

CONTAMINANTS OF POTENTIAL CONCERN WHITE PAPER

1. OVERVIEW

Selection of the appropriate contaminants of concern is critical to preparing an assessment that is representative of Hanford Site risks and useful for making remedial action decisions. The risk assessment must be able to differentiate between background materials, non-site-related materials, and contaminants directly related to waste sites. Several comments on the Draft A 100 Area and 300 Area Component of the River Corridor Baseline Risk Assessment (RCBRA) report questioned the ability of the risk assessment to accurately present risks directly related to the area of study. In response to these comments, a process was developed to identify and focus the contaminants of potential concern (COPCs) identified for the risk assessment evaluation. This process is needed so that the ecological and human health risk assessments focus on relevant Hanford Site contaminants and provide risk managers with the information they require to make informed remedial action decisions.

Problems associated with including non-Hanford Site process analytes in the risk assessment were noted during development of the Draft A report. For example, six “naturally occurring” radionuclides (potassium-40, isotopic radium, and isotopic thorium) were included in the initial assessment, but the calculated risks from these radionuclides were high enough to mask the impact of the site-related constituents and had to be presented separately. While it is prudent to distinguish between natural and anthropogenic radioactivity, this should be accomplished using a deliberate, systematic process based on established selection criteria. The lack of such a process resulted in an unnecessarily complex report and was not well received. There were several other comments on the need to identify Hanford Site COPCs and the effect of reporting risks for non-Hanford site-related analytes. The process proposed in this white paper addresses many comments on the Draft A report.

The purpose of this white paper is to provide information on the regulatory guidance relative to COPC selection and also present a logical, transparent, and scientific approach to COPC selection that is consistent with regulatory guidance. Ultimately the contents of this paper will be included in Draft B of the RCBRA report. The approach addresses both ecological and human health risk. There are parallel COPC selection processes described for each assessment that are basically the same but differ in sample media emphasized.

The process presented in this white paper builds upon approaches and methods for COPC selection presented in the Tri-Party approved 100 Area and 300 Area Component of RCBRA Sampling and Analysis Plan (SAP); (DOE-RL 2005a). The SAP outlines a process for focusing contaminants based on comparing mean concentrations at waste sites or operational areas to background or reference sites using conclusions and data summaries from limited field investigation (LFI), cleanup verification package (CVP), Hanford Site monitoring, and related projects (Figure 2-1 in DOE-RL 2005a). The process is also consistent with applicable regulations and guidance. The COPC refinement process includes evaluation of radiological half-life and essential nutrients, evaluation of detection status, comparison of Hanford Site data to background and reference site data, evaluation of potential toxicity, and a chemical-specific narrative analysis.

Defining the COPC list will aid in understanding whether risks are related to Hanford waste sites. Further information on data sources is provided in Section 3, and information on the graphical and statistical comparisons to background are provided in Section 4.

2. REFINED COPC SELECTION PROCESS

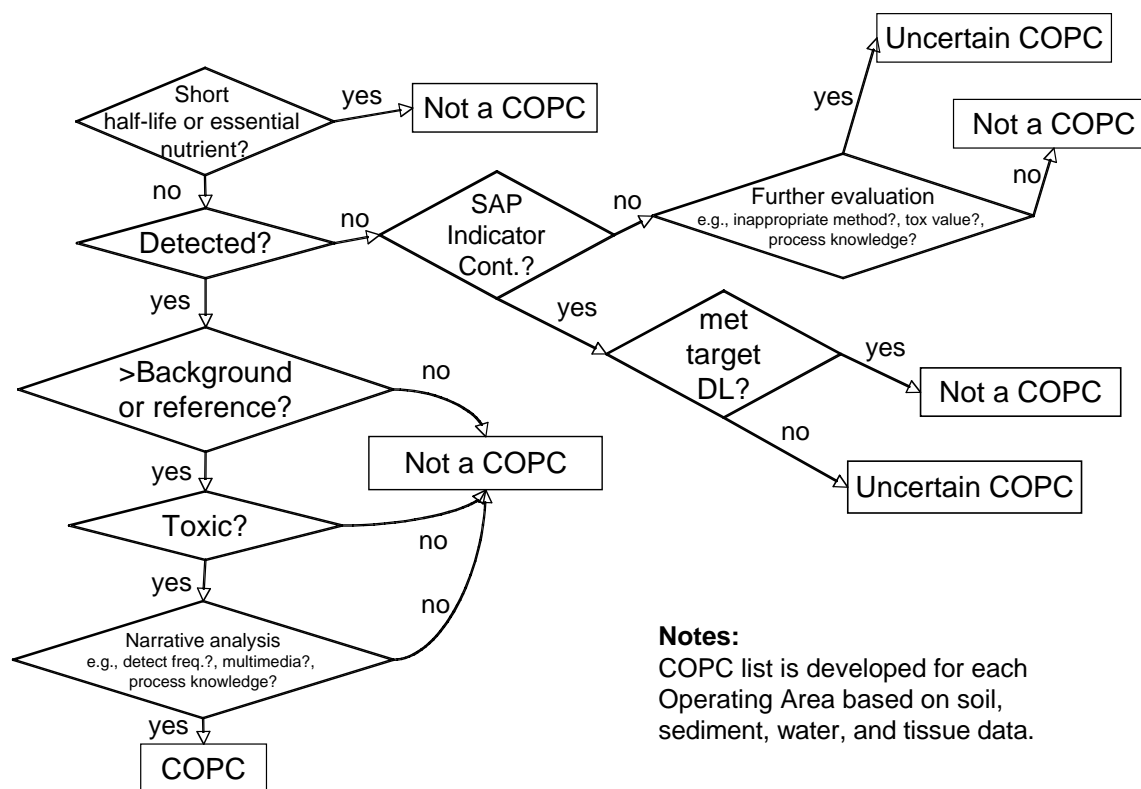
The refined COPC selection process for the Draft B risk assessment report is consistent with U.S. Environmental Protection Agency (EPA) guidance (see Attachment A for excerpts). Inherent in these approaches are some basic considerations and requirements for approval of the process prior to implementation.

The process proposed for Draft B of the 100 Area and 300 Area Component of the RCBRA includes the following components:

- Radiological half-life and essential nutrient status (remedial action goals [RAGS]; Chapter 5 step 9)
- Quantitative assessment based on
 - Data quality review of detection status, quantitation limits and qualifiers for all data (RAGS; Chapter 5 steps 1-6)
 - Comparison to Washington State background and Hanford Site background for soil and groundwater (RAGS; Chapter 5 step 7)
 - Comparison to reference site data for soil, biotic tissues, sediment, and pore water (RAGS; Chapter 5 step 7).
- Qualitative assessment based on
 - Degree of toxicity evaluation, drawing upon human health and ecological toxicity information (RAGS; Chapter 5 step 9)
 - Detailed narrative assessment drawing from one or more of the following; results of the quantitative COPC assessment, spatial or temporal trends in concentrations in a sample medium, differences in detection status or concentration among environmental media, operational process knowledge, Hanford Site monitoring data, detection frequency, contaminant lists from CVP/remaining site verification package (RSVP) reports, LFI/qualitative risk assessment (QRA) information, and indicator contaminants listed in Tables 1-1 and 1-2 in the 100 Area and 300 Area Component of the RCBRA SAP (RAGS; Chapter 5 step 9).

Proper application of this process will include detailed quantitative analyses and the results of this process will be documented in an analyte-specific narrative in Draft B of the report. The process as applied to the ecological and human health risk assessments is summarized in Figure 1.

Analytes identified as COPCs will be the focus of the risk assessment report. Some of the remaining COPCs, particularly non-detect analytes with detection limits exceeding target practical quantitation limits (PQLs) and analytes eliminated by process knowledge, will be evaluated as uncertainties in the risk assessments. Figure 1 shows the steps involved with COPC refinement and these steps are considered in further detail below.

Figure 1. COPC Refinement Process Overview.

2.1 Eliminate Short Half-life Radionuclides or Essential Nutrients

Among radionuclides, primary radionuclides¹ having a half-life exceeding 3 years are retained as COPCs. Radionuclides with shorter half-lives would not be associated with historical Hanford Site operations due to decay over the period of time since operations ceased and, as such, are excluded from the quantitative risk calculations. These include the following:

- Actinium-228; 6-hour half-life (accounted for in the decay chain of radium-228)
- Beryllium-7; 53-day half-life
- Cesium-134; 2.1-year half-life
- Lead-212; 11-hour half-life (accounted for in the decay chain of thorium-228)
- Lead-214; 27-minute half-life (accounted for in the decay chain of radium-226)
- Manganese-54; 312-day half-life
- Radium-224; 3.7-day half-life (accounted for in the decay chain of thorium-228)
- Ruthenium-106; 1.0-year half-life
- Thorium-234; 24-day half-life (accounted for in the decay chain of uranium-238).

¹ Radionuclides with short half-lives (less than 30 days) that are within the decay chain of a primary radionuclide are accounted in the dose conversion factor or slope factor of the parent nuclide.

One additional radionuclide, thorium-228 (half-life of 1.9 years), also has a half-life of less than 3 years. However, because the activity of thorium-228 rapidly reaches equilibrium with its parents (radium-228 and thorium-232), thorium-228 is not excluded from consideration based on its short half-life.

As stated in Section 5.9.4 of EPA (1989a), essential nutrients that are present at relatively low concentrations and are toxic only at high concentrations need not be considered in a quantitative risk assessment. The elements sodium, magnesium, calcium, and potassium are among those cited in EPA (1989a) as examples of such nutrients.

2.2 Data Quality Review and Detection Status

Understanding the quality of data, as well as data utility and uncertainties, is imperative to performing a defensible human health and ecological risk assessment. Data collected must meet the needs and satisfy requirements of the risk assessment as stated in the *100 Area and 300 Area Component of the RCBRA Sampling and Analysis Plan (SAP)* (DOE-RL 2006). Because this risk assessment report considers both RCBRA project-specific and historical data, these data types are evaluated separately as described in the following subsections.

2.2.1 RCBRA-Specific Data Quality

Analytical quality assurance requirements prescribed in the Quality Assurance Project Plan (QAPP) of the SAP (DOE-RL 2006, Section 2) pertain specifically to the data collected for the investigation areas identified for the RCBRA risk assessment effort. The QAPP specified the analytical performance requirements for soil, sediment, water, and biota collected under direction of the SAP. For each media type the analytical methods, required detection limits, precision, and accuracy were specified. Some of the data reported for the RCBRA samples were not usable for a variety of criteria; these unusable data are not considered quantitatively in the report, but they are considered qualitatively as data gaps.

Analytical results for soil, sediment, water, and biota collected for the RCBRA investigation are evaluated against the quality criteria specified in the QAPP (DOE-RL 2006, Section 2). As a measure of data quality, analytical results identified as non-detects in the RCBRA data set (i.e., results for soil, sediment, water, and biota qualified by the laboratory, reviewer, or validator as “U,” and subsequently assigned detect status of “FALSE”) are compared to the laboratory required detection limits prescribed in the QAPP. Non-detect results reported at values higher than the prescribed detection limit are identified for additional consideration (uncertain COPC). These results, where the non-detect result exceeded the target PQL¹, will be acknowledged as uncertainties in the risk analysis.

Analytes that are non-detected, but either have no target PQL in the RCBRA SAP or meet the target PQL, are further evaluated in the report. This further evaluation leads to identifying the contaminant as an “uncertain COPC” or “not a COPC” based on whether the appropriate method was used for analysis, information on process knowledge from waste site sampling, and a toxicity evaluation. Detection status is evaluated for the spatial scales considered for calculation of representative concentrations (see Section 2 of the Representative Concentration white paper).

¹ A value approximately five to tenfold higher than the analytical method detection limit corresponding to a reasonably high degree of confidence in the reported result.

2.2.2 Quality of Data from Non-RCBRA Sources

Non-RCBRA data sources included CVPs and RSVPs, project-specific investigations conducted by other contractors (e.g., 100-B/C Pilot Project [DOE-RL 2005b], 100-NR-2 investigation [PNNL 2005]), and historical groundwater and environmental monitoring reports. Data collection, analysis, and quality requirements for each of these sources were specified in their project-specific planning documentation, which may have differed from requirements of the RCBRA SAP (DOE-RL 2006). Evaluation of analytical results against project-specific quality planning documentation will not be performed for non-RCBRA data sources; such data are presumed to meet the minimum quality project-specific criteria. Results from non-RCBRA sources will not be compared to the quality criteria listed in Section 2.0 of the SAP (DOE-RL-2006). Their appropriateness for use in the RCBRA will be determined by evaluating individual results against more general quality usability criteria that were presented in the Draft A version of the report.

2.3 Background

Background comparisons are possible for selected media and analytes based on the availability of appropriate data. Background data are available for soil. In the case of soil data, background data include natural background for Washington State (eastern region) for selected inorganic chemicals and Hanford Site background for selected radionuclides and inorganic chemicals. For analytes with background data from multiple sources, the Washington State (eastern region) data will be used in preference to Hanford Site background. The data behind the 90th percentile of soil background are available for statistical comparisons.

The statistical tests proposed for the background comparisons include the following:

- Wilcoxon rank-sum (WRS) test or Gehan test (tests for shift in central tendency of the data). These tests are used to detect differences in the interquartile range of the data (25th percentile to the 75th percentile), or the range represented by the box in box and whisker plots. These tests detect an overall shift in site data compared to background.
- Quantile test (tests for difference in upper percentiles versus background). This test evaluates the upper range data sets and is sensitive to a small number of observations from one or more sites being greater than background. Thus, the Quantile test complements the WRS test and provides a greater chance to detect differences among a small set of sample results.
- Slippage test (tests for difference in upper range of data compared to background). The Slippage test evaluates the highest values and is sensitive to detecting as few as one elevated measurement in the site data that are greater than background. Thus, the Slippage test complements the WRS test and the Quantile test and will detect a difference in concentration from a single sample or a single site.
- Chi-square assessment of detection frequency (greater frequency of detections compared to background). The Chi-square assessment is useful for analytes that are not frequently detected. It can be used in concert with other tests depending on the detection frequency.

These tests are proposed to be run in tandem to provide a comprehensive comparison of site data to background. Additional information on these statistical tests is provided in Section 5. If the site data are not significantly greater than background (5% significance level) for all relevant tests, then the analyte is not a COPC (see Figure 1).

Background comparison tests are planned for RCBRA soil data and CVP/RSVP soil data. The RCBRA soil data and most of the CVP data are composite samples and thus represent the average concentration in the sampled area. Some potential bias is introduced by comparing background samples that are generally discrete or grab samples to RCBRA and CVP data. However, running a suite of statistical tests is the most rigorous way of comparing contaminated site analyte concentrations to background, and this approach has been recommended in several papers (e.g., EPA 1989, Gilbert and Simpson 1990, 1992). The tandem testing approach represents an enhancement over the comparison of site data to a single statistic of background (e.g., the 90th percentile), as all fixed percentiles will be exceeded with some frequency.

Background comparisons for soil will be made for three data groups – remediated backfill and native soil waste sites (both with RCBRA-collected data and CVP/RSVP-collected data), and waste sites only with data collected during the CVP/RSVP closeout process. The CVP/RSVP data will be compared to background for each of the six operating areas by pooling data across all waste sites in each area. Although data will be pooled, the statistical tests discussed above are capable of detecting significant differences associated with a single site or a single sample. Performing the statistical comparisons on pooled data has advantages by simplifying data analyses and data presentation; pooled analyses also better supports the scale of decision making for the CERCLA RODs. The specific comparisons considered for each data group are the following:

- Remediated backfill soil data (compare pooled data for 10 waste sites to background data)
- Native soil data (compare pooled data for 10 waste sites to background data)
- CVP/RSVP soil data (163 waste sites organized into 6 operating areas; one comparison for each operating area).

The results of each test will be documented and this quantitative information will be evaluated along with data plots and other information in the process knowledge component of COPC refinement. The frequency of soil data greater than the 90th percentile of background will be tabulated to provide additional information for COPC refinement.

2.4 Evaluation of Reference Sites

Reference site comparisons will evaluate detected analytes in soil, sediment, porewater, and tissue. Although reference site data and background data are expected to be similar for most analytes, reference site comparisons are prepared and tabulated separately to recognize comments provided by some reviewers on Draft A of the report and in workshops. Reference sites are necessary for constituents and media not included in the various sources for background levels. Fundamentally, some reviewers have questioned the legitimacy of reference sites that are located on the current or former Hanford Site. Based on agreements with the Tri-Parties, the Draft B report will expand assessment and analysis of reference sites. The expanded analysis

will include comparison of reference soil concentrations to background concentrations.

The Draft B report will also use additional reference site data collected for other projects:

- Inter-Areas project (four additional near shore aquatic locations)
- Central Plateau Ecological Risk Assessment (two additional locations entirely off the Hanford Site)
- Surface Environmental Surveillance Program and Near Facility data (onsite and offsite) (these data are currently being reviewed to determine sample media, analytical suites, the number of locations, and their proximity to the Hanford Site).

The statistical tests proposed for the reference site comparisons are identical to those described for the background comparisons in Section 2.3. If the site data are not significantly greater than reference sites (5% significance level) for all relevant tests, then the analyte is not a COPC (see Figure 1).

Reference site comparison tests are planned for RCBRA remediated waste site soil and tissue and CVP/RSVP soil data. The RCBRA soil data and most of the CVP data are composite samples and thus represent the average concentration in the sampled area. Because reference site samples are also generally composite samples, there is no clear bias introduced by comparing these data.

Reference site comparisons will include four data groups (remediated backfill, native soil, near-shore aquatic, and CVP/RSVP waste sites). The specific sample media considered for each data group are the following:

- Remediated backfill (compare pooled data for 10 waste sites to pooled data for 5 reference sites)
 - Soil, plant, invertebrate, and small mammal
- Native soil (compare pooled data for 10 waste sites to pooled data for 5 reference sites)
 - Soil, plant, invertebrate, and small mammal
- Near-shore aquatic (compare pooled data for 50 operational area locations to pooled data for 11 reference sites)
 - Sediment, pore water clams, and aquatic macroinvertebrates
- CVP/RSVP (163 waste sites organized into 6 operating areas; 1 comparison for each operating area).
 - Soil

The results of each test will be documented at each scale, and this quantitative information will be evaluated along with data plots and other information in the process knowledge component of COPC refinement.

2.5 Degree of Toxicity

Analytes that have limited toxicity are not evaluated as COPCs. This evaluation is based on the toxicological literature and observations on the measured environmental concentrations. In practice, this evaluation will focus on detected analytes for which appropriate toxicity criteria have not been identified. With respect to the human health risk assessment, these criteria include cancer slope factors and chemical reference dose values published by EPA. An approach consistent with recent EPA-published toxicity information (e.g., Eco-SSLs [<http://www.epa.gov/ecotox/ecossl/index.html>]) will be used for the ecological risk assessment.

2.6 Narrative Analysis

The narrative analysis is proposed as a qualitative assessment to determine whether a detected analyte should be retained as a COPC in one or more sample media. The narrative analysis may encompass several topics to include the following:

- Waste site disposal history and Hanford Site processes (documented in the 100/300 Work Plan, DQO, and SAP)
- Presence and relative concentrations of an analyte in primary receiving media
- Evaluation of potential transport from primary to secondary environmental media (including an assessment of the potential for bioaccumulation)
- Frequency of detection in a sample medium, consistency of detection among different sample media, and spatial distribution within and between environments
- Conclusions and data summaries from LFI/QRA, CVP/RSVP, Hanford Site monitoring data, and related projects (100-B/C pilot and 100-NR-2).

The results of the qualitative assessment will be documented in an analyte-specific narrative with supporting data and citations.

Indicator contaminants were summarized in DOE-RL (2005a). The COPCs associated with the 163 CVP/RSVP waste sites included in this assessment are defined in the various CVP and RSVP reports documenting the interim actions at these sites. COPCs identified for the groundwater monitoring wells are discussed in the annual environmental surveillance reports. This information forms the waste-site or operating area-specific process knowledge and is documented in Tri-Party approved planning documents or reports.

For example, principal contaminants at the liquid waste sites include tritium, carbon-14, cobalt-60, nickel-63, strontium-90, cesium 137, europium-152, europium-154, europium-155, plutonium-239/240, isotopic and total uranium, hexavalent and total chromium, and nitrate. At the solid waste burial grounds, principal radioactive contaminants are similar to those described for the liquid waste sites. Nonradioactive contaminants at the solid waste burial grounds include boron, cadmium, lead, nickel, lithium aluminum alloy, mercury, palladium, uranium, and zirconium.

Three sample media are identified as being particularly associated with residual contamination

from past Hanford Site operations. The two primary receiving media of past Hanford Site releases are subsurface soil beneath liquid waste sites and burial grounds and groundwater. These media are characterized by the CVP/RSVP soil data and groundwater data, respectively. A third receiving medium, surface soils, may have been affected by other release mechanisms, such as upwardly migrating contamination or surface deposition of contaminants from other sources such as overland flow and stack emissions. These surface soils are characterized primarily by the RCBRA upland environment multi-increment soil data.

Additional detected constituents occur in other sampled media, including terrestrial and aquatic biota. For sediments, pore water and associated aquatic organisms in particular, elevated concentrations of certain constituents may reflect the persistence of chemicals released in liquid waste streams directly to the Columbia River. However, since soil and groundwater were affected by the same waste streams as those discharged to the Columbia River, the list of COPCs in soil or groundwater is generally inclusive of potential Hanford Site-related constituents in other media. One exception is whether any bioaccumulative contaminants are identified based on Hanford Site processes and are also detected in biological tissues.

In summary, process knowledge is an integrative approach to identifying COPCs that are associated with Hanford Site waste sites. Thus, this is a critical step in addressing numerous comments on the Draft A report in terms of its complexity, and the difficulty in understanding whether potential risks are related to Hanford Site activities. Focusing the risk assessment on COPCs associated with past Hanford Site operations will therefore provide risk managers with better support for the remedial investigation/feasibility study (RI/FS) and final RODs for the six operating areas.

3. DATA SOURCES

The comparisons use a variety of data sources that represent background, non-waste site sampling, and reference sites. These data sources are summarized in Table 1, and only data representing the same medium will be compared.

Table 1. Summary of Data Resources Used for Comparison to Waste Site and 100 Area and 300 Area RCBRA Investigation Area and Near-Shore Aquatic Location Data. (2 Pages)

Category	Resource (source code is bolded)	Description of Data Extracted
Washington State background	Natural Background Soil Metals Concentrations in Washington State, Charles San Juan, Toxics Cleanup Program, Olympia, Washington, Publication 94-115 , October 1994	Medium: Soil. Non-radionuclides.
Area background	DOE/RL-92-24 , 1995, Hanford Site Background: Part 1, Soil Background for Nonradioactive Analytes, Rev. 4, 2 vols., U. S. Department of Energy, Richland Operations Office, Richland, Washington	Medium: Soil. Data were obtained from HEIS. Non-radionuclides only. Coordinates were not provided.

Table 1. Summary of Data Resources Used for Comparison to Waste Site and 100 Area and 300 Area RCBRA Investigation Area and Near-Shore Aquatic Location Data. (2 Pages)

Category	Resource (source code is bolded)	Description of Data Extracted
Area background	DOE/RL, 1995 , Hanford Site Background: Evaluation of Existing Soil Radionuclide Data, DOE/RL-95-55 , Rev. 0, U.S. Department of Energy, Richland Operations Office, Richland, Washington	Medium: Soil. Data were hand entered. Radionuclides only. Coordinates provided.
Area background	DOE/RL-96-12 , 1996, Hanford Site Background: Part 2, Soil Background for Radionuclides, Rev. 0, U.S. Department of Energy, Richland Operations Office, Richland, Washington	Medium: Soil. Some data were obtained from HEIS and some data were hand entered. Radionuclides only.
Reference	DOE/RL-2005-42 , 100 Area and 300 Area Component of the RCBRA Sampling and Analysis Plan, including Appendix E for Inter-Areas Shoreline Assessment	Media: Soil (discrete and MIS), Biota, Toxicity Bioassays, Dosimeters. Both radionuclides and non-radionuclides
Reference	DOE/RL-2007-50 , Central Plateau Ecological Risk Assessment	Media: Soil (discrete and MIS), Biota, Dosimeters. Both radionuclides and non-radionuclides

4. COMPARING RESULTS BETWEEN OR AMONG GROUPS

This section presents statistical and graphical methods available for comparing two data sets to determine if one data set is significantly elevated over another (Gilbert and Simpson 1990, 1992). Other methods in this section include a visual tool (“Boxplot”) for comparing sites to background or reference sites. A visual inspection of the data is a good starting point for any statistical analysis, followed by an evaluation of the underlying assumptions of the comparison tests about to be used.

4.1 Boxplots

Box and Whisker plots, also known as boxplots (Box, G. E. P., W. G. Hunter, and J. S. Hunter, 1978, *Statistics for Experimenters*, John Wiley and Sons, New York, New York) are useful in situations where a picture of the distribution is desired, but it is not necessary or feasible to portray all the details of the data. A boxplot (Figure 2a) displays several percentiles of the data set. It is a simple plot, yet provides insight into the location, shape, and spread of the data and underlying population. A simple boxplot contains only the 0th (minimum data value), 25th, 50th, 75th and 100th (maximum data value) percentiles. A more complex version included here plots potential outliers as an “x” (Figure 2b). These are points that lie further than 1.5 times the interquartile range (75th percentile minus the 25th percentile) from either end of the central box. Since the boxplot is compact (essentially one-dimensional), several can be placed on a single graph. This allows comparison of the locations, spreads, and shapes of several data sets or different groups within a single data set.

A simple boxplot (Figure 2a) divides the data into four sections, each containing 25% of the data. The length of the central box indicates the spread of the central 50% of the data, while the length of the whiskers shows the breadth of the tails of the distribution. The boxplot also demonstrates the shape of the data in the following manner. If the upper box and whisker are approximately

the same length as the lower box and whisker, then the data are distributed symmetrically. If the upper box and whisker are longer than the lower box and whisker, then the data are right-skewed. If the upper box and whisker are shorter than the lower box and whisker, then the data are left-skewed.

Figure 2a. Simple Boxplot.

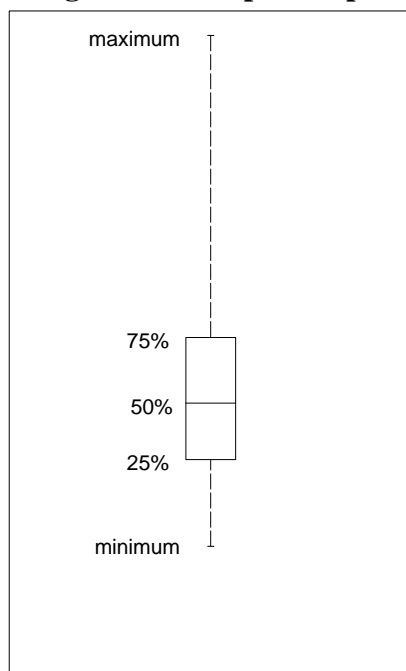
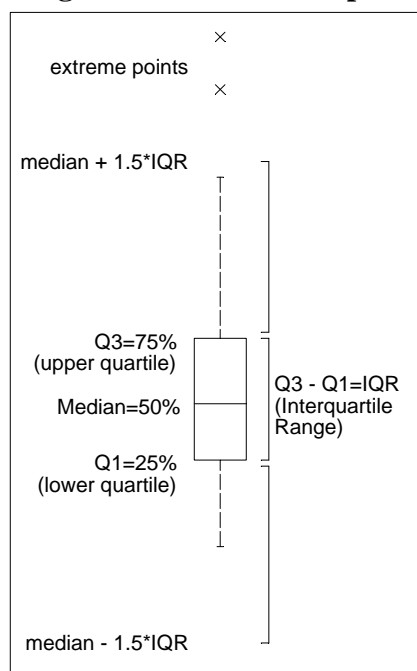


Figure 2b. Outlier Boxplot.



4.2 Comparisons Approach

The statistical tests used for dataset comparisons are the WRS test, slippage test, and chi-square test and are described below.

The WRS test, *Statistical Methods for Environmental Pollution Monitoring* (Gilbert 1987), performs a test for a difference between two population centers (means or medians). This is a nonparametric method that relies on the relative rankings of data values. Knowledge of the precise form of the population distributions is not necessary. The WRS test has less power than the two-sample *t*-test when the data originate from a normal distribution, but the assumptions are not as restrictive. This version of the WRS test uses the Gehan approach for ranking (Gehan 1965, Millard and Deverel 1988). When some of the data are “censored” or reported as below a detection limit, the Gehan approach assigns ranks to the combined set of detects and non-detects in a statistically robust manner. The Gehan approach defaults to standard ranks when the dataset contains no censored data (non-detects). The Gehan ranking approach is recommended in EPA-sponsored workshops and publications due to its broad applicability to environmental data sets (Gilbert and Simpson 1990, 1992).

The slippage test (Gilbert and Simpson 1990) is a nonparametric test appropriate for comparing between datasets with low detection rates. This test is based on the maximum observed

concentration in the background data set and the number (“n”) of site concentrations that exceed the maximum concentration in the background set (Gilbert and Simpson 1990). The result (p-value or significance level) of the slippage test is the probability that “n” site samples (or more) exceed the maximum background concentration by chance alone. The test accounts for the number of samples in each data set (number of samples from the site and number of samples from background) and determines the probability of “n” (or more) exceedances if the two data sets came from identical distributions. This test is similar to a “hot measurement” test, a comparison of the maximum site value to an upper percentile or threshold calculated from the background data, in that it evaluates the largest site measurements. It is more useful than a “hot measurement” comparison because it is based on a statistical hypothesis test, not simply on a statistic calculated from the background distribution.

The chi-square test, *Biostatistical Analysis* (Zar 1974), is used in conjunction with 2x2 contingency tables to test for differences in categorical data. It represents the measure of association between two dichotomous variables and so indicates the strength of correlation, or lack thereof indicating independence. In the case of detection rate, it tests whether the frequency of occurrence of detects in a dataset is independent of the category of that dataset, (e.g., whether the dataset originates from a waste site or background locations). The chi-square test does not consider the magnitude of the analyte – only the frequency of detection. The box plots must be reviewed to determine if the difference in detection frequency has any impact on the range of waste site versus reference site concentrations.

5.0 REFERENCES

Box, G. E. P., W. G. Hunter, and J. S. Hunter, 1978, *Statistics for Experimenters*, John Wiley and Sons, New York, New York

DOE-RL, 2005b, *100-B/C Pilot Project Risk Assessment Report*, Volumes I-III, DOE/RL-2005-40, Rev.0, U.S. Department of Energy, Richland Operations Office, Richland, Washington.

DOE-RL, 2006, *100 Area and 300 Area Component of the RCBRA Sampling and Analysis Plan*, DOE/RL-2005-42, Rev. 1, U.S. Department of Energy, Richland Operations Office, Richland, Washington.

EPA, 1989, *Risk Assessment Guidance for Superfund*, Human Health Evaluation Manual, Part A, Interim Final, OSWER Directive 9285.701A, U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, D.C.

Gehan, E. A., 1965, “A Generalized Wilcoxon Test for Comparing Arbitrarily Singly-Censored Samples,” *Biometrika*, Vol. 52, Nos. 1 and 2, pp. 203-223.

Gilbert, R. O., 1987, *Statistical Methods for Environmental Pollution Monitoring*, Von Nostrand Reinhold Company, Inc., New York, New York.

Gilbert, R. O. and J. C. Simpson, 1990, “Statistical Sampling and Analysis Issues and Needs for Testing Attainment of Background-Based Cleanup Standards at Superfund Sites,” in *Proceedings of the Workshop on Superfund Hazardous Waste: Statistical Issues in*

Characterizing a Site: Protocols, Tools, and Research Needs, U.S. Environmental Protection Agency, Arlington, Virginia.

Gilbert, R. O. and J. C. Simpson, 1992, "Statistical Methods for Evaluating the Attainment of Cleanup Standards," Volume 3: *Reference-Based Standards for Soils and Solid Media*, Pacific Northwest Laboratory, Richland, Washington.

Millard, W. P. and S. J. Deverel, 1988, "Nonparametric Statistical Methods for Comparing Two Sites Based on Data with Multiple Nondetect Limits," *Water Resources Research*, Vol. 24, No. 12, pp. 2087-2098.

PNNL, 2005, *Aquatic and Riparian Receptor Impact Information for the 100-NR-2 Groundwater Operable Unit*, Pacific Northwest National Laboratory, Richland, Washington. (URL: http://www.washingtonclosure.com/projects/EndState/docs/100-NR-2_Report.pdf).

Zar, J.H., 1974, *Biostatistical Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey.

Attachment A

Excerpts from Regulation or Guidance

EPA guidance is provided in RAGS Part A Chapter 5

“After a site sampling investigation has been completed (see Chapter 4), a large quantity of analytical data is usually available. Each sample may have been analyzed for the presence of over one hundred chemicals, and many of those chemicals may have been detected. The following nine steps should be followed to organize the data into a form appropriate for a baseline risk assessment:

- (1) gather all data available from the site investigation and sort by medium (Section 5.1);
- (2) evaluate the analytical methods used (Section 5.2);
- (3) evaluate the quality of data with respect to sample quantitation limits (Section 5.3);
- (4) evaluate the quality of data with respect to qualifiers and codes (Section 5.4);
- (5) evaluate the quality of data with respect to blanks (Section 5.5);
- (6) evaluate tentatively identified compounds (Section 5.6);
- (7) compare potential site-related contamination with background (Section 5.7);
- (8) develop a set of data for use in the risk assessment (Section 5.8); and
- (9) if appropriate, further limit the number of chemicals to be carried through the risk assessment (Section 5.9). Prior to conducting any of these steps, the EPA remedial project manager (RPM) should be consulted to determine if certain steps should be modified, added, or deleted as a result of site-specific conditions. Also, some of the steps may be conducted outside the context of the risk assessment (e.g., for the feasibility study). The rationale for not evaluating certain data based on any of these steps must be fully discussed in the text of the risk assessment report.”