This series of knowledge sharing articles is a project of the
Standardized Biofilm Methods Laboratory in the CBE

KSA-SM-14

Using *R* to calculate confidence intervals for the repeatability and reproducibility standard deviations and the intra-laboratory correlation coefficient of a disinfectant test method

[Key Words: confidence coefficient, parameter, point estimate, R, uncertainty assessment]

A collaborative study of a disinfectant test method is a multi-laboratory assessment of the method. It is conducted by replicate testing of the same disinfectant treatment in each of several laboratories. The goal is to assess the extent to which the test results are repeatable within a laboratory and reproducible among laboratory. The key summary statistics from a collaborative study are based on two response variables for a disinfectant test, the mean log density for untreated control carriers in a test (*TestLD*) and the log reduction measure of disinfectant efficacy (LR). For either of these response variables, investigators typically are interested in the following parameters: the repeatability standard deviation (σ_r), the reproducibility standard deviation (σ_R), and the intra-laboratory correlation coefficient (ρ), which also is the fraction of the total variance attributable to the variance among laboratories ($\rho = (\sigma_R^2 - \sigma_r^2) / \sigma_R^2$).

Comprehensive guidelines for the statistical analysis of a collaborative study are provided in Hamilton, et al. (2013). The terminology and mathematical notation used here are from previous Knowledge Sharing Articles, KSA-SM-03, KSA-SM-07, KSA-SM-10, and especially KSA-SM-13. In practice, the computationally intensive components of the formulas (KSA-SM-10) are calculated using a statistical computer package, such as R (R Development Core Team 2010). KSA-SM-13 presents instructions for using R to calculate point estimates (denoted by S_r, S_R, and r) of the parameters σ_r, σ_R, and ρ. The estimates are based on the variance components from fitting a one-factor, random effects linear model to the response, where the one random factor is "laboratory." The technique used to fit that model may be different for an unbalanced study than for a balanced study; regardless, the analysis is known as "analysis of variance" (ANOVA). A study is "balanced" when each of the L laboratories repeats the test exactly K separate times and a study is "unbalanced" when the number of repeats differs among laboratories, in which case the number of tests by laboratory ℓ is denoted by K_ℓ, $\ell = 1, ..., L$.

When applied to the *TestLD* response variable, the point estimates of σ_r, σ_R, and ρ are CS_r, CS_R, and CS_{lab}^2/CS_R^2, which are used to measure resemblance for control carriers. When applied to the LR response variable, the point estimates of σ_r, σ_R, and ρ are S_r, S_R, and S_{lab}^2/S_R^2, which are used to measure repeatability and reproducibility for the disinfectant test method (KSA-SM-03; KSA-SM-13).

This article shows how to use R for calculating a confidence interval (CI) estimate for each of σ_r, σ_R, and ρ. Numerical examples are provided. A review of CIs is presented in Appendix A. Various different statistical techniques (formulas or algorithms) are available for calculating CIs. After reviewing the statistical literature, we selected the CI techniques recommended by Burdick et al. (2006). Each of the recommended techniques has a sound theoretical foundation, is relatively simple to calculate, and has been evaluated via simulation studies. The strengths and limitations of the selected CI techniques are described in the Discussion section. As in KSA-SM-13, this article is applicable to quantitative and semi-quantitative laboratory disinfectant test methods (KSA-SM-02).

Using R to calculate confidence intervals for σ_r, σ_R, and ρ

Getting started

See Appendix A of [KSA-SM-13](#) for acquiring and installing [R](#) on your computer. Only the version of R running on the MS-Windows operating system is discussed here. Start R so you have an R window in which you can run commands and observe the output. Click on *File* at top-left in the window; then click on the menu item *Change dir...* to point R to the "working directory" you want to use. Note that R displays the prompt character > to indicate it is ready to run a command.

Examples

The examples discussed below require two data files ("`TestLD_TSM.txt`" and "`LR_NaOCl_TSM.txt`") which can be downloaded via [KSA-SM-14](#) on the Knowledge Sharing Articles webpage. If you plan to work through the examples, download these two data files and save them in R's working directory. If you want to analyze your own data, note that the CI calculations require that the data are in a tab-delimited text file. The first row of the data file should contain the column labels for all variables in the data file, including both the laboratory and the response. In this article, we are interested only in the responses *TestLD* and LR.

The example data sets used to illustrate the calculations come from the multi-lab study of the quantitative Three Step Method ([AOAC official method 2008.05 2008](#)) using spores of *Bacillus subtilis* on glass carriers ([Tomasino et al. 2008](#)). In that balanced study, all 8 labs performed the same number of tests, where 3 disinfectant treated carriers and 3 untreated control carriers were observed in each test. The "`TestLD_TSM.txt`" data file lists $K=9$ *TestLD* values for each of the $L=8$ laboratories. For the LR response, we shall use results from testing the medium efficacy level of the sodium hypochlorite treatment (NaOCl). The "`LR_NaOCl_TSM.txt`" data file lists $K=3$ LR values in each of the $L=8$ laboratories.

Let *Lo* and *Up* denote the lower and upper endpoints of a CI and let 1-α (or 100·[1-α]%) denote the associated confidence level; i.e., α is the specified CI error rate. Each of *Lo* and *Up* is calculated from the study data according to a formula that was derived by statisticians to provide an interval that will contain the true parameter value 100·(1-α)% of the time in the long run (when calculating the CI for each of many studies).

We converted the formulas into R code and made the R function *GetCIs* ([Appendix B](#)). When you read the list of commands, please realize that text following a hash (#) designates a non-executable comment; the comment does not continue to the next line unless the next line begins with a hash. Comments are inserted for documentation purposes, especially to note the BQI formula underlying each confidence limit calculation, where BQI in the code denotes [Burdick et al. (2006)](#).

Running the code

First, you must make *GetCIs* available from your working directory. To do that, copy the whole long block of R commands listed in [Appendix B](#), paste it beside the prompt in your R window; if you don't see another prompt after a short wait, hit Enter. Now the code is compiled and the functions (*GetCIs* and a couple of auxiliary functions) are now ready for your use. The examples below show how to run *GetCIs* . Basically, you have to provide the names of the laboratory and response columns in the data file. You can set α to any value between 0.0 and 0.5 via the third input variable (see examples below).

Analyzing TestLD

Download the data file `TestLD_TSM.txt` from the KSA webpage under KSA-SM-14. The data in "`TestLD_TSM.txt`" can now be imported into R as a data object called `d.TestLD` via the `read.table` command, as follows.

```
> # Read in the text file of the TestLD data
> d.TestLD = read.table("TestLD_TSM.txt",header=T)
```

To run an individual command, such as the command listed immediately above, you can either type the command (but not the `>`) into your R window or copy the command and paste it into the R window beside the prompt.

Now you can run the *GetCIs* function on these data.

```
> GetCIs(d.TestLD$Lab,d.TestLD$TestLD,alpha=0.10)
```

The output from *GetCIs* will be displayed in the R window. We recommend that you copy and paste the commands and results into a document file. It is good practice to maintain a statistical-diary document file to record the key analysis steps. When *GetCIs* is run on the data in `TestLD_TSM.txt`, the following output is printed in the R window:

```
Number of labs and the total number of tests:
8 72

First 3 and last 3 lines of data:
   Lab Response
1    1 6.919248
2    1 6.753710
3    1 6.832403
70   8 6.748208
71   8 6.639291
72   8 6.686105

Calculations performed with significance level alpha =
0.1

Number of tests in each lab:
1 2 3 4 5 6 7 8
9 9 9 9 9 9 9 9

Harmonic mean of the number of tests at each lab:
9

Lab means:
       1        2        3        4        5        6        7        8
6.848784 6.946420 7.251723 6.526638 6.999886 6.683945 6.956432 6.689980

Overall mean (=mean of lab means):
6.862976

Lab repeatability SDs
          1          2          3          4          5          6          7
8
0.08644766 0.06305877 0.14012780 0.19254704 0.22680672 0.08209550 0.23745318
0.04218228

Overall repeatability SD (pooled over labs)
0.1518651
```

```
Intermediate ANOVA calculations:
Mean Square Among Labs, Mean Square within labs, Var among labs, interlab correlation
0.463976 0.02306301 0.04899033 0.6799175

Intermediate calculations for approximate modified large sample method in BQI:
G1, G2, H1, H2
0.5023864 0.2351383 2.229751 0.3735407

CONFIDENCE INTERVALS

Overall mean response: Estimate, Lo, Up:
6.862976 6.710888 7.015064

Repeatability SD: Estimate, Lo, Up:
0.1518651 0.1328157 0.1779831

Reproducibility SD: Estimate, Lo, Up:
0.2684275 0.2137969 0.4327334

Intralab correlation coefficient: Estimate, Lo, Up:
0.6799175 0.480646 0.8790057
```

Note that the first part of the printout includes assurance that the correct data were analyzed by the function *GetCIs*. It shows that the analysis is based on the correct number of labs, the correct number of *TestLD* values for each lab, the desired α level, etc. This and other similar checks are prudent when doing statistical analysis (Hamilton, et al. 2013). The next part of the printout provides intermediate statistical calculations, including the mean of *TestLD* for each laboratory, the SD in each laboratory, and key components of the ANOVA. The final section of the printout contains the calculated *Lo* and *Up* endpoints of the 100(1-α)%, two-sided, CIs. Table 1 summarizes the results.

Table 1. Results for the response *TestLD*; collaborative study data of Tomasino et al. (2008). The notation for Estimate is as in the *GetCIs* code (Appendix B).

| Characteristic | Notation | | Results | | |
| | Parameter | Estimate | | 90% confidence interval | |
			Point estimate	Lo	Up
Overall mean	μ	ystar	6.8630	6.7109	7.0151
Repeatability standard deviation	σ_r	sr	0.1519	0.1328	0.1780
Reproducibility standard deviation	σ_R	sR	0.2684	0.2138	0.4327
Intra-laboratory correlation coefficient	ρ	r	0.6799	0.4806	0.8790

In words, for each individual characteristic of the *TestLD* response, one can be 90% confident that the parameter is between *Lo* and *Up*, as in the following two statements:
- The estimate of the mean *TestLD* (μ) is 6.86 and, with 90% confidence, μ is between 6.71 and 7.02.
- The estimate of the intra-laboratory correlation coefficient (ρ) is 0.68 and, with 90% confidence, ρ is between 0.48 and 0.88.

Alternatively, it is often informative to provide one-sided interpretations, as in the following two statements (note that a one-sided interval has a 95% confidence level when based on the *GetCIs* 90% CI endpoint *Up*):

- The estimate (CS_r)of the repeatability standard deviation of *TestLD* (σ_r) is 0.15 and, with 95% confidence, σ_r is no larger than 0.18.
- The estimate (CS_R) of the reproducibility standard deviation of *TestLD* (σ_R) is 0.27 and, with 95% confidence, σ_R is no larger than 0.43.

Analyzing LR

Now consider the results when the response is LR for the medium efficacy level of NaOCl ([Tomasino et al. 2008](#)). The LR values are in the text file "LR_NaOCl_TSM.txt", which is read into R with the commands:

```
> # Read in the text file of the LR data
> d.LR = read.table("LR_NaOCl_TSM.txt",header=T)
```

Run the *GetCIs* function on these data.

```
> GetCIs(d.LR$Lab,d.LR$LR,alpha=0.10)
```

When *GetCIs* is run on the data in LR_NaOCl_TSM.txt, the following output is printed in the R window:

```
Number of labs and the total number of tests:
8 24

First 3 and last 3 lines of data:
   Lab Response
1    1  3.69694
2    1  3.65784
3    1  4.14487
22   8  3.86668
23   8  4.05672
24   8  4.43604

Calculations performed with significance level alpha =
0.1

Number of tests in each lab:
1 2 3 4 5 6 7 8
3 3 3 3 3 3 3 3

Harmonic mean of the number of tests at each lab:
3

Lab means:
       1        2        3        4        5        6        7        8
3.833217 2.662877 4.042740 5.429273 4.345963 4.105833 2.808830 4.119813

Overall mean (=mean of lab means):
3.918568

Lab repeatability SDs
        1         2         3         4         5         6         7         8
0.2706068 0.2354332 0.4290818 0.3943742 0.3064353 0.9115946 0.3589679 0.2898763

Overall repeatability SD (pooled over labs)
0.4480642

Intermediate ANOVA calculations:
Mean Square Among Labs, Mean Square within labs, Var among labs, interlab correlation
```

```
2.302049 0.2007616 0.7004292 0.7772263

Intermediate calculations for approximate modified large sample method in BQI:
G1, G2, H1, H2
0.5023864 0.3915477 2.229751 1.009635

CONFIDENCE INTERVALS

Overall mean response: Estimate, Lo, Up:
3.918568 3.331803 4.505333

Repeatability SD: Estimate, Lo, Up:
0.4480642 0.3495051 0.635183

Reproducibility SD: Estimate, Lo, Up:
0.9493107 0.7156389 1.617874

Intralab correlation coefficient: Estimate, Lo, Up:
0.7772263 0.5249627 0.9286884
```

Table 2 summarizes the results.

Table 2. Results for the response LR when testing the medium efficacy level of NaOCl; collaborative study data of Tomasino et al. (2008). The notation for Estimate is as in the *GetCIs* code (Appendix B).

| Characteristic | Notation | | Results | | |
| | | | | 90% confidence interval | |
	Parameter	Estimate	Point estimate	Lo	Up
Overall mean	μ	ystar	3.9186	3.3318	4.5053
Repeatability standard deviation	σ_r	sr	0.4481	0.3495	0.6352
Reproducibility standard deviation	σ_R	sR	0.9493	0.7156	1.6179
Intralaboratory correlation coefficient	ρ	r	0.7772	0.5250	0.9287

For each individual characteristic of the LR response, one can be 90% confident that the parameter is between *Lo* and *Up*, as in the following two statements:
- The estimate of the mean LR (μ) is 3.92 and, with 90% confidence, μ is between 3.33 and 4.51.
- The estimate of the intra-laboratory correlation coefficient (ρ) is 0.78 and, with 90% confidence, ρ is between 0.53 and 0.93.

One-sided interpretations are shown in the following two statements:
- The estimate (S_r) of the repeatability standard deviation of LR (σ_r) is 0.45 and, with 95% confidence, σ_r is no larger than 0.64.
- The estimate (S_R) of the reproducibility standard deviation (σ_R) is 0.95 and, with 95% confidence, σ_R is no larger than 1.62.

Discussion

Burdick et al. (2006) discussed the advantages and disadvantages of the CI formulas that were programmed into the *GetCIs* code of Appendix B. They identified the types of unusual data sets for which these formulas provide flawed CIs and, for those cases, recommend alternative techniques. Their findings are summarized in this section.

When calculating a CI for the mean μ, the *GetCIs* technique is exact when the study is balanced, but an approximation when the study is unbalanced. In the latter case, the *GetCIs* technique does capture μ with a probability that is at least the nominal $100 \cdot (1-\alpha)\%$. For an unbalanced study when ρ is near zero, the CI may be wider than necessary to achieve the nominal confidence level; i.e., the CI is valid, but conservative.

When calculating a CI for σ_r, the *GetCIs* technique is exact even if the study is unbalanced.

When calculating a CI for σ_R, the *GetCIs* calculations are based on an approximate technique, even if the study is balanced. Statistical investigations have determined that the approximate technique maintains the nominal confidence level for a broad spectrum of studies and data. However, the true confidence level may be less than the nominal level when all of these conditions hold: (i) ρ is small, (ii) the study is extremely unbalanced, and (iii) some laboratories conducted just a single test (no repeats). Under these circumstances, the CI for σ_R generated by *GetCIs* is potentially incorrect; however, these circumstances seldom occur for collaborative studies of disinfectant tests. If you are faced with such data, Burdick et al. (2006) suggest reliable, but computationally intensive, alternative methods.

When calculating a CI for ρ, the *GetCIs* calculations are based on an approximation of an exact CI technique. The approximation is used because the exact technique is computationally intensive. Statistical investigations have determined that the approximate CI technique is reliable unless the study is very unbalanced or ρ is small. In those two cases, the approximate interval produced by *GetCIs* can be much too wide; i.e., the interval is valid, but very conservative.

For a balanced study, the R package VARCOMPCI potentially could be utilized to calculate the confidence intervals of this paper (Civit et al. 2013). Package VARCOMPCI is based on techniques similar to those recommended by Burdick et al. (2006). This investigation may be conducted in a future KSA.

Acknowledgments

References

AOAC International (2008) Official Method 2008.05: *Efficacy of Liquid Sporicides against Spores of Bacillus subtilis on a Hard Nonporous Surface Quantitative Three Step Method (First Action)*. AOAC International, Gaithersburg, MD

Burdick, R.K., Quiroz, J., and Iyer, H.K. (2006) The present status of confidence interval estimation for one-factor random models. *J Statist Planning and Inference* 136:4307-4325

Civit, S., Vilardell, M., Hess, A., Matthew, Z., Ge, Y.,Caballe, A. (2013) Package 'varcompci', (http://cran.r-project.org/web/packages/varcompci/varcompci.pdf)

KSA-SM-02 (2010) *Quantitative, semi-quantitative, quantal, and alternative methods*, Center for Biofilm Engineering at Montana State University, Bozeman (CBE Knowledge Sharing Articles)

KSA-SM-03 (2010) *Desirable attributes of a standardized method*, Center for Biofilm Engineering at Montana State University, Bozeman (CBE Knowledge Sharing Articles)

KSA-SM-07 (2010) *The log reduction (LR) measure of disinfectant efficacy,* Center for Biofilm Engineering at Montana State University, Bozeman (CBE Knowledge Sharing Articles)

KSA-SM-10 (2012) *Assessing resemblance, repeatability, and reproducibility for quantitative methods*, Center for Biofilm Engineering at Montana State University, Bozeman (CBE Knowledge Sharing Articles)

KSA-SM-13 (2013) *Using R to assess Resemblance, Repeatability, and Reproducibility for quantitative and semi-quantitative disinfectant test methods*, Center for Biofilm Engineering at Montana State University, Bozeman (CBE Knowledge Sharing Articles)

Hamilton, MA, Hamilton, GC, Goeres, DM, and Parker, AE (2013) Guidelines for the statistical analysis of a collaborative study of a laboratory method for testing disinfectant product performance. *J AOAC Int* 96(5):1138-1151

R Development Core Team (2010), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 (http://www.R-project.org)

Tomasino, S.F., Pines, R.M., Cottrill, M.P., and Hamilton, M.A. (2008) Determining the efficacy of liquid sporicides against spores of *Bacillus subtilis* on a hard nonporous surface using the quantitative three step method: collaborative study. *J. AOAC Int*. 91(4), 833-852

APPENDIX A: Brief review of confidence intervals

In the example, the analysis of LRs produced $S_R = 0.949$ as the point estimate of the reproducibility standard deviation (Table 2). It is understood that, if we conducted a new collaborative study of exactly the same test method and disinfectant treatment, the resulting S_R would be different from 0.949. Also, it is understood that the true parameter value σ_R is not equal to 0.949. Although 0.949 is our best statistical guess of the numerical value of σ_R, it is important to consider the uncertainty of that guess. The typical way to express that uncertainty is with a confidence interval (CI) estimate. The basic ideas are (i) an interval estimate can correctly capture the parameter whereas a point estimate has a negligible chance of being exactly correct and (ii) a narrower CI indicates less uncertainty in the point estimate.

As in the body of this article, *Lo* and *Up* denote the lower and upper endpoints of the CI and α denotes the chosen error rate so that $1-\alpha$ (or $100\cdot(1-\alpha)\%$) is the associated confidence level. It is understood that, due to chance, any particular study could produce misleading data and the associated interval estimate could miss the parameter. One can be $100\cdot(1-\alpha)\%$ "confident" that the calculated interval is correct for the study at hand. However, one must not forget that a study could be unlucky and the calculated CI could miss the true parameter value. Because *Lo* and *Up* are dependent on the data, they are random variables for which probabilities can be calculated. In the context of estimating the true reproducibility standard deviation σ_R, the task for the statistician is to derive formulas for those endpoints such that Prob{correct interval} = Prob{$Lo \leq \sigma_R < Up$} $\geq 1-\alpha$. A CI for a smaller α (e.g., $\alpha=0.01$) will be wider than a CI for a larger α (e.g., $\alpha=0.10$) when both intervals are calculated from the same data sets. That is, a wider CI is required to capture the true parameter 99% of the time than a CI required to capture the parameter only 90% of the time. Clearly, the formulas for *Lo* and *Up* are functions of the specified α value.

A CI formula also includes an assessment of the statistical uncertainty for the point estimate. For data producing great uncertainty, the CI will be quite wide. For example, consider two collaborative study designs: (i) 2 replicate tests in each of 4 laboratories and (ii) 6 replicate tests in each of 20 laboratories. Point estimates of the parameters for study (i) are affected by greater statistical uncertainty than for study (ii) and the CIs calculated for study (i) typically will be much wider than those for study (ii). Because of this association between uncertainty and width, the CI interval width is a convenient indicator of statistical uncertainty.

The technique used to find the CI is said to be "exact" if the calculated CI contains the true parameter value exactly $100\cdot(1-\alpha)\%$ of the time in the long run (applying the technique to many studies); i.e., Prob{$Lo \leq \sigma_R \leq Up$} = $1-\alpha$. It is not always possible for the statistician to derive formulas for *Lo* and *Up* that guarantee an exact interval. In that case, mathematical approximations are used to produce an "approximate CI." Statistical researchers determine the reliability and domain of applicability of the approximate interval technique.

When an interval misses the parameter, it is either completely below or completely above the parameter; that is, in the context of estimating σ_R, either $Lo < Up < \sigma_R$. or $\sigma_R < Lo < Up$. For the CI techniques presented in this article, the two types of error are equally likely, a fact that makes it easy to turn *Lo* or *Up* into a one-sided, $100\cdot(1-\frac{\alpha}{2})\%$ confidence limit. For example, one can ignore *Lo* and conclude with $100\cdot(1-\frac{\alpha}{2})\%$ confidence that $\sigma_R \leq Up$.

For more about CIs, see the on-line sources Wikipedia–Confidence Interval and *NIST/SEMATECH e-Handbook of Statistical Methods*.

APPENDIX B: R code for calculating confidence intervals for μ, σ_r, σ_R, and ρ

```
############ Beginning of R code ###################
# CONFIDENCE INTERVAL CALCULATIONS FOR A COLLABORATIVE STUDY OF
#   A LABORATORY DISINFECTANT TEST METHOD: via R function GetCIs
# Albert Parker & Martin Hamilton
# Center for Biofilm Engineering (CBE)
# Montana State University - Bozeman
# July, 2013
# Program published in the CBE Knowledge Sharing Article KSA-SM-14
# Analysis of the outcome variable for a collaborative (multi-lab) study
#  of a disinfectant test method. The study can be balanced or unbalanced.
#  Either TestLD or LR is the outcome (response) of interest in most analyses.
# In KSA-SM-13, the responses were fit to a one-factor, random effects model
#  using the REML method. However, the confidence interval (CI) formulas in this KSA
#  are based on ANOVA (or Method of Moments) calculations, following:
#  Burdick, R.K., Quiroz, J., and Iyer, H.K. (2006)
#  "The present status of confidence interval estimation for one-factor random models"
#  Journal of Statistical Planning and Inference 136:4307-4325. (called BQI below)
### OUTPUT
# Some data characteristics to be sure that the right data are being analyzed
# The means and standard deviations within each laboratory
# Some intermediate results that are useful for checking that the code is consistent
#   with the BQI formulas
# The point estimate and 2-sided 100(1-alpha)% CI is printed for:
##  overall mean response
##  repeatability SD
##  reproducibility SD
##  intra-lab correlation coefficient
### DATA FILE
# The inputs to this function are:
# Lab - a vector of lab identifiers
# Response - a vector of either TestLD or LR for the tests
# alpha – a number between 0.0 and 0.5; the default value of alpha is 0.10
# For example, if the text file CollabData.txt has two columns,
# one called "Lab" and the other "LR", and alpha = 0.05 is the desired error rate,
#   then execute this function by
# d.collab = read.table("CollabData.txt",header=T)
# GetCIs(d.collab$Lab,d.collab$LR, alpha=.05)
#####################################################################
######## start of function GetCIs for performing the confidence interval calculations
GetCIs <- function(Lab,Response,alpha=0.10)
{
alpha1 = alpha/2 #alpha1 is the error rate for a one-sided confidence interval
L = length(unique(Lab)) # L = No. of labs
N = length(Lab) # N = Total number of tests across the labs
lab.K = tapply(Response,Lab,length) #K in each lab
lab.mean = tapply(Response,Lab,mean) #mean for each lab
ystar = mean(lab.mean) # mean of lab means estimates overall mean
KH = 1/mean(1/lab.K) # harmonic mean of K per lab
dev = lab.mean-ystar #deviations of lab means from ystar
MSU = KH*sum(dev^2)/(L-1)  #BQI equation 11: MSamong using Unweighted SSamong
#
#ystar  approx CI BQI (12): Overall Mean
L12 = ystar - qt(1-alpha1,L-1)*sqrt(MSU/(L*KH))
U12 = ystar + qt(1-alpha1,L-1)*sqrt(MSU/(L*KH))
#
#sr  exact CI BQI (4): Repeatability SD
#Use ANOVA (MOM) MSE, not REML MSE
lab.var = tapply(Response,Lab,var)
lab.sd = sqrt(lab.var)
Kminus1 = lab.K - 1
MSE = sum(lab.var*Kminus1)/sum(Kminus1)
```

```
sr = sqrt(MSE)
L4 = sqrt(MSE/F1inf(1-alpha1,N-L))
U4 = sqrt(MSE/F1inf(alpha1,N-L))
#
#sR  approx CI BQI (14): Reproducibility SD
sR = sqrt((MSU/KH)+((KH-1)*MSE/KH)) # reproducibility SD using MSU
G1 = 1-Finf2(alpha1,L-1)
G2 = 1-Finf2(alpha1,N-L)
H1 = Finf2(1-alpha1,L-1)-1
H2 = Finf2(1-alpha1,N-L)-1
L14 = sqrt(sR^2 - (1/KH)*sqrt(((G1^2)*MSU^2)+((G2^2)*((KH-1)^2)*MSE^2)))
U14 = sqrt(sR^2 + (1/KH)*sqrt(((H1^2)*MSU^2)+((H2^2)*((KH-1)^2)*MSE^2)))
#
#r  Wald CI BQI (15): intraclass correlation coefficient
varA = max(0,(MSU - MSE)/KH)
r = varA/(sR^2)
L15 = (MSU/(KH*MSE*qf(1-alpha1,L-1,N-L)))-(1/min(lab.K))
L15 = L15/(1+L15)
L15 = max(0,L15)
U15 = (MSU/(KH*MSE*qf(alpha1,L-1,N-L)))-(1/max(lab.K))
U15 = U15/(1+U15)
### end of calculations
#
### Provide output:
d.collab = cbind(Lab,Response)
colnames(d.collab)=c("Lab","Response")
rownames(d.collab) = 1:length(Lab)
rc = dim(d.collab)
#
### Print data checks and analysis results
cat("Number of labs and the total number of tests:","\n")
cat(c(L,N), "\n")
cat("\n")
#
cat("First 3 and last 3 lines of data:", "\n")
st = as.integer(rc[1] - 2)
sf = as.integer(rc[1])
print(d.collab[c(1:3,st:sf),])
cat("\n")
#
cat("Calculations performed with significance level alpha = ","\n")
cat(alpha,"\n")
cat("\n")
#
cat("Number of tests in each lab:","\n")
print(lab.K)
cat("\n")
#
cat("Harmonic mean of the number of tests at each lab:","\n")
cat(KH,"\n")
cat("\n")
#
cat("Lab means:","\n")
print(lab.mean)
cat("\n")
#
cat("Overall mean (=mean of lab means):","\n")
cat(ystar,"\n")
cat("\n")
#
cat("Lab repeatability SDs","\n")
print(lab.sd)
cat("\n")
```

```r
#
cat("Overall repeatability SD (pooled over labs)","\n")
cat(sr,"\n")
cat("\n")
#
cat("Intermediate ANOVA calculations:","\n")
cat("Mean Square Among Labs, Mean Square within labs, Var among labs, interlab
correlation","\n")
cat(c(MSU, MSE, varA, r),"\n")
cat("\n")
#
cat("Intermediate calculations for approximate modified large sample method in
BQI:","\n")
cat("G1, G2, H1, H2","\n")
cat(c(G1,G2,H1,H2),"\n")
cat("\n")
#
cat("CONFIDENCE INTERVALS","\n")
cat("\n")
#
cat("Overall mean response: Estimate, Lo, Up:","\n")
cat(c(ystar, L12, U12),"\n")
cat("\n")
#
cat("Repeatability SD: Estimate, Lo, Up:","\n")
cat(c(sr,L4,U4),"\n")
cat("\n")
#
cat("Reproducibility SD: Estimate, Lo, Up:","\n")
cat(c(sR,L14,U14),"\n")
cat("\n")
#
cat("Intralab correlation coefficient: Estimate, Lo, Up:","\n")
cat(c(r,L15,U15),"\n")
cat("\n")
}
######## END OF GetCIs
#
#
######### NECESSARY FUNCTIONS
# These two quantile look-up functions are used in the formulas:
## F1inf ##
F1inf = function(p,df1)
  {
  qc2 = qchisq(p,df1)/df1
  qc2
  } # end of F1inf
## Finf2 ##
Finf2 = function(p,df2)
  {
  qc1 = df2/qchisq(1-p,df2)
  qc1
  } # end of Finf2
##################### End of R code #####################
```