TITLE: **Methods for Confirmation of Non-Certified Laboratory Datasets and Background Calculation**

DATE EFFECTIVE: March 2009

HISTORY: New addition to the Technical Guidance Compendium

KEYWORDS: Phase II environmental site assessment, confirmation, certified laboratory, background calculation

RULE/ AUTHORITY: OAC 3745-300-07(E)(1)(d); OAC 3745-300-07(H)

QUESTION: How do you compare datasets to demonstrate that the two are not significantly different when confirming non-CL data and calculating background concentrations?

BACKGROUND: The VAP currently allows the inclusion of non-certified lab (CL) data sets to be used as part of a demonstration that the VAP property meets applicable standards. To use the non-certified data set, the data must be confirmed with a minimum of eight samples or a sample number amounting to 10 percent of the original (non-CL) data set, whichever is greater. The data between the two sets must be shown that the two are not statistically different. The VAP rules provide the certified professional (CP) some leeway regarding the methodology used for the confirmation demonstration.

The wording in the current rule states the following:

"*In order to confirm the non-CL data set, the volunteer must quantitatively compare the non-CL dataset and the CL dataset and demonstrate the two are not significantly different, utilizing any generally accepted approach*."

This allows the CP the flexibility to choose a statistical approach (e.g., t-test, nonparametric test, etc.) or any other measure of similarity that may be appropriate.

The rule allows the same flexibility when choosing a statistical approach to determine the background concentration as an applicable standard. Eight soil sampling points must be taken to represent a background level within each zone or soil horizon in order to determine if background is significantly different that the concentrations of COCs in identified areas.

ANSWER: A number of statistical approaches may be used to calculate the significant difference between data sets.

The key guidance referenced in the rules is:

U.S. EPA. 2002. "Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites." EPA 540-R-01-003-OSWER 9285.7-41.

The guidance refers to multiple statistical tests such as (but not limited to):

Method	Application	Comments
Descriptive Summary ▸ Mean ▸ Median ▸ Standard deviation ▸ Variance ▸ Percentiles	Preliminary examination of data for comparison with site history and land use activities in the establishment of background. Use as a preliminary screening tool.	Simple and straightforward; less statistical rigor.
Simple Comparisons	Used with very small data sets.	Not recommended
Parametric Tests ▸ Student t-test ▸ Behrens-Fisher Student t-test	Tests require approximate normality of the estimated means. Use if a larger number of data points are available (n > 25). For smaller data sets, examine data for normality or lognormality in distribution.[4]	Statistically robust and used frequently in parametric data analysis.
Nonparametric Tests ▸ Wilcoxon Rank Sum Test (also called the "Mann-Whitney Test") ▸ Gehan Test	Use when data are not normally distributed, as rank-ordered tests make no assumption on distribution.	Statistically robust and used frequently in background estimation.

EPA 540-R-01-003, Sept. 2002, Section 5.1

General descriptions of a few tests are:

- The **Student's *t*-distribution** (or also ***t*-distribution**), which is a probability distribution that arises in the problem of estimating the mean of a normally distributed population when the sample size is small. It is the basis of the popular Student's *t*-tests for the statistical significance of the difference between two sample means, and for confidence intervals for the difference between two population means.

- The **Behrens-Fisher** problem is the problem of interval estimation and hypothesis testing concerning the difference between the means of two normally distributed populations when the variances of the two populations are not assumed to be equal, based on two independent samples.

- The **Mann-Whitney U test** (also called the Mann-Whitney-Wilcoxon (MWW), Wilcoxon rank-sum test, or Wilcoxon-Mann-Whitney test) is a non-parametric test for assessing whether two samples of observations come from the same distribution. The null hypothesis is that the two samples are drawn from a single population, and therefore that their probability distributions are equal.

The VAP also accepts the **Ohio EPA DERR Guidance (DERR-00-RR-039, 6/30/04)** as an acceptable approach when determining the statistical significance between two datasets.

SUMMARY: The VAP allows flexibility of statistical method when comparing data sets for significant differences. The key guidance cited refers to statistical techniques for site versus background concentration determinations; however the same methodology is applicable to a non-CL versus CL datasets comparison. It is up to the discretion of the volunteer to select which statistical methodology best fit the data sets and appropriately apply the test.

OHIO EPA CONTACT: For any questions concerning this issue, please contact the VAP central office at (614) 644-2924.