Hazard Ranking System Issue Analysis: Use of Significance in Determining Observed Release

MITRE

Hazard Ranking System Issue Analysis: Use of Significance in Determining Observed Release

Richard D. Brown

July 1986

MTR-86W101

SPONSOR:
U.S. Environmental Protection Agency
CONTRACT NO.:
EPA-68-01-7054

The MITRE Corporation
Metrek Division
7525 Colshire Drive
McLean, Virginia 22102-3481

Department Approval: Linda J Duncan

MITRE Project Approval: K. W. Garrett

ABSTRACT

This report deals with the question of what criterion should be used to establish that the concentration of a released uncontrolled hazardous substance is "significantly higher than the background." The criterion would be used in determining an observed release under the Hazard Ranking System (HRS), an Appendix to the National Contingency Plan (NCP) for oil and hazardous susbstances created by the Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA). The report examines options for HRS revision including definitions of significance based on the limits of detection and quantitation to discriminate between contaminant levels at or in the vicinity of an uncontrolled hazardous substance facility and background levels. The report also examines the relevance of EPA responses to public comments relating to significance issues, general concepts and tests of significance, detection limits, and the nature and amount of data available for determining an observed release.

ACKNOWLEDGMENT

The author wishes to recognize colleagues whose ideas were incorporated within this report. Jerry Fitzgerald contributed insight on the degree of variation of the limit of detection achieved in analytical laboratories. Channing Johnson developed the concept of utilizing the limit of quantitation as a basis for determining significance when evaluating concentrations near the limit of detection. Thomas Wolfinger identified several factors to be considered when evaluating air data.

TABLE OF CONTENTS

TABLE OF CONTENTS (Concluded)

LIST OF TABLES

1.0 INTRODUCTION

1.1 Background

The Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA) (PL 96-510) requires the President to identify national priorities for remedial action among releases or threatened releases of hazardous substances. These releases are to be identified based on criteria promulgated in the National Contingency Plan (NCP). On July 16, 1982, EPA promulgated the Hazard Ranking System (HRS) as Appendix A to the NCP (40 CFR 300; 47 FR 31180). The HRS comprises the criteria required under CERCLA and is used by EPA to estimate the relative potential hazard posed by releases or threatened releases of hazardous substances.

The HRS is a means for applying uniform technical judgment regarding the potential hazards presented by a release relative to other releases. The HRS is used in identifying releases as national priorities for further investigation and possible remedial action by assigning numerical values (according to prescribed guidelines) to factors that characterize the potential of any given release to cause harm. The values are manipulated mathematically to yield a single score that is designed to indicate the potential hazard posed by each release relative to all other releases. This score is one of the criteria used by EPA in determining whether the release should be placed on the National Priorities List (NPL).

During the original NCP rulemaking process and the subsequent application of the HRS to specific releases, a number of technical issues have been raised regarding the HRS. These issues concern the desire for modifications to the HRS to further improve its capability to estimate the relative potential hazard of releases. The issues include:

- Review of other existing ranking systems suitable for ranking hazardous waste sites for the NPL.
- Feasibility of considering ground water flow direction and distance, as well as defining "aquifer of concern," in determining potentially affected targets.
- Development of a human food chain exposure evaluation methodology.
- Development of a potential for air release factor category in the HRS air pathway.
- Review of the adequacy of the target distance specified in the air pathway.
- Feasibility of considering the accumulation of hazardous substances in indoor environments.
- Feasibility of developing factors to account for environmental attenuation of hazardous substances in ground and surface water.
- Feasibility of developing a more discriminating toxicity factor.
- Refinement of the definition of "significance" as it relates to observed releases.
- Suitability of the current HRS default value for an unknown waste quantity.
- Feasibility of determining and using hazardous substance concentration data.

- Feasibility of evaluating waste quantity on a hazardous constituent basis.
- Review of the adequacy of the target distance specified in the surface water pathway.
- Development of a sensitive environment evaluation methodology.
- Feasibility of revising the containment factors to increase discrimination among facilities.
- Review of the potential for future changes in laboratory detection limits to affect the types of sites considered for the NPL.

Each technical issue is the subject of one or more separate but related reports. These reports, although providing background, analysis, conclusions and recommendations regarding the technical issue, will not directly affect the HRS. Rather, these reports will be used by an EPA working group that will assess and integrate the results and prepare recommendations to EPA management regarding future changes to the HRS. Any changes will then be proposed in Federal notice and comment rulemaking as formal changes to the NCP. The following section describes the specific issue that is the subject of this report.

1.2 Purpose of Report

As a part of the current HRS process, a score is assigned to a known ("observed") release for each pathway (ground water, surface water, or air) for the migration of a hazardous substance away from a facility. The acceptable evidence indicating an observed release primarily is data that show levels of a contaminant,

attributable to the uncontrolled hazardous substance facility, to be significantly higher than the background level. Guidance is not provided in the HRS with respect to the meaning of "significantly" other than to state that it is "in terms of demonstrating that a release has occurred, not in terms of potential effects." The purpose of this report is to examine the issue of "significantly higher than the background" with respect to an observed release and to suggest options for a more definitive method for determining an observed release.

Excluded from discussion in this report are other observed release issues raised in public comment on the proposed HRS and the proposed NPL and its revisions, such as the relationship of the observed level of contamination to either environmental standards, the permitted release of pollutants, the degree of perceived health threat associated with a release, the validity of sampling and analytical protocols, or the presence of substances commonly found in the environment.

1.3 Organization of Report

The rest of this report begins with a review (Chapter 2.0) of the current use of significance in the HRS. Chapter 3.0 is a summary of EPA responses to relevant public comments on the issue of significance. These comments were raised with regard to HRS promulgation and NPL proposals and promulgations. A detailed review of the EPA responses is provided in Appendix A. These two chapters

examine the degree of consistency in, and the pattern of public comments on, EPA's position with respect to significance issues.

Chapter 4.0 examines general concepts of significance, tests of significance, the role of detection limits in determining significance, and the importance of interpreting analytical data for use in the HRS.

Chapter 5.0 builds upon the preceding chapters to suggest options for clarifications and/or refinements to the HRS.

2.0 CURRENT USE OF SIGNIFICANCE IN THE HRS

2.1 Introduction

The Hazard Ranking System (HRS) is Appendix A of the National Oil and Hazardous Substances Contingency Plan (NCP) promulgated on July 16, 1982 (47 FR 31219-31243). The HRS is designed to estimate the potential hazard presented by releases or threatened releases of hazardous substances, pollutants, and contaminants. The HRS is applied to data from an observed or potential release to enable EPA to calculate a "score" representing the relative hazard from such a release.

The calculation of the HRS score for a release involves analyses of five "pathways" of exposure of humans or sensitive environments: 1) ground water, 2) surface water, 3) air, 4) direct contact, and 5) fire and explosion. A composite migration score is developed from scores for each of the first three pathways. This migration score is used to determine the eligibility of a site for placement on the National Priorities List (NPL). The last two pathways are used to identify emergency situations that require removal actions and are not considered in the placement of sites on the NPL (47 FR 31187, July 16, 1982).

2.2 Observed Release

Each of the three pathways used in determining an HRS migration score incorporates a provision for determining the existence of a release. Provisions for determining an observed release for each

pathway are described in the HRS Users Manual (contained in Appendix A of the NCP).

For the ground water pathway, the determination of an observed release is based on "direct evidence of release of a substance of concern from a facility to ground water." The "direct evidence of a release must be analytical." That is, "if a contaminant is measured (regardless of frequency) in ground water or in a well in the vicinity of the facility at a significantly (in terms of demonstrating that a release has occurred, not in terms of potential effects) higher level than the background level, then quantitative evidence exists, and a release has been observed." Certain "qualitative evidence of a release (e.g., an oily or otherwise objectionable taste or smell in well water) constitutes direct evidence only if it can be confirmed that it results from a release at the facility in question" (47 FR 31224, July 16, 1982).

For the surface water pathway, "direct evidence of release must be quantitative evidence that the facility is releasing contaminants into surface water." The "quantitative evidence could be the measurement of levels of contaminants from a facility in surface water, either at the facility or downstream from it, that represents a significant (in terms of demonstrating that a release has occurred, not in terms of potential effects) increase over background levels." The direct evidence is not dependent on the frequency of measurement (47 FR 31233, July 16, 1982).

For the air pathway, "the only acceptable evidence of release is data that show levels of a contaminant, at or in the vicinity of the facility, that significantly exceed background levels (regardless of the frequency of occurrence)." "Data based on transitory conditions due to facility disturbance by investigative personnel are not acceptable" (47 FR 31236, July 16, 1982).

3.0 ISSUES RAISED RELEVANT TO SIGNIFICANCE

This chapter is a summary of EPA responses to significance issues raised in public comments pertaining to HRS promulgation and NPL proposals and promulgations. A detailed review of the EPA responses is provided in Appendix A.

3.1 Interpretation of Data

Many commenters raised the issue that, in some instances, the levels reported for an observed release did not constitute a threat to human health, particularly when subsequent sampling detected no contaminants.

The EPA position is that the HRS assigns a value for an observed release when there is evidence that substances have migrated from a site, indicating that more may do so in the future, and not because the actual release is a health threat. An observed release is scored whenever contaminants are detected beyond their place of deposition in concentrations "significantly higher" than background levels.

3.2 Amount of Data

Comments relevant to the amount of data required to document a HRS score, including documentation of an observed release, are conflicting. Some commenters felt data requirements were overly extensive, others felt that more data should be required.

The EPA response consistently has been that the current data requirements strike a balance between providing enough data for

decision-making and the cost and time required to collect data. It is important to note that placement on the NPL is only the first step in considering a site for remedial response under CERCLA. A more detailed investigation, subsequent to NPL listing, could indicate the hazard to be more or less of a threat to human health or the environment. Should subsequent data collection and analysis indicate that remedial response is not needed, the site could be removed from the NPL. It is important to quickly list on the NPL those sites which may pose serious threats to human health or the environment. The time required for acquisition of extensive data prior to listing may inhibit remedial response actions which need to be conducted in a timely manner.

The question of the amount of data required for scoring an observed release is important, if statistical tests of significance are utilized to document an observed release. Generally, a large number of samples must be collected to provide a high degree of confidence in statistical tests. However, when the HRS was originally proposed, EPA stated that it was designed so that a site could be scored without having to spend a lot of resources (e.g., time and money). It was intended as a low-cost screening tool to determine which sites the Agency intended to further investigate (47 FR 31186-31187, July 16, 1982). Consequently, large amounts of data are rarely available for use in scoring an observed release.

Sometimes only two values are available, one representing background and one representing the site.

3.3 Definition of Significance

Currently, data that show levels of a site-related contaminant to be significantly higher than the background level constitutes the most commonly reported acceptable evidence of an observed release. Guidance is not provided in the HRS with respect to the meaning of "significantly" other than to state that it is "in terms of demonstrating that a release has occurred, not in terms of potential effects." Commenters have not questioned the lack of guidance, but have questioned the validity of data used to score observed releases in instances where the amount of contaminant detected is near the limit of detection.

As a general response to comments, EPA stated that, if a contaminant is observed in background and site samples and sample concentrations are within 10 to 20 percent, the Agency cannot state conclusively that an observed release has occurred. Sometimes the presence of a pattern of clustering of low and high values among background and site samples, respectively, can demonstrate observed release.

The Agency also has stated that any detectable concentration of a substance of concern above background levels can constitute evidence of an observed release. That is, any variation above background can be considered an observed release if, in the judgment

of the sampling and analysis authority (ultimately EPA), the variation indicates a release from the site. However, based on the wording of the HRS, the inferred meaning of "above background" could be the concept of "significantly above background" based on "quantitative evidence." This inference still leaves open a question as to the meaning of quantitation and at what point one concentration is considered to be "significantly higher" than another concentration.

The Agency has attempted to clarify the HRS wording by stating that in cases in which a specific contaminant is not detected in background, any measurable quantity of the contaminant in the site samples is considered to be "significantly higher" than the background and provides the basis for scoring an observed release. This clarification still leaves open a question as to the meaning of a "measurable" quantity.

4.0 CONSIDERATIONS RELATING TO USE OF SIGNIFICANCE IN THE HRS

This chapter examines various considerations relevant to revision of the method of scoring an observed release under the HRS. It begins (Section 4.1) with a discussion of general concepts of significance, followed by a review (Section 4.2) of the current use and relevance of tests of significance in hazardous waste management. Section 4.3 discusses the currently accepted definitions of the limit of detection and limit of quantitation and their relationship to concepts of significance. Section 4.4 examines the data available for scoring an observed release, their relationship to the concepts of limit of detection and limit of quantitation, and their appropriateness for scoring an observed release.

4.1 General Concepts of Significance

Use of the term "significance" generally is understood to infer meaning. If something is significant, it is considered important and has meaning relevant to a normal state of affairs, events, or facts. Often, determination of significance is judgmental, subjective, and is considered an expression of common sense.

In the field of statistics, significance takes on a defined and inflexible definition within a highly organized context of the nature of assumptions, the type of mathematical distribution, and the particular number and disparity of values. When an observation is very unlikely to have arisen by chance alone, it is considered "statistically significant". That is, it would be highly improbable

that the observation occurred by chance and the difference between the observation and normal occurrences must be accepted as a real difference.

Statistical significance is based on probability. That is, falling short of absolute certainty, it is expressed as a degree of confidence in the reality of the difference between sets of values. A difference that arises on the basis of pure chance only one time in one thousand sampling efforts is considered highly significant. The level of such probability is expressed as $p=0.001$; i.e., there is only one chance in a thousand of the result arising by chance alone.

A decision to apply a statistical test of significance to data used to determine an observed release must be evaluated within the HRS context. In the abstract world of statistics, the only valid foundation (in practice) for making such a determination is on the basis of a statistical test of significance. However, in the evaluation of environmental data, practical considerations must also be weighed, many of which cannot be represented precisely in mathematical terms. Common sense must be applied when considering factors which cannot be adequately accounted for statistically when few data are available. Such factors include the natural variability of the contaminant of concern within the environment, the probable lack of any sources of the contaminant outside of the site, the degree of precision and accuracy associated with the

measurement (including influence of sample collection and processing), and the representativeness of particular kinds of environmental data. Normally, variability associated with laboratory instrumentation and techniques is taken into consideration in the reporting of analytical results (see Section 4.3). The natural variability of a contaminant may be larger than variability attributable to the analytical process.

4.2 Tests of Significance

For purposes of determining when a release is "significant" under the RCRA (The Resource Conservation and Recovery Act) program, EPA requires use of the Cochran's Approximation to the Behrens-Fisher Student's t-test (published as Appendix IV of 40 CFR 264). The Agency currently is considering another method, the averaged replicate t-test, to determine significance under the RCRA program.

A problem associated with incorporating such statistical tests of significance within the HRS, which is different from the RCRA context, is that the data available for scoring are highly variable with respect to amount, source(s), and design of sampling programs. Often, pre-HRS site investigation plans do not require that samples be obtained systematically at specified times and places or contain randomization schemes to ensure unbiased results. In general, sampling programs do not incorporate an experimental design to produce data compatible with the application of a statistical test for significance. Statistical tests of significance are useful only

in instances when a large number of samples have been taken, with adherence to a well defined sampling plan, and the reported concentrations are not near the level of detection. It is not within the scope of current pre-HRS data collection efforts to generate the amount of data needed to effectively apply statistical tests.

4.3 Role of Detection Limits in Determining Significance

The Committee on Environmental Improvement of the American Chemical Society (ACS) has defined a detection limit (DL; also termed the limit of detection) as the lowest concentration level that can be determined to be statistically different from that observed in a blank (ACS 1983).

Blanks are used to identify bias in the final results due to contamination. There are two types of blanks. The first, called a laboratory or method blank, is intended to detect bias resulting from inappropriate handling in the analytical process. These blanks are prepared in the laboratory and carried through the same laboratory operations as the samples. The second, or field blank, is designed to detect bias resulting from inappropriate handling during the sampling process. Field blanks are prepared to accompany samples collected in the field and are treated as field samples in all aspects, including exposure to the type of sample bottle, holding for the same time, and treatment with the same preservatives

(EPA 1985a). Note that the ACS definition does not specify what types of blanks are to be used.

The EPA Contract Laboratory Program (CLP), which provides analytical support to the CERCLA program (see 4.4.1), requires analytical laboratories to verify the attainment of certain laboratory detection limits (specifically, instrument detection limits or IDLs) which are determined in part on the use of laboratory blanks. The IDLs must meet specific contract required detection limits established under the Contract Laboratory Program. Field blanks are not used in the determination of IDLs (Peter Isaacson, VIAR, personal communication, July 1986).

Although the ACS definition of a DL has become widely accepted, other definitions have been used by chemists when reporting data. Non-statistical definitions have been used in which the DL is equated to the background (naturally occurring concentration of a substance within the environment), ten percent of the background, or some arbitrarily determined level which does not represent a threat to human health. Various mathematical definitions of the DL range from one to twenty times the standard deviation of net concentrations. The net concentration of a sample is equal to the total analyte value measured for a sample minus the analyte value measured for a blank (Currie 1968).

The question of detection is particularly important to the HRS when an observed release is to be based on a background level which

is below the DL reported by an analytical laboratory. Thus, with respect to the ACS definition of the DL, a question that needs to be answered (when the measured background level is below the DL) is whether a measured sample value is significantly different from that found for the sample blank. The ACS has determined (for normal distributions) that the difference between the value measured for a sample and the value for the blank can be considered to be greater than zero (at a high degree of confidence, i.e., at the 99 percent confidence level) when that difference is greater than three times the standard deviation of such differences observed in laboratory data (net concentrations). Thus, the ACS recommends that the DL value, established by an analytical laboratory, be set at three times the observed standard deviation for a particular series of net concentrations (ACS 1983).

The ACS position with respect to the DL is that a DL value set at less than three standard deviations for such net concentrations lies within the region of questionable detection (and is, therefore, unacceptable). A high degree of confidence that an analyte has been detected exists at and above the three standard deviation value.

The ACS also has defined the level above which quantitative results may be obtained with a specified degree of confidence. This level is termed the limit of quantitation (LQ). Confidence in the apparent analyte concentration increases above the DL and attains a high degree at the LQ. Both the ACS and the National Bureau of

Standards (NBS) agree that the lower limit of concentration (i.e., the lowest amount) of a substance that must be observed before a method is considered to provide quantitative results is the LQ. Both ACS and NBS agree that, by convention, the LQ is equal to ten standard deviations (or 3.3 times the DL). This LQ represents an uncertainty of only plus or minus 30 percent in the measured value at the 99 percent confidence level. The LQ is useful for determining the lower limit of the useful range of measurement methodology (NBS 1985, ACS 1983, EPA 1982a).

The ACS position with respect to the LQ is that a measured value which is at or above the DL, but below ten standard deviations (3.3 times the DL) lies in a region of less-certain quantitation. A value 3.3 times the DL represents the limit of quantitation. A value above the LQ lies within the region of quantitation.

For HRS purposes, sample results at or near the DL (using the ACS definition) are associated with two problems: 1) uncertainty due to measurement variability can approach and even equal the reported sample value, and 2) confirmation of the detection is essentially impossible. Identification of a contaminant is dependent largely on the selectivity of the analytical method and knowledge of the absence of sources of interference. The problems diminish when higher levels of analytes are present. Thus, ACS recommends that "quantitative interpretation, decision making, and

regulatory actions should be limited to data at or above the LQ" (ACS 1983).

The DL and the LQ have been established by convention at the 99 percent confidence level by ACS (1983), NBS (1985), and EPA (1982a,b). For HRS purposes, a lower level of confidence (e.g., 95 percent, resulting in lower values of DL and LQ could be established. The use of 99 instead of 95 percent places importance on achieving a high state of certainty over the possibility of not scoring an observed release on some sites when it would be desirable to score an observed release, i.e., a 95 percent limit accepts "false positives" in order to identify more "true positives."

4.4 Interpretation of Analytical Data

4.4.1 Detection Limits

Most of the data utilized in the HRS scoring procedure are derived through the EPA Contract Laboratory Program (CLP). The Program is directed by the National Program Office, EPA Headquarter's Analytical Support Branch, Office of Emergency and Remedial Response. The DLs utilized by the Program are Contract Required Detection Limits (CRDLs; EPA 1985b, 1984b, 1983b). The CRDLs are minimum DLs required by EPA of analytical laboratories performing analyses of environmental samples under the CLP. The CRDLs are conservative. Many participating laboratories operate at or below the CRDLs on a routine basis. The CLP laboratories are not required to report

their own DLs associated with a particular data set, but are required to relate sample values to the CRDLs (see Section 4.4.2).

In general, the CRDLs are substantially higher than the DLs realized by a particular laboratory. And, since the DLs (and consequently the LQs) depend on the precision (similarity of repetitive measurements) attainable by an individual laboratory, the DLs attained for a particular analytical method can be highly diverse among laboratories. Thus, analyte values reported near the CRDLs can be viewed as being more reliable with respect to quantitation than analyte values observed near the DL associated with a particular laboratory for a given series of analytical measurements. Although difficult to determine, the LQs for the CLP laboratories (and laboratories operating under the CLP guidelines) with the best instrumentation and quality control practices may be near the CRDLs (because the CRDLs are conservative and should be attained readily by CLP laboratories). The CLP is in the process of evaluating CLP data reported to date for the purpose of determining if the CRDLs could be lowered without diminishing the reliability of reported data (Mike Carter, EPA, CLP, personal communication, April 1986).

Some data used for scoring sites are derived outside of the CLP, e.g., data collected and analyzed by a consultant under contract to a site owner. These data are not accompanied by CLP

CRDLs, but DLs of the laboratory performing the chemical analyses usually accompany the data.

4.4.2 Suspect Data

Some data reported under the CLP is considered and identified as suspect (flagged with the letter "J"). This procedure is part of the CLP's quality assurance procedures. The code letter "J" indicates that the associated numeric value is an estimated quantity because quality control criteria were not met. Data reported as suspect under the CLP reflect that the laboratory performance requirements, specified by contract, have not been met or that other factors may have affected the final result, causing the reported data probably to be biased. Examples of factors which may cause data to be labeled as suspect are observed or inferred interferences in laboratory blanks and/or samples, the exceeding of holding times, and the lack of or inappropriate use of procedures for the preparation of laboratory blanks, calibration standards, calibration verification standards, laboratory control standards, and interference check standards. Narrative is to be provided with the labeling to indicate the cause of possible bias.

Under the CLP, if the observed contaminant value is less than the CRDL (but equal to or greater than the measurement capability of the analytical instrument), the CRDL is reported in brackets (instead of the observed contaminant value), e.g. [10]. If an inorganic substance (e.g., arsenic) was analyzed for but not

detected, the CRDL is reported together with a "U" (e.g., 10U). If an organic substance was analyzed for, but not detected, the laboratory's own estimated sample quantitation limit is reported with a "U". Notations such as "10UJ" could indicate that holding times were exceeded, that the possibility of false negatives may exist, and that the laboratory DL for a particular sample may be elevated over the value reported (EPA 1986, 1985c).

Suspect data should be used for HRS scoring with caution. Values reported at concentrations near the CRDL may not be reliable for scoring an observed release, unless the associated narrative indicates that the data can be considered as valid under certain circumstances.

5.0 SUGGESTED REFINEMENTS FOR USE OF SIGNIFICANCE IN THE HRS

Based on the review of EPA responses to public comments and the considerations in Chapter 4.0, the following suggestions for revision of the HRS would be consistent with those responses. The suggestions would augment what has been stated by explaining what is acceptable with respect to the meaning of the HRS concepts of "significantly higher" (e.g., 47 FR 31224, July 16, 1982) and "measurable quantity" (e.g., 49 FR 37078, September 21, 1984).

5.1 Definition of Limit of Detection

Determination of an observed release is dependent on the presence of a measurable quantity of a contaminant in site samples that is significantly higher than background. When the contaminant is not detected in background samples, any measurable quantity of the contaminant in the site samples is presently considered significantly higher than the background and provides the basis for scoring an observed release (49 FR 37078, September 21, 1984). Thus, values of interest in determining an observed release are the DL and the LQ (which is based on the DL).

Since the DL is determined at the laboratory, it is suggested that EPA utilize the DL as reported by the analytical laboratory as the DL used for HRS scoring. Since CLP laboratories usually report only the CRDLs with analytical data, the CRDLs should be used as the DLs for CLP data. Should a CLP laboratory report its DLs for a particular set of data, its DLs should be used in place of the CRDLs.

5.2 Significance When Background Levels Are Below Limit of Detection

This section presents two options for refinement of the HRS scoring of an observed release when background levels are below the limit of detection. The first option, discussed in Section 5.2.1, is based only on the ability to detect contamination in a sample associated with the site. The second option, discussed in Section 5.2.2, is based on the ability to quantitate the level of contamination in a sample associated with the site. These two options are compared in Section 5.2.3.

5.2.1 Option 1. Presence of Contamination

The ACS has defined the DL as the lowest concentration level that can be determined to be statistically different from a blank. That is, the level of contamination at the DL or higher can be differentiated from that associated with the experimental error or "noise" associated with the analytical protocol. For Option 1, the DL is the basis for the determination of an observed release against a background of "none detected" for a particular contaminant.

When background levels of a contaminant are below their DL, the contaminant is considered to be not detected. The actual background concentration of the contaminant could be very near the DL and also very close to the observed site concentration of the contaminant (should the site value be near the DL). Under these conditions, the results of replicate analyses may be statistically indistinguishable, given the unreliability of measurements near the DL. However, the

detection of contamination can be reliable (at the 99 percent confidence level) at and above the DL.

Although one cannot be certain about the actual level of contamination when the concentration of the site-related sample is near the DL, site contamination can be considered to be present (compared to a background of "none detected"). Thus, it can be said that the migration of a hazardous substance has occurred. The determination that migration has occurred is the only requirement under the current HRS concept of an observed release (i.e., migration has been demonstrated).

5.2.2 Option 2. Quantitation of Contamination

The ACS has defined the region of certain quantitation at 3.3 times the DL. For Option 2, this relationship is the basis for the determination of an observed release against a background of "none detected" for a particular contaminant. For practical use in scoring, the HRS could be revised to indicate that any contaminant value of three or more times the DL (in a sample associated with the site) reasonably represents a measurable quantity, given that the contaminant is not detected in the background sample(s). This (3 times the DL reported by the analytical laboratory) should provide a reasonable and adequate margin to account for uncertainty, especially in light of the conservative nature of the CRDLs (which have been specified by the CLP for many of the most frequently observed substances at hazardous waste sites). Under this option,

one could be highly confident that the site related sample concentration represents a level of contamination that is higher than that existing in background (even though the background contaminant level is not detectable).

The above refinement would clarify EPA's current position that any measurable quantity of a contaminant that is significantly higher than background would provide the basis for scoring an observed release. The lower limit of the measurable quantity would be the LQ for a particular set of data. By definition, any value at or above the LQ would be measurable and significantly higher than any value below the DL for a particular set of data.

5.2.3 Comparison of Options

A site is more likely to be scored for an observed release using Option 1 than Option 2. Under Option 1, the concept of significance is based on the ability to reliably detect contamination. Option 1 is consistent with EPA's comments upon promulgation of the first update of the NPL which infer that detection represents a significantly higher concentration than a concentration which is not detected (49 FR 37078, September 21, 1984). The logic behind Option 1 vs. Option 2 is that, if there reasonably appears to be a demonstration of the migration of a hazardous substance on a site (independent of the contaminant concentration), it would be prudent to score for an observed release.

As currently practiced this would result in listing the site on the NPL so that more definitive data can be obtained through detailed site investigations.

Under Option 2, the concept of significance is based on the ability to quantitate the amount of contamination. This Option is consistent with the viewpoint that quantitation is implied in EPA comments relating to the existence of an observed release; e.g., the presence of "quantitative evidence" (47 FR 31233, July 16, 1982), a "measurable quantity" (49 FR 37078, September 21, 1984), a "quantity higher than background" (48 FR 40665, September 8, 1983), and a "significantly higher concentration" than in background (49 FR 37078, September 21, 1984). The logic behind Option 2 vs. Option 1 is that, in order to determine that one value is higher than another, the basis for decision-making must be quantitative; i.e., one must have a high degree of confidence in the analyte concentration. Such confidence does not exist near the limit of detection.

5.3 Significance When Background Levels Are Above Limit of Detection

This section presents two options for refinement of the HRS scoring of an observed release when background levels are above the DL. The first option, discussed in Section 5.3.1, is based only on the ability to quantitate the concentration of the site related sample. The second option, discussed in Section 5.3.2, is based on the need to attain a degree of separation between site and background

levels before an observed release is assigned. These two options are compared in Section 5.3.3.

5.3.1 Option 1. Minimum Difference

This option for assigning an observed release when the background level of the contaminant of concern is at or above the DL requires that the site concentration must be above the background level and at least equal to or above the LQ (3 times the DL). This option maintains a statistically justifiable separation between the two values when the background level is near the DL; i.e., one can be highly confident the site related sample concentration is higher than background. Confidence in the background concentration increases as it increases above the DL. When both background and site related sample concentrations are above the LQ, a high degree of confidence (99 percent level) is associated with each of the sample values. One can be highly certain that the site level is higher than background (but cannot be certain that both values are not reflective of general background contamination).

5.3.2 Option 2. Margin for Certainty

This option is summarized in Table 5-1. The option is based on the premise that when a contaminant is observed in background, as well as in site samples, certainty of an observed release is questioned by a possibility that the site concentration is merely a reflection of widespread contamination in the general area of the

TABLE 5-1

OPTION 2. MINIMUM DIFFERENCE, BETWEEN BACKGROUND AND SITE SAMPLE CONCENTRATION, REQUIRED BEFORE ASSIGNING AN OBSERVED RELEASE WHEN BACKGROUND LEVEL OF CONTAMINANT IS ABOVE THE LIMIT OF DETECTION

When Background Level Is:	Site Level Must Be At Least:
≥ DL and<2DL	lesser of 3 times background or 4 times DL*
≥ 2DL and<LQ	2 times background
≥ LQ	2 times background

*The value of 4 times DL is placed as a limit to avoid the 3 times background value from exceeding the lower value of the next category, when the background value approaches 2DL.

site. The likelihood of this hypothesis being correct increases as background levels increase.

This option distinguishes between background and site sample concentrations used for assigning an observed release, even when both values are within the region of quantitation. This option provides for a margin of certainty to assure that the site value is likely to represent an observed release, even in light of the contaminant being observed in background.

5.3.3 Comparison of Options

A site is more likely to be scored for an observed release under Option 1 in contrast to Option 2. Option 1 simply treats one sample concentration as higher than another, given the higher (site value) is within the region of quantitation. Option 2 provides an additional margin to account for the possibility that the site value is a reflection of some general contamination represented by the background sample.

5.4 Special Considerations

Situations occur when measurements of several contaminants and/or a number of measurements of the same contaminant appear to indicate that site-related concentrations are above background (whether background is above or below the DL), yet data do not meet the criteria set forth in Sections 5.2 or 5.3. Under such situations, one might conclude based on experience that a pattern exists which indicates that an observed release has occurred. Under

these situations, it is suggested that determination of an observed release be based only on the results of an appropriate statistical test, given that enough data are available to apply a comparison test where the results would be associated with a high (e.g., 99 percent) degree of confidence. The analysis should be documented and verified by a trained statistician and/or chemist as appropriate.

A characteristic, not strongly associated with other media, that confounds the interpretation of air data, is that the representativeness of the data depend on highly variable environmental conditions such as temperature, pressure, wind speed, and stability. Thus, a single pair of measurements is highly unlikely to be representative of an observed release.

Special atmospheric conditions can result in a lower or higher detection or emission level at the time of sampling. Special atmospheric conditions include high wind speeds, low temperature, high relative humidity (including precipitation), flat and open terrain, and an unstable atmosphere. The reader should note that detailed information on the effects of atmospheric conditions on airborne emissions from uncontrolled hazardous waste sites is contained in another report under preparation.

To a lesser extent, similar special conditions are associated with surface water, and even less with ground water. Sometimes, owing to limitations in sampling timeframes and the availability of sampling personnel and equipment, sampling during these conditions

cannot be avoided. Special consideration should be given to laboratory data reflecting background or site concentrations generated from samples acquired under such conditions.

5.5 Use of Suspect Data

The use of suspect data should be avoided, if possible. At times, data which do not meet all contract (CLP) requirements are released to facilitate the progress of projects requiring the data. The contract laboratory may be required to confirm the data or to reanalyze samples. Under these circumstances, an attempt should be made to obtain valid data through the CLP.

If it is necessary to use suspect data, the CLP or the non-CLP laboratory should be contacted to determine how the action, which served as the reason for labeling the data as suspect, introduced bias and influenced the level of confidence. Assurance from the CLP or non-CLP laboratory must be documented with respect to the validity of the data, and reliability of the associated DL, for use in scoring an observed release.

6.0 REFERENCES

American Chemical Society (ACS). 1983. Principles of Environmental Analysis. Anal. Chem. 55: 2210-2218.

Bureau of National Affairs, Inc. (BNA). 1986. Statistical Test Called Seriously Flawed. Environment Reporter 1-24-86: 1780-1781.

Currie, Lloyd A. 1968. Limits for Qualitative Detection and Quantitative Determination. Anal. Chem. 40: 586-593.

National Bureau of Standards (NBS). 1985. Principles of Quality Assurance of Chemical Measurements. NBSIR 85-3105 (PB85-177947). U.S. Department of Commerce, Gaithersburg, Maryland.

U.S. Environmental Protection Agency (EPA). 1986. Laboratory Data Validation: Functional Guidelines for Evaluating Inorganics Analysis. Office of Emergency and Remedial Response, U.S. Environmental Protection, Washington, D.C.

U.S. Environmental Protection Agency (EPA). 1985a. Choosing Cost-effective QA/QC Programs for Chemical Analysis. EPA 600/4-85-056. Environmental Monitoring and Support Laboratory, Cincinnati, Ohio.

U.S. Environmental Protection Agency (EPA). 1985b. Statement of Work for Organics Analysis (Revised). Contract Laboratory Program, Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, D.C.

U.S. Environmental Protection Agency (EPA). 1985c. Laboratory Data Validation: Functional Guidelines for Evaluating Organics Analyses. Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, D.C.

U.S. Environmental Protection Agency (EPA). 1984a. Support Document for the Revised National Priorities List - 1984. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, D.C.

U.S. Environmental Protection Agency (EPA). 1984b. Statement of Work for Inorganics Analysis. Contract Laboratory Program, Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, D.C.

U.S. Environmental Protection Agency (EPA). 1983a. Support Document for the National Priorities List. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, D.C.

U.S. Environmental Protection Agency (EPA). 1983b. Statement of Work for Dioxin Analysis. Contract Laboratory Program, Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, D.C.

U.S. Environmental Protection Agency (EPA). 1982a. Design of 301(h) Monitoring Programs for Municipal Wastewater Discharges to Marine Waters. EPA 430/9-82-010 (PB83-153809). Office of Water Program Operations, U.S. Environmental Protection Agency, Washington, D.C.

U.S. Environmental Protection Agency (EPA). 1982b. Methods for Organic Chemical Analysis of Municipal and Industrial Wastewater. EPA 600/4-82-057. Environmental Monitoring and Support Laboratory, Cincinnati, Ohio.

APPENDIX A

DETAILED REVIEW OF EPA RESPONSES TO COMMENTS RELATED TO SIGNIFICANCE

The following is a review of EPA responses on significance issues raised in public comments pertaining to HRS promulgation and proposed and final NPL and NPL revisions.

A.1 Preamble to the HRS

When the National Contingency Plan was promulgated on July 16, 1982 (47 FR 31180-31243), EPA responded to public comments on the proposed plan. Included in the EPA response was discussion of certain comments which have some bearing on the use of significance in determining an observed release. Some of these comments related to the amount of data needed to support an HRS score, others focused on the validity of scoring one-time or minor releases.

A.1.1 Amount of Data Needed to Determine an Observed Release

Many comments questioned the data requirements of the HRS. The frequent criticism was that the HRS fails to accurately distinguish between the degree of hazard presented by different releases; the result being that the HRS might give high scores to releases that otherwise should not be included on the NPL (47 FR 31186). Some commenters suggested that the data required could be very expensive to acquire and slow the remedial action process. Others suggested that more factors should be considered or that existing factors

should be considered at a higher level of detail through more extensive data collection (47 FR 31187).

The EPA response concerning the extent of data required in the HRS was that the role and importance of the HRS and NPL must be kept in perspective. The NPL is merely the first step in considering a release for Superfund-financed remedial response (CERCLA established Superfund to investigate and clean up abandoned hazardous waste disposal sites). After a release is included on the NPL, a subsequent remedial investigation would acquire more extensive information which could indicate the hazard to be more or less significant than originally thought (47 FR 31186-31187).

With respect to conflicting comments on the amount of data needed for the HRS, EPA felt that the amount of information to be collected must be balanced against the cost and time required to obtain the data. Overall, the extent of information required must be consistent with the costs of data collection, the large number of releases which need to be investigated, and the resources available for implementing the NCP. EPA has tried to minimize the information required for the HRS, so that releases which have not been extensively investigated are not eliminated from the NPL. The Agency determined that the current HRS data requirements are adequate without being unduly burdensome or costly (47 FR 31187).

A.1.2 One-time or Minor Releases

Some commenters asserted that the frequency and quantity of releases are not considered in the determination of an observed release. Thus, one-time or minor releases would be treated as equivalent to a frequent or chronic source of release (47 FR 31188).

The EPA response reflects the context within which an observed release is viewed by the Agency. The determination of an observed release indicates that the likelihood of a release is 100 percent. The fact that some substances have been released from a site is a good indication that substances at the site can escape and increases the likelihood of a substantial release. Gathering more extensive data than currently required (e.g., for the purpose of differentiating between a minor occurrence vs. a frequent or substantial problem) would add inordinately to the cost and time needed to score releases. Such extensive monitoring is more appropriately considered during investigations subsequent to the listing of a site on the NPL (47 FR 31188).

A.2 Preambles to the NPL and NPL Updates

A.2.1 NPL

The NPL was proposed on December 30, 1982 (47 FR 58476-58485). The NCP was amended to include the NPL on September 8, 1983 (48 FR 40658-40673).

In the preamble to the final NPL, EPA responded to public comments concerning low level observed releases. The comments

focused on situations where values used for scoring observed releases were "low" (e.g., below regulatory limits specified under the Safe Drinking Water Act).

The EPA response explained that concentrations of substances migrating in the environment tend to show extreme variation through time and space. Given that only periodic sampling is feasible in most instances (to gather data in support of an NPL listing), requiring contaminants to exceed certain levels before assigning an observed release could exclude many sites from the NPL which may be endangering the public. The Agency explained that the HRS scoring instruction for an observed release is based on the fact that the observed release factor is considered for the purpose of estimating the likelihood that substances can migrate from a site. When a release is observed in any quantity, as long as the concentration is above background level, that likelihood is 100 percent (48 FR 40665).

A.2.2 NPL Update 1

On September 8, 1983, EPA proposed the first update to the NPL (48 FR 40674-40682). The NPL was amended to include the update on September 21, 1984 (49 FR 37070-37082).

Some commenters reiterated the concerns raised when the NPL was proposed with respect to assigning values for observed releases when measured concentrations of the substances involved were below regulatory limits. The EPA response on this topic remained unchanged from that given in the preamble for the final NPL (49 FR 37078).

Some commenters questioned the validity of one-time or low level releases when subsequent sampling showed lower concentrations or the absence of any contaminants at the time of sampling. In response, EPA explained that values are assigned based on the data even if subsequent sampling failed to detect the same contaminants. Such an approach recognizes that many releases vary in concentration through time or occur sporadically. Thus, negative results during one or more sampling intervals cannot refute a finding, when based on valid sampling and analyses, that an observed release has occurred (49 FR 37078).

Several commenters took issue with the use of significance with respect to observed releases to ground waters. The commenters questioned the validity of sampling and analytical data used to establish observed releases, particularly in instances where the amount of contaminant detected in a sample is near the detection limit of the appropriate analytical method.

In response, the Agency explained its method of establishing background levels and determining significantly higher concentrations. In cases in which a specific contaminant is not detected in some site samples, the background level of that contaminant is assumed to be some unknown value less than the detection limit. Any measurable quantity of contaminant in the site samples is considered significantly higher than the background and

provides the basis for scoring an observed release. The validity of this assumption was cited as being supported by the statistical analysis used to establish the detection limits for the method used in analyzing the data (49 FR 37078).

In situations in which a specific contaminant is detected in all site samples, an observed release is sometimes more difficult to determine than when the substance is not detected in background samples. Generally, there are insufficient numbers of samples from a site to apply conventional statistical tests for significance. The scorer often must rely on inspection of the data to evaluate whether an observed release has occurred. If the data cluster into a group of high values and a group of lower values, particularly if the high values are attributed to sampling locations that are apparently downgradient of a site, an observed release is confirmed. If the analytical data from only one sampling location are significantly higher than from all other locations, an observed release also has occurred. However, if the contaminant concentrations are similar among background and monitoring wells (e.g., within a 10 to 20 percent range), EPA cannot state conclusively that an observed release has occurred. In addition, low concentrations (e.g., less than 10 parts per billion) of phthalates and other substances commonly found in ground water are carefully examined along with any other evidence that might tend to corroborate or disprove that a release has occurred (49 FR 37078).

A.3 Support Documents for the NPL and NPL Updates

EPA's responses to site-specific public comments on the proposed NPL and NPL updates, including an explanation of any score changes, are contained in support documents (available in EPA dockets in headquarters and regional offices). This section is a review of the comments and the Agency's responses, contained in the support documents, which are relevant to the use of significance in determining an observed release.

A.3.1 NPL

Based on the support document for the NPL (EPA 1983a), three commenters (pp. 3-79, 4-57, and 11-7) felt that additional data showing a decreased level of contamination over time should justify a lower HRS score. Another commenter (p. 3-76) felt that low concentrations of only one contaminant should not be the basis for an observed release. One commenter (p. 2-16) felt that additional data showing an increased level of contamination over time should justify a higher HRS score. The response to these comments was that the quantity of a substance detected is not relevant to scoring for observed release, as long as the concentrations are greater than background levels. Any observed release is evidence of the ability of substances to migrate from a site.

A.3.2 NPL Update 1

The support document for the first NPL revision is EPA 1984a. In a response (p. 3-44) to a comment questioning the rationale for

determining an observed release of airborne PCBs, the word "substantially" was used to characterize the significant difference between background and site-related contaminated levels (e.g., "...the on-site PCB concentration is substantially above the background level..."). Although the background and site-related data were not collected simultaneously, a three month monitoring of background levels was determined sufficient to establish the range of background concentrations.

Two commenters (pp. 5-11 to 5-12 and 6-51 to 6-52) stated that EPA's reliance on a single set of analytical data containing levels near the limit of detection is inappropriate as a basis to support NPL listing. The response stated that an observed release is scored whenever substances of concern are detected in concentrations higher than background levels regardless of frequency. An additional response to one of the comments (pp. 5-11 to 5-12) noted that many of the substances were measured at levels substantially above detection limits. Some substances were measured at elevated levels compared to background concentrations which were below the detection limit.

One commenter (p. 11-1 to 11-2) inferred that the low levels of PCBs observed in ground water should not be considered evidence that the compounds have migrated away from the site. The response was

that an observed release was scored because the PCBs were detected in concentrations higher than background levels (in this case, over tenfold higher).