6630 U-004-307 .6

FERNALD RISK ASSESSMENT POLICY, GUIDELINES FOR DETERMINING CONTAMINANTS OF POTENTIAL CONCERN - (USED AS A REFERENCE IN OU 2 RI REPORT)

08/26/93

93-3
FERMCO DOE-FN
11
POLICY

FERNALD RISK ASSESSMENT POLICY

The following is an approved policy for the preparation of environmental risk assessments for the Fernald Project. All risk assessments must be prepared consistent with the policy.

Policy Name: **Guidelines for Determining Contaminants of Potential Concern**

Summary/Comment: **This policy provides the requirements for determining CPCs and must be followed for all project risk assessments.**

Policy Number	93-3
Approval Date	08/26/93
EPA Approval Obtained	YES
Additional Discussion with EPA Required	NO
Change to Risk Assessment Work Plan Addendum	YES
Contact Person	Matt Hnatov / Joe Prince

Approval:

[signature] 08/26/93

Environmental Risk Assessment Manager **Effective Date**

August 19, 1993

GUIDELINES FOR DETERMINING CONTAMINANTS OF CONCERN

The following process will be used to determine contaminants of concern for the Fernald project. Consistent with EPA direction and FERMCO policy, the following terms will be used in all risk assessments.

- Constituents of Potential Concern (COPCs)- all constituents which are determined or assumed to be above background following statistical analysis of the data.
- Contaminants of Potential Concern (CPCs)- all contaminants which remain a concern after the toxicological pre-screening procedure has been applied to the CPCs.
- Contaminants of Concern (COCs)- all contaminants which remain a concern after fate and transport evaluations, exposure point calculations, risk calculations, and post screening evaluation (i.e.; at the end of the baseline risk evaluation process).

The process of determining COCs involves statistical screening, pre-screening, and post-screening procedures as defined below.

I. Statistical Analysis Procedure for determining constituents of potential concern.

Statistical comparison of data collected on-site (soils, sediment, sludge, surface water, perched groundwater and groundwater data down-gradient of the site) versus background data will be accomplished as follows:

A. Determination of the proper statistical analytical procedure hinges on the datasets being analyzed. The method of choosing the statistical analysis to determine COPCs is shown in fig. 1. Procedures which will be used include:

1. t-Tests (ANOVA for multiple comparisons) on log-transformed data - If the datasets can be shown to be lognormally distributed (which is often the case with environmental analyte concentration data) and the proportion of non-detects is not too excessive and the log-variances are similar, the data will be log-transformed and t-Tests will be run to compare site data to background data. If there is sufficient data Maximum Likelihood Estimates (MLE) will

be used to estimate the means and standard deviations used in the t-Tests. MLEs have been shown to be more accurate and less sensitive to higher proportions of non-detects than simple substitution methods.[1] The accuracy of the results depends on the level of adherence to the assumptions of log-normality of the data. Probability plots and non-parametric tests such as Kolmogorov-Smirnov test or the Shapiro-Wilk test (Shapiro-Francia test for datasets with n>50) will be used to assess the Normality of the log-transformed data. Bartlett's test for equality of variances will be used to assess comparability of log-variances.

When there are limited data or when there are few or no non-detects and the detection limits are low in relation to the quantifiable concentrations, the parameter estimates will be almost identical to those estimated using the simple substitution method (most commonly replacing non-detects with one half of the sample quantification limit). If are fewer than 15% non-detects the simple substitution method (1/2 SQL) will be used in the calculation; this procedure is quicker and the results are nearly identical.

If the probability level of the t-Test (or the ANOVA) is statistically significant (at approximately the 95% confidence level) then we conclude that there is sufficient evidence to indicate that the site constituent concentrations are different from background and therefore the constituent will be considered a CPC.

2. t-Tests (ANOVA for multiple comparisons) on untransformed data - Standard t-Tests (ANOVAs) procedures will be employed if the data can not be shown to be lognormally distributed but can be identified as normally distributed and the variances (site and background) are not statistically significantly different from each other. MLE estimates or estimates using simple substitution will used under the same conditions as above except that the data will not be transformed.

3. Parameter estimation techniques based on other distributions (e.g. Weibull, F, exponential, etc.) may be used if there is strong statistical and

---

1 Statistical Support Document for the 40 CFR, Part 503 - Final Standards for the Use or Disposal of Sewage Sludge, Volume I.

literature evidence that these distributions better describe the data. Although these distributions are not typical for environmental data, sometimes radionuclide data follows one of these distributions.

4. Non-parametric ANOVA (Kruskal-Wallis or Mann-Whitney U) - Non-parametric procedures will be used if the data are significantly divergent from the normal or lognormal distributions. (With these non-parametric ANOVA procedures the data values are replaced by there relative ranks. Although there is some loss of statistical power, this protects against deviations from Normality which would make the results of a t-Test or parametric ANOVA invalid.) The Kruskal-Wallis is used with multi-category data (e.g. multiple groundwater wells with multiple rounds of data) and the Mann-Whitney U is used for two-category data. (The Mann-Whitney U test and the Wilcoxon Rank Sum are directly related, and are equivalent for all practical purposes.)

5. The distribution of the data from small datasets can not be adequately assessed. Different procedures need to be employed to determine if the constituent should be treated as a CPC.

   a. If there are fewer than five samples and the constituent was detected at least once, a comparison of the maximum detected value to the background 95th percentile will be performed as per I.B. below. In cases where the constituent was not detected process knowledge and professional judgement will be employed to determine if it is probable that the constituent should be present even though it was not detected. Please see Section II for discussion.

   b. If a constituent was detected fewer than seven times but at least four times the CPC determination technique outlined in I.A.4 above will be used.

6. Other techniques - When there are greater than 50% non-detects and MLE parametric estimates fail to yield meaningful results other techniques must be employed to determine if a constituent should be considered a CPC. With high proportions of non-detects the procedure(s) becomes simply a comparison of the distribution of detection limits NOT of

the distributions of the concentration data. One non-parametric procedure that can be used is the Test for Proportions. This test compares the proportion of non-detects between two populations. This procedure categorizes all results into two categories, detect or non-detect, for both datasets and calculates an exact probability (based on the binomial distribution) that the proportions of detect/non-detect are the same. Obviously, this is a drastic analytical technique. But, with a high proportion of non-detects, statistical procedures often yield practically meaningless results. Good scientific and statistical judgement should be used to judge the validity of any result.

B. An additional statistical procedure that will be used to screen for CPCs is the comparison of individual data points to 95th percentile (calculated based on the appropriate distribution) of the uncontaminated background data for each constituent not already identified as a CPC. This technique will be used to protect against "hot spots" being "averaged out". If any site constituent concentration exceeds the calculated 95th percentile then this constituent is added to the CPC list. This procedure will be used only to include in the CPC list constituents not identified previously. It will not be used to remove a constituent from the CPC list.

C. Other dataset comparison consideration:

1. For meaningful results, the science of statistical analysis relies on the adherence to the assumptions of the statistical procedure being employed. Numbers can be run through almost any statistical procedure, but if the assumptions of the procedure are not met the analysis results are meaningless.

2. Parametric procedures should be used when possible. Parametric procedures, those based on the Guassian (Normal) distribution, are more statistically powerful then non-parametric procedures since information about the distribution is known beforehand. Most non-parametric procedures rely on ranks or counts of data and not the observed values as do parametric procedures. Usually, moderate deviations from Normality can be tolerated by most parametric procedures.

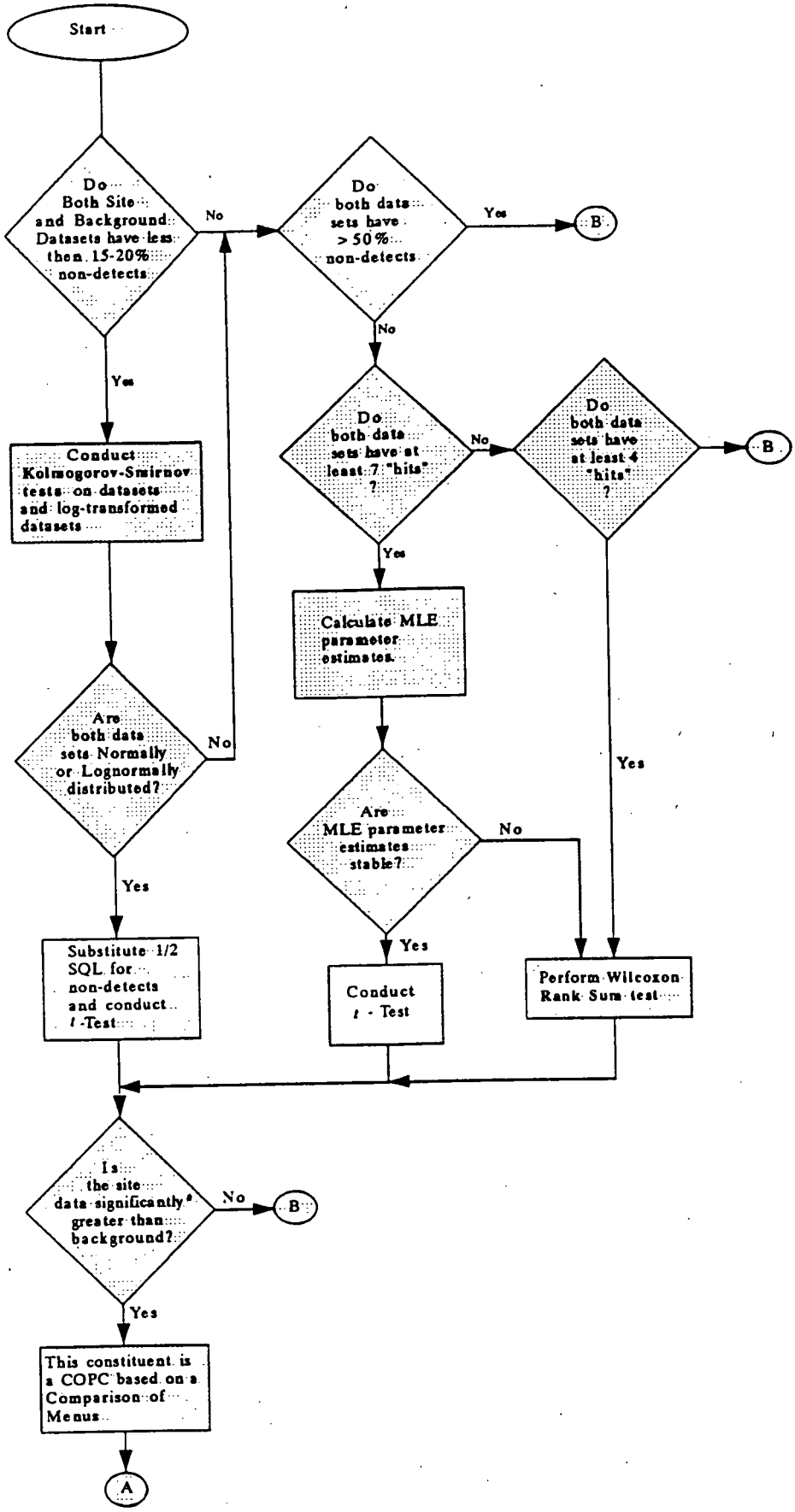
3. Non-parametric procedures such as Kruskal-Wallis, Mann-Whitney or Wilcoxon Signed Rank substitute ranks for observed values. This fact allows for "distribution-free" analysis but the loss of infor-

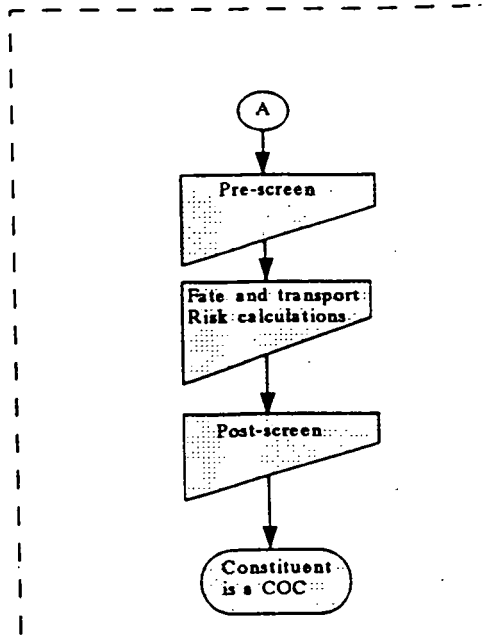
mation reduces the reliability of the results. Generally, non-parametric procedures require larger sample sizes than do parametric procedures to attain the same accuracy level. This should not preclude their use. Non-parametric procedures yield more reliable results than do parametric procedures on data that violate the assumptions of normality.

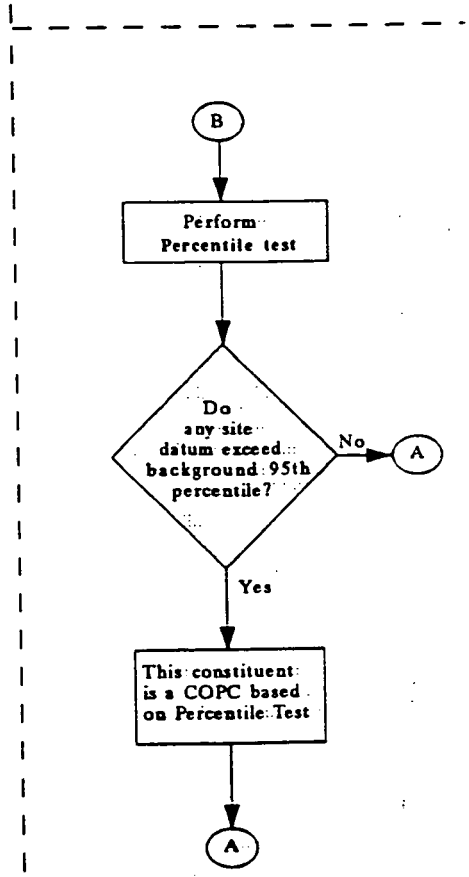
4. Interpretation of statistical results can often be misleading to the non-statistician. The interpretation of a single statistical test (assuming all of the proper assumptions have been met) may be straight forward. However, there are many factors that make interpretation anything but trivial. Some major problems to statistical analysis and interpretation are:

   a. High proportion of non-detects.
   b. Non-symmetrically distributed data.
   c. Apparent bimodal (or polymodal) data - often indicating that the data were probably collected from different populations.
   d. Highly spatially correlated data (non-random distribution of data).
   e. Large differences in variability between site and background data.
   f. Individually, statistical procedures may fail to detect significant differences, but a combined analysis of all constituents may yield highly significant results. Example, say there are ten potential CPCs. Now say that all ten potential CPCs demonstrate site mean concentrations greater, BUT not significantly greater (at the 95% level), than background. If there were no differences between site and background we would expect about an even split between which mean concentration were higher (site or background). Assuming that the populations of site and background were the same (no contamination) the chance that all ten samples means were greater than the background means is approximately 1 in 1000. Clearly, there is evidence here to warrant additional analyses.

Figure 1







Footnotes:
a. the significance level should not be set at an absolute level (e.g. 5%). Any constituent with a questionable significance level should be carried through to post-screening.

II. Toxicological Screening Procedure for Determining Contaminants of Potential Concern (CPC's) and Contaminants of Concern (COC's).

A. Screening Procedure (post-statistical analysis).

1. Evaluate contaminant concentration source terms, frequency of detection persistence and distribution among media.

2. If a contaminant is identified in one medium but not found in any other media or is found at unusually high (unrealistic) concentrations, consider removal unless additional evaluation requires it be included as a CPC.

3. Delete essential micronutrients and macronutrients that are non-toxic at the levels identified (e.g., Na, K, Mg, Ca, Fe, Zn, Mg, Se, etc.).

4. Identify and remove those results identifying non-specific classes of chemical compounds [e.g., Total Organic Carbon, (TOC), Total Petroleum Hydrocarbons (TPH), Polynuclear Aromatic Hydrocarbon's (PAH's), Chlorinated Hydrocarbons (CH's), etc.].

5. Identify and remove those compounds known to be derived from off-site sources (e.g., off-site factory discharges, auto exhausts, various herbicides, insecticides, etc.). However, consider their inclusion if contribution to site risk is significant (e.g., 2,3,7,8-dioxin) and explain relevance.

6. Identify and remove compounds that are ubiquitous in nature and considered to be non-toxic (e.g., Al, Si, Cl). Compare inorganics with background levels; identify background risk. Assess toxicity due to simultaneous exposure from multiple media/pathways.

7. Identify and remove those volatile compounds with vapor pressures less than 10 mm Hg @ 20 degrees C. Consider meteorological conditions.

8. Compounds known to degrade to non-toxic products due to physical, biological, chemical activities (e.g., photolysis, biodegradation, pump and treat, etc.) or that may have been produced by chemical synthesis during analysis (Aldol-condensation) can

be removed if no other requirements for inclusion exist.

9. Consider removing compounds whose potential for toxicity is low (>5.0 Gms/Kg, B. Wt.) and whose point source concentrations are less than 1.0 ppm.

10. Compare the list of contaminants (remaining and removed) with the values in the Region III EPA screening document.

11. Review the list of chemicals removed and identify those whose toxic effects are exerted upon a common target organ, or have synergistic or additive effects. Examine toxicity, concentration and any additive/synergistic effects due to concomitant exposure. Consider structure-activity relationships or other chemical similarities. Assess possible contribution to site risk and consider returning to CPC list. Discuss over/under estimation of risk due to such grouping.

12. Provide scientific rationale and/or documentation to support decision to remove contaminants.

14. Class A/B carcinogens must be included as COC's. Evaluate PAH's under TEF and BAP procedures.

B. Post-Screening Procedure

1. If assessment of a contaminant results in a risk of $\leq 10^{-7}$ following exposure point calculations and risk calculations, consider removal.

2. Using the Concentration-Toxicity screen; calculate the chemical scores and relative risk. Identify those contaminants, on a percentage basis, in a particular medium that will most likely contribute the major risk for that scenario.

   a. Chemical contaminants whose relative contribution to the total risk from a medium is less than 1% (RAGS), may be eliminated. After removal however, you should;

      1. Re-examine the chemicals discarded. Is a particular target organ impacted by those chemicals? Consider possible additive or synergistic effects?

2. If a particular target organ is affected by such chemicals, total the relative percentage contributions for each of the chemicals or congeners. If the total risk of all equals or exceeds 5% of the total risk, re-enter those chemicals and identify the reason for their inclusion.

III. If you have **any** questions regarding use of these guidelines or have a situation not covered by the above, please call Matthew Hnatov, Environmental Statistician (6934) or Dr. Joseph Prince, Senior Toxicologist (8973) in the Risk Assessment Office for assistance.