Clinical Chemistry 62:5
725–736 (2016)

Biological Variation: The Effect of Different Distributions on Estimated Within-Person Variation and Reference Change Values

Thomas Røraas,[1,2*] Bård Støve,[3] Per Hyltoft Petersen,[1] and Sverre Sandberg[1,2,4]

BACKGROUND: Good estimates of within-person biological variation, CV_I, are essential for diagnosing and monitoring patients and for setting analytical performance specifications. The aim of the present study was to use computer simulations to evaluate the impact of various measurement distributions on different methods for estimating CV_I and reference change value (RCV).

METHOD: Data were simulated on the basis of 3 models for distributions of the within-person effect. We evaluated 3 different methods for estimating CV_I: standard ANOVA, ln-ANOVA, and CV-ANOVA, and 3 different methods for calculating RCV: classic, ln-RCV, and a nonparametric method. We estimated CV_I and RCV with the different methods and compared the results with the true values.

RESULTS: The performance of the methods varied, depending on both the size of the CV_I and the type of distributions. The CV-ANOVA model performed well for the estimation of CV_I with all simulated data. The ln-RCV method performed best if data were ln-normal distributed or CV_I was less than approximately 12%. The nonparametric RCV method performed well for all simulated data but was less precise.

CONCLUSIONS: The CV-ANOVA model is recommended for both calculation of CV_I and the step-by-step approach of checking for outliers and homogeneity in replicates and samples. The standard method for calculation of RCV should not be used when using CVs.

The observed variation of a measurand in blood, plasma, or urine within an individual in a steady-state situation is caused by within-person biological variation, CV_I, and analytical imprecision, CV_A *(1)*. Good estimates of CV_I are essential for diagnosing and monitoring patients and for setting analytical performance specifications. Usually, nested ANOVA or a similar method is used when estimating CV_I *(2)*. The use of such methods depends on 2 assumptions: homoscedasticity of the CV values for each level of the model and independence of observations *(3)*. When using an ANOVA model for estimating CVs, it is assumed that the observations can be approximated by a linear combination of certain unobservable quantities known as effects *(3)*. Estimating the variance components by use of a nested ANOVA does not assume normality or any other distributions for the model effects *(2–4)*. However, if confidence intervals are to be calculated or statistical significance tests are to be performed on the variance components, assumptions of normal distributions for the effects of the model are usually needed *(3, 5)*. It must be underlined that when addressing the question of distribution, it is the distributions of the individual model effects—i.e., between-person, within-person, and analytical effects—that are assumed to be normal, and not the distribution of the combined results from all individuals. Thus, when studying the distribution of the measurements from individuals, the distribution of the combined normalized residuals from each person should be checked. Data from all individuals studied should be in a steady-state condition: i.e., there should be no trend in the concentrations of the measurand during the study period. This assumption is covered by the independence of observations assumption: i.e., there is no autocorrelation between observations. If data from all individuals are not in a steady state but can be transformed to have a homoscedastic trend, adjustment for the trend is achievable by use of a suitable ANOVA model (analysis of covariance or generalized linear models) or by transforming data or adjusting the interpretation of the results *(5, 6)*.

A common use of CV_I estimates is in the calculation of the reference change value (RCV),[5] used to monitor and diagnose patients *(7)*. A presupposition for the stan-

[1] Norwegian Quality Improvement of Primary Care Laboratories, Haraldsplass Deaconess Hospital, Bergen, Norway; [2] Department of Global Public Health and Primary Care, University of Bergen, Norway, and [3] Department of Mathematics, University of Bergen, Norway; [4] Laboratory of Clinical Biochemistry, Haukeland University Hospital, Bergen, Norway.
* Address correspondence to this author at: NOKLUS, PO 6165, N-5892 Bergen, Norway. Fax +47-55-97-95-10; e-mail thomas.roraas@noklus.no.

Received November 20, 2015; accepted February 23, 2016.
Previously published online at DOI: 10.1373/clinchem.2015.252296

[5] Nonstandard abbreviations: RCV, reference change value; NP, nonparametric.

dard RCV calculation is that the expected percentage change between samples follows a normal distribution (or that an appropriate transformation to achieve this can be found). The percentage difference between normally distributed variables is not a normally distributed variable itself, however, as it involves a ratio of 2 normally distributed variables, as previously described by Marsaglia *(8, 9)*. As stated by Fraser *(10)*, there are gaps in our knowledge of RCV. The difficulty of estimating the distribution of individual model effects was discussed previously by Harris and DeMets in 1972 *(11)*.

The aim of the present study is to use simulation experiments to evaluate the impact of different measurement distributions on different methods for estimating CV_I and RCV.

Method and Statistics

STANDARD MODEL WITH SDs

When estimating biological variation, model effects are usually assumed to be the between-person, within-person, and analytical effects. The standard model for collecting samples for a biological variation experiment has a nested design: collect *S* samples from *P* individuals and analyze the samples in *R* replicates. This gives the following model:

$$Y_{psr} = \mu + G_p + I_{ps} + A_{psr} \quad \text{(Model 1)}$$

$$(p = 1, \ldots, P;\ s = 1, \ldots, S;\ r = 1, \ldots, R),$$

where Y_{psr} is the measured value of the single replicate *r* for sample *s* from individual *p*, μ is the population mean, G_p is the between-person effect, I_{ps} is the within-person effect, and A_{psr} is the error effect. It is assumed that G_p, I_{ps}, and A_{psr} are mutually independent and completely uncorrelated random variables with mean of zero and variances σ_G^2, σ_I^2, and σ_A^2, respectively. These individual variances are called the "components of variance" *(3)*.

MODEL WITH CVs

In this article, it is assumed that CVs are constant, and the model is thus based on changes in percentages and is multiplicative. The model is consequently formulated as

$$Y_{psr} = \mu \times (1 + G_p) \times (1 + I_{ps}) \times (1 + A_{psr}), \quad \text{(Model 2)}$$

where $\mu_p = \mu \times (1 + G_p)$ is the homeostatic set point for individual *p*.

If it is assumed that the variables are normally distributed, $G_p \cong N(0, \sigma_G)$, $I_{ps} \cong N(0, \sigma_I)$, and $A_{psr} \cong N(0, \sigma_A)$, it is the same model as described by Fraser and Harris *(2)*, who discussed both the SD- and CV-based models. Using ANOVA for estimating the variance components is merely decomposing the sum of squares. It follows from the rule of $var(X + Y) = var(X) + var(Y)$, which assumes the components to be independent of each other.

TRUNCATED DISTRIBUTIONS

The normal distribution, by definition, can include values from minus infinity to infinity, and the ln-normal distribution can include values from zero to infinity. Because real measurements do not have such a wide range of outcomes, we used a method based on inverse transformation sampling for generating truncated random variables *(12)*. The simulated normal distributions were truncated to within 3SD for the normal distribution, and for the simulated ln-normal data with a mean of 1, we used truncation to exclude values >2. These choices were made to avoid values <0, also when mirroring the ln-normal data.

MODELS USED IN SIMULATIONS

To assess the consequences of using different methods for estimating CV and RCV, we simulated data on the basis of the following 3 models:

$$Y_{psr} = \mu \times (1 + G_p^T) \times I_{ps}^{TN} \times (1 + A_{psr}^T), \quad \text{(Model 3.1)}$$

$$Y_{psr} = \mu \times (1 + G_p^T) \times I_{ps}^{T\ln N} \times (1 + A_{psr}^T), \quad \text{(Model 3.2)}$$

$$Y_{psr} = \mu \times (1 + G_p^T) \times I_{ps}^{TM\ln N} \times (1 + A_{psr}^T). \quad \text{(Model 3.3)}$$

The between-person effect G_p^T and analytical effect A_{psr}^T are truncated normally distributed, and the within-person effect I_{ps}^{TN}, $I_{ps}^{T\ln N}$, and $I_{ps}^{TM\ln N}$ are the truncated normal, ln-normal, and mirrored ln-normal (the ln-normal distribution mirrored around the mean 1), respectively. The nontruncated distributions can be simulated as in Forbes et al. *(13)*:

$$I_{ps}^N \cong N(1, SD_I), \quad \text{(Formula 1.1)}$$

$$I_{ps}^{\ln N} \cong \exp\{N[-0.5 \times \ln(CV_I^2 + 1), \sqrt{\ln(CV_I^2 + 1)}]\}, \quad \text{(Formula 1.2)}$$

$$I_{ps}^{M\ln N} \cong 2 - \exp\{N[-0.5 \times \ln(CV_I^2 + 1), \sqrt{\ln(CV_I^2 + 1)}]\}. \quad \text{(Formula 1.3)}$$

Visually, the truncated distributions are similar to the nontruncated distributions for the degrees of truncation used here.

We simulated data with a nested design following the scheme of 15 individuals with 10 samples measured in duplicate. For each of the 3 within-person effect distributions, 16 000 data sets were simulated, each with a random CV_I between 1% and 30%. We simulated data following 2 scenarios, with CV_A one-quarter of CV_I and

with $CV_A = CV_I$. The CV_G was either 10%, 30%, or 40%, where 30% is the average CV_G in the database of Ricos et al. *(14)*. The population mean was set to 100.

ESTIMATION OF WITHIN-PERSON CV_I

To evaluate the impact of the 3 different distributions for the within-person effect, we estimated the CV_I by 3 different methods: (*a*) The standard ANOVA is based on a nested ANOVA performed on the raw data, where the resulting SD_I is divided by the total mean to estimate the CV_I. (*b*) In the ln-ANOVA method, data are ln-transformed [$Y^*_{psr} = \ln(Y_{psr})$], the nested ANOVA is performed on the transformed results, and the results are transformed back to the original scale. The estimated components from the ANOVA performed on the ln-transformed data (Y^*_{psr}) are then transformed back and become CV values on the original scale. (*c*) The CV-ANOVA method is based on the CV transformation with normalization of each person's data by dividing by that person's mean value, and then performing the ANOVA. The result from the ANOVA is then given as CV values but does not provide an estimate of the between-person variation CV_G, since each individual has a mean value of 1. The CV_G can be estimated by the standard ANOVA method.

ESTIMATING RCVs

By definition, the 2-sided 95% RCVs are the expected $P_{2.5}$ and $P_{97.5}$ percentiles for the percent changes between samples. For evaluation of different methods for estimating RCVs in the present study, the true RCVs are needed as references. The true RCV limits have no closed-form solution for the combination of distributions discussed in this article and have to be estimated by simulations *(8)*. For each of the 16 000 random CV_I values with the accompanying CV_A, we simulated 100 000 differences $\{d_i\}$:

$$d_i = 100\% \times \frac{O_2 - O_1}{O_1},$$

(Formula 2)

where $O_i \cong 100 \times (1 + I_{dist}) \times (1 + A)$, where *dist* is the distribution evaluated (truncated normal, ln-normal, or mirrored ln-normal) and A is the analytical error effect. For each of these 16 000 sets of 100 000 differences $\{d_i\}$, we estimated the $P_{2.5}$ and $P_{97.5}$ percentiles and used them as "true" RCV limits, 1 pair for each data set, with which the estimated values were compared.

To evaluate the impact of the 3 different within-person effect distributions, we estimated the RCV limits by 3 different methods.

First, classic RCV is typically used when the model effects are assumed to be normally distributed, as in Method 1 for calculation of CV_I. The CV_I in question and the analytical CV_A are used directly in the RCV formula:

$$RCV(\alpha) = \mp Z_\alpha \times \sqrt{2} \times \sqrt{CV_A^2 + CV_I^2}.$$

(Formula 3)

This method results in symmetrical limits.

Second, ln-RCV is typically used when model effects are assumed to be ln-normal distributed, as described by Fokkema et al. *(15)*, and corresponds to the CV_I calculation in Method 2, where the limits were transformed by use of the exponential function. This method results in asymmetrical limits on the back-transformed values:

$$RCV_{ln}(\alpha) = 100\% \times \exp(\mp Z_\alpha \times \sqrt{2} \times \sqrt{CV_{lnA}^2 + CV_{lnI}^2}).$$

(Formula 4)

Here, CV_{ln} is the CV on the ln-transformed data, so $CV_{lnI} = \sqrt{\ln(1 + CV_I^2)}$.

Third, the nonparametric method (NP-RCV) is based on calculating the observed percentage changes in the concentration of the measurand between subsequent samples from each individual. The percentiles from the empirical distribution of these combined changes are used to estimate the RCV limits for the given population, thus this is a nonparametric approach. (The differences are calculated as in Formula 2.)

QUANTILE REGRESSION

In ordinary least-square regression, the aim is to estimate the conditional mean response depending on the predictors. In contrast, the goal of quantile regression is to estimate the conditional quantiles including the median *(16)*. We performed quantile regression with the response being the ratio between the estimates and the true values for each of the methods for both CV_I and RCV, with the predictor being the true CV_I. The 2.5th, 50th (median), and 97.5th percentiles were estimated. Results from the quantile regressions are indicators of systematic errors (median) and expected range of estimates (2.5th and 97.5th percentiles) of the methods.

SOFTWARE

We performed simulations and plots with a combination of C++ 11 (www.cplusplus.com/articles/cpp11/) and R version 3.2.3 *(17)* with the help of the Rcpp package version 0.12.2 (http://www.rcpp.org) on a CentOS 7 operating system (www.centos.org).

We generated the random numbers with the C++ 11 standard library *(18)*. A random seed was generated by a hardware entropy source by use of *std::random_device rd*, which was then used to initialize the Mersenne twister *std::mt19937_64 generator(rd())*.

Results

ESTIMATION OF WITHIN-PERSON BIOLOGICAL VARIATION, CV_I

For each data set simulated from models 3.1, 3.2, and 3.3 and with both $CV_A = (1/4)CV_I$ (Fig. 1) and $CV_A = CV_I$ (Fig. 2), CV_I values were estimated with the 3 CV methods. The CV-ANOVA model performed well for the estimation of CV_I with all simulated data.

ESTIMATION OF RCV LIMITS

Figs. 3, 4, 5, and 6 show comparisons of the RCV estimates with the true RCVs and lower and upper limits, respectively, for the 3 models investigated and with either $CV_A = (1/4)CV_I$ or $CV_A = CV_I$. For the RCVs, the true CV_A was used in the calculations for the ANOVA methods, and this gives a slight advantage to these methods over the nonparametric method. The ln-RCV method performed best if data were ln-normal distributed or the CV_I was less than approximately 12%, whereas the nonparametric method performed well for all distributions.

Discussion

SIMULATIONS

By use of computer simulations, assumptions of the data can be controlled and used to test the performance of methods on different distributions and experimental designs. At the same time, it is difficult to simulate data that mimic real life. The routinely used distributions such as normal and ln-normal are not really suitable in the sense that they go to infinity in one or both directions, and for computer simulations these distributions can easily generate extreme results, which in turn might lead to unreasonable conclusions. This happens rarely for small CVs but more frequently for larger CVs. Truncated distributions are more suited for simulating real-life data. These are available in statistical software such as the *msm* package for R *(19)*, or they can be simulated by use of the inverse transformation method *(12)* used in the present article.

CV_I ESTIMATION

The distribution of the within-person effect is important when estimating the CV_I. As can be seen in Figs. 1 and 2, the most robust method seems to be CV-ANOVA (Figs. 1C and 2C), as this method shows the least bias and narrowest range on the basis of the percentiles. This method should perform well for all types of distributions of individual effects. For evaluating the homogeneity of the individual CVs in the data set, data have to be standardized. By dividing each patient's results by that patient's mean, the individual variations in CVs can be estimated and then easily used to calculate the collective CV_I with an ANOVA on the normalized data. This makes the CV-ANOVA method suitable for the step-by-step approach of checking for outliers and homogeneity in replicates and samples. ln-ANOVA (Figs. 1B and 2B) performs the worst, except when measurement data are perfectly ln-normal, and should therefore be avoided for most data sets. Standard ANOVA has a wider range in estimates than CV-ANOVA and has no advantages over the CV-ANOVA method for estimating CVs.

The CV_G has no influence on CV_I estimation for the ln-ANOVA and CV-ANOVA methods; however, the standard ANOVA performed slightly worse with increased CV_G, as shown with CV_G of 10% and 40% (see Supplemental Figs. 1 and 2, which accompany the online version of this article at http://www.clinchem.org/content/vol62/issue5).

If using the standard ANOVA, one would still need to standardize data for the homogeneity and/or outlier check, so the first part of the calculations needed for the CV-ANOVA should be performed anyway to compare the normalized individual CVs.

RCV ESTIMATION

RCVs are influenced by both within-person effect and analytical error effect. These effects have a distribution with both a variance and a hidden shape, such as normal or ln-normal. The distribution of the individual measurements and the distribution of the difference between measurements are not equivalent. The first article on RCV used SDs and not CVs and also incorporated correlation between measurements *(7)*. Later, the calculations usually used CVs rather than SDs, without correcting for this change in the RCV formula, giving erroneous results. As Fokkema et al. showed *(15)*, ln-transformation gives correct RCVs in case of ln-distributed measurements, but for other distributions the method is not necessarily valid. The difference in percentage between observations will always have a skewed distribution, even for symmetrical normal distributions of the measurements *(8)*. If the difference is divided by the first observation, the distribution is skewed to the right.

From Figs. 3B, 4B, 5B, and 6B, it can be seen that for the ln-RCV method, the range of estimates grows wider for higher CVs. The bias stays fairly low for the mirrored ln-normal distribution when the CV_A is relatively small but becomes larger when $CV_A = CV_I$. From Fig. 1B and 2B, an even larger bias is to be expected, but although the CV_I is overestimated, the skewness of the distribution between samples is underestimated, and these errors counteract each other. The RCV is relatively larger for the case of $CV_A = CV_I$, so even if the range of values seem smaller in percentage in Figs. 4 and 6 compared with Figs. 3 and 5, it is actually larger, and this is in accordance with the comparison of Fig. 1 and 2. RCVs

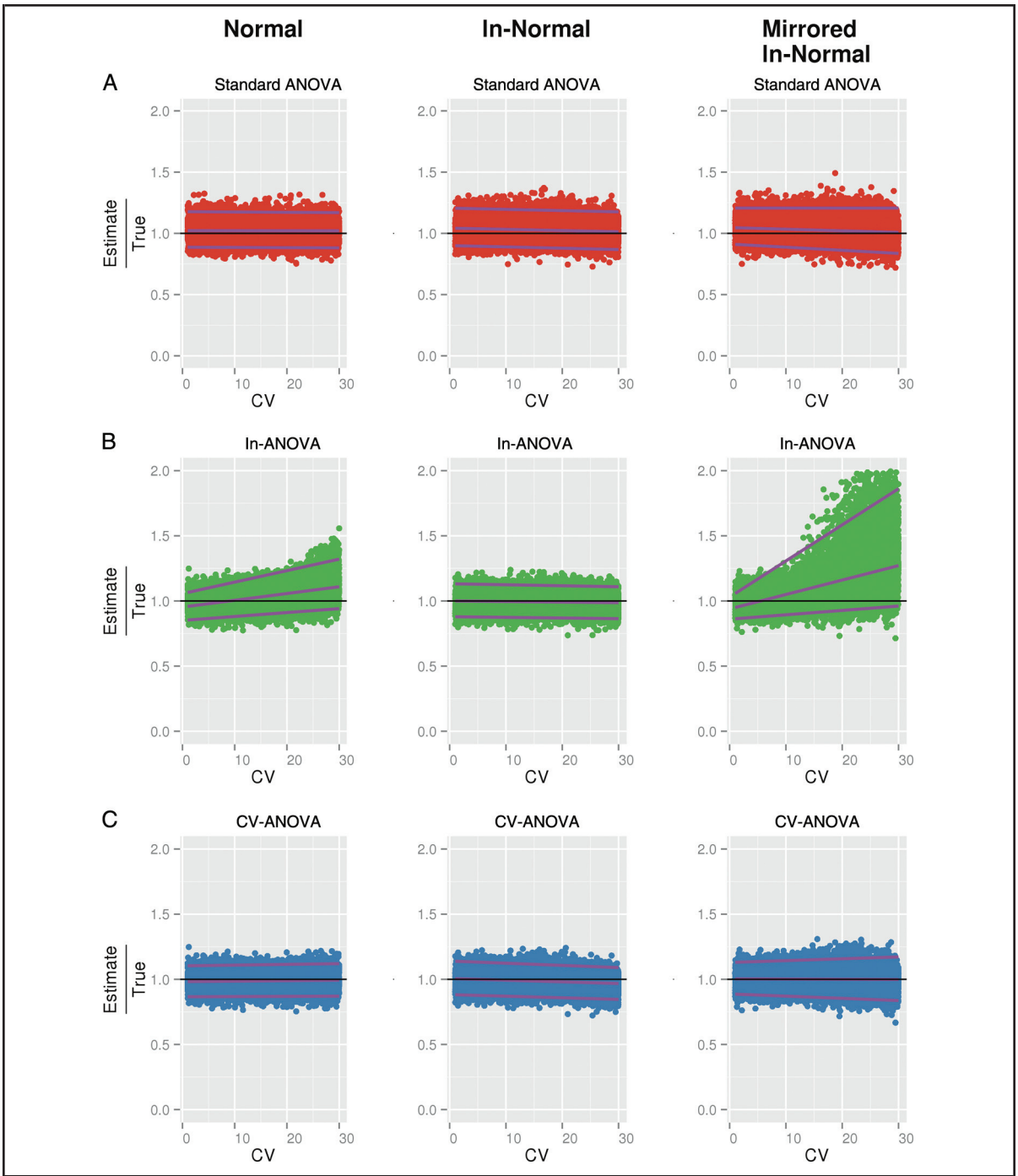

Fig. 1. Ratio between the estimated CV_I and the true CV_I when $CV_A = (1/4)CV_I$.

Standard ANOVA (A), ln-ANOVA (B), and CV-ANOVA (C) are used for estimation of the CV_I. Data are simulated with normal, ln-normal, and mirrored ln-normal distribution of the within-person effect. The 2.5%, 50%, and 97.5% quantiles show systematic error and the expected range of estimates.

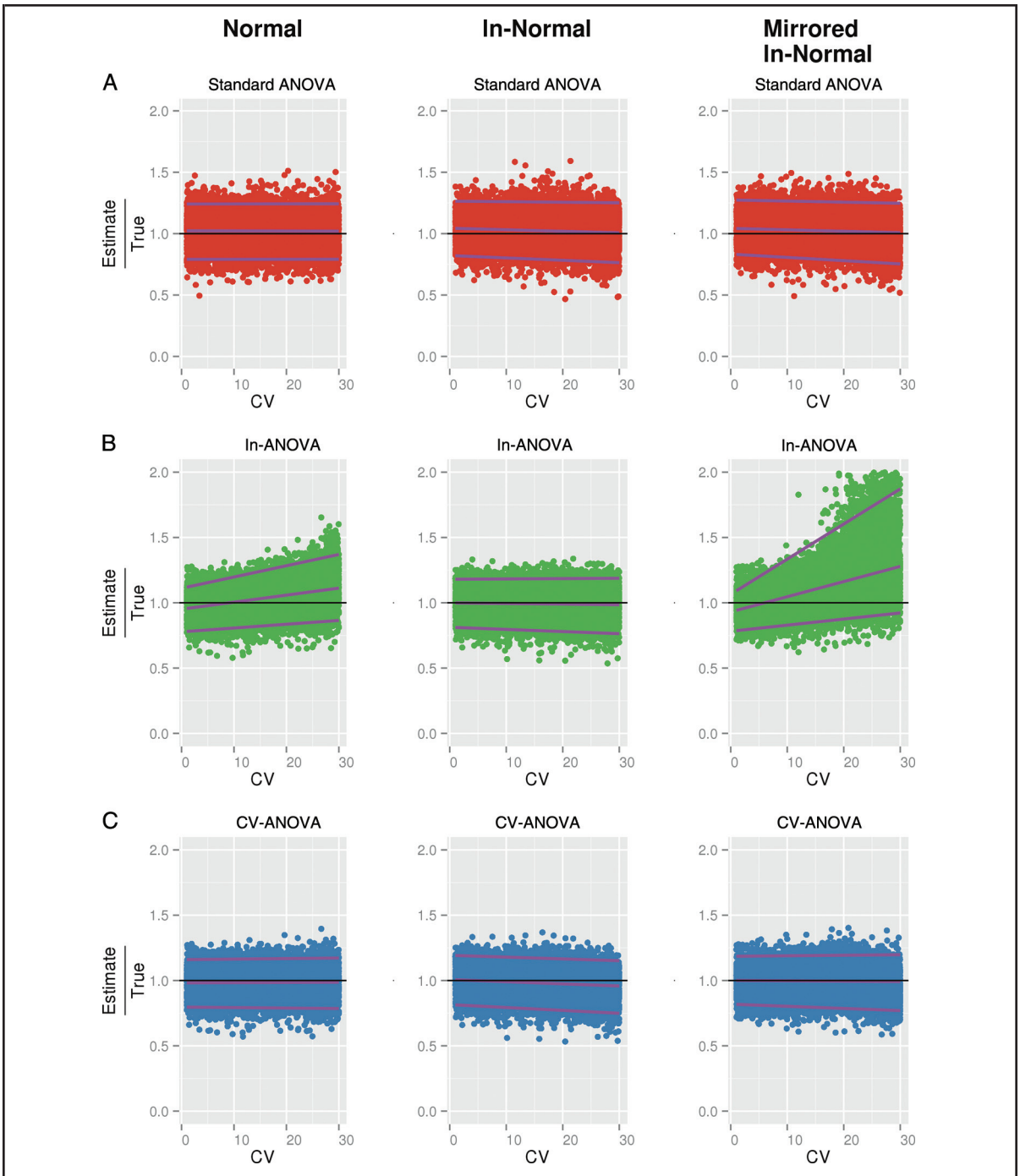

Fig. 2. Ratio between the estimated CV_I and the true CV_I when $CV_A = CV_I$.

Standard ANOVA (A), ln-ANOVA (B), and CV-ANOVA (C) are used for estimation of the CV_I. Data are simulated with normal, ln-normal, and mirrored ln-normal distribution of the within-person effect. The 2.5%, 50%, and 97.5% quantiles show systematic error and the expected range of estimates.

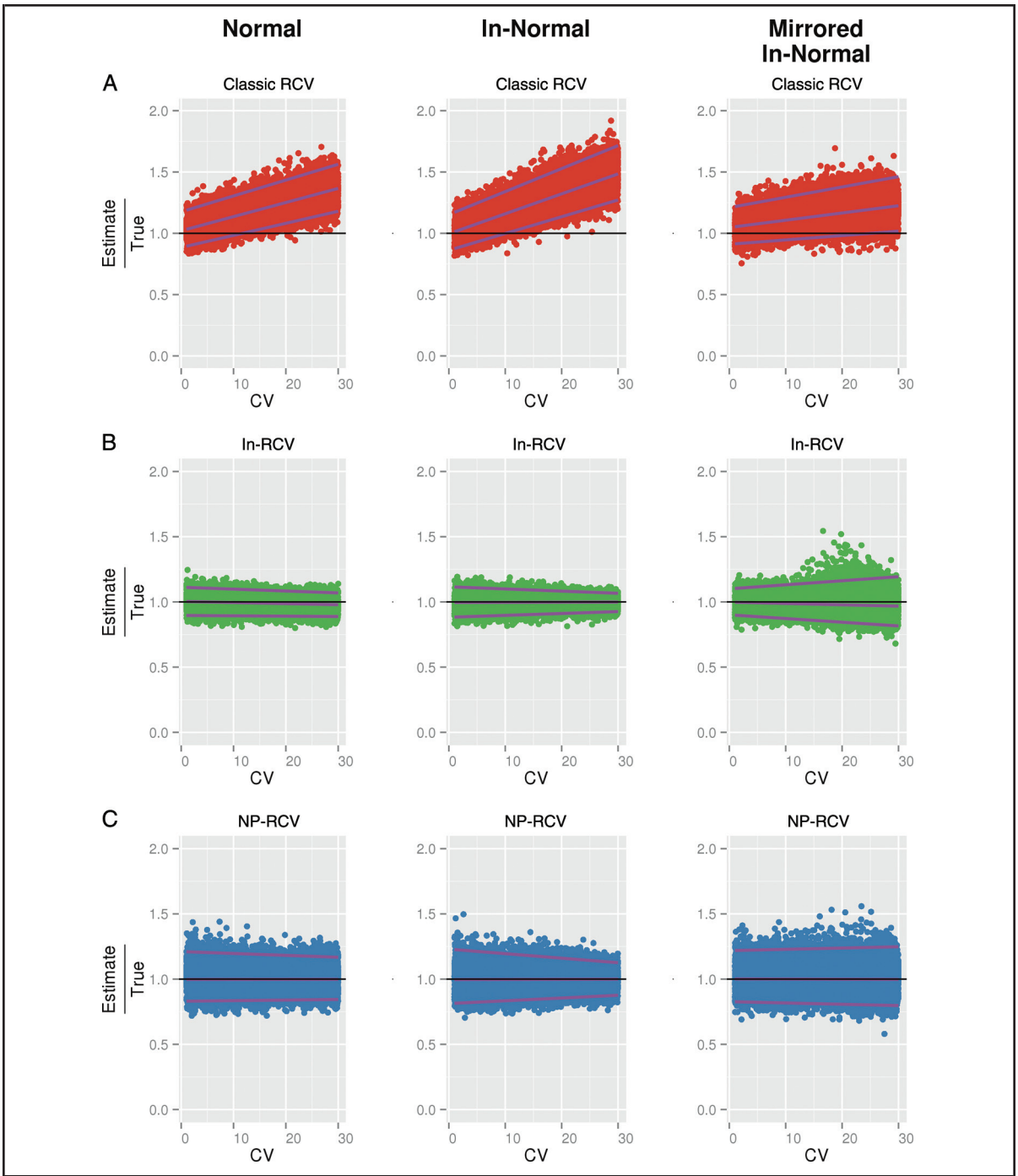

Fig. 3. Ratio between the estimated lower RCV and the true RCV when $CV_A = (1/4)CV_I$.

Classic RCV (A), In-RCV (B), and NP-RCV (C) are used for estimating the RCV. Data are simulated with normal, ln-normal, and mirrored ln-normal distribution of the within-person effect. The 2.5%, 50%, and 97.5% quantiles show the systematic error and expected range of estimates.

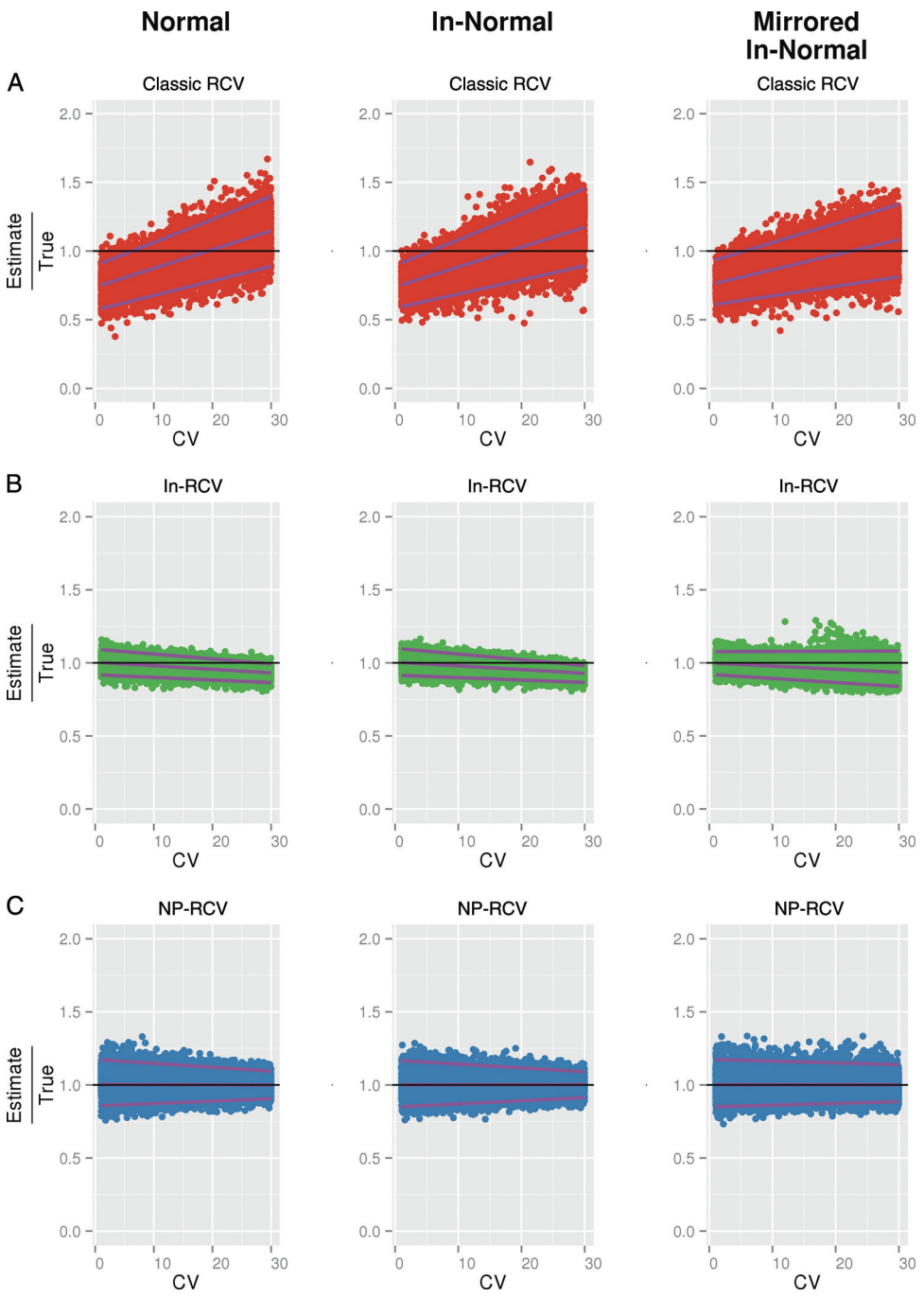

Fig. 4. Ratio between the estimated lower RCV and the true RCV when $CV_A = CV_I$.

Classic RCV (A), In-RCV (B), and NP-RCV (C) are used for estimating the RCV. Data are simulated with normal, ln-normal, and mirrored ln-normal distribution of the within-person effect. The 2.5%, 50%, and 97.5% quantiles show the systematic error and expected range of estimates.

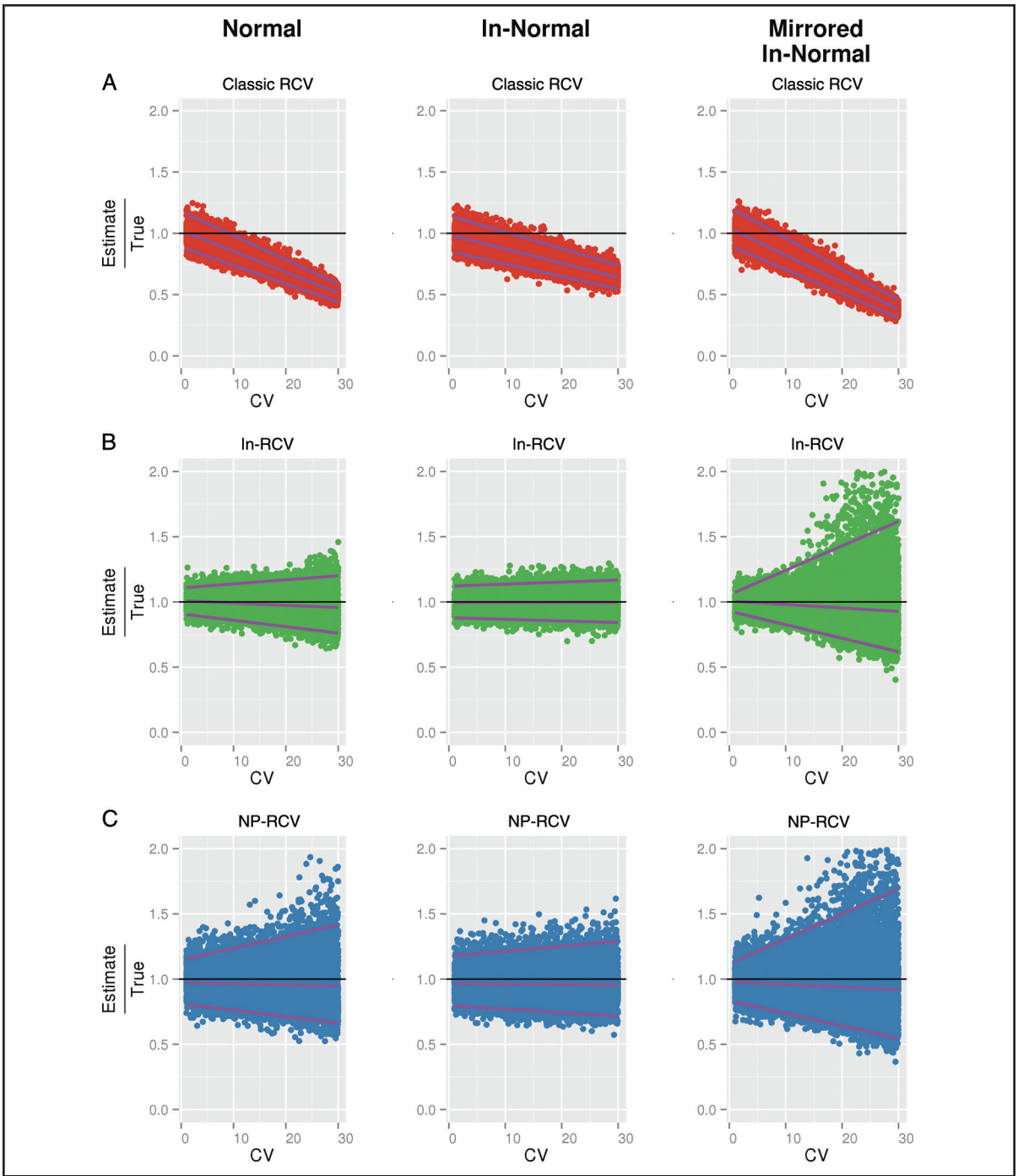

Fig. 5. Ratio between the estimated upper RCV and the true RCV when $CV_A = (1/4)CV_I$.

Classic RCV (A), In-RCV (B), and NP-RCV (C) are used for estimating the RCV. Data are simulated with normal, In-normal, and mirrored In-normal distribution of the within-person effect. The 2.5%, 50%, and 97.5% quantiles show the systematic error and expected range of estimates.

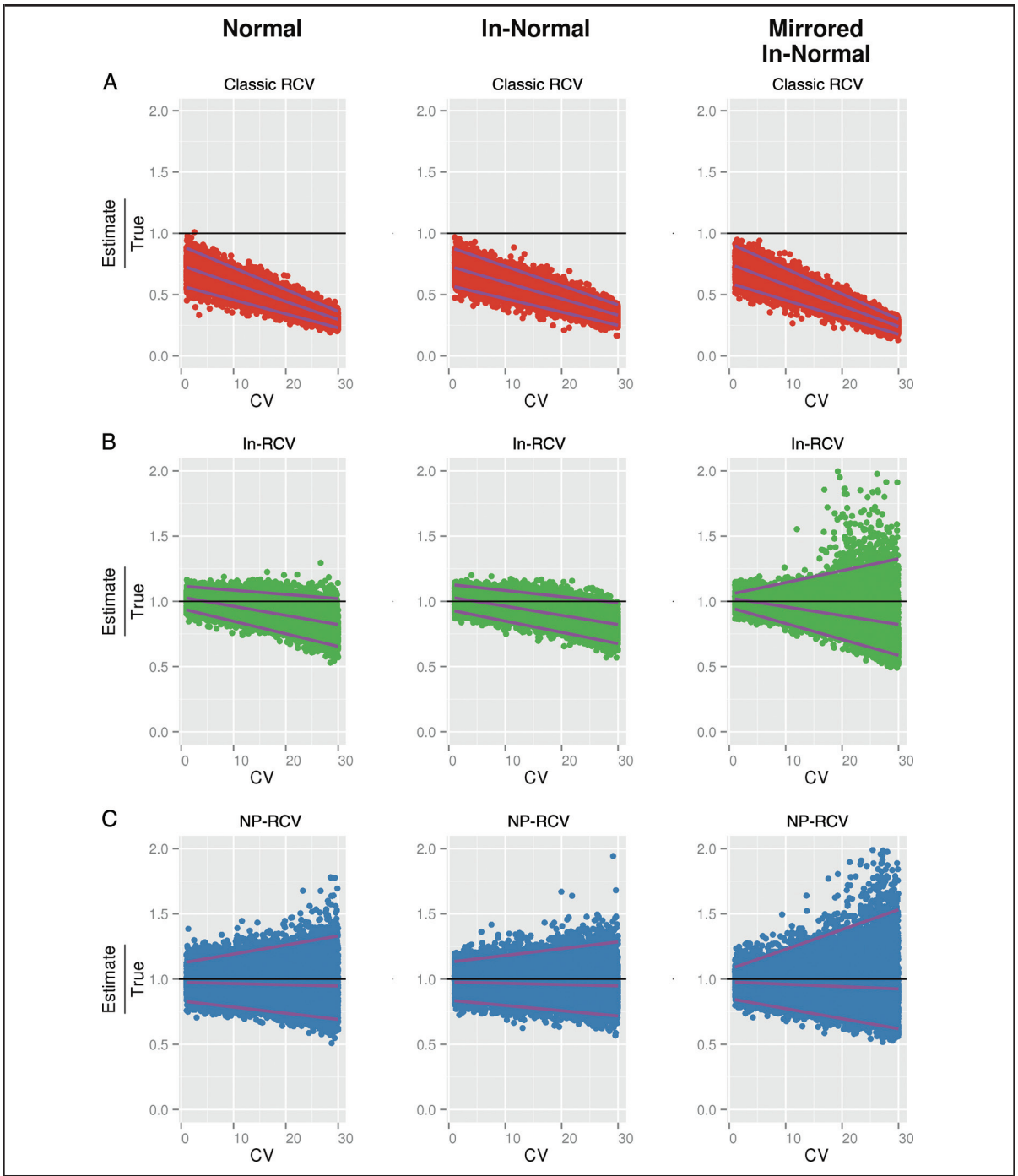

Fig. 6. Ratio between the estimated upper RCV and the true RCV when $CV_A = CV_I$.

Classic RCV (A), In-RCV (B), and NP-RCV (C) are used for estimating the RCV. Data are simulated with normal, ln-normal, and mirrored ln-normal distribution of the within-person effect. The 2.5%, 50%, and 97.5% quantiles show the systematic error and expected range of estimates.

for normally distributed measurements are in fact larger than RCVs for ln-normally distributed measurements with the same CVs, so Formula 4 will underestimate these.

CV_G has no influence on the RCV estimation for the ln-RCV and NP-RCV methods, as exemplified with simulations of CV_G 10% and 40% (see online Supplemental Figs. 3–6).

If the measurements can be assumed to be ln-normal distributed, or the CV_I is relatively small, the ln-RCV method is recommended, whereas the nonparametric method will give the best estimates of RCV for other distributions.

GENERAL NOTES

The current study was done to investigate the performance of commonly used methods for estimating within-person CV and RCV, with the goal of evaluating the degree of bias when researchers estimate both CV_I and RCV. The distribution of true values over time from 1 individual is dependent only on the within-person effect. This effect might follow a normal, ln-normal, or other distribution. The distribution of measurements from 1 individual is dependent on both the distribution of the within-person and analytical error effects. If the ratio between CV_A and CV_I is <0.25, the within-person component will dominate the shape of the distribution of measurements. For a higher ratio, the shape of the distribution will be more evenly influenced by the within-person effect and analytical error effect. As shown in an earlier publication *(20)*, the CV_I estimate is more uncertain with a relatively high CV_A, and this effect can also be seen when comparing Figs. 1 and 2.

For the estimates from the ANOVA model to be to be applicable and representative for the investigated population, the assumptions of homoscedasticity and a steady state for the individuals are important. Homoscedasticity is checked by the Bartlett test if the normality assumption holds *(21)*. If the model effects cannot be assumed to be normally distributed, homoscedasticity is better tested with, for instance, the Fligner–Killeen test *(21)*.

The models investigated in the present article are balanced. For balanced data, the ANOVA estimators are equivalent to the maximum likelihood and restricted maximum likelihood estimators. For unbalanced data, the ANOVA estimators might be more biased depending on the degree of unbalance and the distributions of the model effects.

The RCV, in the form generally used, is based on percentage difference between 2 measurements, and the distribution of these differences is not the same as the distribution of individual measurements. Isolating and estimating the individual distribution functions for the model effects is difficult if not impossible, and approximations are needed. As seen in Figs. 3–6, the ln-transformation for estimating the RCVs is sufficient for CV_I values below approximately 12% for the simulated data. However, the CV_I used in calculating this RCV is not the correctly estimated CV_I from the CV-ANOVA but the overestimated CV_I estimated by the ln-ANOVA as in Figs. 1 and 2.

If SDs are the measures of variability, which give homoscedasticity for each individual person's components, and the within-person and analytical error factors can be assumed to be normally distributed, the classic RCV formula is correct, but then the changes are not in percentages but in the unit of the measurand. When percentage change is computed, the variations should be measured as CVs, and the ln-normal distribution is more appropriate.

The use of ln-transformation has been recommended *(22)* as a standard method for nonsymmetrical distributions when calculating RCVs, but as shown in the present article, the question is more complicated, and one needs to be careful in making categorical recommendations. If the effects in the multiplicative model are ln-normal, the logarithm of this model is an additive model with normally distributed effects. The ln-ANOVA is perfectly suited for analysis of this model and for performing tests, constructing CIs, and calculating RCVs. The CVs, CIs, and RCVs can be transformed back to the original scale afterward. If the effects do not all follow a ln-normal distribution and the CVs are large (say, $>10\%$) specifically for the within-person and analytical error effects, the uncertainty of the estimated parameters is increased.

If the analytical component is small, it will have less influence on the estimation of CV_I, and also RCV. The influence of CV_G on the estimation of CV_I and RCV can be ignored, at least when CV_G is $<50\%$. If the distribution of measurements is a combination of different-shaped distributions and the ratio between CV_A and CV_I is larger—for example, when the ratio changes from 0.25 to 1.0—the estimation of RCVs is difficult, as it is generally difficult to estimate the individual effects. Because the percentiles for percentage changes do not have closed-form solutions, approximations are needed.

In the Ricos et al. database *(14)*, of the 370 included analytes, 189 (51%) have a $CV_I <12.5\%$; in this interval the 3 distributions evaluated in the present article are more or less identical and, therefore, the RCVs as well. However, 13% of the analytes have a $CV_I >30\%$. From these numbers and Figs. 3–6, it can be seen that it is very difficult to get good estimates of the RCVs, unless a transformation to a ln-normal distribution can be achieved or the data set is very large. A transformation of data to a normal distribution by use of the Box–Cox transformation is usually possible *(23)*. Thus a transformation to normality for the ln-transformed data might

exist, but this multistep procedure is neither appropriate nor recommended in most cases.

Although the advantages and disadvantages of using RCV as a tool in laboratories have been presented *(10, 24–26)*, the difficulty of estimating the correct RCVs should be added to the list of disadvantages.

It is important to evaluate the distribution of both the measurements and the changes between measurements, and for the investigator to find the method best suited for a specific set of data. This study suggests CV-ANOVA for the estimation of CVs. If the measurements can be assumed to be ln-normal distributed or the CV_I is less than approximately 12% for the models investigated, the ln-RCV method is recommended, whereas the nonparametric RCV method is recommended for other distributions. RCV limits, when working with CVs, are in general not symmetrical. Symmetrical RCV limits likely indicate faulty assumptions regarding the distribution of changes between observations. Ideally, larger data sets would allow for better estimations of the distributions, and therefore better estimates of the RCVs.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures or Potential Conflicts of Interest: *No authors declared any potential conflicts of interest.*

Role of Sponsor: No sponsor was declared.

References

1. Harris EK, Kanofsky P, Shakarji G, Cotlove E. Biological and analytic components of variation in long-term studies of serum constituents in normal subjects. Clin Chem 1970;16:1022–7.
2. Fraser CG, Harris EK. Generation and application of data on biological biological variation in clinical chemistry. Crit Rev Clin Lab Sci 1989;27:409.
3. Sahai H, Ojeda MM. Analysis of variance for random models. Vol. 1. Balanced data. Boston: Birkhäuser; 2004.
4. Burdick RK, Borror C, Montgomery D. Design and analysis of gauge R&R studies: making decisions with confidence intervals in random and mixed ANOVA models. ASA-SIAM Series on Statistical and Applied Probability. Philadelphia: ASA; 2005.
5. Burdick RK, Graybill F. Confidence intervals on variance components. 1st ed. Statistics: textbooks and monographs. Vol. 127. New York: Marcel Dekker; 1992.
6. Kristoffersen AH, Petersen PH, Sandberg S. A model for calculating the within-person biological variation and likelihood ratios for analytes with a time-dependent change in concentrations; exemplified with the use of D-dimer in suspected venous thromboembolism in healthy pregnant women. Ann Clin Biochem 2012;49:561–9.
7. Harris EK, Yasaka T. On the calculation of a 'reference change' for comparing two consecutive measurements. Clin Chem 1983;29:25–30.
8. Marsaglia G. Ratios of normal variables and ratios of sums of uniform variables. J Am Stat Assoc 1965;60:193–204.
9. Marsaglia G. Ratios of normal variables. J Stat Soft 2006;16:1–10.
10. Fraser CG. Reference change values. Clin Chem Lab Med 2012;50:807–12.
11. Harris EK, DeMets DL. Effects of intra- and inter-individual variation on distributions of single measurements. Clin Chem 1972;18:244–9.
12. Robert CP. Simulation of truncated normal variables. Stat Comput 1995;5:121–5.
13. Forbes C, Evans M, Hastings N, Peacock B. Statistical distributions. 4th ed. Hoboken (NJ): Wiley; 2011.
14. Westgard QC. Desirable biological variation database specifications. http://www.westgard.com/biodatabase1.htm (Accessed November 2015).
15. Fokkema MR, Herrmann Z, Muskiet FAJ, Moecks J. Reference change values for brain natriuretic peptides revisited. Clin Chem 2006;52:1602–3.
16. Koenker R. Quantile regression. 1st ed. Cambridge (UK): Cambridge University Press; 2005.
17. R: a language and environment for statistical computing. Vienna, (Austria): R Foundation for Statistical Computing.
18. The Standard: Standard C++. https://isocpp.org/std/the-standard (Accessed January 2016).
19. Jackson CH. Multi-state models for panel data: the msmPackage for R. J Stat Soft. 2011;38.
20. Roraas T, Petersen PH, Sandberg S. Confidence intervals and power calculations for within-person biological variation: effect of analytical imprecision, number of replicates, number of samples, and number of individuals. Clin Chem 2012;58:1306–13.
21. Fligner MA. Scale tests. Encyclopedia of statistical sciences. Hoboken (NJ): Wiley & Sons; 2004.
22. Klersy C, d'Eril GVM, Barassi A, Palladini G, Comelli M, Moratti R, et al. Advantages of the lognormal approach to determining reference change values for N-terminal propeptide B-type natriuretic peptide. Clin Chim Acta 2012;413:544–7.
23. Ichihara K, Boyd JC. An appraisal of statistical procedures used in derivation of reference intervals. Clin Chem Lab Med 2010;48:1537–51.
24. Plebani M, Lippi G. Biological variation and reference change values: an essential piece of the puzzle of laboratory testing. Clin Chem Lab Med 2012;50:189–90.
25. Fraser CG, Lippi G, Plebani M. Reference change values may need some improvement but are invaluable tools in laboratory medicine. Clin Chem Lab Med 2012;50:963–4.
26. Fuentes-Arderiu X, Padró-Miquel A, Rigo-Bonnin R. Disadvantages of using biological variation data for reference change values. Clin Chem Lab Med 2012;50:961, author reply 963–4.