University of Groningen

Confidence intervals for intraclass correlation coefficients in variance components models

Demetrashvili, Nino; Wit, Ernst C.; van den Heuvel, Edwin R.

Published in:
Statistical Methods in Medical Research

DOI:
[10.1177/0962280214522787](https://doi.org/10.1177/0962280214522787)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Demetrashvili, N., Wit, E. C., & van den Heuvel, E. R. (2016). Confidence intervals for intraclass correlation coefficients in variance components models. *Statistical Methods in Medical Research*, 25(5), 2359-2376. <https://doi.org/10.1177/0962280214522787>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Confidence Intervals for Intraclass Correlation Coefficients in a Nonlinear Dose–Response Meta-Analysis

Nino Demetrashvili* and Edwin R. Van den Heuvel

Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, The Netherlands

*email: n.demetrashvili@umcg.nl

SUMMARY. This work is motivated by a meta-analysis case study on antipsychotic medications. The Michaelis–Menten curve is employed to model the nonlinear relationship between the dose and D_2 receptor occupancy across multiple studies. An intraclass correlation coefficient (ICC) is used to quantify the heterogeneity across studies. To interpret the size of heterogeneity, an accurate estimate of ICC and its confidence interval is required. The goal is to apply a recently proposed generic beta-approach for construction the confidence intervals on ICCs for linear mixed effects models to nonlinear mixed effects models using four estimation methods. These estimation methods are the maximum likelihood, second-order generalized estimating equations and two two-step procedures. The beta-approach is compared with a large sample normal approximation (delta method) and bootstrapping. The confidence intervals based on the delta method and the nonparametric percentile bootstrap with various resampling strategies failed in our settings. The beta-approach demonstrates good coverages with both two-step estimation methods and consequently, it is recommended for the computation of confidence interval for ICCs in nonlinear mixed effects models for small studies.

KEY WORDS: Beta distribution; Maximum likelihood; Nonlinear regression; Restricted profile likelihood; Variance component.

1. Nonlinear Dose Response Meta-Analysis Case Study

Dopamine D_2 receptor is linked to psychopathology of many diseases including schizophrenia. A dose–response relationship on the occupancy of the D_2 receptor can help clinicians to select an appropriate dose for schizophrenic patients. Studies suggest that effective relief of psychotic symptoms is associated with the blockade of at least 65% of the striatal dopamine D_2 receptor (Lako et al., 2013). Furthermore, the dose–response curves for different antipsychotic medications make it possible to discuss dose equivalents between these medications. Therefore, an individual participant data meta-analysis was conducted on eight antipsychotic medications by Lako et al. (2013) to study the relationship between the prescribed dose and the response. Critical to the evaluation of a meta-analysis is the variation around the curves and whether this variation is due to factors within studies (e.g., patient characteristics) or factors between studies (e.g., measurement systems). An intraclass correlation coefficient (ICC) can determine which source is dominating, but only if we would have a narrow and appropriate confidence interval on the ICC. Therefore, our article focuses on estimating the ICC, with accompanying confidence intervals.

The response O_{ij} in the meta-analysis is dopamine occupancy on patient $j = 1, \dots, n_i$ in study $i = 1, \dots, I$. The total number of observations is $n = \sum_{i=1}^I n_i$. The occupancy O_{ij} of a patient is expressed as percentage and it is calculated with respect to an average occupancy of a healthy control group (Lako et al., 2013).

We transformed the occupancy by $y_{ij} = \ln(1 - O_{ij})$ for every antipsychotic separately. The transformation makes the measurement of a patient and the average measurements on the healthy control group additive, as described by Lako et al. (2013). Assuming Michaelis–Menten curve (Michaelis and Menten, 1913) for the relationship between the dose and original occupancy together with our proposed transformation leads to a nonlinear mixed effects model of the form:

$$y_{ij} = \mu(x_{ij}; \beta) + u_i + \epsilon_{ij}, \quad (1)$$

with mean vector $\mu(\mathbf{x}; \beta) = (\mu(x_{i1}; \beta), \dots, \mu(x_{in_i}; \beta))$ given by:

$$\mu(x; \beta) = \ln \left(\frac{\beta_2 + (1 - \beta_1)x}{\beta_2 + x} \right), \quad (2)$$

where $x_{ij} \geq 0$ is the administered dose, $\beta = (\beta_1, \beta_2)$ unknown parameters with $\beta_1 \in (0, 1]$ the maximum occupancy for the population of patients (E_{max}) and $\beta_2 > 0$ the dose that is associated with 50% occupancy (EC_{50}); u_i is the additive random effect of unexplained between-study variation, and ϵ_{ij} is the unexplained within-study (interindividual) variation. The u_i and ϵ_{ij} are assumed mutually independent and have distributions $u_i \stackrel{iid}{\sim} N(0, \sigma_s^2) \forall i$ and $\epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_p^2) \forall i, j$, where σ_s^2 and σ_p^2 are the variance components for the study and patient, respectively. More details on the interpretation of E_{max} and EC_{50} can be found for example in Goutelle et al. (2008).

Table 1*Descriptive statistics on the sample size for the meta-analysis of eight antipsychotic medications*

Medication	Total number of observations	Number of studies	Study size (min–max)	Median study size
Haloperidol	90	16	1–12	6.5
Risperidone	138	14	3–20	8.5
Olanzapine	136	11	3–22	10
Clozapine	114	17	1–16	5
Quetiapine	57	8	2–16	5.5
Aripiprazole	16	1	16	16
Ziprasidone	25	2	10–15	12.5
Amisulpride	62	5	5–29	9

The ICC is the correlation coefficient on repeated measurements (individuals in the same study i), that is, $\text{ICC} = \text{corr}(Y_{ij1}, Y_{ij2})$. It can be interpreted as the proportion of variation unrelated to individuals in the total amount of variation, as shown below:

$$\text{ICC} = \frac{\sigma_S^2}{\sigma_S^2 + \sigma_P^2}. \quad (3)$$

Our data on eight antipsychotic medications consisted of 74 studies with 638 patients in total. Each study contributed a different number of patients. A detailed overview of the sample sizes are listed in Table 1.

We fitted model (1) with relationship (2) to the occupancy data per medication separately. Estimates of σ_S^2 , σ_P^2 , ICC, and standard errors for the estimated ICC (see Section 3) are shown for three medications in Table 2. The results are produced using four estimation procedures: (1) maximum likelihood (ML), (2) second-order generalized estimating equations (GEE2), (3) a two-step iterative estimation procedure which is based on restricted profile likelihood and implemented in macro %NLINMIX by Vonesh (2012); (4) a two-step estimation procedure which is based on a subsequent application of the nonlinear least square (NLS) and the restricted maximum likelihood (REML) estimation methods. The third and fourth method produce the estimates of regression parameters β and variance components $\theta = (\sigma_S^2, \sigma_P^2)$ in separate steps,

while the first two approaches produce estimates of all parameters $\gamma = (\beta, \theta)$ jointly. More details on estimation is given in the Section 3. For illustrative purposes, we present the fitted Michaelis–Menten curves in Figures 1 and 2 for haloperidol and risperidone using the fourth estimation procedure. All results were produced with SAS (version 9.2).

2. Background and Relevance

Nonlinear random and mixed effects models have been employed in the literature to describe the dose–response relationship in a meta-analysis, see Rota et al. (2010), Aune et al. (2011), Rui et al. (2012), Qu et al. (2013). It is crucial to quantify the heterogeneity between studies and possibly investigate its source. Assessing the heterogeneity is important in order to interpret the results of the overall dose–response relationship appropriately. The typical measure that is used for nonlinear dose–response meta-analysis is $I^2 = \sigma_S^2 / (\sigma_S^2 + \sigma_P^2)$ (Higgins and Thompson, 2002) which coincides with our ICC in (3). Recently, two generic approaches for constructing the confidence intervals for ICCs of the form (3) have been published by Demetrashvili, Wit, and Van den Heuvel (2014). These confidence intervals for ICCs have been constructed for balanced and unbalanced linear random and mixed effects models which are estimated with REML. To the best of our knowledge, there is no such approach available for constructing the confidence intervals for ICCs in nonlinear mixed effects models. The only suitable methods that can be given considera-

Table 2*Estimates of variance components, ICC, and standard error of ICC for model (1)*

Antipsychotic	Estimation	$\hat{\sigma}_S^2$	$\hat{\sigma}_P^2$	ICC	$\hat{\tau}_{\text{ICC}}$
Haloperidol	ML	0.34419	0.16235	0.67949	0.10812
	GEE2	0.34419	0.16235	0.67949	0.18496
	NLS&REML	0.36572	0.16298	0.69173	0.10555
	NLS&REML (iterative)	0.38068	0.16376	0.69921	0.10455
Clozapine	ML	0.01138	0.06896	0.14165	0.09798
	GEE2	0.01138	0.06896	0.14165	0.11036
	NLS&REML	0.01308	0.06889	0.15957	0.10296
	NLS&REML (iterative)	0.01327	0.06947	0.16038	0.10320
Amisulpride	ML	0.09354	0.14230	0.39663	0.18507
	GEE2	0.09354	0.14230	0.39663	0.13282
	NLS&REML	0.11899	0.14178	0.45630	0.20518
	NLS&REML (iterative)	0.11648	0.14521	0.44511	0.20008

Figure 1. Michaelis–Menten curve for haloperidol.

tion are the bootstrap (Efron, 1979) and large sample normal approximation (delta method) which uses the first-order Taylor series approximation. The goal of this work is to apply the generic beta-approach of Demetrashvili et al. (2014) to nonlinear mixed effects models of the form (1) using different estimation methods and compare it with the delta method and bootstrap. Application of several estimation methods was motivated by the fact that the REML (which is known to be less biased than ML in linear mixed effects models) is not well defined in nonlinear mixed effects models. The accuracy of the beta-approach depends on accuracy of both, the estimated variance components and their variance-covariances. Our focus is on settings with small sample sizes (e.g., ≤ 10 studies, patients). In Section 3, we describe the estimation procedures and briefly outline the beta-approach along with the generic alternative approaches: the delta method and the bootstrap. In Section 4, we present the simulation study results. In Section 5, we provide the results for our case study.

Figure 2. Michaelis–Menten curve for risperidone.

3. Methods

3.1. Estimation

We briefly describe the four estimation methods. One of them is the ML (Davidian and Giltinan, 1995; Vonesh, 2012). Considering (1) with relationship (2), the marginal distribution of response is $\mathbf{y}_i \sim N_{n_i}(\mu_i, \Sigma_i(\theta))$, where $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})$ is a $n_i \times 1$ response vector; $\mu_i(\mathbf{x}_i; \beta) = \mu_i = (\mu(x_{i1}; \beta), \mu(x_{i2}; \beta), \dots, \mu(x_{in_i}; \beta))$ is a $n_i \times 1$ mean vector; $\Sigma_i(\theta)$ is the $n_i \times n_i$ compound symmetry covariance matrix with $(\sigma_s^2 + \sigma_p^2)$ on the diagonal and σ_p^2 at the off-diagonal. The log-likelihood over all I studies is:

$$\begin{aligned} \ell(\beta, \theta | \mathbf{y}_1, \dots, \mathbf{y}_I) = & -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^I \ln |\Sigma_i(\theta)| \\ & - \frac{1}{2} \sum_{i=1}^I [(\mathbf{y}_i - \mu_i)(\Sigma_i^{-1}(\theta))(\mathbf{y}_i - \mu_i)^T], \end{aligned} \quad (4)$$

where T denotes the transpose. Note, that using the ML all parameter estimates are obtained jointly. No closed-form solution exists for parameters γ , but they can be obtained numerically (e.g., optimization algorithm of Nelder and Mead in optim function in R). We carried out ML estimation using the procedure NLMIXED of SAS (NLMIXED, 2008). Standard errors of the estimates are based on the inverse Hessian matrix $\Omega^{-1}(\gamma)$, where Ω is defined as the negative of the second derivative matrix of log-likelihood function (4). Using quasi-Newton optimization technique (TECHNIQUE=QUANEW) of SAS, the second-order derivatives are approximated.

The second estimation procedure is GEE2, developed by Prentice and Zhao (1991). Underlined models do not require specification of the entire distribution of response, but only the first two moments of the response. Namely, under specification $E(\mathbf{y}_i) = \mu_i$, $\text{Var}(\mathbf{y}_i) = \Sigma_i(\beta, \theta)$, the GEE2 and ML estimates $\hat{\gamma}$ of parameter γ are the same. The difference between ML and GEE2 estimation is in the variance-covariance of the estimate, $\text{var}(\hat{\gamma}) = \Omega(\hat{\gamma})$. In ML the model-based covariance $\Omega(\hat{\gamma})$ is used while in GEE2 an empirical sandwich estimator $\Omega_R(\hat{\gamma})$ is used. The later is known as the robust estimator and it is estimated as:

$$\Omega_R(\hat{\gamma}) = \Omega(\hat{\gamma}) \left(\sum_{i=1}^I \mathbf{U}_i(\hat{\gamma}) \mathbf{U}_i(\hat{\gamma})^T \right)^{-1} \Omega(\hat{\gamma}),$$

where $\Omega(\hat{\gamma}) = (\sum_{i=1}^I \mathbf{D}_i(\hat{\gamma})^T (\mathbf{V}_i^{-1}(\hat{\gamma})) \mathbf{D}_i(\hat{\gamma}))^{-1}$ is the model-based covariance and $\mathbf{U}_i(\hat{\gamma}) = \mathbf{D}_i^T(\hat{\gamma}) (\mathbf{V}_i^{-1}(\hat{\gamma})) (\mathbf{s}_i - \mu_i(\hat{\gamma}))$ is the log-likelihood score function. The $\mathbf{s}_i$ and $\mu_i(\gamma)$ are given by:

$$\mathbf{s}_i = \begin{pmatrix} \mathbf{y}_i \\ (\mathbf{y}_i - \mu_i(\mathbf{x}_i, \beta))^2 \end{pmatrix} \quad \mu_i(\gamma) = \begin{pmatrix} \mu_i(\mathbf{x}_i, \beta) \\ \sigma_i(\beta, \theta) \end{pmatrix},$$

where $\sigma_i^T(\beta, \theta) = (\sigma_{i11}, \sigma_{i12}, \dots, \sigma_{i22}, \dots, \sigma_{in_i n_i})$ is a vector of elements of the upper triangular variance-covariance matrix

$\Sigma_i(\beta, \theta)$. The $\mathbf{D}_i(\gamma)$ and $\mathbf{V}_i(\gamma)$ are given by:

$$\mathbf{D}_i(\gamma) = \begin{pmatrix} \mathbf{D}_i(\beta) & 0 \\ 0 & \mathbf{D}_i(\theta) \end{pmatrix} \quad \mathbf{V}_i(\gamma) = \begin{pmatrix} \Sigma_i(\beta, \theta) & 0 \\ 0 & \Gamma_i(\beta, \theta) \end{pmatrix},$$

with $\mathbf{D}_i(\beta) = \partial(\mu_i)/\partial(\beta^T)$ and $\mathbf{D}_i(\theta) = \partial(\sigma_i(\beta, \theta))/\partial(\theta^T)$ the first-order derivative matrices of $\mu_i(\gamma)$. The weight matrix $\mathbf{V}_i(\gamma)$ is the variance-covariance of $\mathbf{s}_i$ with 0 off-diagonal elements. Thus, $\Gamma_i = \text{var}[(\mathbf{y}_i - \mu_i(\mathbf{x}_i, \beta))^2]$. GEE2 estimation can be conducted using NLMIXED procedure of SAS (NLMIXED, 2008), specifying the EMPIRICAL option (Vonesh, 2012).

The third estimation technique is an iterative two-step procedure developed by Vonesh and implemented in macro %NLINMIX (details are given by Vonesh (2012), pp. 28–33, 158–159). In short, step one starts with initial $\tilde{\beta} = (\tilde{\beta}_1, \tilde{\beta}_2)$ for $\beta = (\beta_1, \beta_2)$ and calculates the pseudo-response vector as follows:

$$\tilde{\mathbf{y}}_i = \mathbf{y}_i - \mu_i(\mathbf{x}_i, \tilde{\beta}) + \tilde{\mathbf{X}}_i \tilde{\beta},$$

where $\tilde{\mathbf{X}}_i = \partial(\mu_i)/\partial(\beta^T)|_{\beta=\tilde{\beta}}$. In a second step the following pseudo-marginal linear model is formed:

$$\tilde{\mathbf{y}}_i = \tilde{\mathbf{X}}_i \beta + \mathbf{e}_i$$

for which it is assumed that the residual has distribution $\mathbf{e}_i \sim N(\mathbf{0}, \Sigma_i(\tilde{\beta}, \theta))$. Then, the maximization of the restricted profile log-likelihood $\ell_{\text{REPL}}(\theta, \tilde{\mathbf{y}}_i)$ is carried out with respect to θ :

$$\begin{aligned} \ell_{\text{REPL}}(\theta, \tilde{\mathbf{y}}_i) = & -\frac{(N-p)}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^N \ln |\Sigma_i(\theta)| \\ & - \frac{1}{2} \sum_{i=1}^N \mathbf{r}_i^T \Sigma_i^{-1}(\theta) \mathbf{r}_i + \frac{1}{2} \ln \left| \sum_{i=1}^N \tilde{\mathbf{X}}_i^T \Sigma_i^{-1}(\theta) \tilde{\mathbf{X}}_i \right|, \end{aligned} \quad (5)$$

where p is the rank of $\tilde{\mathbf{X}}_i$ (here $p = 2$) and $\tilde{\mathbf{r}}_i = \tilde{\mathbf{y}}_i - \tilde{\mathbf{X}}_i \tilde{\beta}$ are the pseudo-residuals. Maximization of (5) results in estimate $\hat{\theta}_{\text{REPL}}(\tilde{\beta})$. Then the estimated generalized least square estimator is computed as:

$$\begin{aligned} \hat{\beta}(\hat{\theta}_{\text{REPL}}) = & \left(\sum_{i=1}^N \tilde{\mathbf{X}}_i^T \Sigma_i^{-1}(\tilde{\beta}, \hat{\theta}_{\text{REPL}}(\tilde{\beta})) \tilde{\mathbf{X}}_i \right)^{-1} \\ & \times \sum_{i=1}^N \tilde{\mathbf{X}}_i^T \Sigma_i^{-1}(\tilde{\beta}, \hat{\theta}_{\text{REPL}}(\tilde{\beta})) \tilde{\mathbf{y}}_i. \end{aligned}$$

The observed estimates in step two are then used in step one to create iteratively new estimates. This continues until convergence is achieved, namely the difference between successive estimates becomes less than the pre-defined tolerance level. The asymptotic variance-covariance of the estimates of variance components is based on an inverse Hessian matrix evaluated at $\hat{\theta}_{\text{REPL}}$, that is, $\hat{\mathbf{\Omega}}(\hat{\theta}_{\text{REPL}}) = \mathbf{H}^{-1}(\hat{\theta}_{\text{REPL}})$ where $\mathbf{H}(\theta)$

is the Hessian matrix. The Hessian matrix is the negative of the second derivatives of the log-likelihood with respect to the variance components θ . Technically, the third estimation procedure is carried out using NLIN and MIXED procedures of SAS (MIXED, 2008).

The fourth method of estimation applies standard analysis techniques in two steps. First, the parameters β' and σ_R^2 are estimated using nonlinear regression model $y_{ij} = \mu(x_{ij}, \beta') + \varepsilon_{ij}$, with relationship (2) and residual error $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_R^2) \forall i, j$, but without the additive random effect of between-study variation. Then, the residuals $r_{ij} = y_{ij} - \mu(x_{ij}, \hat{\beta}')$ are formed with the nonlinear regression parameter estimates $\hat{\beta}'$. In a second step, a simple linear random effects model $r_{ij} = u_i + \epsilon_{ij}$ with additive random effect of between-study variation $u_i \stackrel{iid}{\sim} N(0, \sigma_S^2) \forall i$ and random effect of within-study variation $\epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_P^2) \forall i, j$ is fitted to the residuals r_{ij} to obtain the estimates of variance components $\hat{\sigma}_S^2$ and $\hat{\sigma}_P^2$. Technically, at a first step the NLS estimates of regression parameters β' (Seber and Wild, 2003) are obtained using NLIN procedure and then the residuals r_{ij} are computed. At a second step, these residuals are fitted to one-way random effects model using MIXED procedure (method=REML) to obtain $\hat{\sigma}_S^2$ and $\hat{\sigma}_P^2$. In fact, this step is based not on classical REML, but rather REML version of pseudolikelihood (Davidian and Giltinan, 1995). Regarding asymptotic variance-covariance of variance components, they are based on an inverse Hessian matrix. Brief information about this was already given in a third estimation method when $\hat{\mathbf{\Omega}}(\hat{\theta}_{\text{REPL}})$ was introduced.

3.2. Confidence Interval for ICC: Beta-Approach

The distribution of the ICC estimator is approximated with a beta distribution, $\text{ICC} \sim \text{Beta}(a, b)$ with parameters $a > 0$ and $b > 0$. If ICC is an estimate of the mean and $\hat{\tau}_{\text{ICC}}^2$ is an estimate of the variance of ICC , the method of moment estimates for a and b become:

$$\begin{aligned} \hat{a} &= \frac{\text{ICC}[\text{ICC}(1 - \text{ICC}) - \hat{\tau}_{\text{ICC}}^2]}{\hat{\tau}_{\text{ICC}}^2} \\ \hat{b} &= \frac{(1 - \text{ICC})[\text{ICC}(1 - \text{ICC}) - \hat{\tau}_{\text{ICC}}^2]}{\hat{\tau}_{\text{ICC}}^2}. \end{aligned}$$

Demetrashvili et al. (2014) proposes the use of first-order Taylor expansion for the variance of ICC :

$$\hat{\tau}_{\text{ICC}}^2 = \frac{\hat{\sigma}_P^4}{(\hat{\sigma}_S^2 + \hat{\sigma}_P^2)^4} \hat{\tau}_{\sigma_S^2}^2 + \frac{\hat{\sigma}_S^4}{(\hat{\sigma}_S^2 + \hat{\sigma}_P^2)^4} \hat{\tau}_{\sigma_P^2}^2 - \frac{2\hat{\sigma}_S^2 \hat{\sigma}_P^2}{(\hat{\sigma}_S^2 + \hat{\sigma}_P^2)^4} \hat{\tau}_{\sigma_S^2, \sigma_P^2}^2 \quad (6)$$

with $\hat{\tau}_{\sigma_S^2}^2$ and $\hat{\tau}_{\sigma_P^2}^2$ the estimated variance of estimators $\hat{\sigma}_S^2$ and $\hat{\sigma}_P^2$, respectively, and $\hat{\tau}_{\sigma_S^2, \sigma_P^2}^2$ the estimated covariance of the estimators $\hat{\sigma}_S^2$ and $\hat{\sigma}_P^2$.

The approximate 100%(1 - α) confidence interval on the ICC in (3) using the beta distribution is now given by the lower and upper confidence limits:

$$\text{LCL}_B = B_{a,b}^{-1}(\alpha/2) \quad \text{UCL}_B = B_{a,b}^{-1}(1 - \alpha/2),$$

with $B_{\hat{a},\hat{b}}^{-1}(q)$ the q th-quantile of the Beta(a, b) distribution. Note that the parameter estimates $\hat{a}$ and $\hat{b}$ are somewhat altered when the ICC is estimated by zero (see details in Demetrashvili et al., 2014).

3.3. Confidence Interval for Transformed ICC: Delta Method

The large sample approximation or the delta method can be considered as an alternative approach to our beta-approach. Since the distribution of the ICC is known to be skewed, we conducted the delta method on a transformed ICC, namely $\eta = \ln(\sigma_S^2) - \ln(\sigma_P^2) = \ln(\text{ICC}) - \ln(1 - \text{ICC})$. Based on the derivatives $\partial(\eta)/\partial(\sigma_S^2) = 1/\sigma_S^2$ and $\partial(\eta)/\partial(\sigma_P^2) = -1/\sigma_P^2$, an estimate of the variance of $\hat{\eta}$ using the first-order Taylor's approximation becomes:

$$\hat{\tau}_{\hat{\eta}}^2 = \frac{1}{\hat{\sigma}_S^4} \hat{\tau}_{\hat{\sigma}_S^2}^2 + \frac{1}{\hat{\sigma}_P^4} \hat{\tau}_{\hat{\sigma}_P^2}^2 - \frac{2}{\hat{\sigma}_S^2 \hat{\sigma}_P^2} \hat{\tau}_{\hat{\sigma}_S^2, \hat{\sigma}_P^2} \quad (7)$$

with $\hat{\tau}_{\hat{\sigma}_S^2}^2$ and $\hat{\tau}_{\hat{\sigma}_P^2}^2$ the estimated variance of ML estimators $\hat{\sigma}_S^2$ and $\hat{\sigma}_P^2$ respectively, and $\hat{\tau}_{\hat{\sigma}_S^2, \hat{\sigma}_P^2}$ the estimated covariance of the $\hat{\sigma}_S^2$ and $\hat{\sigma}_P^2$. SAS uses the standard error of (7) combined with a t -value with $(I - 1)$ degrees of freedom to construct the approximate 100%(1 - α) confidence interval on η . For an ICC estimate equal to zero, no confidence interval is calculated.

3.4. Confidence Interval for ICC: Bootstrap Method

We employed the nonparametric percentile bootstrap as another alternative to our beta-approach. We investigated the coverages of bootstrap confidence intervals under three resampling strategies. Namely, we (1) resampled entire studies by maintaining the subjects within studies, (2) resampled all observations by ignoring the grouping structure of the studies, (3) resampled entire studies and then resampled subjects within the studies (two-stage resampling). In all three strategies the resampling was done with replacement and with equal probability from the study data set to create a bootstrap data set of the same size.

4. Simulation Study

4.1. Settings

The simulation study reports three parts. In a first part, we investigate the bias in estimating the ICCs. In a second part, we examine the coverage probabilities (CP) of the proposed beta-approach for the four estimation methods. In a third part, we investigate the coverages of the alternative approaches. All parts are conducted for the same sets of variance components and sample sizes.

For variance components the following parameters are specified: the variance component for study is set at $\sigma_S^2 = 0.001$ and the variance component for patient is varied accordingly $\sigma_P^2 = \{0.009, 0.004, 0.0023, 0.0015, 0.001, 0.00067, 0.00043, 0.00025, 0.00011\}$ to generate ICC values of 0.1, ..., 0.9 (0.1), respectively. We assume a balanced design, that is, equal number of patients per study. Settings on the pairs (I, J) of sample sizes, with I being the number of studies and J being the number of subjects, in a study are (5, 5), (5, 8), (5, 20), (5, 30), (8, 5), (10, 5), (10, 10), (20, 10), (30, 10). The doses on patients were simulated from a uniform distribution, $x \sim \text{Uniform}(1, \text{maxd})$, where maxd is the maximum

Figure 3. Bias of ICC estimators.

dose and it is set equal to 21. We set the maximum occupancy equal to $\beta_1 = 0.8$ and the dose for half-maximum occupancy ED_{50} equal to $\beta_2 = 16$. We simulated the data according to nonlinear mixed effects model using the formula: $y_{ij} = (\beta_1 x_{ij}) / (\beta_2 + x_{ij}) + \sigma_S z_i + \sigma_P z_{ij}$ with z_i and z_{ij} mutually independently distributed from a standard normal distribution, ignoring the log-transformation from our case study.

Below we only show the results for the set of pairs with small sample sizes.

4.2. Results: Bias

Biases of ICC using four estimation methods are depicted in Figure 3. Note that the ML and GEE2 differ only in the calculation of the standard errors and do not differ in parameter estimates. Results show that in most cases the estimates of ICC are biased downwards using all estimation procedures. The bias is generally larger using ML/GEE2 procedure than using two-step procedures. The bias is smallest using two-step iterative estimation method of Vonesh. Differences between the biases of estimators are more prominent for small sample sizes (e.g., (5, 6), (5, 8), (8, 5)).

4.3. Results: Coverages of Beta-Approach

With all estimation approaches the number of simulations is 10,000 except when we implemented %NLINMIX. With %NLINMIX we carried out 1000 simulations (due to capacity of our computer). The marginal CPs for the two-sided 95% confidence intervals using the beta-approach are presented in Table 3. Results show that ML, and particularly GEE2 lead to poor coverages. In contrast, both two-step procedures lead

Table 3

Marginal coverage probabilities for two-sided 95% confidence intervals using beta-approach

I	J	ICC	CP _{ICC} (Beta)			
			ML	GEE2	NLS & REML	NLS & REML (iterative)
5	5	0.1	0.968	0.929	0.963	0.973
		0.2	0.954	0.914	0.956	0.957
		0.3	0.946	0.881	0.950	0.939
		0.4	0.949	0.850	0.953	0.958
		0.5	0.952	0.806	0.961	0.930
		0.6	0.951	0.775	0.964	0.953
		0.7	0.918	0.762	0.972	0.947
		0.8	0.914	0.739	0.956	0.952
		0.9	0.909	0.746	0.957	0.936
5	8	0.1	0.979	0.950	0.972	0.970
		0.2	0.940	0.885	0.975	0.975
		0.3	0.977	0.839	0.978	0.972
		0.4	0.901	0.794	0.976	0.971
		0.5	0.874	0.771	0.950	0.947
		0.6	0.897	0.724	0.936	0.949
		0.7	0.895	0.725	0.933	0.941
		0.8	0.894	0.729	0.937	0.923
		0.9	0.914	0.753	0.946	0.924
8	5	0.1	0.962	0.939	0.961	0.949
		0.2	0.957	0.930	0.959	0.949
		0.3	0.966	0.909	0.966	0.966
		0.4	0.970	0.862	0.972	0.970
		0.5	0.967	0.843	0.975	0.976
		0.6	0.936	0.829	0.955	0.955
		0.7	0.927	0.821	0.944	0.945
		0.8	0.923	0.815	0.946	0.940
		0.9	0.929	0.835	0.952	0.942
10	5	0.1	0.959	0.945	0.958	0.958
		0.2	0.964	0.937	0.965	0.964
		0.3	0.974	0.918	0.977	0.983
		0.4	0.976	0.880	0.978	0.977
		0.5	0.947	0.857	0.959	0.954
		0.6	0.935	0.853	0.950	0.943
		0.7	0.931	0.850	0.947	0.938
		0.8	0.933	0.849	0.951	0.948
		0.9	0.932	0.852	0.950	0.937
10	10	0.1	0.995	0.968	0.994	0.995
		0.2	0.975	0.877	0.979	0.972
		0.3	0.921	0.841	0.943	0.946
		0.4	0.909	0.827	0.932	0.932
		0.5	0.915	0.833	0.933	0.933
		0.6	0.919	0.834	0.936	0.951
		0.7	0.919	0.843	0.937	0.940
		0.8	0.925	0.847	0.941	0.944
		0.9	0.935	0.855	0.949	0.934

Table 4

Coverage probabilities for two-sided 95% confidence intervals using bootstrap and delta methods

I	J	ICC	CP _{ICC} (Bootstrap)			CP _{η} (Delta)	
			Resample studies	Resample observ.	Resample two-stage	MLE marginal	MLE cond.
5	5	0.1	0.940	0.960	1.000	0.527	0.991
		0.2	0.870	0.950	1.000	0.685	0.996
		0.3	0.820	0.910	1.000	0.805	0.998
		0.4	0.790	0.890	1.000	0.873	0.999
		0.5	0.870	0.870	0.980	0.927	0.999
		0.6	0.860	0.890	0.990	0.959	0.998
		0.7	0.800	0.840	0.930	0.980	0.999
		0.8	0.770	0.830	0.880	0.991	0.998
		0.9	0.780	0.810	0.860	0.987	0.989
5	8	0.1	0.830	0.960	1.000	0.626	0.997
		0.2	0.850	0.910	1.000	0.794	0.999
		0.3	0.800	0.900	1.000	0.886	1.000
		0.4	0.850	0.860	0.990	0.932	0.999
		0.5	0.800	0.710	0.940	0.963	1.000
		0.6	0.800	0.810	0.950	0.985	0.999
		0.7	0.790	0.730	0.890	0.991	0.999
		0.8	0.710	0.630	0.870	0.997	1.000
		0.9	0.730	0.680	0.840	0.967	0.968
8	5	0.1	0.910	0.900	0.960	0.616	0.975
		0.2	0.880	0.900	0.980	0.807	0.990
		0.3	0.860	0.890	1.000	0.905	0.993
		0.4	0.890	0.860	0.980	0.957	0.995
		0.5	0.900	0.840	0.990	0.978	0.995
		0.6	0.830	0.880	0.990	0.990	0.996
		0.7	0.910	0.870	0.970	0.994	0.996
		0.8	0.830	0.850	0.920	0.969	0.969
		0.9	0.790	0.860	0.910	0.961	0.961
10	5	0.1	0.950	0.870	0.990	0.665	0.969
		0.2	0.940	0.830	0.990	0.852	0.987
		0.3	0.860	0.880	0.970	0.941	0.991
		0.4	0.920	0.830	1.000	0.975	0.993
		0.5	0.850	0.760	0.990	0.988	0.994
		0.6	0.900	0.850	1.000	0.994	0.995
		0.7	0.900	0.830	0.970	0.974	0.975
		0.8	0.840	0.880	0.970	0.962	0.962
		0.9	0.770	0.850	0.920	0.958	0.958
10	10	0.1	0.920	0.890	1.000	0.847	0.991
		0.2	0.850	0.850	0.980	0.964	0.994
		0.3	0.800	0.790	0.990	0.990	0.997
		0.4	0.840	0.790	0.970	0.994	0.996
		0.5	0.840	0.680	0.930	0.975	0.976
		0.6	0.860	0.750	0.960	0.964	0.964
		0.7	0.870	0.660	0.920	0.956	0.956
		0.8	0.810	0.640	0.890	0.952	0.952
		0.9	0.770	0.70	0.880	0.956	0.956

to equally accurate coverages which are generally quite close to the nominal value of 95%.

4.4. Results: Coverages of Bootstrap and Delta Method

To compare the coverages demonstrated by the beta-approach to the ones produced by the bootstrap and the delta method, we conducted the simulation studies using the settings described in Section 4.1. We conducted $B = 999$ bootstrap replications for 100 simulations and applied only the two-step

NLS&REML estimation method (due to its speed and accuracy). The CPs for two-sided 95% confidence intervals on ICCs using nonparametric percentile bootstrap are presented in Table 4. Clearly, the two single-stage resampling strategies lead to liberal CPs, and the two-stage resampling strategy leads to conservative CPs. Consequently, none of the resampling strategies in combination with the percentile method work satisfactorily to obtain accurate confidence intervals under the investigated settings.

Table 5
Approximate 95% confidence intervals of ICC based on model (1) using the beta-approach

Antipsychotic	Estimation method	$\hat{\text{ICC}}$	$\hat{\text{LCL}}$	$\hat{\text{UCL}}$	Width
Haloperidol	NLS&REML	0.692	0.468	0.875	0.407
	NLS&REML (iterative)	0.699	0.477	0.880	0.403
Clozapine	NLS&REML	0.160	0.019	0.407	0.388
	NLS&REML (iterative)	0.160	0.020	0.408	0.388
Amisulpride	NLS&REML	0.456	0.094	0.851	0.757
	NLS&REML (iterative)	0.445	0.094	0.835	0.741

In Table 4, we also present the coverages from 10,000 simulations for two-sided 95% confidence intervals on η using the delta method. For the delta method we present the marginal and conditional CP on η in Table 4. Conditional coverages exclude all those simulations in which the η was undefined. The marginal coverages contain all simulations counting the undefined estimates for η as being non-covered. Results show that the delta method produces poor coverages for the investigated small sample sizes. They are conservative for the conditional coverages and liberal for the marginal coverages. Only for pairs (10,10) did the conditional coverages start to approach the nominal 95%. We also observe that for this pair the conditional CPs approach the marginal CPs for higher ICCs. Clearly, the number of simulations with a zero ICC estimate vanishes with an increase in ICC and sample size. With higher numbers of studies and patients (≥ 30), the results using the delta method with ML estimation method become more accurate (not shown here).

5. Case Study (continued)

Recall that in Section 1, for every antipsychotic we assessed the heterogeneity of multiple studies involved. In this section, we provide and interpret the confidence intervals for ICCs produced using the beta-approach with the best two estimation methods. The confidence intervals along with their widths are presented in Table 5. Results show that these two estimation methods lead to almost identical confidence limits.

The heterogeneity for haloperidol is substantial. This implies that one group of patients from one study have a tendency to fall on one side of the estimated curve and another group on the other side. For a clinician it becomes more complicated to select the correct dose, since the clinician would either over-dose or under-dose the patients. For clozapine, the heterogeneity is very small, which means that the curve would indicate the correct dose on average. For amisulpride this is still unknown due to moderate ICC and large confidence interval.

6. Discussion

To quantify the impact of heterogeneity and its extent on the meta-analysis, it is important to have an accurate estimate of ICC and its confidence interval. In this article, we applied the generic beta-approach to construct a confidence interval of ICC for nonlinear mixed effects models with an additive random effects structure. We examined the coverages of the beta-approach using four estimation methods: ML, GEE2, NLS in

combination with REML (in iterative and non-iterative way). We also studied the biases of ICC estimators with these four methods. We compared the beta-approach with two alternative generic approaches used for the construction of confidence intervals: the delta method and bootstrapping.

The most accurate estimation method in terms of bias in the ICC is the two-step iterative approach. The largest bias is demonstrated by ML/GEE2 in most cases. Concerning the coverages of the confidence intervals using the beta-approach, both two-step approaches (iterative and non-iterative) provide equally accurate results. The semi-parametric method of moments (GEE2) provides highly inaccurate coverages. Even though the ML provides better coverages than GEE2 overall, it is still quite unsatisfactory in combination with the beta-approach. To summarize the results of the alternative methods, neither bootstrap with any of the resampling strategies nor the delta method is recommended for computation of confidence intervals of ICCs, in particular for small sample sizes.

Limitations of our beta-approach is that it is intended for ICCs of the form, $\sum_{q=1}^Q \sigma_q^2 / (\sum_{q=1}^Q \sigma_q^2 + \sum_{p=Q+1}^P \sigma_p^2)$. Concerning the estimation methods, an iterative two-step estimation method is computationally expensive. In contrast, the non-iterative estimation method is fast but the properties of the estimators are unknown. A strength of the beta-approach is that it can be applied to any additive variance components model with balanced and unbalanced designs for linear models (Demetrashvili et al., 2014), and now also for nonlinear mixed effects regression models. Another strength is the existence of a closed-form expression, which makes sample size calculation feasible when a certain length for the confidence interval on ICC is required. Furthermore, it works particularly well for grouped (or dependent) data when the number of groups and the number of subjects per group is limited. Finally, the beta-approach also provides the confidence intervals for zero ICC estimates, which is not provided by the delta method.

The percentile bootstrap method with three resampling strategies did not work appropriately in our case. However, there are more options that could have been explored, either with respect to resampling strategies for clustered data or the computation of bootstrap confidence intervals. Some of these alternative options are described for instance in Ukoumunne et al. (2003) and Field and Welsh (2007). We did not examine widely used, but computationally expensive bias-corrected accelerated percentile method proposed by Efron (1987) either. It is not obvious whether these alternative bootstrap methods would perform better uniformly

since bootstrapping the repeated measures is not straightforward (see Davison and Hinkley (1997) on pp. 100–102 and the discussion of Demetrashvili et al. (2014)). Given these points, more research should be conducted on bootstrap methods for grouped data.

In our work, we applied the delta method to only one particular form of transformation, namely $\eta = \ln \sigma_s^2 - \ln \sigma_p^2$. However, there are other alternatives, like the well-known Fisher's Z-transformation. It is expected though that Fisher's normalizing transformation may lead to conservative coverages since this was demonstrated for ICC's in one-way balanced and unbalanced random effects models with small and moderate sample sizes (see Donner and Wells, 1986; Demetrashvili et al., 2014).

In conclusion, the beta-approach in combination with either two-step estimation methods is recommended for the computation of a confidence interval for ICC in a nonlinear mixed effects model, particularly for the mix of small sample sizes.

ACKNOWLEDGEMENTS

We would like to express our gratitude to the referee, the associate editor and the editor for their critical and useful comments, which helped to improve this paper substantially.

REFERENCES

- Aune, D., Lau, R., Chan, S. M. D., Vieira, R., Greenwood, D. C., Kampman, E., and Norat, T. (2011). Nonlinear reduction in risk for colorectal cancer by fruit and vegetable intake based on meta-analysis of prospective studies. *Gastroenterology* **141**, 106–118.
- Davidian, M. and Giltinan, D. M. (1995). *Nonlinear Models for Repeated Measurement Data*. St. Edmunds, Suffolk: Chapman & Hall.
- Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and Their Application*. Cambridge: Cambridge University Press.
- Demetrashvili, N., Wit, E. C., and Van den Heuvel, E. R. (2014). Confidence intervals for intraclass correlation coefficients in variance components models. *Statistical Methods in Medical Research*, doi: 10.1177/0962280214522787.
- Donner, A. and Wells, G. (1986). A comparison of confidence interval methods for the intraclass correlation coefficient. *Biometrics* **42**, 401–412.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* **7**, 1–26.
- Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association* **82**, 171–185.
- Field, C. A. and Welsh, A. H. (2007). Bootstrapping clustered data. *Journal of the Royal Statistical Society, Series B* **69**, 369–390.
- Goutelle, S., Maurin, M., Rougier, F., Barbaut, X., Bourguignon, L., Ducher, M., and Maire, P. (2008). The Hill equation: A review of its capabilities in pharmacological modelling. *Fundamental & Clinical Pharmacology* **22**, 633–648.
- Higgins, J. and Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* **21**, 1539–58.
- Lako, I. M., van den Heuvel, E. R., Knegtering, H., Bruggeman, R., and Taxis, K. (2013). Estimating dopamine D_2 receptor occupancy for doses of 8 antipsychotics: A meta-analysis. *Journal of Clinical Psychopharmacology* **33**, 675–681.
- Michaelis, L. and Menten, M. L. (1913). Die Kinetik der Invertinwirkung. *Biochemische Zeitschrift* **49**, 333–369.
- MIXED (2008). *SAS/STAT[®] 9.2 User's Guide: The MIXED Procedure (Book Excerpt)*. Cary, North Carolina: SAS Institute.
- NLMIXED (2008). *SAS/STAT[®] 9.2 User's Guide: The MIXED Procedure (Book Excerpt)*. Cary, North Carolina: SAS Institute.
- Prentice, R. L. and Zhao, L. P. (1991). Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. *Biometrics* **47**, 825–839.
- Qu, X., Jin, F., Hao, Y., Li, H., Tang, T., Wang, H., Yan, W., and Dai, K. (2013). Magnesium and the risk of cardiovascular events: A meta-analysis of prospective cohort studies. *PLoS ONE* **8**, e57720.
- Rota, M., Bellocco, R., Scotti, L., Tramacere, I., Jenab, M., Corrao, G., La Vecchia, C., Boffetta, P., and Bagnardi, V. (2010). Random-effects meta-regression models for studying nonlinear dose-response relationship, with an application to alcohol and esophageal squamous cell carcinoma. *Statistics in Medicine* **29**, 2679–2687.
- Rui, R., Lou, J., Zou, L., Zhong, R., Wang, J., Xia, D., Wang, Q., Li, H., Wu, J., Lu, X., Li, C., Liu, L., Xia, J., and Xu, H. (2012). Excess body mass index and risk of liver cancer: A nonlinear dose-response meta-analysis of prospective studies. *PLoS ONE* **7**, e44522.
- Seber, G. A. F. and Wild, C. J. (2003). *Nonlinear Regression*. Hoboken, New Jersey: John Wiley & Sons.
- Ukounmunne, O. C., Davison, A. C., Gulliford, M. C., and Chinn, S. (2003). Non-parametric bootstrap confidence intervals for the intraclass correlation coefficient. *Statistics in Medicine* **22**, 3805–3821.
- Vonesh, E. F. (2012). *Generalized Linear and Nonlinear Models for Correlated Data: Theory and Applications Using SAS*. Cary, North Carolina: SAS Institute.

Received March 2014. Revised September 2014.

Accepted November 2014.