	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 1 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

The following is a discussion of statistical schemes that have been used and may be appropriate for evaluation of microbiological QC data (duplicates, replicates, daily reference slides for quantifiable analyses). In all cases, initial control limit values should be determined and then updated on a specified frequency as more data is collected. Use of control charts is recommended to facilitate trending.

Consideration for All QC Evaluations

Microbial and particulate loading within a sample will influence level of variability detected in repeat counts for many types of analyses. Samples that are either lightly or heavily loaded with microorganisms for a given matrix (*e.g.,* bulk, air, surface, etc.) will typically display more variability in repeat counts than samples that have optimal loading ranges for counting, especially if only a portion of the sample preparation is assessed.

Amount of sample assessed will also influence variability detected in repeat counts. Repeat counts on sample preparations that have been 100% completely assessed should yield closer numbers than repeat counts on samples that have had a smaller portion assessed (i.e., <100%).

It should also be noted that statistical treatment may not be meaningful at the lowest concentrations for some analyses (*e.g.,* indicating 1 count vs 3 counts or 1 versus 0 on a direct exam may not be a reason for failed QC whereas for air culture it might be). Statistical treatment should be limited to organism concentrations high enough to indicate true failures if calculations exceed acceptance criteria.

After becoming experienced with performing sample preparations and analyses for a given technique, do the following.

1. As mentioned above, load concentration may affect the variation of repeat counts. Decide whether or not the technique requires statistics to be performed separately for different loading ranges (i.e., similar to NIOSH PCM 7400 method).

2. If decision to item 1 is yes, then define the concentration categories/ranges that statistics shall be performed on separately. Remember, category concentration ranges for a given technique are arbitrary, and they should be defined after some observations and empirical data have been obtained. As more QC data are obtained, the loading categories may have to be redefined and changed.

	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 2 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

3. Perform the QC statistics for the concentration range(s) of a given technique. QC statistics will be explained in the next section of this document.

4. Decide how many data points (i.e., 10, 20, 30, etc.) within a concentration range of a technique are required to establish acceptance criteria. The number of data points required to establish a baseline comparison is arbitrary but should be at least 10; 20-30 points is better.

5. After determining statistics based on total counts/concentrations, decide if the same statistics should be performed on the individual dominant microorganisms.

Statistical Comparisons for QC Evaluations

The following are statistical approaches that may be considered and may be appropriate for your laboratory. Other approaches may be equally suitable.

A. NIOSH PCM 7400 Model for Data Processing[Ref 1] (may be most appropriate for air direct exam data for reference slides and reanalysis counts.)

   1. A set of reference slides, containing a range of spore types and concentrations, are read daily by analysts (lab-generated slides containing only one spore type are no longer allowed as they are not representative of real samples and taxon ranking cannot be done). Data collected should be spores/trace and spores/m^3 for the number of spores of each type and total spores. After 10-20 values for each slide are obtained, acceptance criteria for at least total count can be calculated by determining the mean and standard deviation (StdDev) of each data set (Statistical evaluation of concentrations of dominant spore types may also be done but are not mandatory. See discussion of Organism ID and Ranking for requirements). Warning limits for each slide are typically mean +/-2 StdDev while control limits are mean +/-3 StdDev.

   2. Variability or coefficient of variation (Cv), also referred to as relative standard deviation (Sr), can be calculated from the reference slide data (Cv = StdDev/Mean). Cv values should be evaluated to see if there is a linear correlation between Cv value and spore concentration. If there is, pooled Cv values may be calculated for various spore concentration ranges (e.g. 5-50, 50-200, >200 spores counted, or whatever ranges

	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 3 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

seem appropriate). If there is no obvious Cv correlation to spore concentration, one pooled Cv value can be used for the lab. Pooled Cv is calculated as follows:

$$Cv_{pool} = \sqrt{[(Cv_1{}^2 + Cv_2{}^2 + ... Cv_n{}^2)/(n)]}$$

where n = number of Cv values pooled

3. Cv values can then be used in the NIOSH 7400 formula for determining acceptability of the 5% duplicate (intra-analyst) and replicate (inter-analyst) analyses (total spore count/trace). The formula is as follows:

$$\left| \sqrt{X_1} - \sqrt{X_2} \right| \le 2.77 (mean\sqrt{X})(pooled\ Cv/2)$$
$$where\ mean\sqrt{X} = (\sqrt{X_1} + \sqrt{X_2})/2$$

This formula can be simplified to:

$$\left| \sqrt{X_1} - \sqrt{X_2} \right| \le 0.69 (\sqrt{X_1} + \sqrt{X_2})(pooled\ Cv)$$

As long as the left side of this equation is less than or equal to the right side, the duplicate/replicate results are considered acceptable.

B. <u>QC for Intra/Inter Analyst Comparisons</u>

<u>The following are statistical approaches that may be considered and may be appropriate for your laboratory. Other approaches may be equally suitable.</u>

1. <u>The NIOSH 7400 formula from Step 3</u> above can be used for intra/inter analyst acceptability for all other quantifiable analyses as well (quantified direct exam swabs or tape lifts and all culturable analyses), provided a way is found to determine appropriate Cv values without the use of reference slides. One possible way is to determine mean, StdDev and Cv for each replicate or duplicate pair. The Cv values can then be pooled (total or by organism concentration) to produce Cv values to be used in the formula shown in Step 3 above. Duplicate and replicate values to use should be organisms/sample rather than final concentrations based on air volumes, areas, or weights unless consistent values are used for these parameters.

	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 4 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

2. Calculations Based on Relative Percent Difference or RPD (Variations of this are used by many labs. This calculation can be used for any intra/inter analyst data set.)

    a. RPD is calculated as the absolute difference between a pair of readings divided by the mean and multiplied by 100 to change result to % or $((|X_2 - X_1|)/\text{Mean})*100$.

    b. There are correct and incorrect methods for determining RPD control limits, as shown in the Steve Lerman discussion attached at the end of this document as Appendix A. (Arbitrary selection of an acceptance limit (e.g.+/-100%) is not statistics based).

    c. The easiest correct approach is to keep the sign on the RPD value rather than using absolute values. This then allows for use of 2-sided 2 and 3StdDev limits (Steve's Figure 3).

    d. Keeping the sign on the RPD values when calculating the StdDev and then using 1-sided 2 and 3StdDev limits to track absolute value of RPD is also commonly used (Steve's Figure 4).

    e. The FDA Office of Regulatory Affairs Laboratory Manual Volume II's method of using a single control limit at 3.27 times the mean Std Dev is also acceptable (Steve's Figure 5).

    f. As with the Cv discussion above, there may be a concentration dependence whereby more than one RPD database and set of control limits may be needed if there is a wide range of appropriate RPD values.

3. Calculations Based on Standard Deviation

    a. There are statistical ways to estimate StdDev for a series of duplicate/replicate measurements. The following is perhaps easiest as it does not require use of outside tables[Ref 2]:

    $$\text{StdDev}_{est} = \sqrt{(\sum d^2/2N)}$$

    where d = difference between the 2 values in each pair
    N = number of pairs

    Acceptance criteria can then be set at $\text{StdDev}_{sple} \leq 3\text{StdDev}_{est}$.

    b. From the same statistics source, the following formula calculates uncertainty from the estimated StdDev:

	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 5 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

$$U = t(StdDev_{est})/N$$

          where t = value from Student t table

          N = number of pairs used in Step b above

4.   <u>QC for Cultures</u>

a. Standard Methods[Ref 3] uses the range criterion (Ŕ) to determine if duplicate or replicate counts are acceptable or not. The Ŕ is calculated from historical data by summing all differences between logarithms of each duplicate count and dividing by number of duplicate counts. After Ŕ is calculated and established, it can be used to determine whether subsequent duplicate or replicate counts are acceptable or not by using the following formula

$$|\log x_1 - \log x_2| > 3.27 \, \acute{R} \text{ then reject acceptability}$$

        where x_1 = 1st count

                 x_2 = 2nd count

b. The Environmental Laboratory Approval Program Certification Manual stated that duplicate (intra-analyst) counts for cultures should not have more than 5% difference, and that replicate (inter-analyst) counts should not have more than 10% difference to be acceptable. This is not a statistical determination but may be a place to start as a default value.

<u>Protocols for Taxonomic ID and Ranking</u>

<u>The following are approaches that may be considered and may be appropriate for your laboratory. Other approaches may be equally suitable.</u>

1. At present a common approach is to require identification of some percentage of the prevalent organism types in a sample to agree (e.g. top 2, 2 out of top 3, 3 out of top 5, etc). Looking at statistical concentration comparison of top 1-2, in addition to total concentration of all organisms, may be appropriate, but is not mandatory with the current policies.

2. Another method of looking at organism ranking is use of the Spearman Rank Order Correlation Coefficient (r_s). The following formula is used to determine if rankings within 2 sets of results are similar or related:

	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 6 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

$$r_S = 1 - ((6(\sum D^2)/(N(N^2-1))$$

where D = difference in an organism's ranking in the 2 data sets

N = number of organisms in each data set

		Code: QSG-GMC-001
		Revision 3
	Quality System Guidance	Released: 05/15/2023
		Page 7 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

The closer r_S is to 1, the closer the agreement in the rankings. A cutoff value for acceptable values for r_S would need to be established (at least >0.5, probably higher). Spearman ranking alone is probably not enough but combined with statistical concentration comparison of top 1-2 organisms may be appropriate.

References:

1. Abell, M.T., Shulman, S.A., and Baron, P.A. 1989. The Quality of Fiber Count Data, Appl Ind Hyg. Vol. 4 No. 11 pp. 273-285.
2. Taylor, John Keenan, 1987, "Quality Assurance of Chemical Measurements", p22.
3. Standard Methods for the Examination of Water and Waste Water 20th Edition pp. 9-10.

	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 8 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

Appendix A
STATISTICAL QUALITY CONTROL TECHNIQUES FOR DUPLICATE QC DATA

Steve Lerman
SIL Consulting, Inc.
3 Allan Gate
Plainview, NY 11803-6112
(516) 433-3412

Most Laboratory QC Managers are well familiar with the techniques for generating traditional control charts for reference and spike samples by taking the mean and standard deviation of the data, and plotting warning and control limits at $\pm2s$ and $\pm3s$ respectively (see Figure 1 and Table 1). Please note that only six observations have been used for illustrative purposes. Statistically based limits should typically be established using a minimum of 20 to 30 observations collected under similar analytical conditions.

Establishing and plotting control limits for duplicate data, however, is another matter entirely. Here, we see what is essentially the same approach - calculating the mean and standard deviation of the difference, or relative % difference (% RPD), between the two results, then setting up control limits as one would for reference and spike samples. The % RPD is the absolute value of the difference divided by the average, to allow data of different magnitudes to be statistically combined. Although this approach is widely used, some applications of this approach are incorrect.

There is no significance to using the mean of the RPD as a mean - there is no trend line for duplicate data; if there is one, it would be zero. Instead, the RPD can be treated in the same manner as the difference between the values and the mean for reference and spike data, i. e., the data can be taken and calculated as a standard deviation. We will take the data in Table 1, and calculate it several ways to illustrate what we mean.

Table 2 and Figure 2 show the data calculated and graphed using an erroneous approach. Although the appearance looks comforting, it is meaningless. There are no lower control limits for duplicate data, and the mean is not the desired target.

Probably the simplest approach is the one recommended by Paul Britton, the Statistical Data Manager for EPA's proficiency program and Results Advisor for ASTM Committee D19 on water[1]. His approach is similar to the erroneous one,

	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 9 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

except that the sign of the RPD is kept. Using this approach, the resulting distribution is Gaussian, and we can look for trends in the first vs. second result. The mean converges at zero, as we would expect. Table 3 and Figure 3 show this approach.

There are other valid approaches. Taylor[2] uses the root-mean-square approach, where he takes the root-mean-square sum of the RPD value and uses that value as a standard deviation (this is the same as calculating a standard deviation). He then takes 2x and 3x that value as control limits - the trend line is zero. This produces a 1-way control chart, which is totally suitable for duplicate data. Figure 4 and Table 4 show such a graph using the data in Table 1.

A similar approach is used by the FDA Office of Regulatory Affairs Laboratory Manual Volume II[3]. They calculate the mean of the RPD but use that as a measure of standard deviation. They take 3.27 times this value as the single control limit, thus also producing a 1-way control chart, but with a single control limit. 2.51 times this value may also be used as a warning limit. Figure 5 and Table 5 show such a graph using the data in Table 1, and with both control limits.

All of these protocols provide statistically similar and valid results, giving the Laboratory QC Manager a wide variety of options for establishing and plotting control limits for duplicate QC data.

---

[1] ASTM Subcommittee D19.02 Meeting.

[2] Taylor, John K., Quality Assurance of Chemical Measurement, pp. 22-23, Lewis Publishers, Chelsea, MI, 1987.

[3] Food and Drug Administration Office of Regulatory Affairs Laboratory Manual Volume II, Document ORA-LAB.5.9, Rev# 02, pp. 9-11.

	Quality System Guidance	**Code:** QSG-GMC-001
AIHA LAP		Revision 3
		Released: 05/15/2023
		Page 10 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

Figure 1
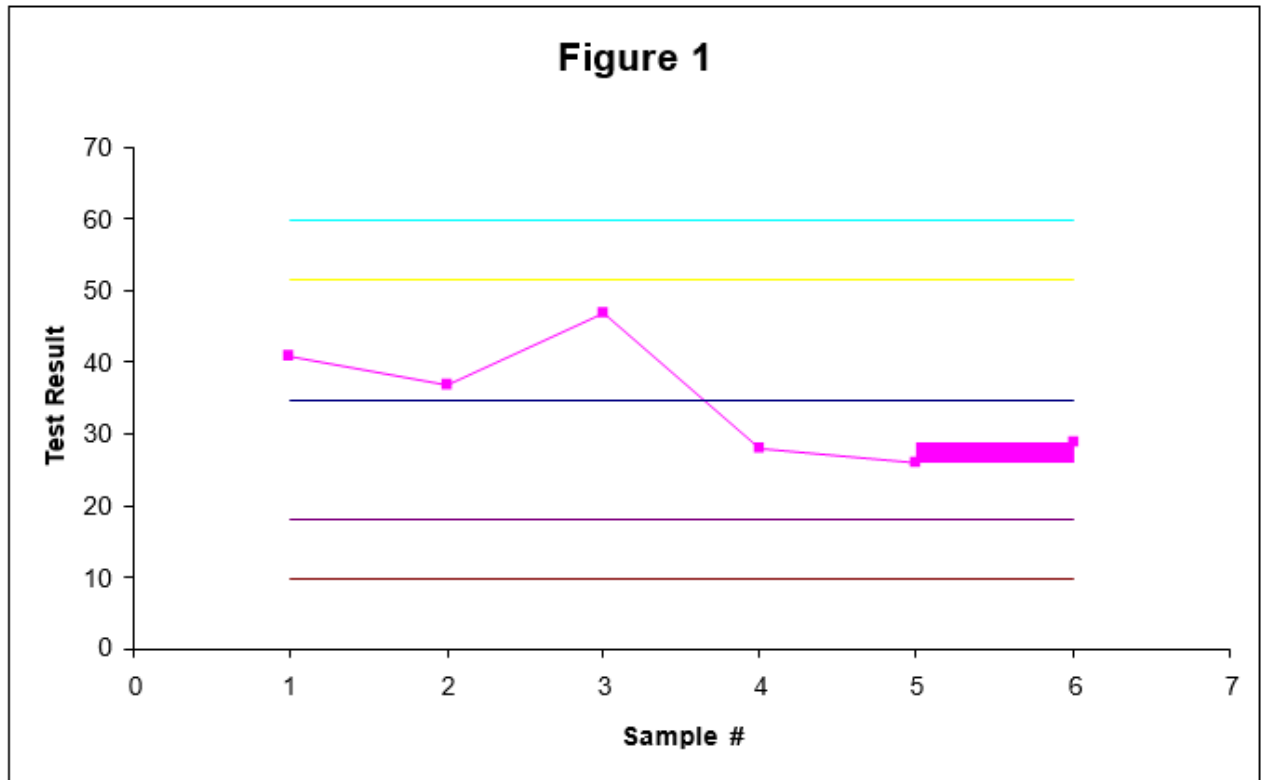
Table 1

Result #	Result 1
1	41
2	37
3	47
4	28
5	26
6	29
Mean	34.7
Std. Dev.	8.4
+2s	51.4
+3s	59.7
-2s	17.9
-3s	9.6

Figure 2
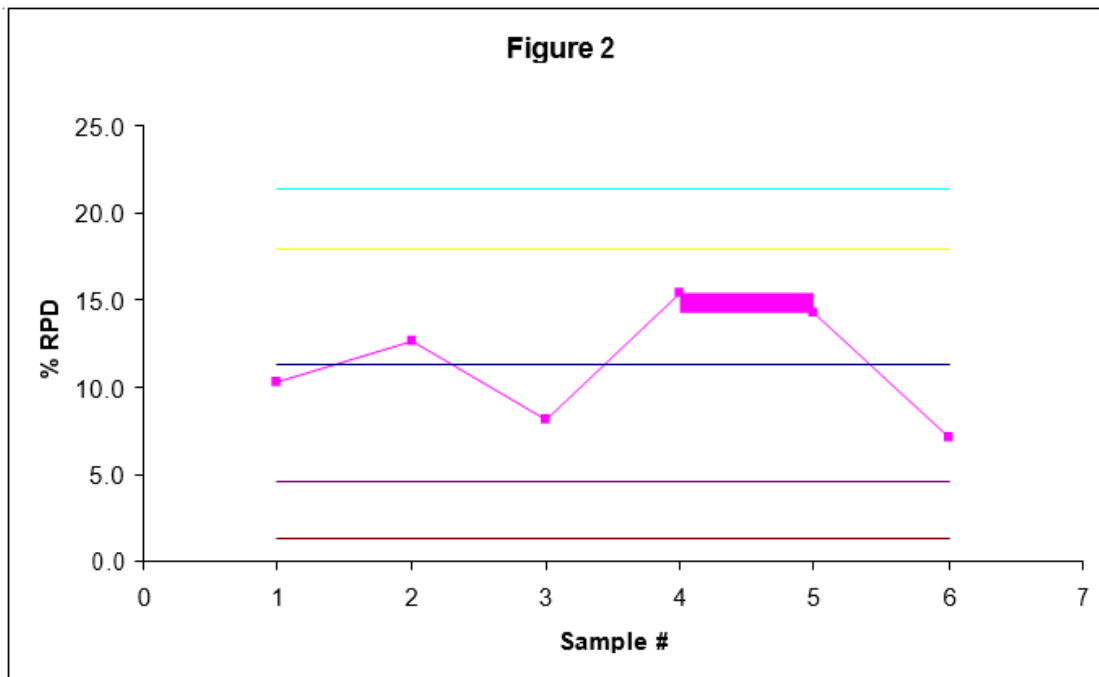
Table 2

Result #	Result 1	Result 2	% RPD
1	41	37	10.3
2	37	42	12.7
3	47	51	8.2
4	28	24	15.4
5	26	30	14.3
6	29	27	7.1
		Mean	11.3
		Std. Dev.	3.3
		+2s	18.0
		+3s	21.3
		-2s	4.6
		-3s	1.3

	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 12 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

Figure 3
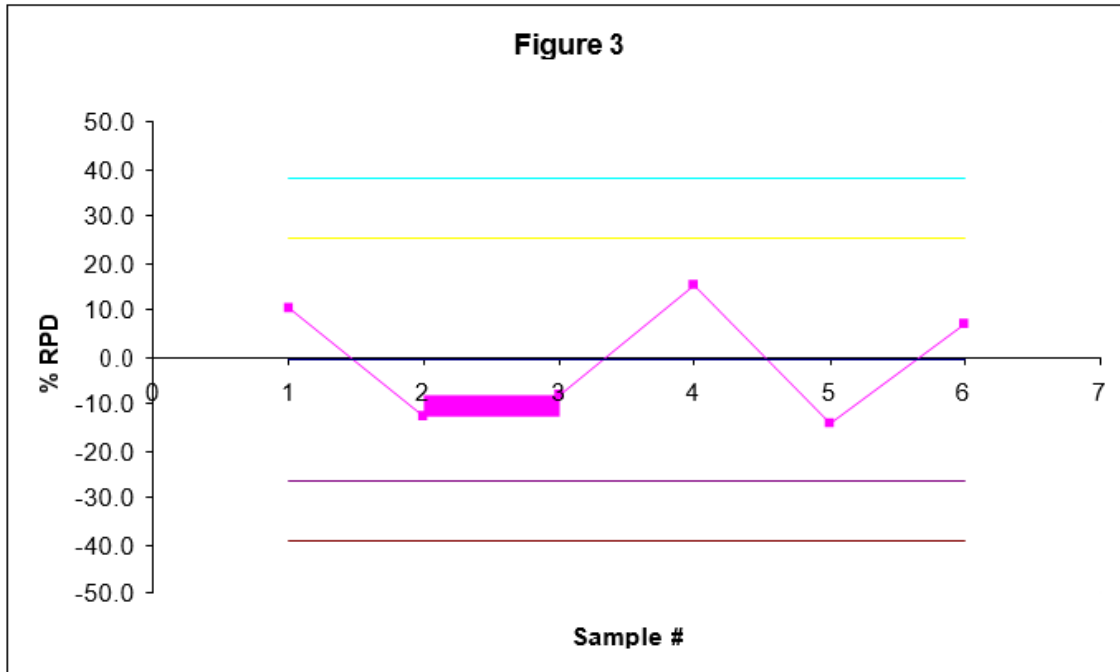
Table 3

Result #	Result 1	Result 2	% RPD
1	41	37	10.3
2	37	42	-12.7
3	47	51	-8.2
4	28	24	15.4
5	26	30	-14.3
6	29	27	7.1

	Mean	-0.4
	Std. Dev.	12.8
	+2s	25.3
	+3s	38.1
	-2s	-26.0
	-3s	-38.9

	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 13 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

Figure 4
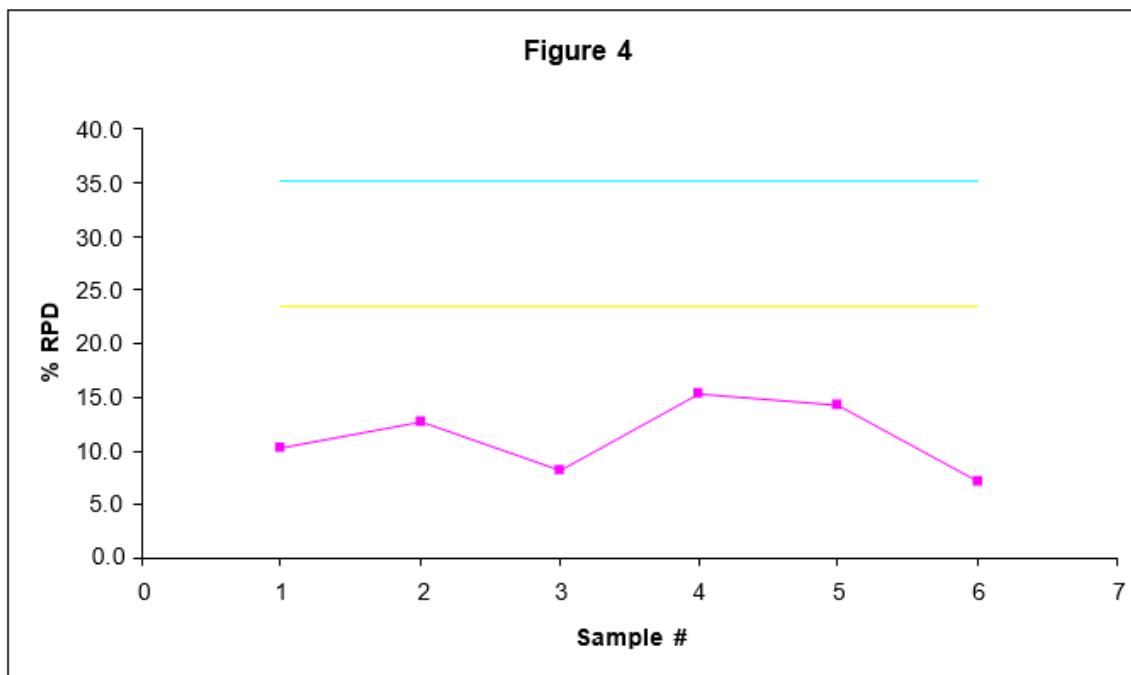
Table 4

Result #	Result 1	Result 2	% RPD
1	41	37	10.3
2	37	42	12.7
3	47	51	8.2
4	28	24	15.4
5	26	30	14.3
6	29	27	7.1
		RMS	11.7
		+2*RMS	23.4
		+3*RMS	35.2

	Quality System Guidance	**Code:** QSG-GMC-001
		Revision 3
		Released: 05/15/2023
		Page 14 of 14

Title: Guidance on Statistical Analysis of EMLAP QC

Figure 5
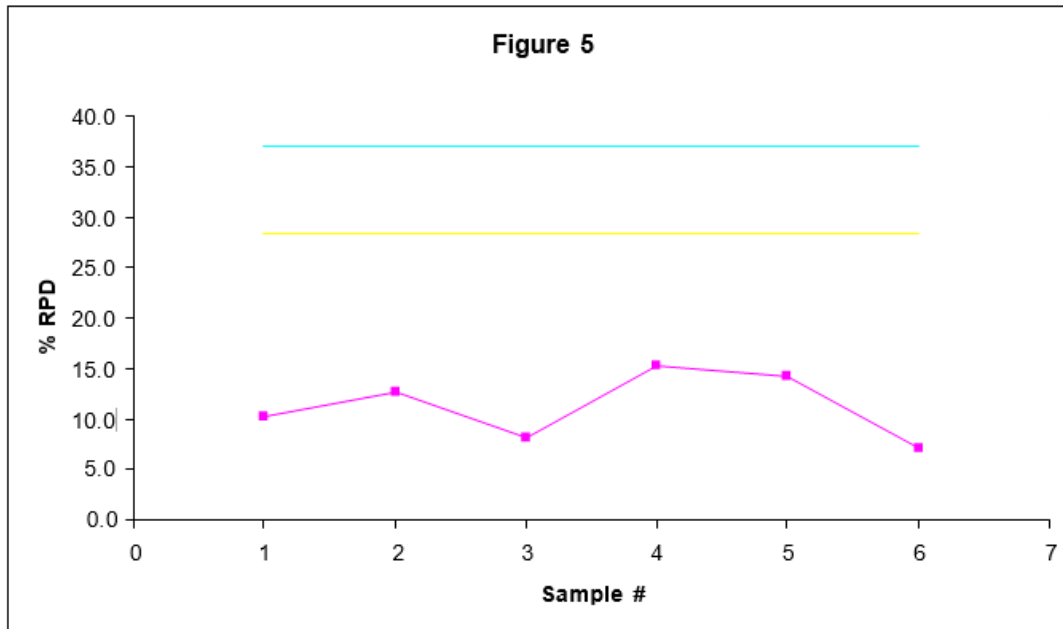
Table 5

Result #	Result 1	Result 2	% RPD
1	41	37	10.3
2	37	42	12.7
3	47	51	8.2
4	28	24	15.4
5	26	30	14.3
6	29	27	7.1
		Mean	11.3
		+2.51*Mean	28.4
		+3.27*Mean	37.0