Regulatory Proposal

PRO2005-04

Guidance for Setting Pesticide Maximum Residue Limits Based on Field Trial Data

This consultation document provides guidance for setting pesticide maximum residue limits (MRLs) or tolerances for plant commodities based on field trial residue data. The Canadian term "maximum residue limit" and the American term "tolerance" are used interchangeably in this document and has legal meaning in each jurisdiction. These guidelines are a product of the NAFTA Tolerance/MRL Harmonization Workgroup, which was charged with formulating a statistically based, scientifically defensible methodology for establishing tolerances. One of the goals of harmonizing the tolerance-setting process is to reduce trade barriers. However, such harmonization is predicated on the assumption that the same data sets are available to the regulatory bodies establishing the tolerances. As such, harmonization should encourage the sharing of field trial data between regulatory bodies.

The first section of this document provides background information on the impetus for establishing the workgroup and introduces the statistical basis of the methodology proposed by the workgroup. The second section describes the decision algorithm for establishing the appropriate statistically based tolerance value. The third section provides standard operating procedures (SOPs) for utilizing an Excel® spreadsheet produced by the workgroup that incorporates the statistical calculations and decision algorithm. The spreadsheet is available at www.pmra-arla.gc.ca/english/pdf/mrl/method_calc.xls. The fourth section will discuss supplemental topics to include rounding. An appendix is also included that provides a sample calculation using the tolerance spreadsheet.

Please provide any comments on this publication within 45 days of the date of release of this document to Publications at the address below.

(publié aussi en français)

28 September 2005

Health Canada   Santé Canada

Canada

NAFTA Guidance Document
for
Guidance for Setting Pesticide Tolerances
Based on Field Trial Data

United States Environmental Protection Agency
Office of Pesticide Programs

and

Health Canada
Pest Management Regulatory Agency

Executive Summary

This document provides guidance for setting pesticide tolerances or maximum residue limits (MRLs)[1] for plant commodities based on field trial residue data. These guidelines are a product of the NAFTA Tolerance/MRL Harmonization Workgroup, which was charged with formulating a statistically based, scientifically defensible methodology for establishing tolerances. One of the goals of harmonizing the tolerance-setting process is to reduce trade barriers. However, such harmonization is predicated on the assumption that the same data sets are available to the regulatory bodies establishing the tolerances. As such, harmonization should encourage the sharing of field trial data between regulatory bodies. The first section of this document provides background information on the impetus for establishing the workgroup and introduces the statistical basis of the methodology proposed by the workgroup. The second section describes the decision algorithm for establishing the appropriate statistically based tolerance value. The third section provides standard operating procedures (SOPs) for utilizing an Excel® spreadsheet produced by the workgroup that incorporates the statistical calculations and decision algorithm. The fourth section will discuss supplemental topics to include rounding. Also included is an appendix that provides a sample calculation using the tolerance spreadsheet.

---

[1]   The Canadian term "maximum residue limit" and the American term "tolerance" are used interchangeably in this document.

Table of Contents

I. Introduction and Background

A. Impetus for Harmonization

This document provides guidelines and standard operating procedures (SOPs) for establishing tolerances based on field trial residue data. The guidelines and SOPs result from the activities of the NAFTA Tolerance/MRL Harmonization Workgroup, which was charged with formulating a statistically based, scientifically defensible methodology for determining tolerances or maximum residue limits (MRLs), the maximum legal limit for a pesticide residue in a food or feed crop. The ultimate goal of the Tolerance/MRL Harmonization Project was to propose a standard methodology that could be used by the United States and Canada as part of our joint review program such that the same or similar data sets result in the same recommendation for tolerance or MRL levels. Such harmonization could significantly reduce trade barriers and should encourage the sharing of field trial data by government regulatory bodies. Although the workgroup is formally considered a NAFTA project, observers from the European Union (EU), the European Commission (EC), and the California Department of Pesticide Regulation (DPR) participated in the meetings and provided valuable input and discussion. The workgroup recognizes that any proposed procedures would serve as guidance for establishing tolerances and not fixed rules.

The current method for setting tolerances is based on identifying the highest residue from a field trial data set and then "rounding up." There is little guidance regarding how much higher the tolerance should be than the highest field trial residue. Different reviewers and different agencies/governments can have different views regarding the level at which tolerances should be set. However, there is a consensus that the level at which tolerance values are set should balance the probability associated with wrongly seizing a legally treated crop versus the probability of being incapable of detecting illegal use with a reasonable degree of confidence. Therefore, the workgroup agreed early on that any proposed methodology must recognize this "balancing act" that is implicit in establishing tolerance levels. A balance must be reached between setting tolerances high enough so as to avoid seizing crops which were legally treated and setting them low enough such that residues from illegal use can be detected. In addition, the harmonization workgroup decided that any developed methodology should be reasonably simple to use by non-statisticians (requiring only calculators or spreadsheets), should aim to be reasonably well-accepted, and should be based – as far as possible – on accepted statistical principles and methods.

B. Statistical Basis

Current tolerance-setting methods ("rounding up") can result in widely divergent tolerance proposals depending upon the size of the sample and the judgement and biases of the specific reviewer and regulatory bodies. The methodology used in current practice does not have a well-grounded statistical basis, and is unduly and inappropriately dependent on the number of field trials that are considered. More specifically, the

maximum value present in a data set is highly variable from one sample to another and are expected to increase monotonically as the sample size grows. The result is that residues from the same parent distribution (i.e., same shape, mean, and standard deviation) can result in dramatically different tolerances in terms of both magnitude and nominal expected percentile depending upon the number of field trials submitted: the fewer field trials/samples submitted, the lower the tolerance and the lower the nominal percentile at which the tolerance is established[2].

There is no intrinsically defensible reason why, *a priori*, a tolerance for commodity with generally fewer required field trials (i.e., a minor crop) should be established at a nominally lower percentile than the tolerance for a commodity with a larger required number of field trials (i.e., a major crop). Ideally, in order to be more defensible in a statistical sense, the tolerance would be established at some consistent nominal percentile of the expected population distribution regardless of the number of field trials. This percentile-based method would lead to a tolerance that was less significantly influenced by the number of samples collected and would be far less variable from sample to sample than estimates based on a single highest observation that is the tolerance-setting method now used.

Any established tolerance-setting procedure should consider as part of that procedure the number of field trials submitted so as to appropriately ensure, with a sufficient desired degree of confidence, that a legally applied pesticide will not result in residues which exceed the tolerance. The greater the number of field trials, the greater the confidence one can have that some given nominal percentile value would adequately cover legal pesticide applications. This suggests that tolerances should be established, in part, based on statistical upper confidence limit (UCL) of a nominal percentile. Establishing a tolerance, in this way explicitly considers one aspect of the balancing that is an intrinsic part of tolerance setting in that it helps to ensure that the tolerance is set at a level that is sufficiently high so as to preclude a reasonable possibility for a legally applied pesticide to result in residues that exceed the tolerance. If a tolerance were set, for example, at the 95% UCL of the 95th percentile, one could be 95% confident that an established tolerance exceeds at least 95 percent of the residues that would result if a pesticide were applied at the maximum application rate and the crop was harvested at the minimum pre-harvest interval (PHI) according to label use instructions.

However, there is another side to this "balancing act." This other side requires that tolerances not be established so high that there is little probability of detecting misuse. In order to reach such a balance, one should avoid establishing a tolerance at a level which far exceeds a given (high end) nominal percentile of the distribution – regardless of the upper confidence limit on the lower percentile. For example, it might be desirable to avoid setting a tolerance at levels greater than the expected 99th percentile of a

---

[2] To verify this, the workgroup compared random samples of various sizes from the same lognormal distribution (i.e., a lognormal distribution with the same mean and standard deviation). The workgroup observed that in general the maximum value increased as the sample size increased.

distribution of residue values. Both sides of this "balancing act" can be combined to establish "dual criteria", which, together, meet the desire to avoid setting tolerances so low that legal use practices can result in illegal (tolerance- or MRL-exceeding) residues, but not so high that illegal uses would only rarely be detected. This can be done by establishing a guideline such that tolerances should be established at the minimum of either the 95% UCL of the 95th percentile or the 99th percentile.

This is in fact the primary focus of the workgroup's proposed methodology, which came to be called "the 95/99 rule." That is, establish the tolerance as the minimum of either the 95% UCL on the 95th percentile or the point estimate of the 99th percentile based on the assumption that field trial residues are distributed lognormally. Other methods for calculating tolerances were considered to supplement the 95/99 rule when certain circumstances (e.g., non-lognormal distribution, small sample size, or extensive number of residues values which are below the LOD) result in underlying assumptions not being met or tolerance values that are impractically high. To the extent that the methodology can or might be accepted by the EU, Codex, or governmental regulatory bodies that are not part of NAFTA, the workgroup took into consideration the current methods used by these bodies.

One of the methods considered by the workgroup was that used informally by the California DPR, which came to be called "the California method". The California method consists of adding three arithmetic standard deviations to the arithmetic mean of the field trial residue values. For a normal distribution, the mean plus three standard deviation represents (the point estimate of) the 99th percentile. Although the DPR workgroup members did not believe the residue values to be normally distributed, they found on an empirical basis that the method produced reasonable estimates and believed that it captured some upper percentile estimate. Indeed, if the California method is interpreted with respect to the Chebychev Inequality, which states that *for any distribution, at least* $(1 - 1/k^2)\%$ *of the distribution is within k standard deviation of the mean*, then the mean plus three standard deviations represents at least the 89th percentile.

As previously mentioned, the calculations involved in the 95/99 rule are based upon the assumption that the field trial residue values are lognormally distributed. As this assumption may not be met for every data set, the workgroup sought an alternative distribution-free approach to calculating an upper percentile based tolerance level. With the Chebychev interpretation, the California method seemed a reasonable alternative to the 95/99 rule for data sets where the validity of the lognormal assumption could be questionable. Due to the generality of the Chebychev Inequality, the California method may not capture as high of an upper percentile value as that sought by the 95/99 rule. However, to attain an upper bound on the 95th percentile, the Chebychev Inequality requires that 4.5 standard deviations be added to the mean, not three. The workgroup agreed that "backing off" the upper percentile value was a reasonable trade off for the generality (i.e., validity with respect to any distribution) of the inequality.

In addition to non-lognormal data sets, the workgroup felt that the methodology needed to specifically address small data sets (i.e., fewer than ca. 15 samples). For small data sets, the estimates of the mean and standard deviation of the population based on the sample may not be very accurate. To reduce the number of parameters being estimated for the lognormal distribution from two (mean and standard deviation) to just one, an additional assumption is made for these small data sets with respect to the relationship between the mean and standard deviation. Specifically, the assumption was made that the standard deviation is equal to the mean (i.e., the coefficient of variation [CV] is equal to one). Empirically, a coefficient of variation of one appeared to be a not unreasonable assumption for field trial residue data. Based on the assumptions that the population is lognormal and a coefficient of variation equal to one, the 95th percentile can be calculated as 3.9 times the median (i.e., 50th percentile). To ensure the tolerance value would be sufficiently high so as not to be exceeded by a legally treated crop, the 95% upper prediction limit (UPL) on the median is calculated and then multiplied by 3.9. This calculation was termed "the UPLmedian95 method." For small sample sizes, this method could be used as an alternative to the 95/99 rule when the estimate of the standard deviation is considered to be inaccurately high.

The EU is one of the few regulatory bodies that have developed statistically based methods for establishing tolerances. One of the methods formulated by the EU (termed EU Method I) is conceptually equivalent to the first part of the 95/99 rule (i.e., it is a 95% UCL on the 95th percentile). However, this EU method relies on the assumption that field trial residues are distributed *normally*, rather than lognormally. The workgroup members agreed that a default assumption of lognormality is more appropriate than an assumption of normality for field trial residue data. The assumed form of the underlying distribution is important as the calculation of the 95% UCL on the 95th percentile will be systematically **larger** if residue values are assumed to be from a lognormal distribution than if they are assumed to be from a normal distribution. Additionally, the EU also employs the Dixon Q-test to determine if high residue values are outliers. As with the EU Method I, the Dixon Q-test relies on the assumption of normality. Consequently, residue values that are consistent with a lognormal distribution (i.e., a right-skewed distribution) might be considered "outliers" on the assumption of a normal distribution. In other words, if the distribution of residue values is assumed to be lognormal (as opposed to normal), then fewer residue values would be considered outliers.

The other method employed by the EU is a "distribution-free" estimate referred to as EU Method II. The method consist of doubling the 75th percentile of the field trial residue values. The support for this method seems to be empirical-based. At the time, the EU considered the 75th percentile to be reasonable upper percentile value that could be reliably estimated for smaller data sets. Doubling the 75th percentile was considered a "safety factor" that ensured the tolerance would not be exceeded. Generally the EU calculates MRLs using both Method I and Method II, and selects the smaller of the two as the tolerance value.

The above methods (the 95/99 rule, EU Method I and II, and the California method) do not encompass all the methods considered, but they do represent the primary methods upon which the workgroup focused. The next section of this document explains the decision algorithm developed by the workgroup that considers the circumstances under which some of the above methods would be employed.

II.   Flowchart/Decision Algorithm

The flowchart in Figure 1 describes the methodology formulated by the NAFTA MRL/Tolerance Harmonization Workgroup for calculating statistically based pesticide tolerances for plant commodities based on field trial residue data where critical good agricultural practices[3] (cGAP) are followed. Initially, the workgroup focused on statistical methods for calculating tolerances for "ideal" data sets. The workgroup considers a data set to be "ideal" if it satisfies the following criteria: (1) there are relative few non-detectable residues, (2) the data set is approximately lognormally distributed, and (3) the sample size is large. The three primary decision nodes of the flowchart require that these three aspects of a data set be assessed before setting a tolerance. If a data set is "flagged" for not satisfying one or more of the above criteria, then supplemental procedures/statistical methods may be employed before setting a tolerance. The flowchart outlines the supplemental procedures/statistical methods and considerations that should be employed if a data set deviates substantially from these ideal criteria.

The first step of the methodology requires the user to enter the field trial data into the tolerance spreadsheet. **As the spreadsheet is highly automated, it is important that the user enter the data using a standard format and following the exact "cut and paste" procedures that are detailed in the "Using the Tolerance Spreadsheet" section**.

The next step of the methodology involves reviewing and inspecting the field trial for issues related to non-detectable residues, inconsistent PHIs and/or application rates (see the "Using the Tolerance Spreadsheet" section of this document for additional details).

Once the data have been entered into the spreadsheet and reviewed/inspected, the next steps in the methodology involves determining which method will be used to establish the tolerance or MRL and whether or not the data need to be supplemented. For simplicity, we will first consider the case when a data set is "ideal." If a data set is not flagged for any of the three criteria (i.e., the outcome of all three primary decision nodes is "No"), then "the 95/99 rule" (see the "Statistical Basis" section) should be used to set the tolerance. In other words, for an ideal data set, the tolerance would be set as the minimum of either the 95% upper confidence limit on the 95th percentile or the point estimate of the 99th percentile.

---

[3]   Conditions under which a pesticide is applied at the maximum application rate and the crop is harvested at the minimum PHI according to label use instructions are termed "critical good agricultural practices."

If a data set is flagged for containing many non-detects (i.e., the outcome of the "Flag for ND Issue?" decision node is "Yes"), then the user should considering using maximum likelihood estimation (MLE) techniques to supplement the data set. The workgroup suggests that no more than 10–15% of the data be non-detects. The workgroup observed that for data sets with less than 10–15% censoring, there seemed to be little difference in the tolerance values calculated from the original data sets and their corresponding MLE-supplemented counterparts. If MLE techniques have been used to supplement a data set, the supplemented data set should continue through the other decision nodes.

If a data set is flagged for not being reasonably lognormal based on the conclusion of the approximate Shapiro-Francia test statistic (i.e., the outcome of the "Flag for Lognormal Issue?" decision node is "Yes"), then the user should visually inspect the lognormal probability plot to determine if the data set is reasonably lognormal (see the "Using the Tolerance Spreadsheet" section for more detail). If the user's visual inspection of the probability plot confirms the conclusion drawn from the approximate Shapiro-Francia test statistic (i.e., the outcome of the "Visually Inspect Probability Plot, OK?" decision node is "No"), then the "Mean + 3SD method" (i.e., the "California method") should be used to calculate the tolerance for the data set. The tolerance value calculated using the California method represents an upper bound estimate of the 89th percentile for any distribution (see the "Statistical Basis" section). However, if the user does not agree with conclusion drawn from the Shapiro-Francia test statistic (i.e., the user determines from the probability plot that the data set is reasonably lognormal) or the conclusion drawn from the Shapiro-Francia test statistic indicates that the data set is reasonably lognormal, then the user should continue to the last decision node.

If the data set is flagged for having too few samples (i.e., the outcome of the "Flag for # Points Issue?" decision node is "Yes"), then the "UPLMedian95th method" should be calculated. The workgroup recommends that data sets with fewer than 15 samples be considered small. The value calculated by the UPLMedian95th method represents upper estimate of the 95th percentile based on the median residue value (see the "Statistical Basis" section). The value calculated using the UPLMedian95th method should then be compared to the value calculated from the 95/99 rule (both methods assume the residue data are lognormally distributed) and the minimum of these two values should be used as the tolerance or MRL value.
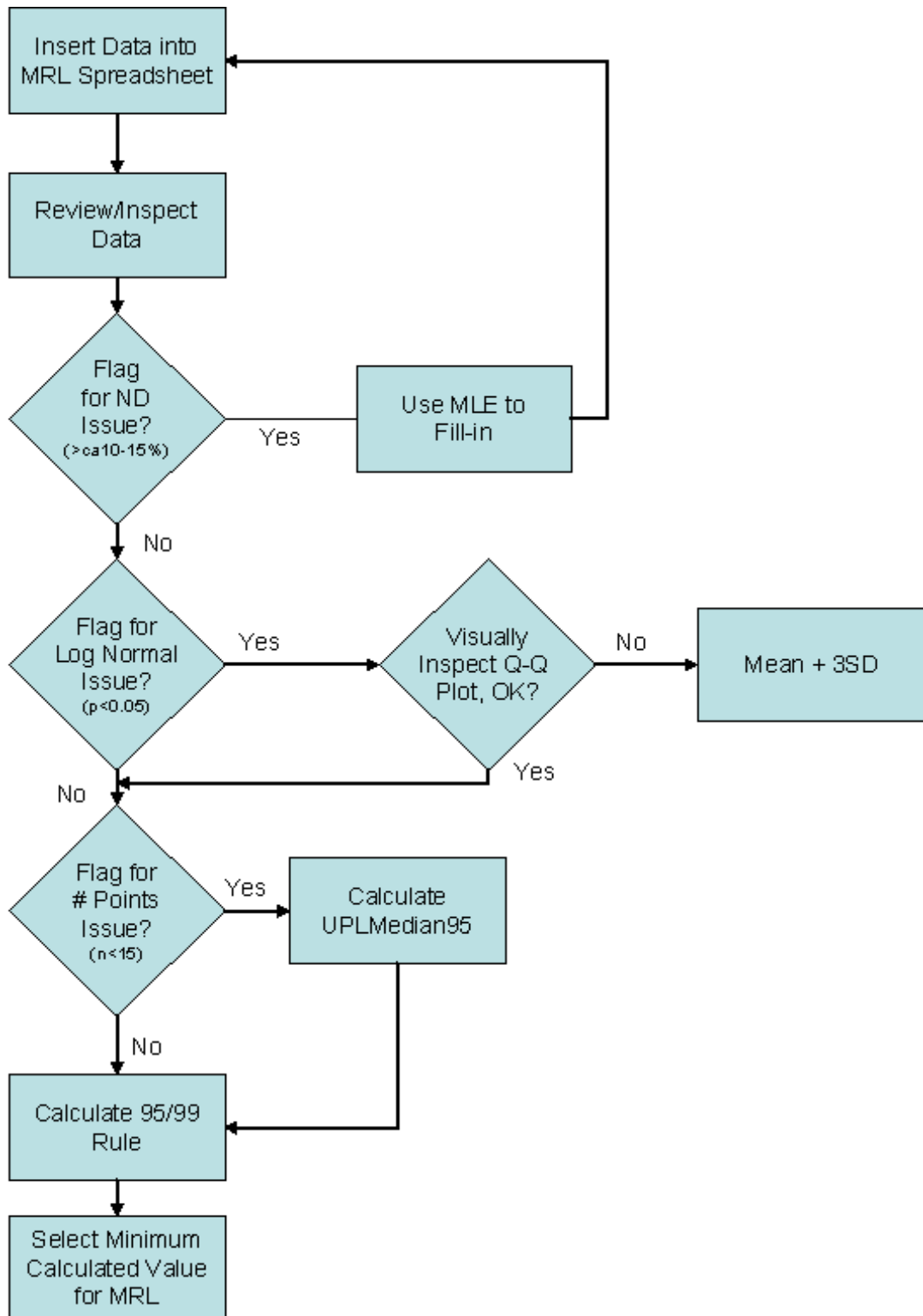
Figure 1   Spreadsheet Algorithm for Estimating Pesticide Tolerance/MRLs

III.   Using the Tolerance Spreadsheet

A.   Inspect and Adjust Data

Before a data set is actually entered into the tolerance spreadsheet, the field trial data should be inspected to determine which residue values are to be included. Sometimes a field trial study will include samples that were collected at various pre-harvest intervals (PHIs) and/or with multiple application rates. Generally, all samples within a data set should be within ±25% of the maximum application rate and ±25% of the minimum PHI as indicated on the label. The workgroup strongly recommends that only one of these parameters (application rate or PHI) be allowed to deviate from cGAP as indicated on the label[4]. Although the workgroup supports a case-by-case approach to adjusting values, it strongly cautions against using field trial data where either the application rate or PHI is not within ±25% of cGAP.

Additionally, the analyst should note the field trial samples with residue values below the laboratory's limit of quantitation (LOQ) and those that are below the limit of detection (LOD), if possible[5]. Residues below the LOQ are sometimes referred to as non-quantifiables (NQs), whereas residues below the LOD are sometimes referred to as non-detects (NDs). Although there is a distinction between the NQs and NDs these values are often treated in a similar fashion. At this point, the analyst should also decide which the values to impute for the NQs (or NDs). Commonly used values include $LOQ/\sqrt{2}$ (or $LOD/\sqrt{2}$), LOQ/2 (or LOD/2), and LOQ (or LOD). Although the values imputed for NQs (or NDs) can influence estimates of the mean and standard deviation, these effects are considered minimal if a relatively small fraction of the data (less than ca. 10%) is censored (i.e., below LOQ)[6].

B.   Enter Data into Spreadsheet

Once the data have been inspected and adjusted for deviations from cGAP (if necessary and deemed appropriate), the data should be entered (as a column) into the "Data Library" worksheet in the tolerance/MRL spreadsheet. The data should be entered using a standard format with the following information:

(1)   the regulator to which the field trial data were submitted (e.g., EPA or PMRA)

---

[4]   In some case, the effect of a PHI longer than minimum PHI may *counteract* the effect of application rate higher than the maximum rate. If one were to use such data set, the residue should not be adjusted.

[5]   An easy way to differentiate these values is to change the colour of the values to red.

[6]   Maximum likelihood estimation (MLE) techniques provide better estimates of the mean and standard deviation of censored data sets. Although this guidance document does not provide instructions for obtaining MLE-based estimates, the workgroup has developed a spreadsheet for calculating MLE estimates that will be made available.

(2)   the name of the chemical used in the field trial study

(3)   the name of the crop to which the chemical was applied in the field trial study

(4)   the nominal PHI (or PHI range) for the field trial data set, as multiple PHIs and/or application rate can be included in field trial study

(5)   the nominal application for the field trial data set, again as multiple PHIs and/or application rate can be included in field trial study

(6)   the name of the company or group submitting the field trial data.

This identifying information is important for uniquely identifying data sets, especially when being shared by multiple regulatory bodies. An additional "Citation" worksheet is included in the spreadsheet for providing additional detail pertaining to the field trial study[7].

Once the data set has been entered on the "Data Library" worksheet in the appropriate format, the entire data set (including the identifying information) should be copied to the "Data" worksheet. This is the only step required of the reviewer. There is no need to sort the residue values. The rest of the procedures/calculations in the tolerance spreadsheet are performed automatically. **As the spreadsheet is highly automated, it is important for the reviewer to follow the exact procedures outlined (and illustrated) hereafter. Any deviation from these instructions could result erroneous calculations.**

---

[7]   The information entered on the "Data Library" worksheet is automatically copied to the "Citation" worksheet. However, any additional information must be directly entered on the "Citation" worksheet.

The following steps should be followed for using the spreadsheet.

1. Copy the entire data set from the "Data Library" worksheet.

2.   Go to the "Data" worksheet and select cell B1, then paste the data set.

NOTE: Usually, there will already be a data set on the data worksheet. The newly pasted data set will either be longer or shorter than the previous data set. Steps 3a to 5a detail the procedures to follow if the newly pasted data set is longer than the previous data set, and steps 3b to 4b detail the procedures to follow if the newly pasted data set is shorter than the previous data set.
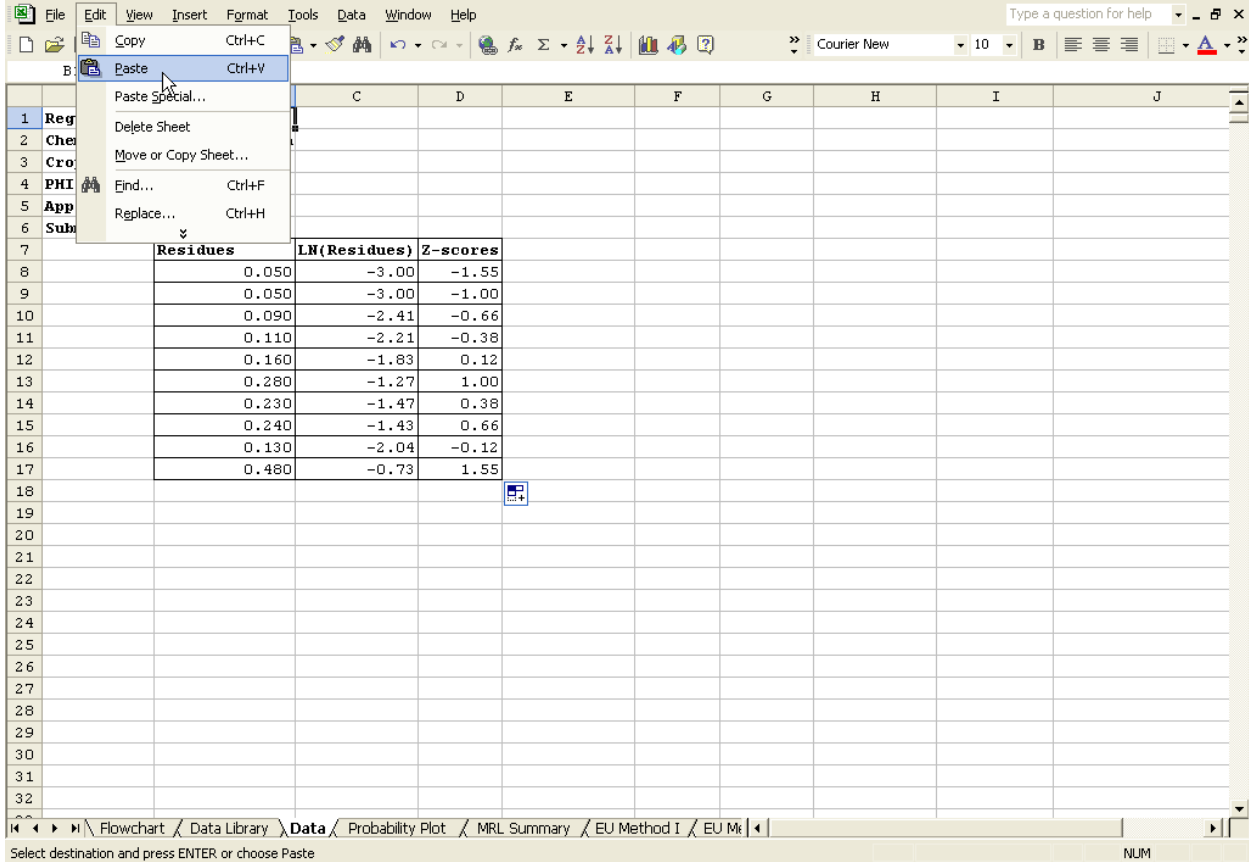
3a. **If the newly pasted data set is longer than the previous data set**, select the last cell in both column C and D, and then place the cursor over the bottom right corner of the last cell in column D to activate the "drag and fill" procedure in Excel® indicated by the black crosshair.

4a.   Next hold down the left mouse button and drag the crosshair straight down until the bottom edge of the shadow box is even with the last residue value from the newly pasted data set.

5a. Once the left mouse button is released, the shadow box fills in the appropriate formulae for the cells in the shadow box. This concludes the manual operations required by the users in order to obtain tolerance calculations.

NOTE: The "drag and fill" outlined in the previous section is equivalent to copying the last cell in both column C and D, and pasting them into the blank cell of columns C and D to the right of the newly pasted data set.

3b. **If the newly pasted data set is shorter than the previous data set**, select the cells in columns B, C, and D that extend past the last residue value of the newly pasted data set.

	A	B	C	D
1	Regulator:	EPA		
2	Chemical:	Pymetrozine		
3	Crop:	LeafLettuce		
4	PHI:	0-1 Day		
5	App. Rate:			
6	Submitter:			
7		Residues	LN(Residues)	Z-scores
8		0.616	-0.48	0.19
9		0.574	-0.56	0.06
10		0.838	-0.18	0.59
11		0.316	-1.15	-0.31
12		1.314	0.27	1.13
13		0.977	-0.02	0.74
14		0.789	-0.24	0.45
15		0.757	-0.28	0.31
16		1.403	0.34	1.40
17		0.492	-0.71	-0.06
18		1.938	0.66	1.87
19		1.036	0.04	0.92
20		0.136	-2.00	-0.45
21		0.374	-0.98	-0.19
22		0.055	-2.90	-1.87
23		0.055	-2.90	-1.40
24		0.060	-2.81	-1.13
25		0.083	-2.49	-0.92
26		0.110	-2.21	-0.74
27		0.110	-2.21	-0.5
28				
29				
30				
31				
32				

Flowchart / Data Library \ **Data** / Probability Plot / MRL Summary / EU Method I / EU Me

Select destination and press ENTER or choose Paste   Sum=-21.697   NUM

4b.   Next press the "Delete" key on the keyboard. This concludes the manual operations required by the users in order to obtain tolerance calculations.

NOTE: Erroneous calculations may result if the delete function from the edit menu or from the right mouse click menu is employed.

	A	B	C	D	E	F	G	H	I	J
1	Regulator:	EPA								
2	Chemical:	Pymetrozine								
3	Crop:	LeafLettuce								
4	PHI:	0-1 Day								
5	App. Rate:									
6	Submitter:									
7		Residues	LN(Residues)	Z-scores						
8		0.616	-0.48	-0.27						
9		0.574	-0.56	-0.45						
10		0.838	-0.18	0.27						
11		0.316	-1.15	-1.21						
12		1.314	0.27	0.90						
13		0.977	-0.02	0.45						
14		0.789	-0.24	0.09						
15		0.757	-0.28	-0.09						
16		1.403	0.34	1.21						
17		0.492	-0.71	-0.66						
18		1.938	0.66	1.71						
19		1.036	0.04	0.66						
20		0.136	-2.00	-1.71						
21		0.374	-0.98	-0.90						
22										
23										
24										
25										
26										
27										
28										
29										
30										
31										
32										

C.   Residues Values Below the LOD and/or LOQ

As mentioned previously, residue values below the LOQ are referred to as non-quantifiables (NQs) and those below the LOD are called non-detects (NDs). Although there is a clear distinction between the two categories, these types of data are treated similarly and are generically referred to as censored data. These censored values should have been identified earlier (preferably in some easily identifiable manner), as recommended in the "Inspect and Adjust Data" section.

The workgroup recommends the use of MLE techniques to estimate the mean and standard deviation of a data set if more than 10–15% of the data set is censored. The workgroup has developed a spreadsheet that implements the "robust method", an MLE-based procedure for estimating the mean and standard deviation of a lognormal

population based on a censored (i.e., below the LOQ or LOD) sample[8]. Based on the MLE estimates, values consistent with the estimated distribution are calculated for the censored residues. The advantage of using this MLE technique is that it utilizes information about the proportion of the sample that is censored in addition to the values of the detectable residues when estimating the mean and standard deviation. However, the workgroup cautions against using MLE techniques for data sets with a large degree of censoring (i.e., >60%), as such methods become less reliable.

D.   Assess Lognormality Assumption

A normal probability plot is a plot of the ordered values of a data set versus their expected normal order statistic values. If the data are normally distributed, they will approximate a straight line. Similarly, when the logarithms of the ordered values are plotted against their expected normal order statistics, an approximate straight line indicates that the data are lognormally distributed.
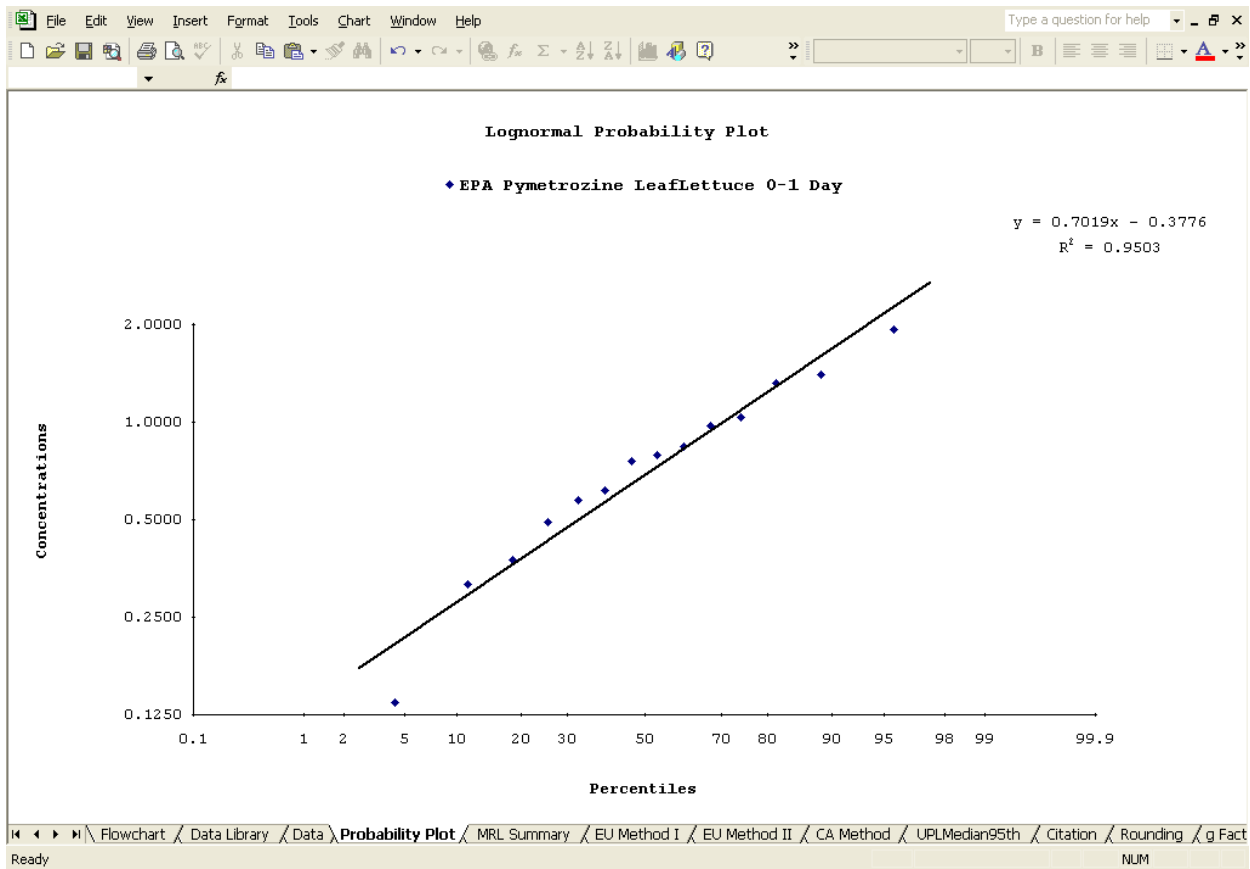
A lognormal probability plot is automatically created in the spreadsheet on the "Probability Plot" worksheet based on the field trial data entered or copied into the "Data" worksheet. The logarithms of the residue values are plotted on the vertical axis, but units are that of the original untransformed data (e.g., ppm); thus, the vertical axis has a logarithmic scale. Additionally, the expected normal order statistics of the log transformed residue values are plotted on the horizontal axis. However, the horizontal axis is labeled with the more informative percentiles corresponding to the normal order statistics[9].

The probability plot allows the analyst to visually assess the assumption of lognormality by judging how well the (log-transformed) residue values approximate a straight line. The probability plot also allows the user to compare the actual value of a residue to its expected value based on the assumption of lognormality. The y-values of the regression line represent the expected values of the residues for the percentiles on x-axis. Thus, if the actual residue value is above (or below) the regression line, then it is higher (or lower) than its expected value. For example in the plot below, the highest residue value (1.94 ppm) represents approximately the 95th percentile and is slightly lower than its expected value (2.28 ppm). However, the lowest residue value (0.14 ppm) represents approximately the 5th percentile and is quite a bit lower than its expected value (0.21 ppm). Overall the data set looks reasonably lognormal.

---

[8]   Although this guidance document does not provide instructions for obtaining MLE-based estimates, the workgroup will eventually make an MLE spreadsheet available and provide detailed instruction for its use.

[9]   The normal order statistics -2.33, -1.645, 0, 1.645, and 2.33 correspond to the 1st, 5th, 50th, 95th, and 99th percentiles, respectively.

Although the visual inspection of the lognormal probability plot is a very important component of assessing the lognormality assumption, the workgroup also felt it was important to include a formal test statistic for assessing lognormality. The Shapiro-Wilk and Shapiro-Francia are two test statistics used to assess normality. By using log-transformed data, the two test statistics can be used to assess lognormality. The complex calculations involved in the computation of these test statistics would be difficult to implement in a spreadsheet. Therefore, the workgroup opted to use the much more easily calculated approximation to these test statistics base on the coefficient of determination (R^2) of the linear regression line from the probability plot.

The normality test statistic based on the coefficient of determination is calculated on the "MRL Summary" worksheet and identified as the "Approximate Shapiro-Francia Normality Test." Below the test statistic is the conclusion as to whether or not to reject the assumption of lognormality of the residue data (i.e., normality of the log-transformed residue data) based on a 5% significance level. The critical values for the test statistic are on the "R Critical Values" worksheet and were obtained from a National Institute of Standards and Technology website (www.itl.nist.gov/div898/handbook/eda/section3/eda3676.htm).

The spreadsheet (MRL Summary worksheet) shows:

	Regulator:	EPA	
	Chemical:	Pymetrozine	
	Crop:	LeafLettuce	
	PHI:	0-1 Day	
	App. Rate:		
	Submitter:		
	n:	14	
	min:	0.14	
	max:	1.94	
	median:	0.77	
	average:	0.83	

NOTE: Tabled values in parentheses indicate 95% upper confidence bounds on the point estimates of the 95th or 99th percentiles. No upper confidence bounds on the 99.9th percentile are provided and these are represented by "(--)". Tabled values that are shown directly without parentheses represent point estimates of the indicated percentile (e.g., 95, 99, or 99.9)

NOTE: Additional details regarding the algorithms and decision rules used by this spreadsheet can be found under the tab labeled "Flowchart" in this workbook.

	95th Percentile	99th Percentile	99.9th Percentile
EU Method I	1.62	1.95	2.32
Normal	(2.09)	(2.56)	(--)
EU Method I	2.12	3.38	5.71
Log Normal	(4.12)	(8.02)	(--)
EU Method II		2.21	
Distribution-Free			
California Method		2.28	
$\mu + 3\sigma$			
UPLMedian95th		5.34	
Approximate Shapiro-		0.9503	
Francia Normality		p-value > 0.05 : Do not reject lognormality assumption	
Test Statistic			

Would you like the above values rounded? (Y or N)==>	N

NOTE: Rounding is performed and significant figures are presented in accordance with the EPA/PMRA guidance document entitled *Guidance for Setting Pesticide Tolerances Based on Field Trial Data* that accompanies this spreadsheet. Details are provided in this document as well as the worksheet tab labeled "Rounding" located in this workbook.

Evaluating the appropriateness of the assumption of lognormality of a data set is important as the conclusion determines which method will be used to calculate a tolerance value. If the lognormality assumption is determined to be unreasonable, then a "distribution-free" method is used to calculate the tolerance; otherwise, methods that assume lognormality are employed.

E. Sample Size Consideration

The sample size of a field trial data set is one of summary statistics reported on the "MRL Summary" worksheet in addition to the minimum, maximum, median, and average. The workgroup agreed that for small data sets, the estimated mean and standard deviation could potentially be inaccurate for a right-skewed distribution such as lognormal. Therefore, for data sets with fewer than 15 samples, another lognormal-based method is considered. In addition to the lognormality assumption, the supplemental method assumes that the mean and standard deviation are equal (i.e., the coefficient of variation is equal to one) and makes use of a calculation of the mean and standard deviation based on the median of the residue values.

F.   Selecting the Appropriate Tolerance Value

This section explains the logic employed by the spreadsheet for selecting the appropriate tolerance value. For more detailed information regarding the statistical methods and the decision algorithm utilized in the spreadsheet, the user should read the "Statistical Basis" and "Flowchart/Decision Algorithm" sections. However, Excel® comments have been included on the "MRL Summary" worksheet to provide the user with short descriptions of the various tolerance values being calculated. To access these comments, the user must hold the cursor over the cells with a red triangle in the upper right hand corner. Also, notes are included on the "MRL Summary" page to provide additional detail about the summary page and the tolerance-setting methodology. A screenshot with the notes and all the comments displayed is provided below.

The "MRL Summary" worksheet contains the results of various calculations performed automatically by the tolerance spreadsheet. The tolerance value suggested by the spreadsheet is shaded gray and coloured red; the other potentially useful methods are in **bold**. For an "ideal data set" where the lognormality assumption is determined to be reasonable and there are at least 15 samples, the minimum of either the 95% upper confidence limit (UCL) of the 95th percentile (cell B18) or the point estimate of the 99th percentile (cell C17) is selected. If there are fewer than 15 samples for the data set

but the assumption of lognormality is determined to be reasonable, then the minimum of the 95% UCL of the 95th percentile (cell B18), the point estimate of the 99th percentile (cell C17), or the upper bound estimate of the 95th percentile based on median (cell C23) is selected. If the assumption of lognormality is determined to be unreasonable, then the distribution-free method that represents an upper bound on the 89th percentile (cell C21) is selected.

The tolerance calculation selected by the spreadsheet is based on the result of the approximate Shapiro-Francia test statistic. **Therefore, it is important for users to decide if the result of the approximate Shapiro-Francia test agrees with their visual examination of the probability plot.** The Shapiro-Wilk, Shapiro-Francia, and any of their approximations allow further departures from normality for smaller sample sizes than for larger sample sizes. If the approximate Shapiro-Francia statistic suggests that lognormality is a reasonable assumption, but the user disagrees, then the distribution free method (cell C21) should be selected. If the approximate Shapiro-Francia statistic concludes that lognormality is not reasonable, but the user disagrees, then for small data set (i.e., fewer than 15 samples) the minimum of the 95% UCL of the 95th percentile (cell B18), point estimate of the 99th percentile (cell C17), and the upper bound estimate of the 95th percentile based on median (cell C23) should be selected; otherwise the minimum of the 95% UCL of the 95th percentile (cell B18) and point estimate of the 99th percentile (cell C17) should be selected.

IV.   Establish Tolerance/MRL

In some cases, it may be necessary to adjust for storage stability loss or analytical recovery rates. However, the workgroup strongly recommends that either one or the other be adjusted for a given data set, but not both. The workgroup also recommends that these types of adjustments be made to the tolerance values calculated by the spreadsheet and not the individual residue values previous to the calculations.

Whether or not it is appropriate to adjust for storage stability loss should be considered on a case-by-case basis. However, the workgroup agrees that field trial data with more than 30% loss should be considered cautiously. If adjusting for storage loss is deemed appropriate, then the tolerance value calculated by the spreadsheet should be divided by one minus the percent loss (expressed in decimal form). For example, if the spreadsheet calculated a tolerance value of 15 ppm and 25% of the analytical sample is lost in storage, the proposed tolerance should be $15 \div (1 - 0.25) = 20$ ppm.

Workgroup members agree that analytical recovery rates between 70% and 120% are considered acceptable and do not recommend adjusting residue values for recovery as laboratories responsible for enforcement will probably have similar recovery rates. However, as with storage stability loss, the workgroup recommends that the decision to adjust for recovery rates be made on a case-by-case basis, but cautions against using field trial data with recovery rates outside of the 70% to 120% range.

Outside of NAFTA, residue classes are frequently used to round up calculated tolerance values. These residue classes can result in final tolerance values that are significantly higher than the calculated values. On the other hand, prescribing a standard decimal place (e.g., tenths, hundredths) or a set number of significant figures to which to round could imply a false degree of precision and would potentially result in a proliferation of tolerance levels. In order to reach a balance between these approaches, the workgroup developed a procedure that specifies the number of significant figures to be used and the nearest unit to which to round based on the magnitudes of the calculated tolerance values. The procedure recommended by the workgroup for rounding up the tolerance values calculated by the tolerance spreadsheet is detailed in Table 1. The rounding procedure was formulated under the assumption that residue values are expressed in parts per million (ppm). Therefore, if the tolerance value is less than 0.01 or more than 1000, then the residue values should be converted to parts per billion (ppb) or parts per thousand, respectively. Once converted to the appropriate unit, the same rounding procedure should be applied. A "yes/no" toggle is provided in the spreadsheet, which allows the user to view the unrounded tolerance values (displayed with two decimal places) or the tolerance values that have been rounded up according to the procedure recommended by the workgroup.

Table 1   Rounding and Significant Figure Conventions for Establishing Tolerances/MRL

Calculated Tolerance or MRL Value	Round Up to the Next Multiple of[10]	Number of Significant Figures
0.01–0.10	0.01	1
0.11–0.50	0.05	2
0.51–2.00	0.1	2
2.01–5.00	0.5	2
5.01–20.00	1	2
20.01–100.00	5	2
100.00–1000.00	10	2

---

[10]   If the calculated value falls exactly on a right-hand-side interval value (e.g., 0.05, 2.0, or 5.0), then it should not be rounded up.

List of Abbreviations

ca.	circa
cGAP	critical good agricultural practices
DPR	California Department of Pesticide Regulation
EC	European Commission
EU	European Union
EPA	United States Environmental Protection Agency
LOD	limit of detection
LOQ	limit of quantitation
MLE	maximum likelihood estimation
MRL	maximum residue limit
NAFTA	North American Free Trade Agreement
ND	non-detect
NQ	non-quantifiable
PHI	pre-harvest interval
PMRA	Pest Management Regulatory Agency
R^2	coefficient of determination
SOP	Standard Operating Procedure
UCL	upper confidence limit

Appendix I Sample Calculation

A.   Inspect and Adjust Data

The field trial data for pymetrozine on leaf lettuce include data for which the application rates were similar (i.e., within ±25% of label maximum). Although not all of the PHIs are within ±25% of the label minimum, zero- and one-day PHIs were considered similar enough to combine. It is also noted that none of the residue values were below the LOD.

B.   Enter the Data into Spreadsheet

The data set, including all identifying information, and residue values are entered into the "Data Library" worksheet.

	FD	FE	FF	FG	FH	FI	FJ	FK	FL
1	EPA	EPA	EPA	EPA	EPA	EPA	EPA	EPA	EPA
2	Pymetrozine	Pymetrozine	Pymetrozine	Pymetrozine	Pymetrozine	Pymetrozine	Pymetrozine	Pymetrozine	Pymetrozine
3	LeafLettuce	LeafLettuce	LeafLettuce	LeafLettuce	LeafLettuce MLE	LeafLettuce	NonBellPeppers	NonBellPeppers MLE	NonBellPeppe
4	0-1 Day	3 Days	5 Days	7 Days	7 Days	9 Days	0-1 Day	0-1 Day	3 Days
5									
6									
7	Residues	Residues	Residues	Residues	Residues	Residues	Residues	Residues	Residues
8	0.616	0.136	0.021	0.059	0.059	0.036	0.120	0.120	0.0
9	0.574	0.195	0.127	0.060	0.060	0.037	0.020	0.008	0.0
10	0.838			0.232	0.232		0.021	0.021	
11	0.316			0.036	0.036		0.130	0.130	
12	1.314			0.020	0.011		0.020	0.012	
13	0.977			0.047	0.047		0.024	0.024	
14	0.789			0.168	0.168		0.020	0.020	
15	0.757			0.102	0.102		0.024	0.024	
16	1.403			0.387	0.387				
17	0.492			0.054	0.054				
18	1.938			0.020	0.016				
19	1.036			0.062	0.062				
20	0.136								
21	0.374								

The new data set has more residue values than the previous data set, so the last cells in columns C and D are copied down to the last value in the data set.

C. Residue Values Below the LOD and/or LOQ

This data set has no censored values, so MLE estimates need not be considered for estimating the mean and standard deviation of the data set.

D. Assess Lognormality Assumption

The probability plot suggests that the lognormal assumption for the data set is reasonable. From the probability plot, we see that the estimated percentiles correspond quite well to the actual values (resulting in an approximate straight line) and that the highest field trial value represents the 95th or 96th percentile.
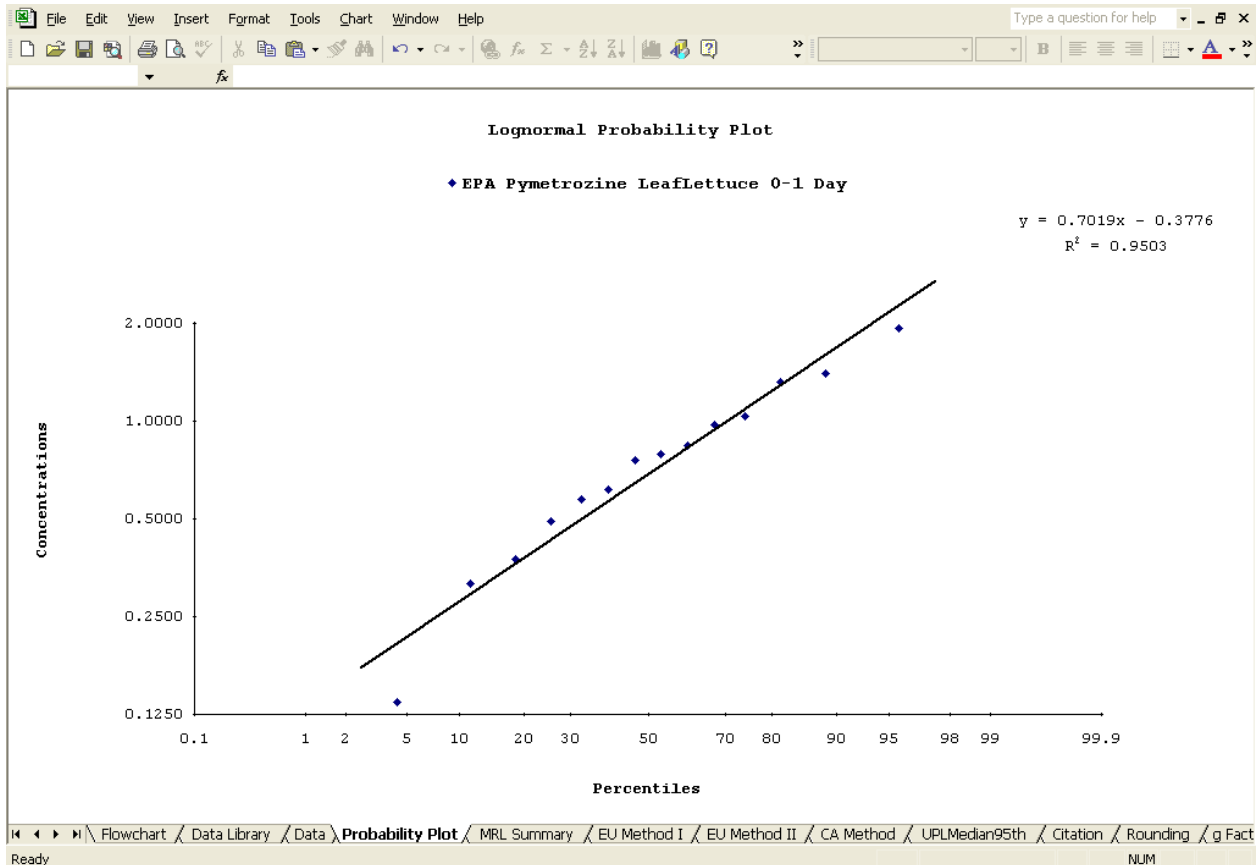
Additionally, the qualitative assessment of lognormality based on the probability plot is verified by the approximate Shapiro-Francia test statistic (from the "MRL Summary" worksheet), the conclusion of which is "Do not reject lognormality assumption."

C25 = `=CORREL(Data!C8:C65536,Data!D8:D65536)^2`

	B	C	D
Regulator:		EPA	
Chemical:		Pymetrozine	
Crop:		LeafLettuce	
PHI:		0-1 Day	
App. Rate:			
Submitter:			
n:		14	
min:		0.14	
max:		1.94	
median:		0.77	
average:		0.83	

	95th Percentile	99th Percentile	99.9th Percentile
EU Method I Normal	1.62 (2.09)	1.95 (2.56)	2.32 (--)
EU Method I Log Normal	2.12 (4.12)	3.38 (8.02)	5.71 (--)
EU Method II Distribution-Free		2.21	
California Method		2.28	
UPLMedian95th		5.34	
Approximate Shapiro-Francia Normality Test Statistic		0.9503	

p-value > 0.05 : Do not reject lognormality assumption

Would you like the above values rounded? (Y or N)==>   N

NOTE: Tabled values in parentheses indicate 95% upper confidence bounds on the point estimates of the 95th or 99th percentiles. No upper confidence bounds on the 99.9th percentile are provided and these are represented by "(--)". Tabled values that are shown directly without parentheses represent point estimates of the indicated percentile (e.g., 95, 99, or 99.9).

NOTE: Additional details regarding the algorithms and decision rules used by this spreadsheet can be found under the tab labeled "Flowchart" in this workbook.

NOTE: Rounding is performed and significant figures are presented in accordance with the EPA/PMRA guidance document entitled *Guidance for Setting Pesticide Tolerances Based on Field Trial Data* that accompanies this spreadsheet. Details are provided in this document as well as the worksheet tab labeled "Rounding" located in this workbook.

Flowchart / Data Library / Data / Probability Plot / MRL Summary / EU Method I / EU Me

E.   Sample Size Consideration

As the data set consists of 14 values, the tolerance value calculated by the UPLMedian95th method (in cell C23) must be considered in the next step when selecting the appropriate tolerance value.

F.   Selecting the Appropriate Tolerance Value

The data set contains no residue values that were below the LOD (or LOQ), so MLE techniques are not necessary to estimate the mean and standard deviation. The probability plot and approximate Shapiro-Francia test statistic both indicate the data set is approximately lognormal, so there is no need to consider a distribution-free method. However, as there are fewer than 15 data points, an alternative tolerance calculation (i.e., UPLMedian95th) based on the median residue must considered. The 95/99 rule, which consists of selecting the minimum of either the 95% upper confidence limit on the 95th percentile (cell B18) or the point estimate of the 99th percentile (cell C17), indicates that the (point estimate of the) 99th percentile should be selected. The

UPLMedian95th tolerance value, in this case 5.34, is greater than the tolerance value selected by 95/99 rule, in this case 3.38.

As indicated by the gray shading and red font, the 99th percentile with a value of 3.38 is the appropriate value to use as the tolerance. This value is smaller than the 95% upper confidence limit on the 95th percentile, which has a value of 4.12, and the alternative tolerance value calculated by the UPLMedian95th method, which has a value of 5.34.

G.   Establish Tolerance/MRL

Now that the appropriate tolerance value has been selected, the analyst should consider if adjustments for storage stability loss and recovery rates are appropriate. Although these are important steps in tolerance-setting process, they are not done automatically in the spreadsheet. For this example, there is no need to adjust for either storage stability loss or recovery rates.

However, the rounding convention recommended by the workgroup has been automated in the spreadsheet. To obtain the rounded tolerance values "Y" (or "y") is entered into cell C30, which serves as a rounding toggle[11]. **According to this guidance document, the tolerance for pymetrozine on leaf lettuce should be established at 3.5 ppm.**

---

[11]   If values other than "Y", "y", "N", or "n" are entered into cell C30, the user will get an error message and the tolerance values will remain unrounded by default.

File Edit View Insert Format Tools Data Window Help   Type a question for help

Courier New   10   B

C30   ƒx   Y

	A	B	C	D	E	F	G	H	I
1		Regulator:	EPA						
2		Chemical:	Pymetrozine						
3		Crop:	LeafLettuce						
4		PHI:	0-1 Day						
5		App. Rate:							
6		Submitter:							
7									
8		n:	14						
9		min:	0.14						
10		max:	1.94						
11		median:	0.77						
12		average:	0.83						
13									
14		95th Percentile	99th Percentile	99.9th Percentile					
15	EU Method I	1.7	2.0	2.5					
16	Normal	(2.5)	(3.0)	(--)					
17	EU Method I	2.5	3.5	6.0					
18	Log Normal	(4.5)	(9.0)	(--)					
19	EU Method II		2.5						
20	Distribution-Free								
21	California Method		2.5						
22	μ + 3σ								
23	UPLMedian95th		6.0						
24									
25	Approximate Shapiro-		0.9503						
26	Francia Normality		p-value > 0.05 : Do not reject lognormality assumption						
27	Test Statistic								
28									
29									
30	Would you like the above values rounded? (Y or N)==>		Y						
31									
32									

NOTE: Tabled values in parentheses indicate 95% upper confidence bounds on the point estimates of the 95th or 99th percentiles. No upper confidence bounds on the 99.9th percentile are provided and these are represented by "(--)". Tabled values that are shown directly without parentheses represent point estimates of the indicated percentile (e.g., 95, 99, or 99.9).

NOTE: Additional details regarding the algorithms and decision rules used by this spreadsheet can be found under the tab labeled "Flowchart" in this workbook.

NOTE: Rounding is performed and significant figures are presented in accordance with the EPA/PMRA guidance document entitled *Guidance for Setting Pesticide Tolerances Based on Field Trial Data* that accompanies this spreadsheet. Details are provided in this document as well as the worksheet tab labeled "Rounding" located in this workbook.

Flowchart / Data Library / Data / Probability Plot \ MRL Summary / EU Method I / EU M

Ready   NUM