NDEP Guidance on Qualifying Data due to Blank Contamination for the Basic Management Incorporated Complex and Common Areas

1. Background and Current Practice

Recent Nevada Division of Environmental Protection (NDEP) guidance specific to qualifying data due to blank contamination is found in the February 26, 2009 *Supplemental Guidance on Data Validation* with additional clarification provided in the March 19, 2009 *Supplemental Guidance on Data Validation*. The February and March 2009 supplements were established by NDEP because of an updated version of the United States Environmental Protection Agency (EPA) National Functional Guidelines (NFG) for Superfund Organic Methods Data Review (June 2008). The 2008 guidance from USEPA included a new algorithm for qualifying volatile organic chemical (VOC) results based on blank contamination, and the NDEP guidance extended this approach to semi-VOCs (SVOCs).

Historically the USEPA NFGs have used a factor (e.g. 5X, 10X) to determine whether results, that have been associated with blank contamination, should be censored. Briefly this rule compares the blank level to the sample level and requires that the sample level be greater than that found in the blank by some factor. For Inorganic Data Review (OSWER 9240.1-45, October 2004) this factor is 10X (see Table 4, page 17 for example), although this factor is only applied in specific cases. The 2009 organic chemical NFGs revised this slightly for VOCs, reducing the multiplication factors to 2X. The March 2009 NDEP guidance recommends in most cases that the censor level should be set to the reported concentration when the difference between the sample and blank level is a factor of 2 or less. Sample concentrations above this factor are not censored, though a J qualifier is applied with an associated reason code to indicate contamination in an associated blank.

The latest USEPA NFG for Inorganic Superfund Data Review (January 2010) continues to use a complicated set of algorithms to determine the censoring approach. The NFGs base their comparison on two sensitivity indicators, a Method Detection Limit (MDL) and a Contract Required Quantitation Limit (CRQL). The CRQL is analogous to the Practical Quantitation Limit (PQL), and the MDL is analogous to the Sample Quantitation Limit (SQL) for the purposes of this NDEP guidance. These algorithms depend upon the blank type, and levels of both blanks and samples compared to the associated MDL and PQL. The 10X factor is part of the algorithm in several cases where a comparison to blank contamination is used. For example, when the preparation blank is greater than the CRQL and the sample is above the blank level a 10X factor is applied. The USEPA NFGs also recommend that the sample values that are less than the CRQL but greater than the MDL should be censored if there is blank contamination within or above this region. This USEPA algorithm and censoring approach is overly complicated and results in censoring data that may bias the data high. NDEP understands that the relative uncertainty around an analyte value is greater when it is below the PQL, but does not believe using a single datum approach to decision making is best for the types of decisions typically made at the Basic Management Incorporated (BMI) Complex and Common Areas.

There are two main conditions to consider when evaluating a single result (datum) for presence of an analyte. For these conditions, the PQL does not have a separate decision framework:

Assuming the Blank and Sample values are greater than the MDL/SQL and:

1. one or more associated Blank values > Sample value, then:
    a. The sample contains NO (significant) native analyte, or,
    b. The sample contains some percentage of native analyte.
2. the Sample value > all associated Blank values, then:
    a. The sample contains some percentage of the native analyte

This discussion recognizes that blank levels need to be compared to sample levels on an equal basis. If dilution factors, different matrices (soil versus water), or sample weights and volumes complicate the comparison, the comparison will need to be performed on the raw data (e.g., counts, areas). In addition, it is recognized that some analytical techniques have a sensitivity that will pick up a fairly static level of background signal. These techniques include High Resolution Gas Chromatography Mass Spectrometry (e.g., HR GC/MS of PCB congeners, and dioxin/furans) and Inductively Coupled Plasma Mass Spectrometry (ICP-MS). This static background level issue is not the typical laboratory contaminant case encountered with phthalates or methylene chloride. In most cases these static levels are much less than important action levels. However, there are cases where laboratories have prevalent contamination that can significantly impact sample data such that many of the sample results are censored based on blank contamination. Examples at the BMI Complex and Common Areas include formaldehyde and in some instances metals using ICP-AES.

Under Case 1, the USEPA NFG would recommend the result to be censored, in most cases at the CRQL (PQL for the BMI Complex and Common Areas). The presumption is that the majority, if not all, of the analyte in the native sample is from blank contamination. Under a datum (single value) decision framework this might be logical. For example, if this single value is compared to an action level. If the action level is greater than the sample concentration, it is reasonable to conclude the true value is very likely below the action level. If instead the action level is below the sample concentration, then the blank and sample concentrations must be investigated in more detail.

For the BMI Complex and Common Areas, this censored datum is rarely considered separately, but used in an analyte distribution. This value is but one of a data set, from which a decision must be made relative to background conditions or action levels. The typical action is then either to add this censored datum to the distribution at the PQL level or perhaps some multiple of this level (e.g., one half the PQL). This results in a biased distribution, often a high bias as most contamination is less than any PQL, <u>and often less than one half the PQL</u>. If instead, the laboratory sample value is reported with an associated qualifier and reason code, a decision can be made with better information. Both the reported value and information about the associated contamination would be provided.

Under Case 2, the EPA NFG would also recommend the result to be censored, unless the sample value is greater than the PQL with sufficient difference between the blank and sample values. The logic here is that the sample value contains some amount of contamination, and is therefore only usable if there is sufficient confidence (some factor is used) that the native amount is (significantly) greater than the blank. When the sample value is greater than the PQL, the relative uncertainly associated with the concentration is considered less than when the sample value is below the PQL. However, in those cases where the value is less than the PQL, but greater than the blank, there is some likelihood that some level of the analyte truly is associated with the native sample. Similar to Case 1, if the sample value from the laboratory is reported with an associated qualifier, then more complete information is provided for decision-making.

2. Reasoning behind Recommended Change in Qualifying Data

Censoring results in loss of data and therefore information. In cases where data quality indicators indicate severe bias, such as low spike recoveries, censoring is often justified. But in the case of blank contamination, the data should not be censored. Following the EPA NFGs, censoring is done *a priori*. This is before a complete understanding is gained of how the data will be used. By not censoring during the data validation step, but understanding the influence of blank contamination and including this information in the data usability and subsequent decision-making, the full complement of data are still used and available for the decision making. Further, the approach recommended in this guidance requires no additional effort in data reporting and attempts to provide a path to using as much of the data as has been collected.

In many instances the approach taken when blank contamination is evident (censor or only qualify) may have little influence on the ultimate decision(s). This is common when the concentrations of most samples are significantly above or below any risk-based action level of interest. Under this scenario, blank contamination is often insignificant with respect to the risk-based decisions that will be made. However, the most critical cases to consider are when the sensitivity of the analytical method is near a risk-based action level, and the blank contamination and sample concentrations are near the limits of detection. For these cases, the full data set needs careful consideration to support a reasonable risk-based decision.

Many types of blank samples may be associated with a set of samples, including field and laboratory blanks (e.g., calibration, preparation). It is impossible to associate a particular blank with a particular sample and it is possible that even though there is contamination of the blanks, this is not true for the samples. Recoveries of laboratory control spikes are one way to assess this. If the recovery is very close to the expected recovery, any contamination in the blanks may not necessarily be associated with samples. EPA guidance has always recommended comparing sample values against the highest blank concentration in cases where more than one blank is associated with the sample. Blank concentrations levels can also change with time, which can be shown using continuing calibration blanks. Blank concentrations can also exhibit some random response. The challenges with adequately characterizing blank samples, and the potential lack of association of blank concentrations with sample concentrations suggest that sample data should not be adjusted or censored, but that the sample concentration should be reported along with a qualifier that indicates blank contamination if appropriate.

There is an additional reason for not continuing to censor data due to blank contamination. This issue involves the relationship between the BMI Complex and Common Areas companies (the Companies) and the commercial laboratories. When there are examples of blank contamination that are unexpected, the typical approach is to just censor the results and continue with the project. This provides no incentive for the laboratory to improve their operations. By not censoring the data, but using it in the decision analysis, the impact of these laboratory practices will become more apparent and hopefully improved upon.

3. Example Data Sets

Results from several BMI Complex and Common Areas projects were reviewed to show instances where censoring of sample results because of blank contamination has resulted in data sets that were adversely impacted. In most instances the effect is to bias the data set high, since the censored level is greater than the sample reported (actual) value. In several cases large numbers of data were censored well above the original reported levels. Data from the Mohawk Sub-Area Soil Investigation (Datasets 52, 52a, 52b) are provided as examples below.

Results from 83 samples collected at the Mohawk site for antimony were all adjusted to due blank contamination. The mean and median values of the actual reported samples (unadjusted) were 0.33 and 0.31 mg/kg respectively. In most cases the values were adjusted up to the quantitation limit of 1 mg/kg, in some cases higher. The resulting mean and median values are 1.33 and 1 respectively. This resulting shift in the distribution to these much higher central values is impacting the comparison of this data set to background values. Antimony data from the 2010 Background Soil Compilation Report, Table 2 shows both the censored (non-detect) data (mean and median values of 0.33 and 0.24 respectively) and detect data (mean and median values of 0.199 and 0.175 respectively) are below these adjusted Mohawk mean and median values. The effect of censoring due to blank contamination is to change the background comparisons decision.

Similarly for boron, the Mohawk uncensored mean and median values are 7.35 and 7.05 respectively with the censored values at 34.4 and 21.75. These censored values are well above the background levels for boron where the mean and median values in the detected data set are 7.85 and 6.6. Similar examples can be shown for mercury, thallium, molybdenum, and selenium. These examples show that when censoring is performed at the PQL for blank contamination inappropriate background comparison decisions can be made.

4. Proposed Changes

NDEP is recommending that instead of censoring data due to blank contamination at the data validation step, the Companies should follow the same approach taken by the laboratories. That is, the sample result should be reported, and a qualifier should be applied to the associated data along with sufficient information to understand the level of contamination relative to that found in the samples. These data will therefore have a J qualifier, with an associated reason code indicating blank contamination is associated with the results. These data will be carried through to the data usability and analysis process.

By using a single common approach across all data sets where contamination is recognized but data are not censored during data validation, data comparability is ensured.

The impact of the blank contamination will be evaluated in the data usability analysis and considered within the context of any decisions. Contamination will need to be considered on an equal basis and dilution factors, different matrices (soil versus water), or sample weights and volumes recognized.

NDEP Recommendation

In cases of blank contamination, report the sample result with a J-flag qualifier and a reason code that notes the blank contamination with sufficient information to understand the level of blank contamination relative to that found in the sample. If the sample result is less than the SQL, then report a non-detect at the SQL.

In effect, report the sample result as normal, whether a detect or a non-detect relative to the SQL, add a J-flag qualifier, and a reason code as described above.