

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

KSA-SM-13

Using R to assess Resemblance, Repeatability, and Reproducibility for quantitative and semi-quantitative disinfectant test methods

[Key Words: statistical software, log reduction]

In [KSA-SM-3](#) the eight desirable attributes of a disinfectant test method were identified. In [KSA-SM-10](#) the mathematical formulas to assess three of these attributes for quantitative and semi-quantitative methods were provided: *resemblance* of the untreated control data, *repeatability* of the response of interest across multiple tests, and *reproducibility* of the response across multiple laboratories. The three key response variables for a disinfectant test are the log density (LD) of organisms on each carrier, the mean LD of the untreated control carriers per test (*TestLD*), and the log reduction (LR) (KSA-SM-02). In this article, we demonstrate how to implement R (R Core Team 2010) package *nlme* (Pinheiro et al. 2009) to perform the appropriate statistical calculations. R is a free software environment for statistical computing and graphics. A future KSA will show how to perform these same calculations using the commercial software Minitab ([Minitab Statistical Software 2010](#)), which provides a ‘point-and-click’ interface, and excellent, easy-to-use, graphing capabilities.

This KSA has two parts. The first part shows how to analyze microbiological data generated from multiple laboratories in order to assess resemblance, repeatability and reproducibility of a method. This first part can be easily adapted to consider data generated by multiple technicians. The second part focuses on the more common situation where an analyst wishes to analyze data from multiple, independent experiments generated by a single lab.

Statistical calculations using the statistical programming language R

New users of R may find it beneficial to have prior computer programming experience. Use of R will require basic knowledge of statistical terminology. Appendix A of this article shows you how to install R on your computer for Windows. After installation of R, start R so you have an R window in which you can run commands and observe the output. All files referred to in the example below should be located in the R working directory on your computer. Appendix B of this article contains a short primer on setting the working directory and importing data into R.

When using R, the symbol `>` is a prompt in the R command window preceding the next R command. The symbol `#` indicates a comment that is not an executable command but is useful for documentation of the analyses. In this article, R commands and corresponding output are presented in `this typeface`; R commands are preceded by the prompt `>`, R output is not. As you read this document, copy and paste the R commands from this document, without the `>` character, directly into your R window. Each line of code must be pasted separately, and then hit enter to execute the code.

Analysis of a multi-laboratory study

Consider the multi-lab study of the quantitative three step method (TSM, AOAC official method 2008.05) using spores of *Bacillus subtilis* on glass carriers (Tomasino et al. 2008). The results from evaluations of resemblance, repeatability, and reproducibility for the TSM were already presented in Tables 2 and 3 of [KSA-SM-10](#). Here, we will show how R can be used to generate the results. The TSM data we will consider are in Appendices B and C of Tomasino et al. (2008). In this collaborative study, each of $L = 8$ labs performed $K = 9$ tests, with each test using $J = 3$ untreated control carriers and $I = 3$ treated carriers. The controls used in all 9 tests are comparable. That is, in all experiments across all labs, the untreated control carriers were manipulated in the same manner, including exposure to the same neutralizer. Consequently, the control data from all 9 tests are used in the resemblance analyses below. In these 9 tests, 3 different disinfectants were applied and so 3 LR values were calculated for each disinfectant treatment. Just as in the example in [KSA-SM-10](#), we will present repeatability and reproducibility analyses of the 24 LR values (3 tests in each of 8 labs) observed for a low efficacy glutaraldehyde treatment.

Resemblance across multiple labs

To import the TSM untreated control carrier log densities (LD, with units $\log_{10}(\text{CFU}/\text{carrier})$) from Appendix B of Tomasino et al. (2008), download the tab-delimited text file called `Controls_TSM.txt` from the CBE website at [KSA-SM-13](#). With this file in R's working directory, the control data can now be imported into R as a data object called `d.control` via the `read.table` command (Appendix B).

```
> # Read in a tab delimited text file of the control data
> d.control = read.table("Controls_TSM.txt",header=T)
```

It is prudent to list the data on the screen to check that it was properly imported; because of space limitations, we only show the first and last few lines that were displayed by R.

```
> d.control
  Lab Test  LD
1     1   1 6.916262
2     1   1 6.848245
3     1   1 6.993236
4     1   2 6.712650
.
.
.
213    8   8 6.556851
214    8   9 6.593588
215    8   9 6.692206
216    8   9 6.772522
```

Observe that the numerical lab identifiers #1-8 are listed under `Lab`, the numerical test (or experiment) identifiers #1-9 are listed under `Test`, and the LDs are listed under `LD`. To initiate the analysis, plot the control carrier LD versus the test number (the number indicates chronological order within a laboratory) by running each of the following commands.

```
> # Load the nlme package into R
> library(nlme)

> # Plot the control LDs to create Figure 1
```

```
> ControlsByLab <- groupedData(LD ~ Test | Lab, data = d.control)
> plot(ControlsByLab)
```

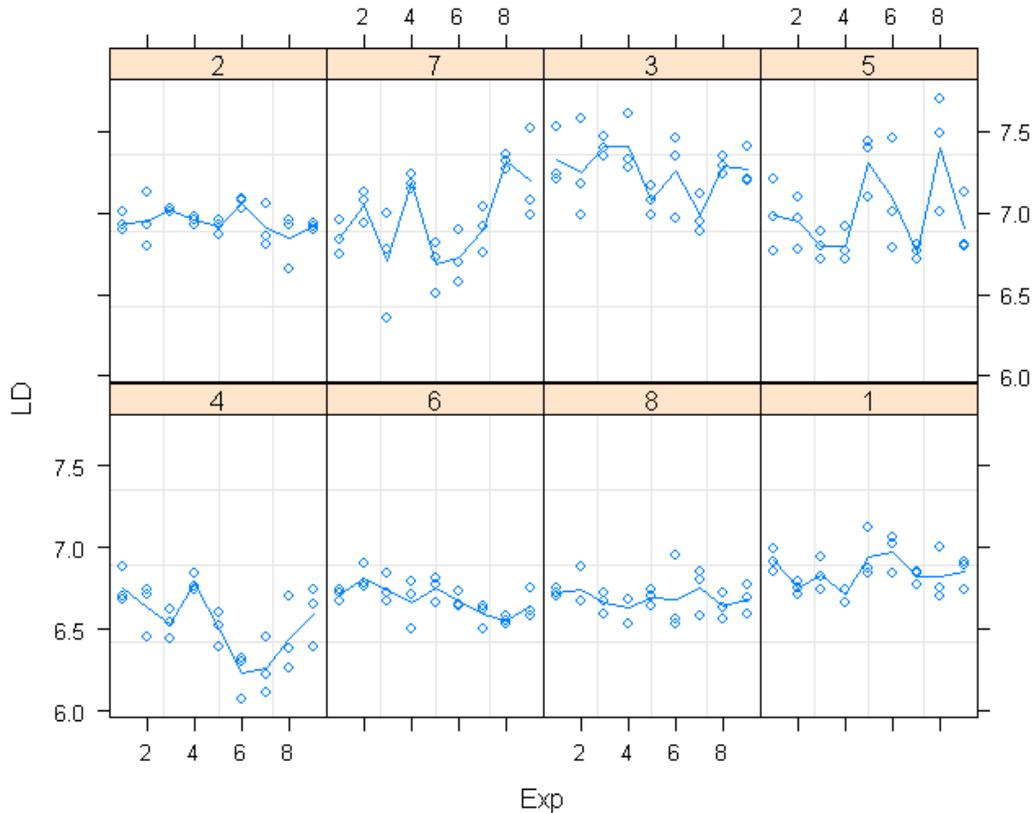


Figure 1. An individual value plot, generated by R, of the control carrier data from the 8-lab study of the TSM. Each point in the figure is the LD for an untreated control carrier. Each of the 8 panes in the figure shows the control data for a single lab, with the lab number shown in the heading at the top of the pane. The horizontal axis in each pane lists the test numbers. The vertical axis has units $\log_{10}(\text{CFU}/\text{carrier})$. The segmented lines in each pane connect the mean LD for each test across all test days for the lab.

The purpose of Figure 1 is not only to visualize the variability of the control LDs, but also to see whether there are systematic trends or cycles, which would indicate that uncontrolled factor(s) might be affecting resemblance of the controls. Observe that R ordered the panes by lab (4, 6, 8, 1, 2, 7, 3, 5) according to the magnitude of the LD values per lab.

The ANOVA with random effects due to lab and “test nested in lab” (described in detail in [KSA-SM-10](#)) is fit to the control LDs as follows.

```
> # Fit a nested, 2-way nested random effects ANOVA model
> m.control = lme(LD ~ 1, random=~1|Lab/Test, data=d.control)
> summary(m.control)
```

```
Linear mixed-effects model fit by REML
Data: d.control
      AIC      BIC    logLik
-101.6162 -88.13363 54.80809
```

```
Random effects:
```

```
Formula: ~1 | Lab
(Intercept)
StdDev: 0.2213407
```

```
Formula: ~1 | Test %in% Lab
(Intercept) Residual
StdDev: 0.1267794 0.1448101
```

```
Fixed effects: LD ~ 1
Value Std.Error DF t-value p-value1
(Intercept) 6.862976 0.0802763 144 85.49194 0
```

The standard deviation (SD) associated with each random effect is listed in the `lme` output after `Fixed effects:` shows that the overall mean of the control LDs across all labs is 6.862976. The `lme` output after `Random effects:`² displays SDs that are easily converted to variance components by squaring (cf. Table 2 of [KSA-SM-10](#)): $CS_{lab}^2 = 0.2213407^2 = 0.04899$, $CS_{test}^2 = 0.1267794^2 = 0.01607$, and $CS^2 = 0.1448101^2 = 0.02097$. R will do these variance calculations for you:

```
> VarCorr(m.control)
          Variance      StdDev
Lab =      pdLogChol(1)
(Intercept) 0.04899171 0.2213407
Test =      pdLogChol(1)
(Intercept) 0.01607303 0.1267794
Residual    0.02096996 0.1448101
```

As explained in [KSA-SM-10](#), it is of interest to consider the repeatability and reproducibility of the mean of the control LDs for each test (*TestLD*). The repeatability and reproducibility SDs for the *TestLD* can be calculated via equations (2) and (3) in [KSA-SM-10](#), respectively, by the following R code.

```
> # Calculate the repeatability and reproducibility SDs for TestLD
> varLD.output=VarCorr(m.control)
> varLD = as.numeric(varLD.output[c(2,4,5),1])
> CS2 = varLD[3]
> CS2_test= varLD[2]
> CS2_lab = varLD[1]
> L = m.control$dim$ngrps[2] # L = 8 labs
> K = m.control$dim$ngrps[1]/L # K = 9 tests of controls per lab
> J = m.control$dim$N / m.control$dim$ngrps[1] # 3 carriers per test

> CS_r = sqrt(CS2/J + CS2_test) # eqn (2) in KSA-SM-10
> CS_r
0.1518651

> CS_R = sqrt(CS2/J + CS2_test + CS2_lab) # eqn (3) in KSA-SM-10
> CS_R
0.2684302
```

¹ Ignore the DF and p-value, they are incorrect for this model!

² In the context of this example, “Residual” variability can be interpreted as coupon-to-coupon variability in each test.

Note that entering `CS_r` at the prompt `>` tells R to output the value 0.1518651 for the repeatability SD for the *TestLD*, CS_r (cf. example in [KSA-SM-10](#)). A practical interpretation of CS_r is that, in a given lab, the *TestLD* for a randomly chosen test typically differs from the true *TestLD* for that lab by 0.152.³ The reproducibility SD is $CS_R = 0.268$, indicating that the *TestLD* for a randomly chosen test in a randomly chosen lab typically differs from the true overall mean *TestLD* across all labs by 0.268.⁴ Using the acceptance criteria given in Table 1 of [KSA-SM-10](#), the quantitative TSM exhibits acceptable resemblance across multiple tests in a single lab, and acceptable resemblance across multiple labs.

One very useful summary of this analysis is to consider the contribution to the total variability of the TSM *TestLDs* (CS_R^2) from each source identified in the model:

```
> # Get the contribution to the variance of TestLD from each source
> c(CS2/J, CS2_test, CS2_lab) / CS_R^2
0.09700941 0.22306698 0.67992361
```

Thus, 68% of the variability of the *TestLDs* is due to among-lab sources (e.g., different geographic locations), 22% is due to among-test sources in each lab (e.g., differences in environmental conditions on each test day), and the final 10% of variance is due to among-carrier sources (e.g., differences in the way the bacteria were distributed across the carriers in each test). This distribution of contributions to the total variance is common for disinfectant test methods.

While an estimate for the true overall mean *TestLD* across all labs is 6.862976, a measure of the uncertainty of this estimate is the SEM, which R reported as 0.0802763 in the `lme` output above. In general, in a balanced study in which each of L labs conducts K tests, the SEM can be calculated using the following formula (Hamilton et al. 2013):

$$SEM = \sqrt{\frac{CS^2}{LKJ} + \frac{CS_{test}^2}{LK} + \frac{CS_{lab}^2}{L}}. \quad (1)$$

The R code for this SEM calculation is provided for illustrative purposes:

```
> # Corroborate R's overall TestLD SEM with the SEM calc. via eqn (1)
> SEM.control = sqrt(CS2/(L*K*J) + CS2_test/(L*K) + CS2_lab/L)
> SEM.control
[1] 0.08027629
```

To calculate a 95% two-sided confidence interval (CI) for the true overall *TestLD* across all labs, a t -distribution with $L - 1$ degrees of freedom should be used (Neter et al. 1996). For the TSM, $L - 1 = 7$. Use the following R commands:

```
> # The overall mean TestLD estimated by lme (REML)
> Mean.control = m.control$coef$fixed[1]
> Mean.control
(Intercept)
```

³ The interested reader can go back to Figure 1 and convince herself that the *TestLD* for each test in a lab is typically about 0.152 from the mean of all control LDs for that lab.

⁴ Again, the interested reader can go back to Figure 1, this time considering the distance between the mean *TestLD* for each lab and 6.86, our best estimate for the true overall mean of the control LDs.

```
6.862976
```

```
> # For a balanced study, the REML estimate is the arithmetic mean
> mean(d.control$LD)
[1] 6.862976
```

```
> # Two-sided 95% CI for the mean of the control carriers
> Mean.control + c(-1,1)* qt(.975,L-1)*SEM.control
[1] 6.673153 7.052799
```

Thus, we are 95% confident that the true mean *TestLD* for the TSM is between 6.67 and 7.05.

The above 95% CI for the overall mean *TestLD* is valid if the LDs are normally distributed and exhibit the same variance within each lab. The following R function generates diagnostic plots for evaluating the normality and homogeneity of variance assumptions, as well as for helping identify outliers (output not shown). To activate the function, copy and paste the following block of commands.

```
check.model <- function(m){
  # Put both diagnostic plots in one graphic
  par(mfrow=c(1,2))
  # Normal probability plot of the standardized residuals
  qqnorm(resid(m,type="p"))
  qqline(resid(m,type="p"))
  # Residual vs. fits plot to check for curvature & constant variance
  plot(fitted(m),resid(m,type="p"))
  abline(0,0) }
```

Now the assumptions for the model can be assessed by running the function in R (output not shown):

```
> # Check homogeneity of variance and normality assumptions
> check.model(m.control)
```

Repeatability and reproducibility of the LR across multiple labs

We will now use R to assess the repeatability and reproducibility of the log reductions (LR, a unitless quantity KSA-SM-02) observed in a collaborative study of a low efficacy glutaraldehyde sporicidal treatment applied via the TSM. First copy and paste the *Glut* rows of the data matrix in Appendix C of Tomasino et al. (2008) into a tab-delimited text file called *LR_Glut_TSM.txt*. Edit that text file so the first row contains the column labels

```
Lab   Chemical   Test Low   Medium   High
```

This data is also available at [KSA-SM-13](#). Import this data file into R by the following commands:

```
> # Read in the tab delimited text file of the LR data for Glut
> d.LR = read.table("LR_Glut_TSM.txt",header=T)
```

For this example, we analyze only the *LowLR* column of LR values. As we did for the control LD data, the LRs are displayed on the screen to check that they were properly imported; because of space limitations here, we only show the first and last few lines that R displayed.

```
> # Check data
```

```

> d.LR
  Lab Chemical Test   LowLR MediumLR HighLR
1   1   Glut   1 -0.02873  4.55688  6.05474
2   1   Glut   2  0.20799  4.80843  6.14700
3   1   Glut   3 -0.22884  4.35228  6.12070
4   2   Glut   1 -0.05047  5.52754  6.25082
.
.
.
21  7   Glut   3  0.41119  3.61482  5.82127
22  8   Glut   1 -0.00600  4.58671  5.66678
23  8   Glut   2  0.03544  4.85763  5.98350
24  8   Glut   3 -0.02767  5.26226  5.94032

```

We next plot the LowLR values. The results of the following R commands are presented in Figure 2. As in Figure 1, R has ordered the results according to the magnitude of the LRs at each lab.

```

> # Plot the LRs
> LR_ByLab <- groupedData(LowLR ~ Test | Lab,data=d.LR)
> plot(LR_ByLab)

```

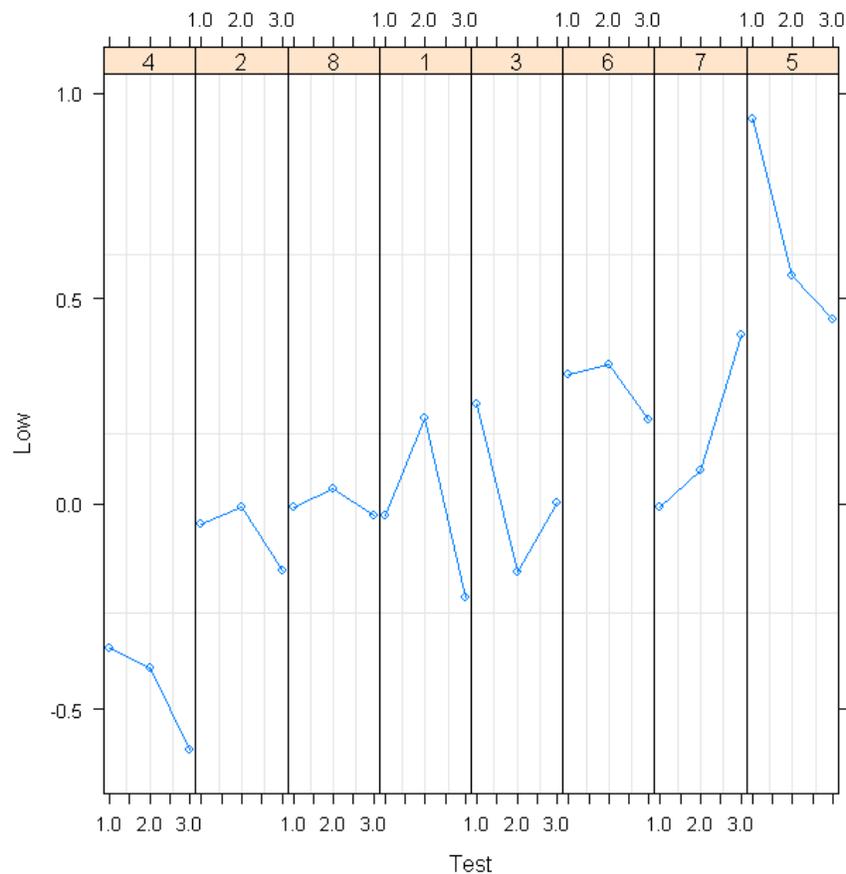


Figure 2. An individual value plot of the LowLR values. Each point in the figure is the LR for a single test. Each of the 8 panes in the figure shows the 3 LRs for a single lab, with the lab number shown in the heading at the top of the pane. The horizontal axis in each pane shows the experimental number. The vertical axis shows the magnitude of the LR. The lines in each pane connect the LRs across the three experimental days for the lab.

The repeatability and reproducibility of the LowLR values across all labs are assessed by an ANOVA with a random effect due to lab ([KSA-SM-10](#)). Enter the following R commands:

```
> # Fit a 1-factor random effects ANOVA model
> m.LR.Low = lme(LowLR ~ 1, random = ~1|Lab, data=d.LR)
> summary(m.LR.Low)
Linear mixed-effects model fit by REML
Data: d.LR
      AIC      BIC    logLik
9.495474 12.90196 -1.747737
```

```
Random effects:
Formula: ~1 | Lab
      (Intercept)  Residual
StdDev:   0.2989825  0.1712544
```

```
Fixed effects: LowLR ~ 1
              Value Std.Error DF   t-value p-value5
(Intercept) 0.07305625 0.1113365 16 0.6561752 0.521
```

```
> # Check the model assumptions: homogeneity of variance and normality
> check.model(m.LR.Low)
```

The `check.model` function must be loaded into memory as described in section *Resemblance across multiple labs*.

The repeatability SD is displayed as the Residual StdDev in the R output after Random effects; i.e., $S_r = 0.1712544$. Consequently, in a given lab, the LR for a randomly chosen TSM test typically differs from the true mean LR for that lab by 0.17.⁶

The R output after Random effects: can be used to calculate the results shown in Table 3 of [KSA-SM-10](#); $S_{lab}^2 = 0.2989825^2 = 0.08939051$ and $S_r^2 = 0.1712544^2 = 0.02932808$. In R:

```
> VarCorr(m.LR.Low)
Lab = pdLogChol(1)
      Variance  StdDev
(Intercept) 0.08939051 0.2989825
Residual    0.02932808 0.1712544
```

The reproducibility SD for the LR for the low efficacy level of glutaraldehyde can be calculated from equation (6) in [KSA-SM-10](#) by the following R code.

```
> # Get reproducibility SD of the LRs for low Glut
> varLR.output=VarCorr(m.LR.Low)
> varLR = as.numeric(varLR.output[c(1,2),1])
> S2_r= varLR[2]
> S2_lab = varLR[1]
```

⁵ Ignore the DF and p-value, they are incorrect for this model!

⁶ Consider the display of LRs in Figure 2. Are they typically about 0.2 from the mean LR for each lab?

```

> L = m.LR.Low$dims$ngrps[1] # L = 8 labs
> K = m.LR.Low$dims$N /L # K = 3 Low efficacy Glut tests per lab

> S_R = sqrt(S2_r + S2_lab) # eqn (6) of KSA-SM-10.
> S_R
[1] 0.3445556

> # Get the contribution to the total variance from each source
> c(S2_r,S2_lab)/S_R^2
[1] 0.2470386 0.7529614

```

Thus, the reproducibility SD is $S_R = 0.345$ which implies that the LR for a randomly chosen test in a randomly chosen lab typically differs from the true overall mean LR for all labs by 0.344.⁷

Using the acceptance criteria given in Table 1 of [KSA-SM-10](#), the quantitative TSM exhibited acceptable repeatability and reproducibility when testing a low efficacy level of glutaraldehyde against *Bacillus subtilis* spores on glass carriers. As is common for disinfectant test methods, the lab-to-lab component contributes most (75%) to the variability of the LRs.

Listed under `Fixed effects`: in the `lme` output above, an estimate of the overall mean LR is 0.07305625 with an SEM of 0.1113365. The general formula for the SEM across all L labs, when K tests are performed at each lab, is (Hamilton et al. 2013):

$$SEM = \sqrt{\frac{S_r^2}{LK} + \frac{S_{lab}^2}{L}}. \quad (2)$$

To calculate a 95% two-sided confidence interval for the true mean LR across all L labs, a t -distribution with $L - 1$ degrees of freedom should be used (Neter et al. 1996). For the TSM study, $L - 1 = 7$. A 95% CI for the mean LR across all labs is calculated by the following R commands:

```

> # A 95% two-sided CI for the true mean LR for low Glut
> 0.07305625 + c(-1,1) *qt(.975,L-1)*0.1113365
[1] -0.1902127 0.3363252

```

Thus, we are 95% confident that the true mean LR for the TSM using a low efficacy glutaraldehyde is between -0.19 and 0.34.

In instances where it is expected that a product will exhibit anti-microbial efficacy, it may be of interest to generate a lower, one-sided, 95% confidence limit for the true mean LR for that product. As an example, let us test (at a 5% significance level) whether the low efficacy level of glutaraldehyde is truly efficacious, on the average. In R:

```

> # A lower 95% one-sided confidence limit for the mean LR
> 0.07305625 - qt(.95,L-1)*0.1113365
[1] -0.1378795

> # p-value for upper one-tailed test for the mean LR
> 1-pt(0.07305625/0.1113365,L-1)

```

⁷ The interested reader may wish to visualize this in Figure 2.

```
[1] 0.2663397
```

Thus, we are 95% confident that the true mean LR is larger than -0.14 and the mean LR is not statistically significantly positive (p-value = 0.27).

Analysis of a study of multiple technicians at a single lab

The statistical approach just shown for analyzing microbial data generated by multiple labs can also be used to analyze data generated by multiple technicians at the same lab. Whereas in the first case it is of interest to estimate lab-to-lab and test-to-test variances when each lab performs repeated experiments, in the latter case one estimates tech-to-tech and test-to-test variances when each technician (or team of technicians) performs repeated experiments at the same lab. The tech-to-tech assessment is performed by following the approach just outlined after replacing “lab data” with “tech data.” Whereas it is common for the lab-to-lab variance to be large in a microbiological method, it is very desirable, even required, for the tech-to-tech variance of the method to be small compared to the other sources of variance.

Analysis of a single-laboratory study

Most analysts will only have access to data from a single laboratory. Even when analyzing a collaborative multi-lab study, it is recommended to separately consider resemblance and repeatability for each lab that participated in the study (Hamilton et al. 2013). From single-lab data, it is not possible to evaluate reproducibility of a method. Resemblance and repeatability can be assessed by analyzing data from multiple experiments all conducted at the single lab. To illustrate the calculations, we will consider the data generated by lab 2 in the TSM study.

Resemblance at a single lab

To evaluate resemblance of TSM control data generated at lab 2, we will analyze the control LDs from the $K = 9$ experiments conducted at lab 2 in the TSM study. The data from the 9 experiments at lab 2 are parsed and visualized by the following R commands:

```
> d.control.lab2=d.control[d.control$Lab==2, ]
> d.control.lab2
  Lab Test      LD
28  2     1 6.893257
29  2     1 7.003225
30  2     1 6.929744
31  2     2 6.792837
.
.
.
51  2     8 6.958434
52  2     9 6.916023
53  2     9 6.900069
54  2     9 6.937974

> # Graphically display the control data from lab 2
> par(mfrow=c(1,1))
> plot(LD ~ Test,data=d.control.lab2)
```

The output from the `plot` command is not shown since it is “pane 2” in Figure 1.

An ANOVA with a random effect due to test is fit to these data from lab 2 using the following R commands:

```
> m.control.lab2 = lme(LD ~ 1, random=~1|Test, data=d.control.lab2)
> summary( m.control.lab2)
Linear mixed-effects model fit by REML
Data: d.control.lab2
      AIC      BIC   logLik
-37.13653 -33.36224 21.56827

Random effects:
Formula: ~1 | Test
      (Intercept)   Residual
StdDev:  0.03124095 0.09487488

Fixed effects: LD ~ 1
              Value Std.Error DF   t-value p-value8
(Intercept) 6.94642 0.0210196 18 330.4735      0

> # Check the model assumptions: homogeneity of variance and normality
> check.model(m.control.lab2)

> # Get repeatability SD of the TestLD for lab 2
> varLD.lab2.output=VarCorr(m.control.lab2)
> varLD.lab2 = as.numeric(varLD.lab2.output[c(1,2),1])
> CS2.lab2 = varLD.lab2 [2]
> CS2_test.lab2 = varLD.lab2[1]
> K = m.control.lab2$dim$ngrps[1] # K = 9 tests of controls at lab2
> J = m.control.lab2$dim$N / K # J = 3 control carriers per test
> CS_r.lab2 = sqrt(CS2.lab2/J + CS2_test.lab2)
> CS_r.lab2
[1] 0.06305879

> # Get contribution to lab 2 TestLD variance due to different sources
> c(CS2.lab2/J, CS2_test.lab2)/ CS_r.lab2^2
0.7545533 0.2454467
```

The previous R output shows that the repeatability SD of the *TestLD* at lab 2 is 0.06. Consequently, the *TestLD* in a typical test from lab 2 differs from the true mean *TestLD* for lab 2 by 0.06. This is excellent repeatability (cf. lab 2 data in Figure 1)! Lab 2's repeatability SD is smaller than the repeatability SD estimated from all 8 labs, $CS_r = 0.15$. Within-test variability among carriers is the major contributor (75%) to the total variability of the control LDs at lab 2.

The output after Fixed effects: in the lme output above shows that the mean *TestLD* at lab 2 is 6.94642 and the SEM is 0.0210196. The SEM for the mean *TestLD* across all K tests at a single lab when each test has J control carriers can be calculated using the following formula (Hamilton et al. 2013):

⁸ Ignore the DF and p-value, they are incorrect for this model!

$$SEM = \frac{CS_r}{\sqrt{K}} = \sqrt{\frac{CS^2}{KJ} + \frac{CS_{test}^2}{K}}. \quad (3)$$

To calculate a 95% two-sided confidence interval for the true mean *TestLD* at a single lab, a *t*-distribution with $K - 1$ degrees of freedom should be used (Neter et al. 1996). For lab 2 (or any other lab in the TSM study), $K - 1 = 8$. A 95% CI for the true mean *TestLD* at lab 2 is calculated by the following R commands:

```
> # A 95% two-sided CI for the true mean TestLD for lab 2
> 6.94642 + c(-1,1) *qt(.975,K-1)*0.0210196
[1] 6.897949 6.994891
```

Thus, we are 95% confident that the true mean TSM *TestLD* at lab 2 is between 6.90 and 6.99.

Repeatability of the LR at a single lab

Next consider the repeatability of the LRs for a low efficacy glutaraldehyde treatment for TSM experiments conducted at lab 2. These LRs are accessed in R by the following commands:

```
> d.LR.lab2=d.LR[d.LR$Lab==2, ]
> d.LR.lab2
  Lab Chemical Test   LowLR MediumLR HighLR
4   2      Glut   1 -0.05047  5.52754 6.25082
5   2      Glut   2 -0.00839  5.27848 6.11558
6   2      Glut   3 -0.16111  4.91897 6.20881

> # Plot the LRs by experimental day
> plot(LowLR ~ Test,data=d.LR.lab2)
```

The output from the preceding commands make clear that there were only $K = 3$ experiments at lab 2 that used the low efficacy level of glutaraldehyde. The graph from the `plot` command is not shown since it is included as “pane 2” in Figure 2. Repeatability of the LRs for the low efficacy level of glutaraldehyde at lab 2 is assessed by simply taking the SD of the K LRs (cf. equation (5) of [KSA-SM-10](#)):

```
> # Repeatability of the TSM LRs for low Glut at lab 2
> sd(d.LR.lab2$LowLR)
[1] 0.07888318
```

Thus, the repeatability SD at lab 2 is 0.07 and the typical LR in lab 2 differs from the true mean LR for lab 2 by 0.07 (cf. Figure 2).

The lab 2 mean LR for the low efficacy glutaraldehyde is found by taking the arithmetic mean of the K LRs.

```
> mean(d.LR.lab2$LowLR)
[1] -0.07332333
```

In general, the SEM of the mean LR across all K tests at a single lab is:

$$SEM = \frac{S_r}{\sqrt{K}}. \quad (4)$$

To calculate a 95% two-sided CI for the true mean LR at a single lab, a t -distribution with $K - 1$ degrees of freedom should be used. For lab 2 efficacy data (or any other lab in the TSM study), $K - 1 = 2$. A 95% CI for the lab 2 mean LR for the low efficacy glutaraldehyde is calculated in R as follows:

```
> # A 95% two-sided CI for the true mean LR at lab 2 for low Glut
> -0.07332333 + c(-1,1) *qt(.975,2)*0.07888318/sqrt(3)
[1] -0.2692800  0.1226334
```

Hence, we are 95% confident that the true lab 2 mean LR for a low efficacy glutaraldehyde tested according to the TSM is between -0.27 and 0.12. These are rudimentary t -distribution calculations, R will do them for us:

```
> t.test(d.LR.lab2$LowLR)

      One Sample t-test

data:  d.LR.lab2$LowLR
t = -1.61, df = 2, p-value = 0.2487
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.2692800  0.1226333
sample estimates:
 mean of x
-0.07332333
```

In instances where it is expected that a product will exhibit anti-microbial efficacy, it may be of interest to generate an upper one-sided CI for the true mean LR at a single lab. Equivalently, one can test whether the true mean LR at the one lab is positive. In R, we calculate a one-sided CI and perform the hypothesis test for the low efficacy glutaraldehyde treatment applied via the TSM to spores at lab 2:

```
> t.test(d.LR.lab2$LowLR,alternative="greater")

      One Sample t-test

data:  d.LR.lab2$Low
t = -1.61, df = 2, p-value = 0.8757
alternative hypothesis: true mean is greater than 0
95 percent confidence interval:
 -0.2063089      Inf
sample estimates:
 mean of x
-0.07332333
```

As suggested by the results above, the low efficacy glutaraldehyde treatment applied by lab 2 via the TSM does not exhibit statistically significant kill on the average. In other words, lab 2's mean LR is not statistically significantly positive. Moreover, with 95% confidence the true mean LR at lab 2 is as small

as -0.21. These conclusions regarding the mean LR for this TSM treatment applied only at lab 2 are similar to the conclusions based on an analysis of all labs pooled together.

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.
- 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(4).
- Minitab Statistical Software (2010) Minitab Inc., State College, PA (<http://www.minitab.com>)
- Neter, J., Kutner, M.H., Wasserman, W., and Nachtsheim, C.J. (1996) *Applied Linear Statistical Models, 4th Ed.*, McGraw-Hill, Boston.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2012) *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-104
- R Development Core Team (2012), *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.

Version date: 7 August 2013
Authors: Albert E. Parker, parker@math.montana.edu
Martin A. Hamilton, mhamilton@biofilm.montana.edu

APPENDIX A: Obtaining and Installing R

1. Go to: <http://cran.r-project.org>. This is the “official” site of the The Comprehensive R Archive Network (CRAN). Bookmark this address. Lots of information (manuals, answers to frequently asked questions, etc) can be downloaded from this site.
2. The first box on this page is labeled *Download and Install R*. In that box, click on the appropriate link. For example, MAC users will click on [Download R for \(Mac\) OS X](#) and Microsoft Windows users will click on the link [Download R for Windows](#). The rest of the instructions in this article are specific to Windows users.
3. On the new page, click on the link [base](#).
4. On the new page, click on the link [Download R 3.0.1 for Windows](#). This will download an exe file to the hard drive on your computer.
5. Exit from your Internet Browser and open Windows Explorer. Go to the folder in which you saved the exe file run the program.
6. You will be guided through the installation by a Setup Wizard.

There are many excellent resources for using R. One interactive site is at <http://www.math.csi.cuny.edu/Statistics/R/simpleR/>, called “Simple R” by John Veranzi.

Special-purpose software routines are bundled as separate “packages.” Some packages are automatically downloaded when base R is downloaded. To download additional packages, run R on your PC and then click on the tab **Packages** at the top of the screen. From the drop down menu, click on **Install package(s) ...** and then choose the package(s) that you want to download. The package that you will need to download for this article is *nlme*.

APPENDIX B: Entering Data into R

Consider the TSM control data considered in this article, which are included in a text file `Controls_TSM.txt`. The first few lines of this data file are shown next:

```
Lab  Test  LD
1    1     6.916262
1    1     6.848245
1    1     6.993236
1    2     6.71265
1    2     6.754626
```

The first line of the file, called the *header*, contains *variable names*. Text data files that are tab or space delimited can be imported into R. This means that the names of the variables in the file cannot have spaces in them (e.g. don't use `Lab Number` in the first line of the file). To get `Controls_TSM.txt` into R, execute the following command:

```
> d.control = read.table("Controls_TSM.txt",header=T)
```

`read.table` is an R *function*, and the parameter `header=TRUE` tells R that the first line of the file contains the variable names of each of the columns of data. You could end up with an error like:

```
Error in file(file, "r") : unable to open connection
In addition:
Warning message: cannot open file `Control_TSM.txt'
```

The above error occurred because `Controls_TSM.txt` was not in the *working directory*. To change the working directory to the one where `Controls_TSM.txt` resides, in R, click on tab **File**→**Change dir ...** and you will see a **Browse for Folder** window appear. In this window, you can directly enter the directory that contains `Control_TSM.txt` on your computer, or you can **Browse** to find the directory. Once you select the directory that contains `Control_TSM.txt`, then click **OK** in the **Browse for Folder** window. Now try to read the data into R again.

```
> d.control = read.table("Controls_TSM.txt",header=T)
```

The R-variable `d.control` which contains the data is called a *data frame*. We could have used any variable name like `TSMControlData` or `TSM.controls`, but we don't like to type much, so we used `d.control`. Note that you cannot have spaces in your R-variable names! Type the variable name at the R prompt to see what the data look like:

```
> d.control
```

To access the individual columns of the data in `d.control`, type

```
> d.control$Lab
 [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2
[38] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3
[75] 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 5 5 5
[112] 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 6 6 6 6 6 6 6 6 6
[149] 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
[186] 7 7 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
```

```

> d.control$Test
 [1] 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6 7 7 7 8 8 8 9 9 9 1 1 1 2 2 2 3 3 3 4
 [38] 4 4 5 5 5 6 6 6 7 7 7 8 8 8 9 9 9 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6 7 7
 [75] 7 8 8 8 9 9 9 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6 7 7 7 8 8 8 9 9 9 1 1 1
 [112] 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6 7 7 7 8 8 8 9 9 9 1 1 1 2 2 2 3 3 3 4 4 4 5
 [149] 5 5 6 6 6 7 7 7 8 8 8 9 9 9 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6 7 7 7 8 8
 [186] 8 9 9 9 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6 7 7 7 8 8 8 9 9 9

> d.control$LD
 [1] 6.916262 6.848245 6.993236 6.712650 6.754626 6.793854 6.738492 6.820560
 [9] 6.938156 6.662758 6.739572 6.740865 7.125215 6.842098 6.871626 6.846281
 [17] 7.058219 7.026422 6.853256 6.845718 6.771253 6.753513 7.005102 6.700389
 .
 .
 [193] 6.668555 6.668894 6.883868 6.722634 6.667283 6.593890 6.528917 6.684682
 [201] 6.677939 6.641113 6.745784 6.701254 6.530200 6.561643 6.955557 6.802401
 [209] 6.855739 6.586485 6.634294 6.726727 6.556851 6.593588 6.692206 6.772522

```

R is case-sensitive! The upper and lower-case letters in the variable name must be EXACTLY as given in the data file or R will not find it. For example,

```

> d.control$lab
NULL

```

In addition to `read.table`, we will be using a few other functions that R has available. For example, `mean()` calculates the mean and `median()` calculates the median. For example:

```

> mean(d.control$LD)
[1] 6.862976
> median(d.control$LD)
[1] 6.817565

```

Oftentimes, it is a good idea to store a result in an R-variable so that you can refer to it later. Then you can type the new variable name to see what is stored in it. For example,

```

> meanTestLD = mean(d.control$LD)
> meanTestLD
[1] 6.862976
> CFU = 10^meanTestLD
> CFU
[1] 7294170

```

The last command shows that R-variables can be used with the mathematical operators `+`, `-`, `*`, `/`, and `^`. To compute the mean *TestLD* for each lab, execute

```

> tapply(d.control$LD, d.control$Lab, mean)
      1      2      3      4      5      6      7      8
6.848784 6.946420 7.251723 6.526638 6.999886 6.683945 6.956432 6.689980

```