Chapter 6
Application of Data to Risk Assessments

This chapter provides guidance for integrating the assessment of data useability to determine the overall level of uncertainty of risk assessment. This guidance builds on each of the previous chapters.

- Chapter 2 explained the risk assessment process and the roles and responsibilities of key participants. Exhibit 5 defined a continuum of level of certainty in the baseline risk assessment result based on the ability of the risk assessor to quantitate or qualify the level of uncertainty associated with the analytical data.
- Chapter 3 defined six data useability criteria and examined preliminary issues that must be considered while planning sampling and analysis activities to increase the certainty of the analytical data collected for the risk assessment.
- Chapter 4 presented strategies for planning sampling and analysis activities based on the six data useability criteria.
- Chapter 5 described how to use each data useability criterion to determine the effect of sampling and analysis issues on data quality and on the useability of data in baseline risk assessment.

The Data Useability Worksheet (Exhibit 63) assists the risk assessor in summarizing data quality across the various assessment phases. This worksheet is the basis for this chapter's discussion of the impact of analytical data quality on the level of certainty of the risk assessment.

6.1 ASSESSMENT OF THE LEVEL OF CERTAINTY ASSOCIATED WITH THE ANALYTICAL DATA

This section explains how to assess the level of confidence in sampling and analytical procedures in the context of the four major decisions to be made by the risk assessor with environmental analytical data. Exhibits in this section apply the data useability criteria, defined in Chapter 3 and appearing on the Data Useability Worksheet, to these four decisions. Data useability criteria affect the level of confidence involved in each decision. The level of certainty in the data collection and evaluation component of risk assessment affects the overall certainty of the risk estimate.

6.1.1 What Contamination is Present and at What Levels?

The risk assessor's first task is to use analytical data to determine what contamination is present at the site and at what levels (i.e., what potential exists for increased risk from the contamination). Exhibit 71 lists the criteria from the Data Useability Worksheet that affect this decision. The most critical analytical data question to be answered before calculating the risk is the probability of false negatives or false positives. False negatives are of greater concern in risk assessment than false positives, since false negatives may result in a decision that would not be protective of human health. False positives cause the calculated risk to be biased high, and are of concern because taking unnecessary action at a site is costly.

> ☛ *The major concern with false negatives is that the decision based on the risk assessment may not be protective of human health.*

Probability of false negatives. False negatives occur when chemicals of potential concern are present but are not detected by the sampling design or the analytical method. The probability of false negatives can be determined by using the following parameters from the Data Useability Worksheet: analytical methods, data review, sampling completeness, sampling representativeness, analytical completeness, analytical precision and accuracy, and combined error.

> ☛ *False negatives can occur if sampling is not representative, if detection limits are above concentrations of concern, or if spike recoveries are very low.*

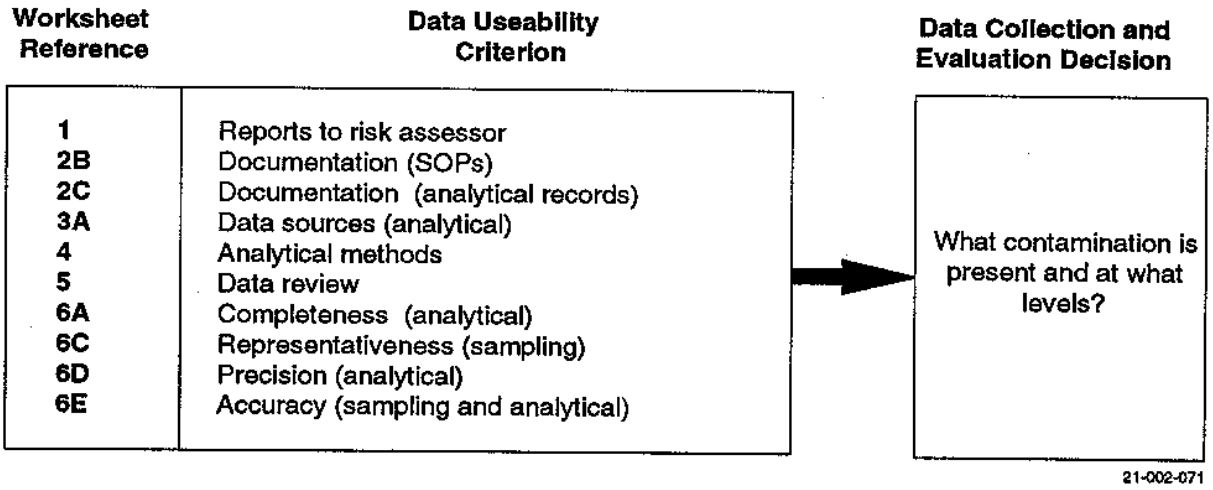
Sampling strategies can increase the probability of false negatives if too few samples were taken or if sections of the site were not sampled. The probability of false negatives increases if sampling of any exposure pathway was not representative.

Knowledge of analyte-specific detection limits is critical to determining the probability of false negatives. Recovery values from spikes, internal standards,

Acronyms

RAGS	Risk Assessment Guidance for Superfund
SAP	sampling and analysis plan
SOP	standard operating procedure

EXHIBIT 71. DATA USEABILITY CRITERIA AFFECTING CONTAMINATION PRESENCE



surrogates, and system monitoring compounds are used to assess the level of accuracy and precision in laboratory data and determine whether the detection limits stated in the analytical methods have been met.

- The probability of false negatives for an analyte is high if the concentration of concern is at or below the detection limit. This probability should have been documented during planning if no analytical methods were found with detection limits below the concentration of concern. If the concentration of concern is very near the detection limit, a false negative can occur because of "drift" in instrument response. This behavior may not be reflected in data from spike recoveries or blanks.
- The probability of false negatives is low if spike recoveries are acceptable, or biased high as documented during data review, and the detection limits are below the concentration of concern for each analyte.
- The probability of false negatives is directly related to the amount of bias if spike recoveries are biased low and detection limits are below the concentration of concern for each analyte. The effect is more pronounced the closer the concentration of concern is to the detection limits.
- The possibility of false negatives should be carefully evaluated whenever sample extracts have been highly diluted (i.e., diluted beyond normal method specifications).

Probability of false positives. False positives occur when a chemical of concern is detected by an analytical method but is truly not present at the site. Assessment of the following parameters from the Data Useability Worksheet can be used to determine the probability of false positives: analytical methods, data review, sampling accuracy, analytical completeness, analytical precision and accuracy, and combined error.

> ☛ *False positives can occur when blanks are contaminated or spike recoveries are very high.*

Sampling and analysis uncertainties connected with false positives can be assessed by examining the results of quality control samples. Blank contamination is the most important indicator of probability of false positives, particularly when accompanied by high spike recoveries. As described in Chapter 5, samples can be contaminated during sampling, storage, or analysis. Field and laboratory blanks identify this problem by determining the level and point of contamination. Sample matrix interferences can also cause false positives. High spike recoveries indicate that matrix interference has occurred.

- The probability of false positives is high if the chemical of potential concern has been detected in any blanks. False positives should be suspected for any sample value less than 5 times the blank concentration (10 times for common laboratory contaminants). High spike recoveries combined with blank contamination increase the likelihood of false positives.
- The probability of a false positive for an analyte is directly related to the amount of bias if chemicals of potential concern are detected in blanks and spike recoveries for the analyte are biased high.

- The probability of false positives is highest when the reported concentration is near the detection limit for an analyte.
- The probability of false positives is low if chemicals of potential concern have not been detected in any blanks and spike recoveries are not biased high.

6.1.2 Are Site Concentrations Sufficiently Different from Background?

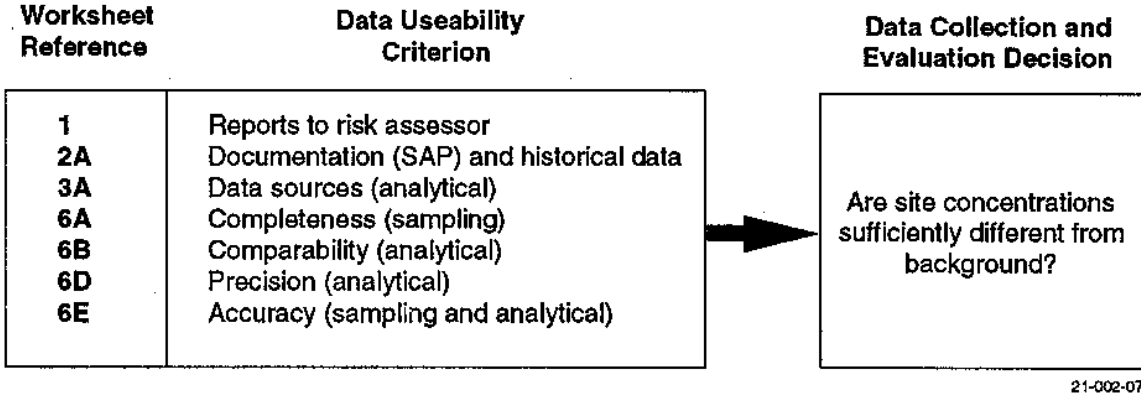
Background samples provide baseline measurements to determine the degree of contamination. Background samples are collected and analyzed for each medium of concern in the same manner as other site samples. They require the same degree of quality control and data review. Background samples differ from other samples in that the sampling points, as defined in the sampling and analysis plan (SAP), are intended to be in an area that has not been exposed to the source of contamination. Historical data, when available, are particularly useful in selecting sampling and analysis techniques used to determine the representative concentrations of chemicals of potential concern in background samples. Historical data can help to delineate physical areas that are background and provide a basis for temporal trends in the concentration of chemicals of potential concern. Exhibit 72 lists the criteria from the Data Useability Worksheet that affect this decision.

As part of the risk assessment process, the risk assessor must determine if background samples are uncontaminated. The entire data collection process will be simplified if chemicals of potential concern are not found in background samples. If chemicals of potential concern are found in the background samples, the risk assessor must determine whether they are at naturally occurring levels, of anthropogenic origin, due to contamination during the sampling process, or are site contaminants.

Both naturally occurring chemicals and anthropogenic chemicals have significance for risk assessment. Naturally occurring chemicals are those expected at a site in the absence of human influence. Metals are naturally occurring chemicals that are often included in risk analysis; they are often present in environmental media in varying concentrations. For example, soils of high organic content, such as humus, would have a low concentration of metals by weight, while soils with a high clay content would contain higher metal levels. Anthropogenic chemicals are defined in RAGS (EPA 1989a) as chemicals that are present in the environment due to man-made, non-site sources (e.g., industry, automobiles). Chemicals of anthropogenic origin may include organic compounds such as phthalates (plasticizers), DDT, or polycyclic aromatic hydrocarbons and inorganic chemicals such as lead (from automobile exhaust). Guidance highlights for background concentration issues for risk assessment are:

- Organic chemicals of potential concern found in background samples should not be considered naturally occurring. They may be present because they are either site contaminants or are of anthropogenic origin. They also could be a result of contamination during sampling.
- The risk assessor may eliminate chemicals from risk assessment calculations if their concentrations fall within naturally ocurring levels and are below the concentration of concern.
- Contamination of background samples is indicated if chemical concentrations are higher than naturally occurring levels. Such contamination may come

EXHIBIT 72. DATA USEABILITY CRITERIA AFFECTING BACKGROUND LEVEL COMPARISON



from anthropogenic sources or from problems in sampling or analysis activities. The risk assessor may include analytical data with other site data or perform a separate risk assessment based on best professional judgment.

- Anthropogenic chemicals should not be eliminated from the risk assessment.
- Statistical analysis may be necessary to determine if site levels are distinctly different from those found in background samples when background results approach site concentration levels.
- Statistical analysis may be necessary where chemicals of potential concern are detected in site samples at very low concentrations. It is difficult to distinguish a difference between background and site sample concentrations at levels close to the detection limit.

☛ *Statistical analysis may determine if site concentrations are significantly above background concentrations when the differences are not obvious.*

6.1.3 Are All Exposure Pathways and Areas Identified and Examined?

The identification and examination of exposure pathways is discussed in detail in RAGS. Exhibit 73 summarizes the criteria that the risk assessor must assess to determine the probable level of certainty that all exposure pathways and areas have been identified and examined.

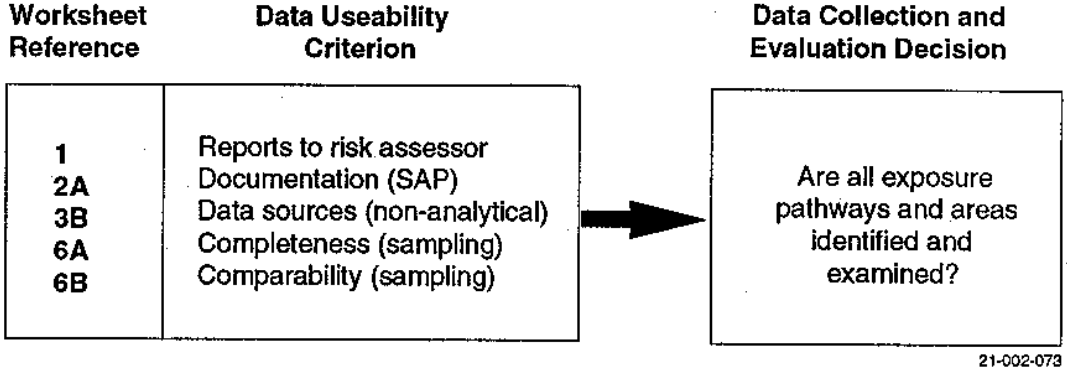
The nature of the exposure pathways and areas to be examined is critical to the selection of a sampling design and analytical methods. If the pathways and areas are not identified properly, the resulting characterization may be inappropriate. The risk assessor should determine which pathways and areas are not adequately assessed and determine the effect on the risk assessment if they are excluded from study. Guidance highlights for exposure pathway identification for risk assessment are:

- Recommend acquisition of additional samples from the inadequately represented exposure pathway or area if feasible. (Sampling considerations presented in Chapter 3 should be re-examined).
- Investigate whether computer simulation modeling is feasible if additional samples cannot be collected from an inadequately represented pathway or area. For example, air flow models could be used to estimate transport of volatile contaminants if the contamination of soil and water at a site is fully characterized but no air samples were obtained.
- Note in the report that the risk could not be determined for a pathway or area, or use simple chemical/physical relationships to estimate exposure if additional samples cannot be collected from an inadequately represented pathway and no simulation models are appropriate. For example, equilibrium partition coefficients can be used to estimate movement in the vadose zone of soil if insufficient data exist to calibrate a groundwater transport model.

6.1.4 Are All Exposure Areas Fully Characterized?

Assessing how well exposure areas have been characterized involves evaluation of completeness, comparability, and representativeness across analytical and sampling data quality indicators. Exhibit 74 lists the criteria from the worksheet that affect this decision. To be fully characterized, the exposure area must have

EXHIBIT 73. DATA USEABILITY CRITERIA AFFECTING EXPOSURE PATHWAY AND EXPOSURE AREA EXAMINATION



been appropriately sampled. Broad spectrum analyses must also have been conducted for the media of concern and analyte-specific methods used where appropriate. The uncertainty in data collection and analysis depends on the evaluation of completeness, comparability and representativeness as discussed in Section 5.6. Based on these indicators, the risk assessor should determine the magnitude of the effect of data confidence on the risk assessment. Guidance highlights for characterization of exposure areas for risk assessment are:

- Use the data but note the level of confidence associated with assessment of the affected exposure area if it is not significant.
- Statistical interpretation procedures (e.g., sensitivity analysis) may be used if the confidence level associated with data for an exposure area is significant but does not warrant resampling and reanalysis.
- If the uncertainty associated with the data is high, the risk assessor may determine that an exposure pathway or area is not fully characterized.

6.2 ASSESSMENT OF UNCERTAINTY ASSOCIATED WITH THE BASE-LINE RISK ASSESSMENT FOR HUMAN HEALTH

The level of certainty in making each of the four decisions discussed in Section 6.1 contributes to the overall uncertainty in data collection and analysis components of risk assessment. The critical factor in assessing the effect of uncertainty on the environmental analytical data component of risk assessment is not that uncertainty exists, but rather that the risk assessor is able to qualify and/or quantitate the uncertainty so that the decision-maker can make informed decisions. The certainty levels for risk assessment, represented in Exhibit 75, are based on the ability to quantitate the uncertainty in analytical data collection and evaluation. However, data collection and evaluation is only one source of uncertainty in the risk assessment. Other components of the risk assessment process, such as toxicity of chemicals and exposure assumptions, influence the four decisions to be made and contribute significantly to the uncertainty of the baseline risk assessment.

The most quantitative level of risk assessment occurs when the uncertainty in data can be determined quantitatively. The next level occurs when the uncertainty can be determined qualitatively, or the impact of the uncertainty is assessed using sensitivity analysis. The least desirable situation occurs when the uncertainty in data is unknown. This situation can occur if the minimum requirements given in Chapter 5 for the data useability criteria have not been achieved.

☛ *The primary planning objective is that uncertainty levels are acceptable, known and quantitatable, not that uncertainty be eliminated.*

EXHIBIT 74. DATA USEABILITY CRITERIA AFFECTING EXPOSURE AREA CHARACTERIZATION

Worksheet Reference	Data Useability Criterion
1	Reports to risk assessor
2A	Documentation (SAP)
2B	Documentation (SOPs)
2C	Documentation (field records)
3A	Data sources (analytical)
3B	Data sources (non-analytical)
6A	Completeness (sampling and analytical)
6B	Comparability (sampling and analytical)
6C	Representativeness (sampling and analytical)
6D	Precision (sampling)

→ **Data Collection and Evaluation Decision:** Are all exposure areas fully characterized?

21-002-074

EXHIBIT 75. UNCERTAINTY IN DATA COLLECTION AND EVALUATION DECISIONS AFFECTS THE CERTAINTY OF THE RISK ASSESSMENT