	0.948	0.945	0.947	0.927
		0.6	0.945	0.941	0.943	0.949	0.918	0.943	0.941	0.947	0.917
		0.8	0.945	0.942	0.942	0.948	0.915	0.947	0.938	0.947	0.916
	0.02	0.3	0.943	0.942	0.948	0.947	0.931	0.945	0.942	0.954	0.929
		0.6	0.945	0.943	0.945	0.950	0.921	0.947	0.937	0.949	0.917
		0.8	0.947	0.946	0.942	0.950	0.915	0.950	0.940	0.949	0.912
	0.04	0.3	0.945	0.943	0.945	0.950	0.930	0.946	0.943	0.950	0.934
		0.6	0.940	0.946	0.946	0.947	0.918	0.951	0.945	0.951	0.922
		0.8	0.944	0.942	0.944	0.949	0.911	0.946	0.946	0.952	0.922
12	0.01	0.3	0.945	0.943	0.945	0.948	0.932	0.946	0.941	0.951	0.918
		0.6	0.944	0.944	0.949	0.946	0.909	0.946	0.938	0.951	0.905
		0.8	0.951	0.950	0.951	0.951	0.903	0.949	0.936	0.950	0.900
	0.02	0.3	0.944	0.951	0.948	0.945	0.924	0.952	0.945	0.951	0.924
		0.6	0.945	0.949	0.949	0.947	0.913	0.952	0.939	0.952	0.908
		0.8	0.949	0.950	0.944	0.949	0.904	0.950	0.941	0.946	0.907
	0.04	0.3	0.945	0.943	0.949	0.945	0.927	0.947	0.942	0.953	0.927
		0.6	0.950	0.947	0.949	0.953	0.916	0.948	0.943	0.949	0.912
		0.8	0.941	0.947	0.950	0.947	0.910	0.951	0.943	0.954	0.915
Mean			0.943	0.943	0.944	0.949	0.923	0.949	0.943	0.950	0.923

that was treated with the j th level of A and the k th level of B , ϕ is an intercept, κ_i is an effect of the i th replicate, α_j is an effect of the j th level of A , β_k is an effect of the k th level of B , $(\alpha\beta)_{jk}$ is an interaction effect of the j th level of A with the k th level of B , a_{ij} is an effect of the j th main plot in the i th replicate, and e_{ijk} is a residual error, $i = 1, 2, \dots, r$; $j = 1, 2, \dots, a$; $k = 1, 2, \dots, n$. It will be assumed that y_{ijk} , a_{ij} and e_{ijk} are realized values of the random variables Y_{ijk} , A_{ij} and E_{ijk} , respectively, where $A_{ij} \sim N(0, \sigma_A^2)$, $E_{ijk} \sim N(0, \sigma_E^2)$, and $Y_{ijk} = \phi + \kappa_i + \alpha_j + \beta_k + (\alpha\beta)_{jk} + A_{ij} + E_{ijk}$.

A simulation study was performed based on the agricultural split-plot example of Section 4.2. Based on estimates obtained using the lme function of the R package nlme, (γ_A, γ_E) was set to (0.006, 0.062), ϕ was set to 77.044, (ρ_2, ρ_3, ρ_4) was set to (0.344, -3.075, -3.944), $(\alpha_2, \alpha_3, \alpha_4)$ was set to (-3.075, -8.050, -4.825), $(\beta_2, \beta_3, \beta_4)$ was set to (-23.275, -3.150, -1.900, 2.325), and $((\alpha\beta)_{22}, (\alpha\beta)_{32}, (\alpha\beta)_{42}, (\alpha\beta)_{23}, (\alpha\beta)_{33}, (\alpha\beta)_{43}, (\alpha\beta)_{24}, (\alpha\beta)_{34}, (\alpha\beta)_{44})$ was set to (2.325, 6.400, 7.975, -1.150, 7.075, 3.950, -5.800, 4.925, 3.700). All other fixed-effects parameters were set to 0. These parameter setting was used for random generation of 10,000 balanced datasets. In addition, 10,000 unbalanced datasets were randomly generated. This was achieved by random selection of five observations that were removed before analysis. A new random selection of five

missing values was made for each of the 10,000 datasets. In the Newton-Raphson procedure, the same starting value and stopping rules were used as for the simulation study of the one-way random-effects model (Section 3.1). For each generated dataset, generalized 95 % confidence intervals were computed using 1000 random samples of the pivotal quantities $\hat{\gamma}_E$ and $\hat{\gamma}_A$.

For the balanced datasets, estimated coverage (i. e., the observed frequency of 95 % confidence intervals covering the true parameter value) was 0.943 and 0.949 for γ_E and γ_A , respectively. For unbalanced datasets, coverage was estimated to 0.950 and 0.942 for γ_E and γ_A , respectively.

4 Examples

4.1 A bioanalytical method validation using a one-way random-effects model

Following the European Medicines Agency's (EMA) guideline on bioanalytical method validation [13], ligand-binding assays or immunoassays should be studied using quality control (QC) samples assayed in at least six runs over several days. Table 5 includes measurements of concentration of a QC sample in a precision study.

Table 5: Measurements of a QC sample.

Run	Concentration (mg/l)	Run	Concentration (mg/l)
1	119	6	109
1	113*	6	103
2	103	7	103
2	91*	7	100
3	118	8	108
3	106*	8	101
4	93	9	98
4	93*	9	94
5	109	10	105
5	110*	10	106

* Removed for the unbalanced example.

Generalized 95 % confidence intervals were computed based on 10,000 random samples, each comprising one observation from each of the random variables U_A , U_E and Z . For each sample, the fiducial generalized pivotal quantities $\hat{\sigma}_E^2$, $\hat{\sigma}_A^2$, $\hat{\mu}$, $\hat{\gamma}_E$ and $\hat{\gamma}_A$ were computed using (4)–(6). The lower 2.5th and the upper 97.5th percentiles of the obtained empirical distributions of these fiducial generalized pivotal quantities are the lower and the upper limits of the generalized 95 % confidence intervals, respectively. The first row of Table 6 reports the obtained limits, i. e., the resulting generalized 95 % confidence intervals for σ_E^2 , σ_A^2 , μ , γ_E and γ_A . Figure 1 is an R script (www.r-project.org) for this analysis.

Five observations were removed (Table 5). In this case, the procedure described in Section 2.4 yielded generalized 95 % confidence intervals as presented on the second row of Table 6.

4.2 An agricultural split-plot experiment analyzed using a mixed-effects model

Federer and King [32] presented an example of a split-plot agricultural field experiment with $r=4$ replicates investigating effects of $a=4$ seedbed preparations and $n=4$ planting methods on maize yield (Table 7).

Table 6: Generalized 95 % confidence intervals. Each example used 10,000 random samples.

	σ_E^2	σ_A^2	μ	γ_E	γ_A
One-way					
Balanced	(10.6, 66.1)	(7.70, 166)	(98.8, 109)	(0.031, 0.078)	(0.027, 0.124)
Unbalanced	(4.37, 65.9)	(3.40, 222)	(99.5, 112)	(0.020, 0.078)	(0.018, 0.142)
Split-plot					
Balanced	(11.3, 28.8)	(-3.51, 10.4)	(64.9, 67.3)	(0.051, 0.081)	(0.000, 0.049)
Unbalanced	(8.53, 23.7)	(-1.88, 17.2)	(64.6, 67.8)	(0.044, 0.074)	(0.000, 0.063)

```

Y <- c(119, 113, 103, 91, 118, 106, 93, 93, 109, 110,
      109, 103, 103, 100, 108, 101, 98, 94, 105, 106)
a <- 10 ; n <- 2; A <- factor(rep(1:a, each = n))
### Sums of squares
x_A <- anova(lm(Y ~ A))$"Sum Sq"[1]
x_E <- anova(lm(Y ~ A))$"Sum Sq"[2]
### Sampling
NSample <- 10000
U_E <- rchisq(NSample, df = a*n - a)
U_A <- rchisq(NSample, df = a - 1)
Z <- rnorm(NSample)
### Pivotal quantities
G.var_E <- x_E/U_E
G.var_A <- x_A/(n*U_A) - G.var_E/n
G.mu <- mean(Y) - Z*sqrt(G.var_A/a + G.var_E/(a*n))
G.CV_A <- sqrt(ifelse(G.var_A > 0, G.var_A, 0))/G.mu
G.CV_E <- sqrt(G.var_E)/G.mu
### Generalized confidence intervals
quantile(G.var_E, probs = c(0.025, 0.975))
quantile(G.var_A, probs = c(0.025, 0.975))
quantile(G.mu, probs = c(0.025, 0.975))
quantile(G.CV_E, probs = c(0.025, 0.975))
quantile(G.CV_A, probs = c(0.025, 0.975))

```

Figure 1: An R script for the balanced bioanalytical validation example.

The proposed method for computing generalized 95% confidence intervals was applied. The fiducial generalized pivotal quantities $\hat{\sigma}_E^2$, $\hat{\sigma}_A^2$, $\hat{\mu}$, $\hat{\gamma}_E$ and $\hat{\gamma}_A$, as specified in (4)–(6), were generated 10,000 times. The 2.5th percentiles of the generated distributions were used as lower generalized 95 % confidence limits, and the 97.5th percentiles were used as upper generalized 95 % confidence limits. The third line of Table 6 presents the obtained generalized 95 % confidence intervals.

For an example of an unbalanced dataset, five randomly selected observations were removed from the dataset (Table 7). Based on generation of 10,000 pivotal quantities $\hat{\gamma}_A$ and $\hat{\gamma}_E$, the method for unbalanced datasets gave generalized 95 % confidence intervals as presented on the fourth row of Table 6.

5 Discussion

This article contributed the idea of using generalized inference methodology for computation of confidence intervals for intra- and inter-subject coefficients of variation in mixed-effects models. A method was developed for unbalanced datasets in mixed-effects models. Burdick et al. [26] studied other functions of variance components than the intra- and inter-subject coefficients of variation. Their methods for unbalanced datasets apply to random-effects models only.

Table 7: Observed yields (bushels per acre) of maize^a [32].

Replicate	Planting method			
	1	2	3	4
Seedbed preparation 1				
1	81.1 ^b	46.2*	78.6	77.7
2	72.2	51.6	70.9	73.6
3	72.9	53.6*	69.8	70.3
4	74.6	57.0*	69.6	72.3
Seedbed preparation 2				
1	74.1	49.1	72.0	66.1
2	76.2	53.8	71.8	65.5
3	71.1	43.7	67.6	66.2
4	67.8*	58.8	60.6	60.6
Seedbed preparation 3				
1	68.4	54.5	72.0	70.6
2	68.2	47.6	76.7	75.4
3	67.1	46.4	70.7	66.2
4	65.6	53.3	65.6*	69.2
Seedbed preparation 4				
1	71.5	50.9	76.4	75.1
2	70.4	65.0	75.8	75.8
3	72.5	54.9	67.6	75.2
4	67.8	50.2	65.6	63.3

a Copyright 2007 by John Wiley & Sons, Inc. All rights reserved.

b This number differs from the source, since it was noted that [32] actually used this number in their analysis.

* Removed for the unbalanced example.

Simulation studies indicated good performance in terms of coverage at the 95 % confidence level, although these studies also verified that the proposed methods are approximate. Since deviations from the nominal confidence level 0.95 were generally not large, the proposed methods can be recommended for practical purposes. However, the proposed methods are sensitive to the assumptions of normal distributions. The simulation studies revealed that the 95 % confidence interval proposed by Shoukri et al. [15] for the intra-subject coefficient of variation in the balanced one-way model tends to be too narrow in small datasets. For such datasets, the generalized 95 % confidence intervals proposed in the present article should be preferred.

The postulated linear mixed-effects model (1) is a fully parametric model, which assumes normal distributions. However in practice, random effects and errors are rarely exactly normally distributed. As noted in Section 2.2, model (1) can be viewed as a special case of the more realistic semiparametric mixed-effects model that was proposed by Tao et al. [29]. For this case, when the distribution of the random effects is unknown, model (1), which assumes the normal distribution, was proposed as a working model. This approach is not unproblematic, since the performance of the method depends on how well the assumption of normality is met. In practice, the proposed generalized confidence intervals may, in the non-normal case, have coverage somewhat other than the nominal. Still, the simulation studies indicate that the proposed generalized confidence intervals for γ_E give better coverage than the Shoukri et al. [15] confidence intervals for γ_E , also when the assumption of normality is not fulfilled. Furthermore, the present article introduced a method for computing confidence intervals not only for the intra-subject coefficient of variation γ_E , but also for the inter-subject coefficient of variation γ_A .

The generalization to more than two variance components is straightforward for balanced datasets. Every variance component has a unique unbiased estimator based on sums of squares [33]. Thus, a

generalized pivotal quantity can be specified for any function of the parameters of a balanced mixed-effects model [24]. The generalization to mixed-effects models with several variance components is more difficult for unbalanced datasets. Although extensions have been made to the special case of unequal cell frequencies in the last stage [27, 33, 34, 35], unbalanced cases are generally less explored in the literature on generalized inference. It is possible that the method presented here can be generalized to more complex mixed-effects models, but since the proposed method involves numerical computations, it might be a challenge to provide an efficient general algorithm.

As pointed out by Hannig et al. [25], when h is a fiducial generalized pivotal quantity and θ is a scalar, then $T(\mathbf{y}, \mathbf{Y}, \boldsymbol{\xi}) = \theta - h(\mathbf{y}, \mathbf{Y}, \boldsymbol{\xi})$ is a generalized test variable as defined by Tsui and Weerahandi [36] for testing the null hypothesis $H_0: \theta \leq \theta_0$ against the alternative $H_1: \theta > \theta_0$. The methodology of the present article can for this reason also be used for hypothesis testing.

Acknowledgement: This research was conducted using the resources of High Performance Computing Center North (HPC2N).

Appendix

$$\begin{aligned} \{(\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T) \mathbf{V} / \sigma_E^2\}^2 &= (\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T) (\mathbf{C} \mathbf{C}^T \rho + \mathbf{I}_N) (\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T) (\mathbf{C} \mathbf{C}^T \rho + \mathbf{I}_N) = \\ &= (\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T) (\mathbf{C} \mathbf{C}^T \rho - \mathbf{C} \mathbf{C}^T \mathbf{D}(\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \rho + \mathbf{I}_N - \mathbf{D}(\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T) (\mathbf{C} \mathbf{C}^T \rho + \mathbf{I}_N) = (\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T) \\ &= (\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T) (\mathbf{C} \mathbf{C}^T \rho + \mathbf{I}_N) = (\mathbf{I}_N - \mathbf{D}(\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T) \mathbf{V} / \sigma_E^2 \\ &\text{where } \rho = \sigma_A^2 / \sigma_E^2. \end{aligned}$$

$$\begin{aligned} (\mathbf{P} \mathbf{V})^2 &= \mathbf{V}^{-1} (\mathbf{I}_N - \mathbf{B}(\mathbf{B}^T \mathbf{V}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{V}^{-1}) (\mathbf{I}_N - \mathbf{B}(\mathbf{B}^T \mathbf{V}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{V}^{-1}) \mathbf{V} = \\ &= \mathbf{V}^{-1} (\mathbf{I}_N - 2\mathbf{B}(\mathbf{B}^T \mathbf{V}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{V}^{-1} + \mathbf{B}(\mathbf{B}^T \mathbf{V}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{V}^{-1} \mathbf{B}(\mathbf{B}^T \mathbf{V}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{V}^{-1}) \mathbf{V} = \\ &= \mathbf{V}^{-1} (\mathbf{I}_N - 2\mathbf{B}(\mathbf{B}^T \mathbf{V}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{V}^{-1} + \mathbf{B}(\mathbf{B}^T \mathbf{V}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{V}^{-1}) \mathbf{V} = \mathbf{P} \mathbf{V} \end{aligned}$$

References

- Owen DB. A survey of properties and applications of the noncentral t-distribution. *Technometrics* 1968;10:445–478.
- Miller EG, Feltz CJ. Asymptotic inference for coefficients of variation. *Commun Stat Theory Methods* 1997;26:715–726.
- Mahmoudvand R, Hassani H. Two new confidence intervals for the coefficient of variation in a normal distribution. *J Appl Stat* 2009;36:429–442.
- Vangel MG. Confidence intervals for a normal coefficient of variation. *Am Stat* 1996;50:21–26.
- Forkman J. Estimator and tests for common coefficients of variation in normal distributions. *Commun Stat Theory Methods* 2009;38:233–251.
- McKay AT. Distribution of the coefficient of variation and the extended “t” distribution. *J R Stat Soc* 1932;95:695–698.
- Forkman J, Verrill S. The distribution of McKay’s approximation for the coefficient of variation. *Stat Probab Lett* 2008;78:10–14.
- Verrill S, Johnson RA. Confidence bounds and hypothesis tests for normal distribution coefficients of variation. *Commun Stat Theory Methods* 2007;36:2187–2206.
- Hayter AJ. Confidence bounds on the coefficient of variation of a normal distribution with applications to win-probabilities. *J Stat Comput Simul* 2015;85:3778–3791.
- Wong AC, Wu J. Small sample asymptotic inference for the coefficient of variation: normal and nonnormal models. *J Stat Plan Inference* 2002;104:73–82.
- Searle SR, Casella G, McCulloch CE. *Variance components*. Hoboken: Wiley, 2006.
- Quan H, Shih WJ. Assessing reproducibility by the within-subject coefficient of variation with random effects models. *Biometrics* 1996;52:1195–1203.
- European Medicines Agency. *Guideline on bioanalytical method validation*. London: The European Medicines Agency, 2011. Available at: <http://www.ema.europa.eu/ema/>.

14. Laidig F, Drobek T, Meyer U. Genotypic and environmental variability of yield for cultivars from 30 different crops in German official variety trials. *Plant Breed* 2008;127:541–547.
15. Shoukri MM, Elukum N, Walter SD. Interval estimation and optimal design for the within-subject coefficient of variation for continuous and binary variables. *BMC Med Res Method* 2006;6:1–10.
16. Weerahandi S. Generalized confidence intervals. *J Am Stat Assoc* 1993;88:899–905.
17. Tian LL. Inferences on the common coefficient of variation. *Stat Med* 2005;24:2213–2220.
18. Behboodian J, Jafari AA. Generalized confidence interval for the common coefficient of variation. *J Stat Theory Appl* 2008;7:349–363.
19. Jafari AA. Inferences on the coefficients of variation in a multivariate normal population. *Commun Stat Theory Methods* 2015;44:2630–2643.
20. Krishnamoorthy K, Lee M. Improved tests for the equality of normal coefficients of variation. *Comput Stat* 2014;29:215–232.
21. Burdick RK, Quiroz J, Iyer HK. The present status of confidence interval estimation for one-factor random models. *J Stat Plan Inference* 2006;136:4307–4325.
22. Chiang AK. A simple general method for constructing confidence intervals for functions of variance components. *Technometrics* 2001;43:356–367.
23. Iyer H, Mathew T. Comments on Chiang (2001). *Technometrics* 2002;44:284–285.
24. Iyer HK, Patterson PL. A recipe for constructing generalized pivotal quantities and generalized confidence intervals. Colorado State University, Dept. of Statistics, 2002. Report No.: 2010/10.
25. Hannig J, Iyer H, Patterson P. Fiducial generalized confidence intervals. *J Am Stat Assoc* 2006;101:254–269.
26. Burdick RK, Borror CM, Montgomery DC. Design and analysis of gauge R&R studies : making decisions with confidence intervals in random and mixed ANOVA models. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2005.
27. Gamage J, Mathew T, Weerahandi S. Generalized prediction intervals for BLUPs in mixed models. *J Multivar Anal* 2013;120:226–233.
28. Shao J. *Mathematical statistics*, 2nd ed. New York: Springer, 2003.
29. Tao NG, Palta M, Yandell BS, Newton MA. An estimation method for the semiparametric mixed effects model. *Biometrics* 1999;55:102–110.
30. McCulloch CE, Searle SR, Neuhaus JM. Generalized, linear, and mixed models [Internet], 2nd ed. Hoboken, NJ: Wiley, 2008. Available at: <http://www.loc.gov/catdir/enhancements/fy0827/2008002724-d.html>. Cited: 19 Aug 2015.
31. Piepho HP, Büchse A, Emrich K. A hitchhiker's guide to mixed models for randomized experiments. *J Agron Crop Sci* 2003;189:310–322.
32. Federer WT, King F. Variations on split plot and split block experiment designs. Hoboken: Wiley, 2007.
33. Khuri AI, Mathew T, Sinha BK. Statistical tests for mixed linear models. New York: Wiley, 1998.
34. Zhou L, Mathew T. Some tests for variance components using generalized p values. *Technometrics* 1994;36:394–402.
35. Li XM, Wang J. A simple method for testing variance components in unbalanced nested model. *Commun Stat Simul Comput* 2011;40:1310–1323.
36. Tsui KW, Weerahandi S. Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. *J Am Stat Assoc* 1989;84:602–607.

Supplemental Material: The online version of this article (DOI: 10.1515/ijb-2016-0093) offers supplementary material, available to authorized users.