Paul E. Helliker
Director

Department of Pesticide Regulation

Gray Davis
Governor

Winston H. Hickox
Secretary, California Environmental Protection Agency

M E M O R A N D U M

TO: Worker Health and Safety Branch Staff **HSM-00011**

[HSM # assigned after original issuance of policy]

FROM: Chuck Andrews **[Original signed by C Andrews]**
Branch Chief
(916) 445-4260

DATE: February 4, 2000

SUBJECT: Worker Health and Safety Branch Policy on the Statistical Analysis for Dislodgeable Foliar Residue Data

---

The attached document outlines procedures that have been adopted as Branch policy for the statistical treatment of dislodgeable foliar residue (DFR) data. These procedures should be followed by staff who analyze DFR studies conducted by the Branch, and by staff who analyze DFR studies conducted by registrants and other investigators. I plan on incorporating this policy in HS-1600, *Guidance for Determination of Dislodgeable Foliar Residue*, and HS-1612, *Guidance for the Preparation of Human Exposure Assessment Documents*.

If you have any questions, please contact your supervisor.

Attachment: *Standard Method for the Statistical Analysis of Dislodgeable Foliar Residue Data.*

Standard Method for the Statistical Analysis of Dislodgeable Foliar Residue Data

Worker Health and Safety Branch
February 4, 2000

This document outlines the policy of the Worker Health and Safety Branch with respect to the statistical treatment of dislodgeable foliar residue (DFR) data. The purpose of the policy is threefold: first, to provide a statistically valid method in order to ensure the high quality of scientific analysis of this data; second, to ensure that all Branch staff will apply a consistent approach when treating DFR data; third, to document publicly the statistical method that will be used by the Branch. This method is intended to be used by field staff and other staff in analyzing DFR studies conducted by the Branch, as well as by exposure assessors and other staff in analyzing DFR studies conducted by registrants and other investigators.

I. DATA PREPARATION

1. Nondetects

   Substitute one-half the limit of detection (LOD) for any nondetected (ND) sample taken postapplication. For background (preapplication) samples, substitute zero if all samples taken the same day under the same conditions were ND.

   If all samples from the final sampling day(s) were ND, drop the day(s) from the analysis. In other words, use only the data through the last day with any detects.

   The WHS Field Team does not apply any correction for percent analytical recovery. Exposure assessors may apply a correction if adequate laboratory data exist to estimate percent recovery.

2. Arithmetic mean of samples for each day

   Take the arithmetic mean of the (usually four) samples taken each day under the same conditions. This day mean is the basic unit of analysis.

   Generally, the logarithm (either base 10 or base *e*) of the arithmetic mean will be taken for the purpose of statistical analysis.

In describing the study, the term *samples* should be used for the samples taken each day. If multiple sites or applications are monitored under similar conditions, the sites or applications are the *replicates*. In general, the term *replicate* refers to an independent repetition of the complete experiment (Milliken and Johnson, 1984, p. 49; Neter *et al*., 1985, p. 899). Independent repetition in a DFR study requires, minimally, a new tank mix.

3. Correction for background

If for the purpose of estimating initial deposition, Day 0 samples are adjusted for background residues, it should be done by subtracting the arithmetic mean of the preapplication samples from the arithmetic mean of the Day 0 samples. If logarithms are to be used, take the log of the difference.

When there is no background residue, it is preferable to estimate initial deposition from the intercept of the dissipation curve.

For the purpose of estimating DFR dissipation, samples are not corrected for background.

II. STATISTICAL ANALYSIS

1. Means and standard deviations

When simple means and standard deviations are presented, they should be the arithmetic statistics, calculated on the untransformed variable (i.e., not on the logs). This is true even when the variable is thought to be lognormally distributed and logs are used in the regression analysis. (There are better ways to estimate the mean and standard deviation of a lognormal distribution, but they are slightly complicated. You may consult a statistician to do these calculations.)

2. Confidence intervals

Normally distributed variable. The familiar formula,

Arithmetic Mean ∀ $t_{(.975;\, n-1)}$*(S.D./%n),

is valid for the 95% confidence interval for the mean of a normally distributed variable.

Lognormally distributed variable. Ordinarily we will assume that DFR is lognormally distributed. The 95% confidence interval for the mean can be found in either of two ways. One is by calculating the arithmetic mean and S.D. of the logs, substituting them in the previous formula, then taking the antilog of the result:

antilog{Arithmetic Mean of logs ∀ $t_{(.975;\, n-1)}$*(S.D.of logs/%n)}.

Alternatively, the CI can be calculated from the geometric statistics:

Geometric Mean*Geometric S.D.$^{(\forall\, t(.975;\, n\text{-}1) / \%n)}$.

3. Dissipation curve

The log-linear regression model

$$\log DFR = \beta_0 + \beta_1{*}(days)$$

or the log-quadratic model

$$\log DFR = \beta_0 + \beta_1{*}(days) + \beta_2{*}(days)^2$$

should be fit to the log of mean DFR for each day (as described in Section I.2. above). If there are replicates (e.g., multiple applications), they are analyzed in one regression analysis. The simpler log-linear model may be used if it adequately describes the data. The log-quadratic model should be used if adding the days-squared term increases R^2 by 0.05 or more over the log-linear model (if you know how to do a stepwise regression, you can use that technique to decide whether to include days-squared; in 26 DFR datasets, the 0.05 rule of thumb gave a good approximation to the results of stepwise regression with the significance level to enter at 0.10). Occasionally it may be necessary to consider other models. If neither model fits well (significance p for overall model > 0.05) or the results seem anomalous, consult a statistician.

Half-life may be reported, but it will generally be more meaningful to give a table of predicted DFR by day after application. Predicted DFR in µg cm^{-2} should be calculated by an unbiased backtransformation of the predicted log (Powell, 1991). The table should give predicted µg cm^{-2} for every day from Day 0 through the last sampling day used in the regression analysis. Normally, prediction should not be extrapolated beyond the last sampling day used in the regression, but exceptions may be made if a specific day after application or a specific DFR level are of interest and lie beyond that day. Prediction limits (usually only the one-sided upper limits) should also be given in the table. Prediction limits are similar to confidence limits, but they apply to individual replicates rather than to the mean.

When the quadratic model is used, predicted DFR may begin to increase at some point. This can happen because of the nature of the quadratic model and/or random variation in the data. (You should consult a statistician to make sure the model is fit correctly.) In such cases, the lowest value reached by predicted DFR will be used as the predicted value for subsequent time points.

Some DFR studies compare deposition and dissipation under different conditions, for example, inside and outside the canopy, or with different application methods. For these studies, more complex regression models are required, and a statistician should be consulted.

Milliken, G. A. and D. E. Johnson. 1984. *Analysis of Messy Data, Volume I: Designed Experiments*, Van Nostrand Reinhold, New York.

Neter, J., W. Wasserman and M. H. Kutner. 1985. *Applied Linear Statistical Models: Regression, Analysis of Variance and Experimental Designs*, 2nd/Ed. Richard D. Irwin, Inc., Homewood, Illinois.

Powell, S. 1991. *Implementation in the SAS Sytem of the Bradu-Mundlak minimum variance unbiased estimator of the mean of a lognormal distribution. In* "16th annual meeting of the SAS Users Group International". SAS Institute, Inc., New Orleans, LA.