Office of Inspector General

Audit Report

Information Technology

Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) Data Quality

Report No. 2002 - P - 00016

September 30, 2002

Inspector General Division Conducting the Audit	**Information Technology Audit Division, Washington, DC**
Regions Covered	**EPA-wide**
Program Office Involved	**Office of Solid Waste and Emergency Response** **Office of Enforcement and Compliance Assurance**
Audit Team Members	**Edward Densmore** **Diane Vigue** **Corey Costango** **Cheryl Reid**

Abbreviations

CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act
CERCLIS	Comprehensive Environmental Response, Compensation, and Liability Information System
EPA	U.S. Environmental Protection Agency
GAO	U.S. General Accounting Office
GPRA	Government Performance and Results Act
NFRAP	No Further Remedial Action Planned
NPL	National Priorities List
OERR	Office of Emergency and Remedial Response
OIG	Office of Inspector General
OMB	Office of Management and Budget
OSWER	Office of Solid Waste and Emergency Response
RCRA	Resource Conservation and Recovery Act
SPIM	Superfund/Oil Program Implementation Manual

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

SEP 30 2002

OFFICE OF
THE INSPECTOR GENERAL

MEMORANDUM

SUBJECT: Final Report: Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) Data Quality
Audit No. 2000-0000776
Report No. 2002-P-00016

FROM: Patricia H. Hill, Director
Business Systems (2421T)

TO: Marianne L. Horinko, Assistant Administrator
Office of Solid Waste and Emergency Response (5101)

John Peter Suarez, Assistant Administrator
Office of Enforcement and Compliance Assurance (2201A)

Michael B. Cook, Director
Office of Emergency and Remedial Response (5201G)

Attached is our audit report entitled "Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) Data Quality." The objective of this audit was to determine whether CERCLIS data was accurate and reliable. Specifically, we evaluated the accuracy, completeness, timeliness, and consistency of the data entered into CERCLIS.

This audit report contains findings that describe problems the Office of Inspector General (OIG) has identified and corrective actions the OIG recommends. This audit report represents the opinion of the OIG and the findings contained in the report do not necessarily represent the final Environmental Protection Agency (EPA) position. Final determinations on matters in this audit report will be made by EPA managers in accordance with established audit resolution procedures.

In this particular audit, the OIG did not measure the audited offices' performance against the standards established by the National Contingency Plan. The findings contained in this audit report relate only to programmatic measures, and cannot be relied upon to create any rights, substantive or procedural, enforceable by any party in litigation with the United States.

Action Required

In accordance with EPA Order 2750, you, as the action official, are required to provide us with a written response to the audit report within 90 days of the final audit report date. For corrective actions planned but not completed by the response date, reference to specific milestone dates will assist us in deciding whether to close this report.

Should you or your staff have any questions regarding this report, please contact Edward Densmore, IT Projects Manager, Information Technology Audit Division, at (202) 566-2565.

Attachment

cc: Johnsie Webster, Audit Coordinator
Office of Program Management (5103)
Greg Marion, Audit Coordinator
Administration and Resources Management Support Staff (2201A)
Michael Ryan
Deputy Chief Financial Officer and Agency Follow-up Official (2710A)
Kathy Sedlak-O'Brien
Agency Audit Follow-up Coordinator (2724A)

EXECUTIVE SUMMARY

INTRODUCTION

Congress passed the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) in 1980. This statute established the Environmental Protection Agency's (EPA's) hazardous substance release reporting and cleanup program, known as the "Superfund" program. The Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) is the official repository for all Superfund site data compiled in support of CERCLA. EPA uses CERCLIS data to track Superfund site activities and for annual Superfund reporting to Congress.

OBJECTIVE

The objective of this audit was to determine whether CERCLIS data, for active and archived sites, was accurate and reliable (timely, complete, and consistent). Although CERCLIS data is used to manage the Superfund program, we did not visit any Superfund sites to determine if any of the sampled actions had in fact been performed. Our verification work was limited to reviewing Superfund site document files at EPA's 10 regions and interviewing responsible Agency officials. We did not review the effectiveness of the Agency's Superfund response and enforcement activities, nor did we determine the impact of any data deficiencies on those activities.

In this particular audit, we did not measure the audited offices' performance against the standards established by the National Contingency Plan (NCP). The findings contained in this audit report relate only to programmatic measures, and cannot be relied upon to create any rights, substantive or procedural, enforceable by any party in litigation with the United States.

RESULTS IN BRIEF

Over 40 percent of CERCLIS data on site actions reviewed was inaccurate or not adequately supported. We identified actions with inaccurate dates, as well as actions not supported by appropriate documentation or without the signature of an approving official on the documentation. As a result, CERCLIS users do not have error free data. The data is used to analyze and report on the Superfund program, as well as track internal EPA measurements of progress in assessing the inventory of sites. Further, EPA does not have a complete official record documenting the history of activities at CERCLIS sites. These weaknesses were caused by the lack of an effective quality assurance process.

Status-related data on sites was often inaccurate. Data on the National Priority List (NPL)[1], non-NPL, and archive (i.e., removed) status codes, were incorrect. In addition, we identified the following issues, primarily at non-NPL sites: (1) inconsistent use of NPL and non-NPL status codes, (2) active sites without any actions entered for at least 10 years, and (3) frequent use of a non-descriptive status code. As a result, users of CERCLIS data did not have accurate and complete information regarding the status and activities of Superfund sites, which can adversely impact planning and management. These weaknesses were caused by the lack of adequate internal controls over CERCLIS data quality.

RECOMMENDATIONS

The report includes 11 recommendations to improve controls over CERCLIS data quality. Recommendations include developing and implementing a quality assurance process for CERCLIS data that requires periodically selecting random samples of CERCLIS data elements and then verifying this data to source documents. In addition, we recommended developing and utilizing exception reports to identify sites: (1) that have not had any actions entered into CERCLIS for a reasonable amount of time, and (2) with a non-NPL status code that indicates an action is needed or ongoing when this particular action has already been completed. We also recommended updating the CERCLIS policies and procedures to adequately address the appropriate use of NPL and non-NPL status codes, as well as when a site should be unarchived or archived.

AGENCY COMMENTS AND OIG EVALUATION

We received comments from Assistant Administrators for EPA's Office of Solid Waste and Emergency Response (OSWER) and Office of Enforcement and Compliance Assurance (OECA). OSWER strongly objects to the study design and report conclusions, stating they do not focus on OSWER's data quality hierarchy and the importance it places on NPL sites. OSWER further states the report "may mislead the public as to the quality of NPL data, when in fact, the margin of the purported 'errors' were found in non-NPL CERCLIS actions." OECA is primarily concerned that the report's findings could slow down EPA's Superfund enforcement program. In addition, OECA states the report's language (1) does not emphasize extensive data quality efforts with respect to post-1990 data, (2) does not accurately portray the status of CERCLIS data for the Superfund enforcement program, and (3) "could lead CERCLA defendants to needlessly question the quality of EPA's CERCLIS" data.

OSWER states it is committed to developing a replacement for the CERCLIS system that will design data quality into the front end, rather than having to be developed at the

[1] A national list of hazardous sites with the most serious threats to human health and the environment.

system's end. OSWER reported it is currently re-engineering CERCLIS and plans to reevaluate and institute data quality processes that will meet the cited recommendations.

We believe the audit methodology is both valid and objective, that the report findings accurately summarize the overall quality of key CERCLIS data elements, and implementing our recommendations will reduce the risk of inaccurate and unsupported data. We tested data elements that EPA uses for Congressional reporting, as well as ones managers told us were important for internal management purposes. Agency officials were given opportunities to comment on our sampling methodology at the start of our verification work. They did not express any concerns that would have resulted in a different methodology. We did not focus on OSWER's "data quality hierarchy," because the first time this was mentioned was in OSWER's September 10, 2002 response memorandum. Also, as explained in Chapter 2, our analysis does not substantiate that NPL site action data is treated with any greater importance than data for non-NPL sites.

In our opinion, the recommendations are needed and will be useful in addressing the causes of the reported weaknesses and should, when fully implemented, reduce the risk of inaccurate and unsupported data. The recommendations can and should be implemented prior to the development of a new CERCLIS.

TABLE OF CONTENTS

APPENDICES

CHAPTER 1

INTRODUCTION

Purpose

The objective of this audit was to determine whether the Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) data, active as well as archived, was accurate and reliable. Specifically, we evaluated the accuracy, completeness, timeliness, and consistency of the data entered into CERCLIS.

Background

Congress passed the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) in 1980. This statute established the Environmental Protection Agency's (EPA's) hazardous substance release reporting and cleanup program, known as the "Superfund" program. CERCLA provides broad Federal authority to respond directly to releases or threatened releases of hazardous substances that may endanger public health or the environment. CERCLA requires the Agency to maintain a list of hazardous sites, known as the National Priorities List (NPL), with the most serious threats to human health and the environment. See Appendix 1 for a description of the stages for processing Superfund hazardous waste sites.

CERCLIS is the official repository for all Superfund site data compiled in support of CERCLA. CERCLIS, implemented in 1987, is an integrated system which holds national site assessment, removal, remedial, enforcement, and financial information. It is a relational database system that uses client-server architecture (each computer or process on the network is either a client or server), installed on separate local area networks, at EPA Headquarters, and at all 10 regional Superfund program offices. A relational database management system stores data in related tables and can be viewed in many different ways.

EPA's Office of Emergency and Remedial Response (OERR) is responsible for system availability and continuity of operations. OERR is part of the Office of Solid Waste and Emergency Response (OSWER).

CERCLIS has over 1,900 users, at EPA Headquarters and its 10 regional offices. CERCLIS data is copied nightly from the regional databases to a national database maintained at OERR, and is also updated nightly with financial transaction records extracted from the Integrated Financial Management System.

CERCLIS data is used by EPA to (1) track activities for each site in the Superfund program; (2) support financial statements and report on the Superfund program; (3) maintain an inventory of reported potentially hazardous waste sites with site descriptions, site investigations, and cleanup activities; (4) project dates and the costs of clean-up and enforcement activities; and (5) support Superfund program and project management processes.

Scope and Methodology

We conducted this audit from April 2000 to November 2001 at EPA's Headquarters and its 10 regional offices. We interviewed personnel in OSWER, including OERR; the Office of Enforcement and Compliance Assurance; and EPA regional Superfund offices. We reviewed applicable Federal policies and guidelines. CERCLIS listed data on 44,007 potential hazardous waste sites, comprised of 11,754 active sites and 32,253 archived (i.e, removed) sites. We obtained three random samples from the June 30, 2000 CERCLIS database to review and analyze data. We conducted this audit in accordance with *Government Auditing Standards*, issued by the Comptroller General of the United States.

We did not evaluate the effectiveness of the Agency's Superfund response and enforcement activities, nor did we determine the impact of any data deficiencies on those activities. Although CERCLIS data is used to manage the Superfund program, we did not visit any Superfund sites to determine if any of the sampled actions had in fact been performed. Our verification work was limited to reviewing Superfund site document files at EPA's 10 regions and interviewing responsible Agency officials.

Also, the OIG did not measure the audited offices' performance against the standards established by the National Contingency Plan (NCP). The findings contained in this audit report relate only to programmatic measures,

and cannot be relied upon to create any rights, substantive or procedural, enforceable by any party in litigation with the United States.

Details on our scope and methodology, including applicable policies, our statistical sampling, and prior U.S. General Accounting Office (GAO) and EPA audit coverage, are found in Appendix 2.

CHAPTER 2

SITE ACTION DATA INACCURATE OR UNSUPPORTED

Over 40 percent of the CERCLIS data on site actions was inaccurate or not adequately supported. Looking at CERCLIS data since its inception, we identified actions with inaccurate dates, as well as actions not supported by appropriate documentation or without the signature of an approving official on the documentation. As a result, CERCLIS users do not have error free data to analyze and report on the Superfund program, as well as track internal EPA measurements of progress in assessing the inventory of sites. Further, EPA does not have an accurate and complete official record documenting the history of activities at CERCLIS sites. These weaknesses occurred because OERR did not establish an effective quality assurance process for CERCLIS data.

Criteria

As detailed in Appendix 2, Public Laws, Office of Management and Budget (OMB) Circulars, and EPA Directives all require accurate and reliable data. For example, EPA Directive 2100, *Information Resources Management Policy Manual*, requires that management ensure the quality (accuracy, adequacy, and reliability) of data. EPA Directive 2160, *Records Management Manual*, requires adequate and proper documentation of transactions. OSWER Directive 9200.3-14-1E, *Fiscal Years (FY) 1987-2000 Superfund/Oil Program Implementation Manuals (SPIM)*, define requirements for site actions, such as documentation, dates, and authorizations.

Inaccurate and Unsupported Site Action Data

Many actions entered into CERCLIS for active sites were inaccurate or not adequately supported. Site actions are activities that have taken place at a site, such as preliminary assessments, site inspections, removals, combined remedial investigations/feasibility studies, potentially responsible party searches, records of decision, remedial actions, and cost recovery decisions not to sue.

We obtained a statistical sample of 221 site actions out of a universe of 38,649. For a list of the types of actions in our sample see Appendix 3. We found that 42.9 percent of the actions in our sample (95 of 221) were in error.[2] In particular, we found that:

- 67 stated an inaccurate start and/or completion date,
- 22 were not supported by required documentation,[3] and
- 16 did not contain required approval signatures.

We determined dates to be inaccurate if they were: (1) required but not input into CERCLIS, or (2) not adequately supported. Actions with inaccurate dates included: preliminary assessments, site inspections, expanded site inspections, Hazard Ranking System packages, combined remedial investigations/ feasibility studies, remedial design/remedial action negotiations, and remedial actions. Start and completion dates for actions are used to analyze and track cleanup progress at sites and determine the amount of time it takes to perform actions. Although we used the SPIM requirements in determining the accuracy of the start and completion dates, OERR officials indicated they are primarily concerned with (1) dates required, but not input into CERCLIS, and (2) dates in the wrong fiscal year.

---

[2] Some actions contained more than one error, but we did not count more than one error per action item when calculating error projections for the CERCLIS database. Therefore, while we found a total of 105 errors, only 95 were used for projection purposes.

[3] Agency officials do not think that unsupported actions should constitute an error. They believe this condition should be addressed as a records management issue and not used to project error rates in the CERCLIS database. While we agree that an unsupported action may indicate a records management issue, it may also be that 1) the supporting documentation is filed correctly but the action was entered to the wrong site or 2) the documentation never existed.

For the 38 sampled action items lacking documentation:

- 22 actions did not have any supporting documentation, although Agency officials were given six months to produce requested documents,

- 16 actions had source document that was not signed by an approving official.

These actions included: preliminary assessments, site inspections, expanded site inspections, removals, and remedial design/remedial action negotiations.

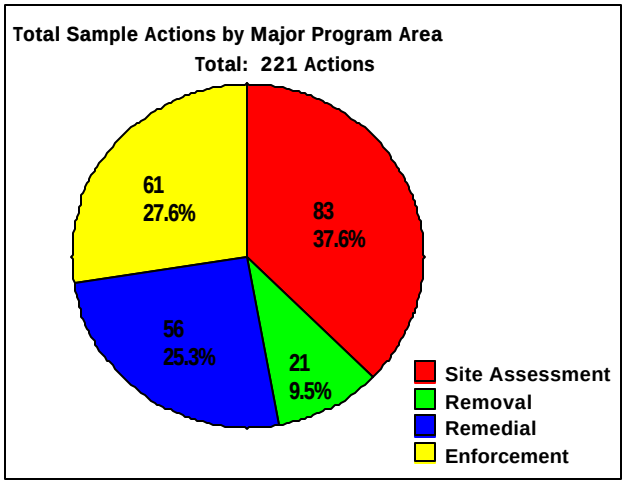
Error Projections for CERCLIS Site Action Data

Based on the 42.9 percent error rate in our sample of actions for active sites, we projected with 95 percent confidence the accuracy of the 38,649 actions entered into CERCLIS, between October 1, 1986 and June 30, 2000. According to our projections, the number of actions in CERCLIS containing errors ranged from 14,064 to 19,239, or a midpoint of 16,615 errors.

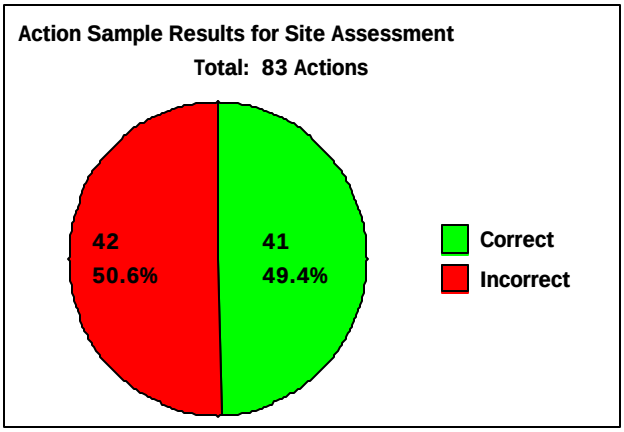
Action Sample Data Presented by Major Program Area

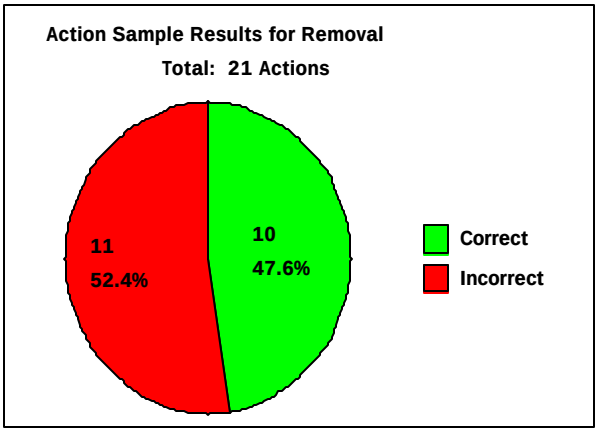
Subsequent to issuance of the draft report, Agency officials indicated they would like to see the action sample presented by four program areas: site assessment, removal, remedial, and enforcement. Because this sample was not stratified by major program areas, no error projections could be made to these individual areas. However, we have depicted the basic distribution of sampled items and resulting data errors by major program area.

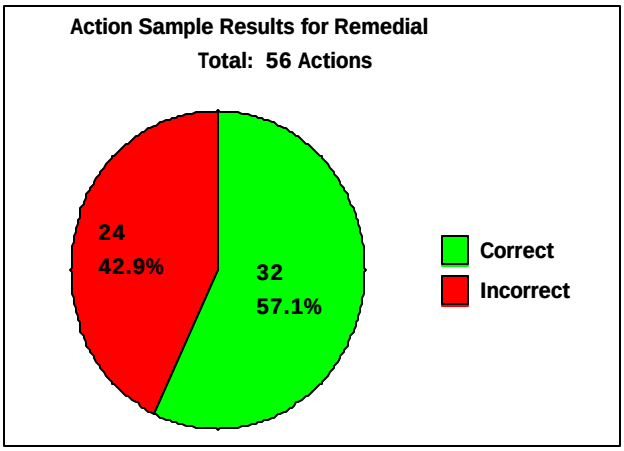
The following chart identifies the distribution of sample action items (i.e., 221 actions) amongst the key program areas.

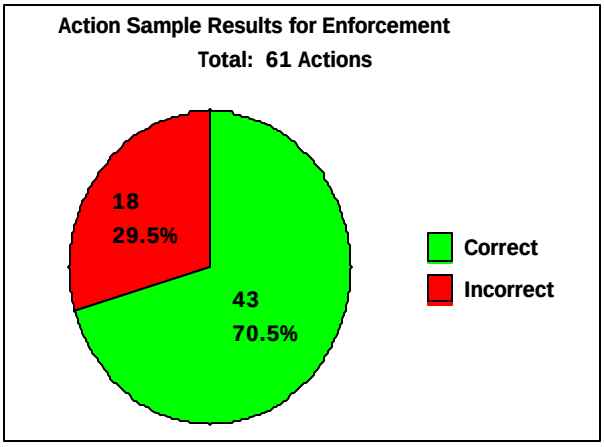


Likewise, the following charts depict the raw results of sampled actions, distributed by major program area.









For more details regarding the types of actions sampled in each major program area and the corresponding results, see Appendix 4.

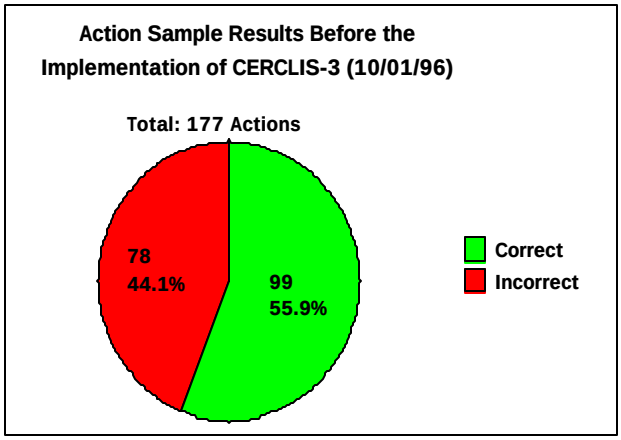
Data Errors Before and After CERCLIS Version 3

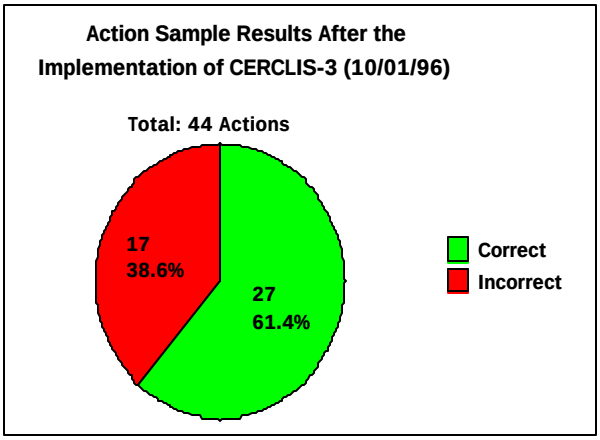
Our examination of sampled action items showed no significant decrease in the percentage of data errors following the implementation of CERCLIS Version 3 (i.e., October 1, 1996). Subsequent to issuance of the draft report, OSWER representatives stated they believed the error rate for the most current data to be significantly lower (i.e., 8 percent or less), and that the audit report was misleading because it did not give credit to OSWER's quality assurance improvements. Our analysis, depicted below, does not support this assertion.

Agency representatives stated the audit failed to account for changes in data systems and business processes from 1987 to 2000. They also emphasized that each new version of CERCLIS represented a new set of business processes and program guidance rules, which differed substantially from earlier versions of the system. OSWER stated these new rules and business processes were needed to keep the system in step with the changes in the program, as well as to improve data reliability. Representatives also noted that management made a conscious decision not to change the data from the predecessor systems nor to retroactively correct any information to fit newer rules or processes.

Although the SPIM changed on a frequent basis, we measured each sample item to the SPIM requirements applicable at the time of the action. Because the sample was not stratified based on the implementation of CERCLIS Version 3, we could not project the rate of error within the entire CERCLIS system. However, using raw data, we can depict the basic distribution of data errors between these two main strata - that is, before and after the implementation of CERCLIS Version 3. Of the 221 sampled action items, 177 predated the implementation of CERCLIS Version 3, and 44 occurred after the start of fiscal 1997.

The following charts show that the percentage of data errors did not substantially decrease after the implementation.



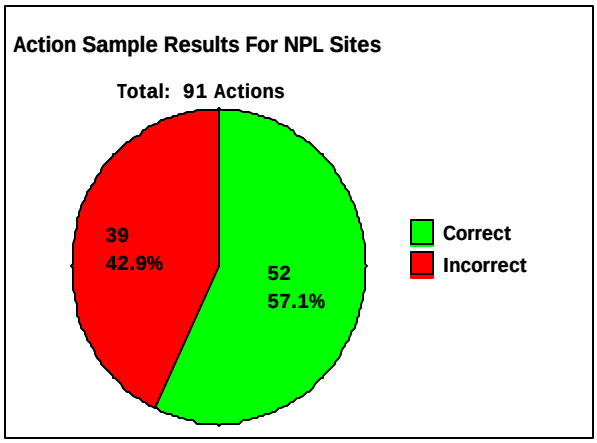


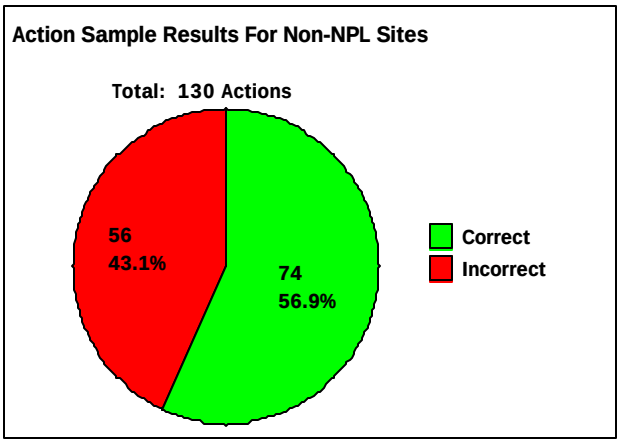
Data Errors For NPL Versus Non-NPL Sites

In response to the draft report, OSWER referenced a 'data quality hierarchy' and noted the importance it places on NPL sites. OSWER states that the report findings "may mislead the public as to the quality of NPL data, when in fact, the margin of the purported 'errors' were found in non-NPL CERCLIS actions."

Our analysis does not support the assertions that: (1) NPL site action data is treated with any greater importance than data for non-NPL sites, or (2) the margin of purported "errors" were found in non-NPL CERCLIS actions. Because the sample was not stratified based on whether actions were for NPL sites versus non-NPL sites, no error projections could be made to the universe of NPL and non-NPL sites. However, using raw data, we can depict the basic distribution of sample action data errors between NPL and non-NPL sites. Of the 221 sampled action items, 91 were related to NPL sites and 130 were related to non-NPL sites.

The following charts show the percentage of data errors are comparable for actions related to NPL sites versus non-NPL sites.





Deficiencies Resulted in Unreliable CERCLIS Site Action Data

As a result, CERCLIS users may not have error free information about site activities. Effective management of the Superfund program requires the availability of accurate information on sites throughout the country. CERCLIS data is used by EPA managers to (1) track and manage activities for sites under the Superfund Program, (2) maintain an inventory of reported potentially hazardous waste sites, (3) project dates for site cleanup and enforcement activities, (4) set funding and workload priorities, and (5) assess progress in achieving Superfund accomplishment goals.

An important function of Superfund managers, at Headquarters, is to report on the national progress of the Superfund program. These managers rely on CERCLIS data to report on accomplishments. Many CERCLIS reports are generated based on the site action data and/or start and completion dates for these actions. Therefore, inaccurate and unsupported data in these fields negatively impacts key management reports. For example:

- The Superfund Accomplishments Report is used by EPA managers to track accomplishment actions, including NPL and non-NPL removals, remedial investigations/feasibility studies, remedial actions and remedial design/remedial action negotiations.

- The Government Performance and Results Act (GPRA) Report is used by managers to track GPRA performance goals and measures, such as the number of removal actions.

- The Superfund Historical Performance Reports provide graphical presentations of progress made at sites, and are used to present an overall picture of Superfund program activities.

Superfund managers also use CERCLIS site action data to perform trend and duration analyses of events. For example, the Superfund Accomplishments Report tracks the time span from the final NPL listing to the remedial investigation/feasibility study, and the time span from the Record of Decision signature to the remedial action start. These durations are calculated based on actual dates. In addition, the Cost Recovery Targeting Report uses the start and completion dates of actions to develop a list of sites that may have potential statutes of limitations expiring for cost recovery at a site. Without accurate site action data, managers cannot rely on CERCLIS to effectively manage the Superfund program.

Without adequate supporting documentation, EPA managers do not have an accurate and complete official record of activities at CERCLIS sites. For example, the documents supporting site inspections provide the soil and water sampling results. The data collected and analyzed as part of a site inspection are used to generate a hazardous ranking system score. This score is used to determine whether a site should be considered for listing on the NPL. Also, the remedial investigation/feasibility study documents important historical information on a site, such as the extent of contamination, identification of preliminary remedial alternatives, and recommendations of a cost effective remedy. These documents, which are part of the administrative record, may be needed for remediation activities, as well as protecting the rights of those involved in the transfer or sale of the land or property involved. According to EPA's Approved Superfund "Administrative Records" Schedule, any judicial reviews concerning the

adequacy of site response actions are limited to the administrative record. Without adequate supporting documentation, EPA managers lack assurance that the data in CERCLIS is accurate.

Ineffective Quality Assurance Process Led to Inadequate Data

The weaknesses noted were caused by OERR's lack of an effective quality assurance process to ensure the accuracy of, and proper support for, CERCLIS data.

EPA Directive 2100 requires management to ensure the quality of its data, and states quality includes such characteristics as accuracy, adequacy, and reliability. CERCLIS data quality requirements are in the SPIM, which states data owners are responsible for complete, current, consistent, and accurate data. In addition, the SPIM assigns data sponsors (the EPA program office or individual responsible for the data element in CERCLIS) with the responsibility for taking an active role in improving the quality of CERCLIS data by periodically conducting focus studies. However, the SPIM does not state what is involved in a focus study or provide guidance on how one should be conducted.

None of the data sponsors we interviewed were aware of the SPIM requirement to conduct focus studies, or what constitutes a focus study. One data sponsor stated that if a focus study entails taking a random sample of data elements and then verifying this data to the source documents in the files, a review like this has not been done since approximately 1992. Although the data sponsors are performing limited quality assurance-related activities, these activities are not detailed or comprehensive enough to provide a high level of assurance that CERCLIS data is accurate, reliable, and adequately supported.

Data sponsors are involved in some data quality activities, such as comparing CERCLIS accomplishment data to Department of Justice reports relating to consent decrees, and regional general counsel reports on settlements. The data sponsors coordinate with regional Superfund officials to resolve any discrepancies they identify. These limited activities may help identify data not entered into

CERCLIS, but probably would not determine whether data in CERCLIS is adequately supported.

Recommendations

We recommend the Director of OERR:

2-1. Develop and implement a quality assurance process for CERCLIS data that includes:

- clearly delineating quality assurance responsibilities, and
- periodically selecting random samples of CERCLIS data elements and then verifying this data to source documents in the site files.

2-2. Review and update (if necessary) the documentation, official date, and signature requirements for CERCLIS site actions in the SPIM so that these requirements are clearly delineated.

Agency Comments and OIG Evaluation

We received comments from Assistant Administrators for OSWER and OECA. OSWER strongly objects to the study design and report conclusions, stating they do not focus on OSWER's data quality hierarchy and the importance it places on NPL sites. OSWER further states the report "may mislead the public as to the quality of NPL data, when in fact, the margin of the purported 'errors' were found in non-NPL CERCLIS actions." OSWER's response notes the audit erroneously concludes that the following conditions constitute inaccurate information: 1) locating paper records in a place other that the primary Regional Office, 2) placing only the month or year in a data field rather than the quarter or fiscal year, and 3) lack of signature on a paper record. OECA is primarily concerned that the report's findings could slow down EPA's Superfund enforcement and lead people to reach inaccurate conclusions about the program. In addition, OECA states that the report's language: (1) does not emphasize extensive data quality efforts with respect to post-1990 data, (2) does not accurately portray the status of CERCLIS data for the Superfund enforcement program, and (3) "could lead CERCLA defendants to needlessly question

the quality of EPA's CERCLIS and other information systems data."

OSWER states it is committed to developing a replacement for the CERCLIS system that will design data quality into the front end, rather than having to be developed at the system's end. OSWER reported it is currently re-engineering CERCLIS and plans to reevaluate and institute data quality processes that will meet the cited recommendations.

We believe the audit methodology was both valid and objective, that the report findings accurately summarize the overall quality of key CERCLIS data elements, and that the proposed recommendations will reduce the risk of inaccurate and unsupported data. We consulted with GAO statisticians throughout the development of our sampling plan and audit methodology, and they concurred with the validity of the methodology used. We tested data elements that EPA uses for Congressional reporting, as well as ones managers told us were important for internal management purposes. Agency officials were given opportunities to comment on our sampling methodology prior to the start of our verification work at EPA's 10 regions. However, they did not express any concerns that would have resulted in a different methodology.

We did not focus on OSWER's "data quality hierarchy," because the first time this was mentioned was in OSWER's September 10, 2002 response memorandum. Also, as explained in this chapter, our analysis does not substantiate that NPL site action data is treated with any greater importance than data for non-NPL sites.

Concerning OSWER's statement that we reached erroneous conclusions, we never state or imply in the report that locating paper records in a place other than the primary regional office constituted an error. However, we do state that if Agency officials did not produce supporting documentation for a sample action within the six months given to them, then we counted it as an error. Also, we do not state that placing only the month or year in a data field

rather than the quarter or fiscal year constitutes inaccurate information. Rather, the report states that dates were determined to be inaccurate if: (1) they were required but not input into CERCLIS, or (2) the dates entered were not in accordance with the SPIM. We chose the SPIM as the standard for judging date accuracy, because Agency representatives questioned the objectiveness of our initial, more flexible test criteria. Finally, the report does not state the lack of a signature on a paper record constitutes inaccurate information. The report merely notes that without adequate documentation: (1) EPA does not have an accurate and complete official record documenting the history of activities at CERCLIS sites, and (2) EPA managers lack assurance that the data in CERCLIS is accurate.

In our opinion, the recommendations are needed and will be useful in addressing the causes of the reported weaknesses and should, when fully implemented, reduce the risk of inaccurate and unsupported data. The recommendations can and should be implemented prior to the development of a new CERCLIS.

CHAPTER 3

SITE STATUS DATA INACCURATE

The status of non-NPL sites in CERCLIS was often inaccurate. All non-NPL sites contain data fields for NPL, non-NPL, and archive status codes, and we discovered these codes were frequently incorrect. We also identified: (1) inconsistent use of the NPL and non-NPL status codes, (2) active sites without any actions entered for at least 10 years, and (3) frequent use of a non-descriptive status code. As a result, users of CERCLIS data do not have complete and error free information regarding the status and activities of many sites, particularly non-NPL sites, which may adversely impact planning and management. These weaknesses were caused by the lack of adequate internal controls over CERCLIS data quality, and an ineffective quality assurance process.

Criteria

As detailed in Appendix 2, Public Laws, OMB Circulars, and EPA Directives all require the accuracy and reliability of data. For example, EPA Directive 2100 requires that management ensure the quality (accuracy, adequacy, and reliability) of data. The SPIM, issued by OSWER, states the data entry requirements and establishes the archiving policy.

Incorrect Site Status Codes

CERCLIS contained inaccurate or incomplete data on the NPL, non-NPL, and archive status codes. Specifically:

- Approximately 40 percent of the active sites (123 of 309) in our aging[4] sample had incorrect NPL or non-NPL status codes. This involved 7 NPL and 116 non-NPL status codes.

[4] The "aging" sample consisted of sites which were maintained as active sites in the CERCLIS database, but for which no action had been entered into the system for at least 10 years.

- **S** Approximately 26 percent of the sites (86 of 333) in our sample of archived sites (i.e., archiving sample) should have been unarchived prior to entering more than 190 actions. An archived site is to be changed to an active site if its condition changes or new information becomes available. Also, we identified 20 percent of the sites (61 of 309) in our aging sample as candidates for archiving.

For a description of the NPL, non-NPL, and archive status codes, see Appendix 5. These status codes are the basis for the reporting logic in CERCLIS-generated reports. The following sections provide details on our sample results, in addition to results of other tests on CERCLIS data using computer-assisted audit techniques.

Incorrect NPL Status Codes

We identified seven sites in our aging sample with incorrect NPL status codes. These sites had an NPL status code of "N" (Not on the NPL) although all of these sites were being addressed as part of an NPL site. According to the SPIM, these sites should have had an NPL status code of "A" (Part of an NPL site).

Due to the problems noted, we performed an additional test against the universe of all active sites (i.e., sites not archived) in CERCLIS. We did this by identifying sites with the following codes: an NPL status code of "N," a non-NPL status code of "AX" (Addressed as part of an NPL site), and an action qualifier of "A" (Addressed as part of an existing NPL site). An action qualifier is a code identifying the priority level or recommendation for further action at a site. This test resulted in the identification of 30 additional sites with an incorrect NPL status code. Based on the non-NPL status code and action qualifier, all of the sites should have had an NPL status code of "A."

Incorrect Non-NPL Status Codes

We identified 116 sites from the aging sample with incorrect non-NPL status codes. For example:

- **!** Sites had a status code of "SS" (Site inspection start needed) when in fact a site inspection was completed.

- Sites had non-NPL status fields that were not filled in but should have been.

- Sites deferred to Resource Conservation and Recovery Act (RCRA) should have had a non-NPL status code of "DR" (Deferred to RCRA), but instead had such codes as "SS," "PS" (Preliminary assessment start needed), and "SX" (Status not specified).

Based on the non-NPL status code inaccuracies found in the aging sample, we performed additional tests against the universe of active sites in CERCLIS. These tests included reviewing different site coding combinations for accuracy. We selected 16 non-NPL status codes that indicate the start of an action is needed or is ongoing (see Appendix 4). We identified 200 sites that indicated an illogical/incorrect relationship between the non-NPL status code and the actions entered into CERCLIS. For example, 44 sites with a non-NPL status code of "PS" (Preliminary assessment start needed) had an action entered in CERCLIS for a completed preliminary assessment. Also, 43 sites had a non-NPL status code of "SS" (Site inspection start needed) but site inspections had been completed according to CERCLIS data.

We also analyzed the relationship between non-NPL and NPL status codes in our aging sample. We identified 25 sites without non-NPL status codes even though they should have had them. Specifically, 17 of the sites had an NPL status code of "A" (Part of another NPL site) and 8 had an NPL status code of "N" (Not on the NPL). According to the SPIM, any site with an NPL code indicating the site is not on the NPL is required to have a non-NPL status code. We performed additional tests to identify the number of active CERCLIS sites that met these two conditions. We found another 388 sites with an NPL status code of "A" or "N" that did not have a non-NPL code and should have had one.

Incorrect Archive Status Codes

We found archiving weaknesses in both our archiving and aging samples. We identified sites listed as archived although ongoing Superfund activity was taking place. We

also found active sites that were not being archived timely when no further interest existed at the site under the Superfund program.

We found that 86 of the 333 sites in our archiving sample should have been unarchived prior to entering site actions into CERCLIS. Actions recorded after the archive date for these 86 sites included site assessment, removal, enforcement, cost recovery, and oversight activities, even though the SPIM indicates these actions should not be taking place after a site is archived. According to the SPIM, an archived site should be unarchived if conditions change or new information indicates that Superfund involvement is warranted.

Although the remaining 247 sites from the archiving sample had actions recorded in CERCLIS after the archive date, these sites were part of a February 1995 EPA archiving initiative. This initiative was an automated effort that turned on the archive indicator and retroactively generated an archive date. This date coincided with the completion date of the last site assessment action entered in CERCLIS with an action qualifier of "N" (No further remedial action planned). However, the system logic did not take into consideration non-site assessment actions (such as remedial and enforcement) entered after the last completed site assessment action with an "N" action qualifier. Therefore, it appears that actions were entered after the site was archived.

We showed an OERR official the types of actions (e.g., site assessment, removal, enforcement, etc.) identified after the archive date and he stated if a site had a removal action that was started and/or completed, the system logic in the archiving initiative would not have archived those sites. He explained that if a site had a removal action after the archive date, the action had been entered in CERCLIS after the site was archived. This official further stated there should not be any archived sites with removal actions started but not completed. We performed additional tests against the universe of all archived sites and identified 58 archived sites

with a removal, removal assessment, or removal negotiation action started but not completed. OERR should review these sites and determine whether they were archived appropriately.

Based on the results of our archiving sample, we conducted another test against the universe of archived sites. We excluded the majority of sites that were part of EPA's 1995 archiving initiative, unless they had actions entered after February 1995 or the sites were archived after that date. We identified 567 additional sites that should have been unarchived prior to entering site actions into CERCLIS based on the SPIM. The actions recorded after the archive date for these sites included site assessment, removal, enforcement, cost recovery, and oversight activities. We asked OERR if there are any actions that can be entered to a site after it is archived without having to unarchive it. Based on our question, a list of actions was compiled and provided to us. However, this list of actions is not contained in CERCLIS policies or guidance stating they can be entered to archived sites without unarchiving the sites.

In addition to sites not being unarchived even though there was ongoing Superfund activity, we found 61 sites in our aging sample that were candidates for archiving. According to the National Oil and Hazardous Substances Pollution Contingency Plan and the SPIM, regional Superfund offices are required to regularly identify and archive sites where no further Superfund program interest exists. Specifically, the SPIM states a site is an archive candidate if any of the following conditions exist:

- The site is deleted from the final NPL.
- The site is removed from the proposed NPL or withdrawn from the final NPL.
- The site has only completed the site assessment process and has either been given a No Further Remedial Action Planned (NFRAP) or Deferred Decision.

- The site has completed both the removal and site assessment process, or has completed the removal process and does not require site assessment work. Additionally, these sites must have completed all related enforcement, cost recovery, and oversight activities.

We identified the 61 sites in our aging sample as archive candidates by analyzing the NPL and non-NPL codes used. We reviewed and discussed these site files with regional Superfund staff, who concurred with our conclusions. The archive candidates from our aging sample included sites that were: (1) deleted from the NPL, (2) listed as having no further remedial action planned, (3) deferred to RCRA, and (4) having waste removed and cost recovery completed, and needing no further remedial assistance.

Using the SPIM archiving criteria, we performed additional tests on the universe of active sites in CERCLIS. Results indicated that 2,351 additional sites met the SPIM archiving conditions, as shown in the following chart:

NPL/Non-NPL Status	CERCLIS Coding	Number of Active Sites
Deleted from the Final NPL	NPL - "D"	206
Removed from the Proposed NPL	NPL - "R"	45
Withdrawn from the Final NPL	NPL - "W"	6
No Further Remedial Action Planned	non-NPL - "NF"	1,203
Deferred to RCRA Program	non-NPL - "DR"	244
Deferred to Nuclear Regulatory Commission	non-NPL - "DN"	1
Removal Only Site	non-NPL - "RO"	611
State Deferral	non-NPL - "SD"	35
Total		2,351

OERR officials should review sites that meet the SPIM archiving conditions, in order to determine whether these archive candidates should be archived.

As evidenced by the results of our tests, sites are not being unarchived when there is ongoing Superfund activity, or are not being archived timely if no further interest exists under the Superfund program. These archiving deficiencies significantly affect the accuracy of site status information in CERCLIS.

Additional Issues Noted

In addition to the weaknesses already discussed, we noted several other issues of concern. Specifically, we found that NPL and non-NPL status codes were not consistently used at some non-NPL sites in CERCLIS; sites were listed as active even though they had no actions entered for at least 10 years; and a non-descriptive code was used too frequently. Details on these issues follow.

Inconsistent Use of Site Status Codes

Along with the use of incorrect site status codes, we identified inconsistent use of NPL and non-NPL status codes by EPA regions. For example:

- Three regions used an NPL status code "O" (Not a valid site or incident) for sites undergoing pre-CERCLIS screening while five other regions used the NPL code "N" (Not on the NPL) for the same purpose.

- Two regions used an NPL status code of "O" for sites that had U.S. Coast Guard removals, while five other regions used an NPL code of "N." Although the NPL status code of "O" is defined in the CERCLIS data dictionary as "not a valid site or incident," it is used as the NPL status code for almost 900 sites. This code is not adequately described in CERCLIS policies and procedures, and clarification on the code's use is needed.

- Three regions used the non-NPL status code "OS" (Other State lead cleanup activities) for 599 sites,

while 4 other regions used this code only 72 times and 3 did not use it at all. Superfund officials indicated that all regions should have some sites with a non-NPL status code of "OS." Without consistent use of this code, site assessment reports will not accurately reflect which and how many non-NPL sites have other cleanup activities being performed under State programs.

Active Sites Had No Actions Entered for at Least 10 Years

None of the sites in our aging sample had any actions entered in at least 10 years. Many of these sites should be archived or reassessed. We identified source documents indicating site activity in 17 percent of the site files (53 of 309 sampled items). These site activities had not been entered into CERCLIS and included: preliminary assessments, site inspections, expanded site inspections, site reassessments, Hazard Ranking System packages, consent decrees, and records of decisions. CERCLIS does not accurately reflect the current status of sites if actions are not being entered timely.

Frequent Use of a Non-Descriptive Status Code

While analyzing the relationship of the NPL and non-NPL status codes in our aging sample, we identified 13 sites with an NPL status code of "N" (Not on the NPL) and a non-NPL status code of "SX" (Status not specified). Through additional testing, we found another 2,190 sites coded as "SX" from the active sites in CERCLIS. These sites accounted for almost 19 percent of the active sites.

Although "SX" is a valid non-NPL status code, it does not clearly describe the status of a site, and we consider its frequent use to be a matter of concern. When we brought this to the attention of CERCLIS managers, they said that non-NPL status codes were initially generated based on the NPL status code, the actions entered, and the action qualifiers. If the data in the fields did not clearly indicate ongoing work or the next steps to be taken at a site, an non-NPL status code of "SX" was automatically assigned.

In addition to the "SX" code, there are 29 other non-NPL status coding options available, many of which are more descriptive. Superfund managers agreed and stated they

have been reviewing some of these sites and changing the non-NPL status codes. During the course of our audit, the number of sites coded as "SX" was reduced to 1,075. According to a Superfund manager, this number has been further reduced to approximately 600 sites. Any site with a non-NPL status code of "SX" should be reviewed and changed to a more descriptive code in a timely manner.

Deficiencies Resulted in Unreliable CERCLIS Site Status Information

As a result of the weaknesses previously discussed, CERCLIS users do not have error free data, particularly about non-NPL sites regarding site status data. The system logic for CERCLIS reports is heavily dependent on the NPL, non-NPL, and/or archive status information. For example, the NPL and non-NPL Site Summary Reports track the major activities of Superfund sites in CERCLIS. The accuracy of these reports is contingent on NPL, non-NPL, and archive status data being correct, as well as the accuracy and timeliness of the site actions entered into CERCLIS. These reports would be of little value if status codes and actions are not timely entered into CERCLIS to reflect site activities that have taken place.

Additionally, the Active Site Inventory Report tracks the status of all active sites in the CERCLIS database. The integrity of this report is based on the archive status field and the site actions being entered accurately and timely. For example, if sites are not getting archived timely, this report is overstating the number of active sites.

Also, the Superfund Accomplishments Report is used to track and assess the status of NPL and non-NPL sites in meeting performance goals. The integrity of this report is also dependent on the accuracy of the NPL and non-NPL status codes. For example, this report uses the NPL status code "A" to report on sites addressed as part of an existing NPL site. The GPRA Report uses the archive flag to determine the number of sites archived per year, as well as the number of sites archived without corresponding assessment decisions. The archive flag and non-NPL status codes are also used in this report to determine the number of sites with: (1) no further remedial actions planned, and (2) assessment work still underway. EPA managers cannot

rely on the site status information in CERCLIS to accurately support the planning and management of the Superfund program. Furthermore, the public, which has access to CERCLIS data via EPA's web site, is being misinformed on the correct status and site activities of many Superfund sites.

Inadequate Internal Controls Led to Inaccurate Data

The weaknesses occurred because OERR did not establish adequate internal controls over CERCLIS data quality. For example, Superfund managers have not developed a process to review older sites that have not had any actions entered into CERCLIS for a number of years. Management has backlog reports that identify the next step for a site in the Superfund process, but these reports do not identify the number of years a site has gone without having any actions entered into CERCLIS. Management needs to determine what is a reasonable amount of time for active sites to sit idle, and then develop exception reports to alert managers as appropriate.

In addition, CERCLIS policies and procedures do not adequately address the appropriate use of NPL and non-NPL status codes. As discussed earlier in this chapter, more specific guidance is needed on such codes as the NPL status code "O" (Not a valid site or incident) and the non-NPL code "SX" (Status not specified). By not clearly defining a code for users in terms of how and when the code is to be used, the code is subject to interpretation, which results in it being inconsistently applied.

Also, the archiving policy needs to be clarified to clearly convey to users when a site should be unarchived or archived. The SPIM contains the archiving policy and states an archived site is to be returned to CERCLIS as an active site if its condition changes or if new information becomes available that indicates additional Superfund involvement is warranted. However, many users were unclear of what actions would warrant unarchiving a site. Additionally, the SPIM states a site should be archived when there are no further site assessment, remedial, removal, enforcement, cost recovery, or oversight activities being planned or conducted at the site. However, there is

no national process to ensure eligible candidates are archived in a timely fashion.

Further, as discussed in Chapter 2, quality assurance-related activities were not sufficiently detailed or comprehensive to provide a high level of assurance that CERCLIS data, for active as well as archived sites, was accurate, reliable, and adequately supported.

Recommendations

We recommend the Director of OERR:

3-1. Develop and utilize exception reports to identify sites:

- that have not had any actions entered into CERCLIS for a reasonable amount of time, and
- with a non-NPL status code that indicates an action is needed or ongoing when this particular action has already been completed.

3-2. Update CERCLIS policies and procedures to adequately address the appropriate use of NPL and non-NPL status codes. Specifically, CERCLIS policies and procedures need to:

- clearly define the use of the NPL status codes "A" (Site is part of an NPL site) and "O" (Not a valid site or incident), as well as the non-NPL status code "OS" (Other Cleanup Activity: State-Lead Cleanup), and
- require managers to review sites with a non-NPL status code of "SX" (Status not specified) in a timely manner and to change this code to one that is more descriptive.

3-3. Continue to review the sites with a non-NPL status code of "SX," and change this code to a more descriptive one.

3-4. Review archived sites with removal actions started but not completed, and determine whether they were archived appropriately. If it is determined that any of these sites are active, then unarchive them.

3-5. Review and archive (if appropriate) any sites meeting the current archiving criteria.

3-6. Update the archiving policy to clarify when a site should be unarchived or archived. Specifically, the policy needs to state what actions, if any, can be made to archived sites without having to unarchive them and what actions would warrant unarchiving a site.

3-7. Implement system edit checks to prohibit entry of unacceptable actions at archived sites.

3-8. Develop and implement a process for having a representative from the various Superfund program areas (e.g., site assessment, removal, enforcement, etc.) concur with the decision to archive a site.

3-9. Emphasize to users the importance of the NPL, non-NPL, and archive status codes and their accuracy, as well as the need to enter all actions defined in the SPIM into CERCLIS in a timely manner.

Agency Comments and OIG Evaluation

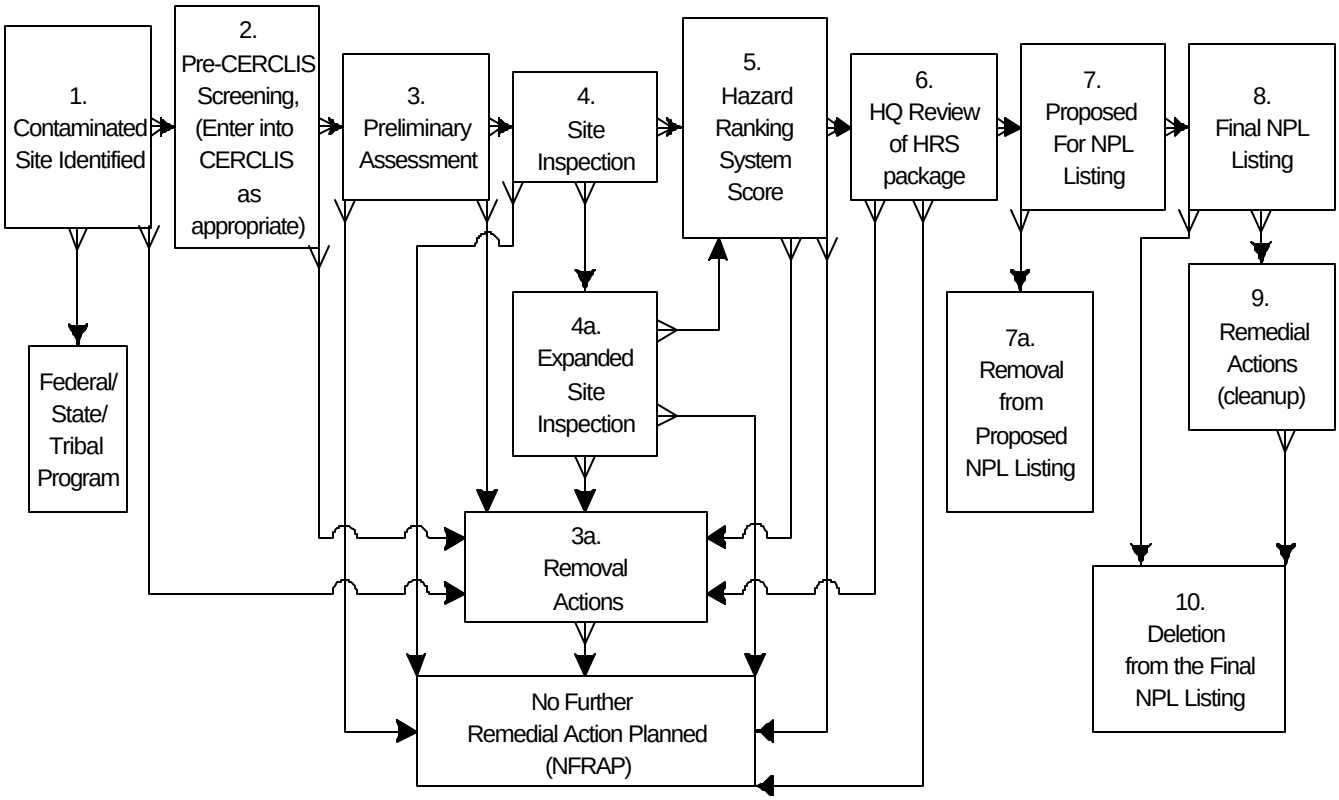
Officials from OSWER and OECA did not specifically respond to the weaknesses and recommendations in Chapter 3. However, both offices offered general comments which have been summarized in the Executive Summary and Chapter 2.

APPENDIX 1

DESCRIPTION OF THE STAGES FOR PROCESSING SUPERFUND HAZARDOUS WASTE SITES

The first step in the Superfund process is the identification of contaminated sites that may pose risks to the public or the environment. Sites may be identified to EPA by State officials, private citizens, or referrals from other regulatory programs. Once identified, EPA has the authority to take action at these sites under its Removal and/or Remedial Program. Data is entered into CERCLIS in order to maintain a listing of the various site actions and to monitor the status of site cleanup efforts.

The following diagram depicts the stages for processing Superfund hazardous waste sites. For a description of each stage, see the corresponding number following the diagram.



1. Sites may be identified by state officials, private citizens, or referrals from other regulatory programs. Not all sites identified are entered into CERCLIS; some are cleaned up under other federal/state/tribal programs.

2. The pre-CERCLIS screening process begins for sites referred to EPA's Superfund program. This process is used to determine whether further site evaluation steps are required under CERCLA, and to minimize the number of sites unnecessarily entered into the CERCLIS inventory.

3. Once a site is entered into CERCLIS, the site generally undergoes a Preliminary Assessment. This involves reviewing existing reports and documentation about the site and analysis of geological and hydrological data, as well as identifying populations and sensitive environments likely to be affected. At this stage, goals are to establish whether a removal action is necessary and determine whether the site poses potential risks to public health. If a Preliminary Assessment determines the site does not present potential risk, the site is typically eliminated from further consideration by designating it as No Further Remedial Action Planned (NFRAP).

    3a. Removal Actions may be performed at any point in the assessment/cleanup process. A removal action is a short-term response intended to stabilize or clean up an incident or site that poses a threat to public health or welfare. Removal actions generally last no longer than 12 months and cost no more than $2 million.

4. For those sites requiring further investigation, a Site Inspection is generally conducted. Although the Preliminary Assessment is typically an off-site review, the Site Inspection involves a hands-on inspection in which soil and/or water samples are collected to better characterize site contamination.

    4a. If additional data are required, an Expanded Site Inspection will be conducted, which can include complex background sampling and the installation of monitoring wells. If the Site Inspection or Expanded Site Inspection determines that releases from the site pose no threat to human health and the environment, the site is designated as NFRAP.

5. A Hazard Ranking System score is developed based on the data collected and analyzed as part of the preliminary assessment and site inspection. This System is EPA's screening tool to determine whether a site should be considered for listing on the NPL. Application of the Hazard Ranking System scoring is the primary mechanism by which EPA places sites on the NPL. It takes into account the likelihood that a site has released, or has the potential to release, hazardous substances into the environment; the toxicity and quantity of the hazardous substances at the site; and the proximity of people and sensitive

environments to the release. The Hazard Ranking System scores up to four pathways of potential human exposure to contamination (i.e., groundwater, surface water, soil, and air) and then combines the individual pathway scores into an aggregate site score. Sites with a score below 28.5 are listed as NFRAP. If the preliminary score equals or exceeds 28.5, the site may be considered for inclusion on the NPL and goes through additional review and screening before it is actually proposed for listing on the NPL.

6. EPA's regional staff submit the Hazard Ranking System package to EPA headquarters for review. After evaluating the data, EPA headquarters, in collaboration with the region and the state in which the site is located, may propose the site for listing on the NPL. Based on the results of the completed Hazard Ranking System package, EPA may also determine that the site should be designated as NFRAP.

7. EPA publishes the list of sites in the Federal Register as a proposed rulemaking, at which time the site listing decisions are generally subject to a 60-day public comment period.

   7a. The Federal Register publishes the proposal to remove the site from the proposed NPL listing, and subsequently publishes the final notice of the removal.

8. At the end of the 60-day period, after public comments have been considered and negotiations with other interested parties have been conducted, EPA compiles a final rule, which is published in the Federal Register. It is only at this point that a site is considered a "final" NPL site.

9. Remedial actions are "long-term" cleanups. Long-term remedial actions permanently and significantly reduce the dangers associated with actual or potential releases of hazardous substances that are serious but not immediately life threatening. Remedial responses can be conducted only at sites on the NPL list.

10. With state concurrence, EPA may delete sites from the NPL when it determines that no further response is appropriate under CERCLA. The Federal Register publishes both the Intent to Delete and Notice of Deletion of the site from the NPL.

APPENDIX 2

DETAILS ON SCOPE AND METHODOLOGY

The focus of this audit was to assess whether CERCLIS data was accurate and reliable. We reviewed system documentation for CERCLIS, such as the file structure, record layouts, user manual, and data element dictionary. We used automated tools, such as the Interactive Data Extraction and Analysis for Windows 3.0 software and the Auditors' Statistical Sampling Estimation Tool, to gather, extract, and analyze CERCLIS data.

We conducted the audit from April 2000 to November 2001 at EPA Headquarters in Washington, DC, and EPA's 10 regional offices. We interviewed personnel in OSWER, including OERR; OECA; and EPA regional Superfund offices. In addition, we performed on-site reviews of Superfund site document files at EPA's 10 regions.

We conducted this audit in accordance with *Government Auditing Standards,* issued by the Comptroller General of the United States. Our audit included tests of management and related internal controls, such as using automated tools to examine the accuracy of the CERCLIS action and site status data. In addition, we reviewed and analyzed policies and procedures specifically related to the audit objectives.

We want to emphasize this was an audit of CERCLIS data quality (i.e., the system), and not specifically of the Agency's implementation of the Superfund program. Although CERCLIS data is used to manage the Superfund program, we did not visit any Superfund sites to determine if any of the sampled actions had in fact been performed. Our verification work was limited to reviewing Superfund site document files at EPA's 10 regions and interviewing responsible Agency officials.

Public Laws, OMB Circulars, and EPA Directives

To accomplish this audit, we reviewed the following documents:

- Public Law 96-510, *Comprehensive Environmental Response, Compensation, and Liability Act of 1980*: This is the law that established the Superfund program, and governs cleanups of both Federal and non-Federal hazardous waste sites.

- Public Law 99-499, *Superfund Amendments and Reauthorization Act of 1986:* This law amended the 1980 CERCLA law and made several important changes. Changes

included revising the Hazard Ranking System to ensure it accurately assessed the relative degree of risk to human health and the environment posed by uncontrolled hazardous waste sites that may be placed on the NPL.

- Public Law 104-13, *Paperwork Reduction Act of 1995:* This law states each agency is responsible for carrying out information resources management activities to improve productivity, efficiency, and effectiveness related to information used within and outside the agency.

- OMB Circular A-123, *Management Accountability and Control:* This circular requires agencies to develop and implement management controls to ensure reliable and timely information is obtained, maintained, reported, and used for decision making.

- OMB Circular A-130, *Management of Federal Information Resources:* This circular requires agencies to protect and ensure the integrity and availability of its information.

- EPA Directive 2100, *Information Resources Management Policy Manual:* This directive requires that management ensure the quality of its data, and notes that quality includes such characteristics as accuracy, adequacy, and reliability.

- EPA Directive 2160, *Records Management Manual:* This directive, which addresses the Agency's records/information management program, requires the Agency to maintain adequate and proper documentation for its transactions.

- EPA Directive 2195, *Information Security Manual:* This directive, which establishes requirements for securing Agency information resources, requires the Agency to ensure that its information systems provide accurate, timely, and credible information.

- EPA OSWER Directive 9200.3-14-1E, *FY's 1987-2000 Superfund/Oil Program Implementation Manuals (SPIM)*: These manuals establish policy for managing the Superfund program, including CERCLIS data entry, and define the requirements for site actions, such as the documentation, dates, and appropriate authorizations needed for actions.

- EPA National Records Management Program, *Approved EPA Records Schedules - Superfund:* This policy establishes retention periods for maintaining official records.

Statistical Sampling Methodology

We obtained three random statistical samples from the June 30, 2000 CERCLIS database to determine whether CERCLIS information was complete, consistent, timely, and adequately

supported by source documents. CERCLIS contained data on 44,007 potential hazardous waste sites (11,754 active and 32,253 archived sites). The active CERCLIS sites had 146,759 actions applied to them. We worked with statisticians from GAO to verify that our sampling methodology was statistically valid based on a 95 percent assurance level with a +/- 5 percent precision for attribute sampling. Attribute sampling is used to estimate the frequency of occurrence of a specific event or item in a universe. We used a commercial software program to determine the sample sizes. The three attribute samples used were as follows:

1. **Action Sample:** Attribute sampling techniques were used to estimate the percentage of error in recording various site actions in CERCLIS (see Appendix 3 for a list of actions in this sample). These actions reflect activities that have taken place at sites in the active CERCLIS inventory. Site action data that differed from, or could not be adequately supported by, source documents were classified as errors. We eliminated planned actions, subactions, and actions not defined in the SPIM from the 146,759 site actions for active CERCLIS sites. Because fiscal year 1987 was the first year of written guidance, we then removed all actions with a completion date prior to fiscal year 1987. For this reason, we also removed all actions with a start date prior to fiscal year 1987 that did not have a completion date. This resulted in a universe of 38,649 site actions. From this, we selected a random sample of 221 site actions. We verified the CERCLIS site action data to the SPIM requirements and to the source documents in the site files.

2. **Aging Sample:** The purpose of this sample was to review active sites that were in the active portion of the database but did not have any actions entered into CERCLIS in at least 10 years. From the 11,754 active sites, we identified a universe of 1,579 sites without an action recorded in CERCLIS in at least 10 years. From this universe, we selected a random sample of 309 sites to determine whether actions had in fact occurred or the site should have been archived.

3. **Archived Sample:** We used the archiving sample to identify sites that were archived but had an action entered into CERCLIS after the archive date. Sites that EPA decides do not warrant further Superfund attention may be assigned archive status. From the 32,253 archived sites in CERCLIS, we identified a universe of 2,503 archived sites with actions entered after the site was archived. We selected a random sample of 333 sites from the universe.

GAO Prior Audit Coverage

Environmental Information - EPA Is Taking Steps to Improve Information Management, but Challenges Remain (GAO/RCED-99-261), issued September 1999: The report noted that EPA has decided to implement several data improvement initiatives. However, the report indicated

that while these were steps in the right direction, they were limited in scope and did not provide the overall strategy needed to ensure the completeness, compatibility, and accuracy of EPA's environmental data. GAO recommended EPA develop an action plan that details the key steps that the Agency needs to take to ensure that EPA's environmental and regulatory data are sufficiently complete, compatible, and accurate.

Hazardous Waste - Unaddressed Risks at Many Potential Superfund Sites (GAO/RCED-99-8), issued November 1998: The report stated CERCLIS inaccurately listed some sites as awaiting an NPL decision although they were not eligible for listing. According to an EPA Superfund program official, the incorrect data entries may have resulted from regional program managers' misinterpretation of EPA's guidance on CERCLIS coding. GAO recommended EPA correct the errors in the CERCLIS database that incorrectly classified sites as awaiting an NPL decision and prevent the recurrence of such errors.

Superfund - Information on the Status of Sites (GAO/RCED-98-241), issued August 1998: The report stated testing was performed on the accuracy of data in EPA's Superfund database of the progress of sites through the cleanup process for a statistically random sample of 98 NPL sites. Based on the sample results, GAO estimated the cleanup status of NPL sites in the database to be 95 percent accurate.

EPA OIG Prior Audit Coverage

Review of the Superfund Annual Report to Congress for Fiscal Year 1998 (Audit Report 2000-P-2), issued October 1999: We reported inconsistencies between CERCLIS data and source documentation for preliminary assessments, site inspections, and removal actions. We suggested OERR make the necessary corrections for greater accuracy.

Superfund Sites Deferred to Resource Conservation and Recovery Act (RCRA) (E1SFF8-11-0006-9100116), March 1999: The report stated that 34 sites recorded in CERCLIS were misclassified. We recommended that CERCLIS be updated to reflect the correct site status.

EPA Had Not Effectively Implemented Its Superfund Quality Assurance Program (E1SKF7-08-0011-8100240), issued September 1998: The report stated EPA managers had not demonstrated commitment to an effective quality assurance program by fully developing and effectively implementing the program to obtain Superfund and other data of known and adequate quality. We recommended OSWER require OERR quality assurance staff to continue performing regional management and technical assessments to ensure that the data quality objectives policy is being adequately implemented in the Superfund program.

Report on CERCLIS Reporting (E1SFF9-15-0023-0100187), issued March 1990: The report stated material errors arose within CERCLIS reports and any information reported by the system

was suspect and should only be employed cautiously. These errors resulted because of the absence of good controls. We recommended the Agency modify the CERCLIS Reports Library to reflect report changes.

APPENDIX 3

LIST OF SITE ACTIONS IN ACTION SAMPLE

Action Code	Action Name
AC	Administrative Order on Consent
AN	Remedial Design/Remedial Action Negotiations
AR	Administrative Records
BB	Potentially Responsible Party Removal
BF	Potentially Responsible Party Remedial Assessment
CA	Consent Agreement (Administrative)
CD	Consent Decree
CO	Combined Remedial Investigation/ Feasibility Study
DD	Cost Recovery Decision Document-No Sue
EE	Engineering Evaluation/Cost Analysis
ES	Expanded Site Inspection
FE	Five Year Remedy Assessment
HR	Hazard Ranking System Package
IN	Inter-Agency Agreement Negotiations
LR	Long Term Response Action
LV	Federal Facility Removal
LW	Federal Facility Remedial Investigation/ Feasibility Study
MA	Management Assistance
ND	Deletion From NPL

Action Code	Action Name
NF	Final Listing on NPL
NG	Negotiation (Generic)
NP	Proposal to NPL
NR	Removed From The Proposed NPL
NS	NPL Responsible Party Search
OF	Operational & Functional
PA	Preliminary Assessment
RA	Remedial Action
RD	Remedial Design
RN	Removal Negotiations
RO	Record of Decision
RP	Non-NPL Potentially Responsible Party Search
RV	Removal
SI	Site Inspection
SS	Expanded Site Inspection/Remedial Investigation
SV	Section 107 Litigation
TA	Technical Assistance
UA	Unilateral Administrative Order

APPENDIX 4

RESULTS OF ACTION SAMPLE

		Total Actions	Correct	Incorrect	Cause of Action Error			Explanation
					No Documents	No Sig.	Data Problem	
Site Assessment								
ES	Expanded Site Inspection	6	2	4	2		2	
HR	HRS Package	3	2	1			1	
ND	Deletion from NPL	2	1	1	1			
NF	Final Listing on NPL	5	5					
NP	Proposal to NPL	3	3					
NR	Removed from the Proposed NPL	1		1	1			
PA	Preliminary Assessment	29	13	16		7	14	5 PA Actions contained more than one error
SI	Site Inspection	33	15	18	3	3	15	3 SI Actions contained more than one error
SS	ESI/RI	1		1			1	
	Subtotals	**83**	**41**	**42**	**7**	**10**	**33**	
Removal								
BB	PRP Removal	6	2	4			4	
EE	Engineering Eval / Cost Analysis	1		1		1		
LV	FF Removal	5	1	4	1		3	
RV	Removal	9	7	2		1	1	
	Subtotals	**21**	**10**	**11**	**1**	**2**	**8**	
Remedial								
AR	Administrative Records	18	11	7	1	3	3	
BF	PRP RA	4		4	1		3	
CO	Combined RI / FS	8	3	5	1		4	
FE	Five Year Remedy Assessment	2	1	1	1			

		Total Actions	Correct	Incorrect	Cause of Action Error			Explanation
					No Documents	No Sig.	Data Problem	
LR	Long Term Response Action	1	1					
LW	FF RI / FS	2	1	1	1			
MA	Management Assistance	1	1					
OF	Operational & Functional	1	1					
RA	Remedial Action	3		3	1	1	2	1 RA Action contained more than one error
RD	Remedial Design	2	1	1	1			
RO	Record of Decision	12	10	2	1		1	
TA	Technical Assistance	2	2					
	Subtotals	**56**	**32**	**24**	**8**	**4**	**13**	
Enforcement								
AC	Admin Order on Consent	7	7					
AN	RD / RA Negotiations	13	3	10	2		8	
CA	Consent Agreement (Admin.)	2	2					
CD	Consent Decree	4	3	1			1	
DD	Cost Rcvry Decsn Doc - No Sue	7	5	2	1		1	
IN	IAG Negotiations	3	2	1	1		1	1 IN Action contained more than one error
NG	Negotiation (Generic)	5	5					
NS	NPL RP Search	4	2	2	2			
RN	Removal Negotiations	2	2					
RP	Non-NPL PRP Search	8	6	2			2	
SV	Section 107 Litigation	2	2					
UA	Unilateral Admin Order	4	4					
	Subtotals	**61**	**43**	**18**	**6**	**0**	**13**	
Sample Totals		**221**	**126**	**95**	**22**	**16**	**67**	10 Actions Contained More Than One Error

APPENDIX 5

DESCRIPTION OF CERCLIS STATUS CODES

NPL Status Code

The NPL status code identifies a site's status with respect to EPA's National Priorities List. Every active and archived site is assigned an NPL status code. The options available for this field are:

A - Site is Part of NPL Site
D - Deleted from the Final NPL
F - Currently on the Final NPL
N - Not on the NPL
O - Not a Valid Site or Incident
P - Proposed for NPL
R - Removed from Proposed NPL
S - Pre-Proposal Site
W - Withdrawn

Non-NPL Status Code

The non-NPL status code indicates the status of a site not on the NPL. EPA conducts site studies and addresses contamination at such non-NPL sites, and the non-NPL status code is used to describe the current status of activity at those sites. There are 30 non-NPL status code options, as follows. Of the 30 codes, 16 indicate an action is needed or is ongoing. Those 16 are marked with an asterisk (*):

AX - Addressed as Part of a National Priorities List Site
* CO - Combined Preliminary Assessment/Site Inspection Ongoing
* CS - Combined Preliminary Assessment/Site Inspection Start Needed
DN - Deferred to Nuclear Regulatory Commission
DR - Deferred to Resource Conservation and Recovery Act
* EO - Expanded Site Inspection Ongoing
* ES - Expanded Site Inspection Start Needed
HN - Hazard Ranking System Package Completed - Further Evaluation Needed
* HO - Hazard Ranking System Ongoing

* HS - Hazard Ranking System Start Needed
* IN - Integrated Removal/Remedial Evaluation Ongoing
* IO - Integrated Expanded Site Inspection/Remedial Investigation Ongoing
* IR - Integrated Removal/Remedial Evaluation Start Needed
* IS - Integrated Expanded Site Inspection/Remedial Investigation Start Needed
NF - No Further Remedial Action Planned
OF - Other Cleanup Activity: Federal Facility-Lead Cleanup
OP - Other Cleanup Activity: Private Party-Lead Cleanup
OS - Other Cleanup Activity: State-Lead Cleanup
OT - Other Cleanup Activity: Tribal-Lead Cleanup
* PO - Preliminary Assessment Ongoing
* PS - Preliminary Assessment Start Needed
RO - Removal Only Site (No Site Assessment Work Needed)
RR - Referred to Removal - No Further Remedial Action Planned
RW - Referred to Removal - Further Assessment Needed
SD - Deferral of National Priorities Listing Decision While States Oversee Response
* SG - Site Inspection Prioritization Ongoing
* SN - Site Inspection Prioritization Start Needed
* SO - Site Inspection Ongoing
* SS - Site Inspection Start Needed
SX - Status Not Specified

Archive Status Code

The archive status code identifies whether a site is archived (i.e., removed) from CERCLIS. Archiving represents a site-wide decision or status indicating that no further interest exists at the site under the Federal Superfund program. It is a comprehensive decision, meaning there are no further site assessment, remedial, removal, enforcement, cost recovery, or oversight activities being planned or conducted at the site. Archived sites are indicated in CERCLIS with the code NFA (No further action).

APPENDIX 6

OSWER COMMENTS TO DRAFT REPORT

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

SEP 10 2002

OFFICE OF
SOLID WASTE AND EMERGENCY
RESPONSE

MEMORANDUM

SUBJECT: OSWER Response to OIG Draft Report "Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) Data Quality" Audit No. 2000-0000776

FROM: Marianne Lamont Horinko
Assistant Administrator

TO: Patricia H. Hill, Director
Business Systems
Office of Inspector General

The Office of Solid Waste and Emergency Response appreciates the opportunity to comment on the subject draft report. See attached comments on the executive summary.

We strongly object to the study design and conclusions, which did not focus on our data quality hierarchy and the importance we place on NPL sites. The audit erroneously concludes that: 1) locating paper records in a place other than the primary Regional Office; 2) placing only the month or year in a data field rather than the quarter or fiscal year; and 3) lack of signature on a paper record constitute inaccurate information. As a result, the audit's recommendations are not helpful to us in terms of program management and may mislead the public as to the quality of NPL data, when in fact, the margin of the purported "errors" were found in non-NPL CERCLIS actions.

While the overall impression from the audit would suggest a flawed information system, we remain confident in the data that we use to manage the program. Unfortunately, the review of a random sample of 1100 data fields and thousands of records provide a somewhat skewed perspective on our ability to understand and manage the Superfund remediation and removal programs. The data on which we depend for national reporting and against which the regions are measured in quarterly reports are sufficiently robust to ensure that sites are appropriately listed on the NPL and that the Hazard Ranking System (HRS) is appropriately applied to evaluations.

As noted in our earlier comments, OSWER has issued successive versions of CERCLIS starting in 1987. Each new version of CERCLIS included new business processes and program guidance rules. These changes were needed to keep in step with changes in the program as well

as to improve data reliability, again focused on our data quality hierarchy. The program made a decision not to change data from predecessor versions nor to retroactively correct any information to fit new rules or processes. We continue to believe that a massive cleanup effort of the aged data at non-NPL sites is not cost effective, nor will it significantly improve our ability to manage the program.

More importantly, we are committed to developing a replacement for the CERCLIS system that will design data quality into the front end, rather than having to be developed at the system's end. We also anticipate streamlining of the data collection, focusing primarily on the data, which will indeed lead to improved program management and evaluation. This replacement system will be designed to reflect management priorities and needs of today, rather than fixing data quality problems of the past. There will be an emphasis on collecting well documented data of known quality and managing the program accordingly. As part of that effort, we will be reviewing the key program data contained in the current system to ensure that data brought forward to the replacement system are data assured before being included. We believe these efforts towards improved data quality planning and implementation are where our limited program resources can most appropriately be focused.

We look forward to continuing our discussion on ways we can work with your office in our ongoing efforts to improve CERCLIS data quality. If you have any questions or need additional information, please contact Mike Cullen at (703) 603-8881 or Johnsie Webster, OSWER Audit Liaison, at (202) 566-1912.

Attachment

EXECUTIVE SUMMARY

INTRODUCTION

Congress passed the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) in 1980. This statute established the Environmental Protection Agency's (EPA's) hazardous substance release reporting and cleanup program, known as the "Superfund" program. The Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) is the ***repository for all Superfund site data compiled in support of CERCLA, including all official as well as unofficial data. EPA uses CERCLIS to track a wide variety of activities including, but not limited to National Priority List (NPL) sites which present hazards to the local community as well as sites which were assessed but found to present no threat to the local community.***

OBJECTIVE

The objective of this audit was to determine whether CERCLIS data, ***including data collected more than two decades ago at sites predominantly not listed on the NPL,*** was accurate and reliable (timely, complete and consistent). We did not review the effectiveness of the Agency's Superfund response activities. Although CERCLIS data is used to manage the Superfund program, we did not visit any Superfund sites to determine if any of the sampled actions had in fact been performed. Our verification work was limited to reviewing Superfund site document files at EPA's 10 regions and interviewing responsible Agency officials.

In this particular audit, the OIG did not measure the audited offices' performance against the standards established by the National Contingency Plan (NCP). The findings contained in this audit report relate only to programmatic measures, and cannot be relied upon to create any rights, substantive or procedural, enforceable by any party in litigation with the United States. Moreover, they are not binding in any enforcement proceeding brought by EPA or the Department of Justice under section 107 of the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) to recover costs incurred not inconsistent with the NCP.

RESULTS IN BRIEF

Over forty percent of the CERCLIS site actions reviewed were

inaccurate or not adequately supported ***(i.e., paper records to support the data were not co-located on-site).*** We identified actions with inaccurate dates, as well as actions not supported by appropriate documentation or without the signature of an approving official on the documentation. As a result, CERCLIS users do not have error free data, particularly those concerning non-NPL site activities. The data is used to analyze and report on the Superfund program, as well as track internal EPA measurements of progress in assessing the inventory of sites. Further, EPA does not have an accurate and complete official record documenting the history of activities at CERCLIS sites. These weaknesses were caused by the lack of an effective quality assurance process.

Also, status-related data on sites was often inaccurate. Data on the National Priorities List (NPL), non-NPL, and archive (i.e., removed) status codes, were incorrect. In addition, we identified the following issues, primarily at non-NPL sites: (1) inconsistent use of NPL and non-NPL status codes, (2) active sites without any actions entered for at least 10 years, and (3) frequent use of a non-descriptive status code. As a result, users of CERCLIS data, including Congress, the public, and EPA management, were being misinformed regarding the status and activities of many ***non-NPL*** sites, which can adversely impact planning and management. These weaknesses were caused by the lack of adequate internal controls over CERCLIS data quality.

It is important to note that the vast majority of errors related to non-NPL sites, and this study was not stratified to assess the relative accuracy of data at NPL sites- - upon which EPA places the highest degree of quality control. EPA's hierarchy for data quality control was not considered as part of this study.

RECOMMENDATIONS

The report includes 11 recommendations to improve controls over CERCLIS data quality. ***We understand that OSWER is currently reengineering CERCLIS and plans to reevaluate and institute data quality processes that will meet the cited recommendations.*** The recommendations include developing and implementing a quality assurance process for CERCLIS data that requires periodically selecting random samples of CERCLIS data elements and then verifying this data to source documents in the site files. In addition, we recommended the Director for Emergency and Remedial Response develop and utilize exception reports to identify sites: (1) that have not had any actions entered into

CERCLIS for a reasonable amount of time, and (2) with a non-NPL status code that indicates an action is needed or ongoing, when this particular action has already been completed. We also recommended updating the CERCLIS policies and procedures to adequately address the appropriate use of NPL and non-NPL status codes, as well as when a site should be unarchived or archived.

APPENDIX 7

OECA COMMENTS TO DRAFT REPORT

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

SEP 19 2002

OFFICE OF
ENFORCEMENT AND
COMPLIANCE ASSURANCE

MEMORANDUM

SUBJECT: Pending Final Report: Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) Data Quality Audit No. 2000-0000776

FROM: John Peter Suarez
Assistant Administrator

TO: Nikki Tinsley
Inspector General

The Office of Enforcement and Compliance Assurance (OECA) has reviewed the pending final report prepared by the Office of Inspector General titled "Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) Data Quality" dated August 16, 2002. I am concerned that the draft, if finalized in its current form, could slow down EPA's Superfund enforcement and could lead people to reach unfair and inaccurate conclusions about the Superfund program.

As a threshold matter, I note that OECA's Office of Site Remediation Enforcement (OSRE) submitted comments on the report to your office in May 2002. Although some changes were made, most comments were not addressed. We still believe those comments were valid and attach them again. We would ask that the draft reports be edited to reflect our comments. Overall, we disagree that in context it is accurate to state in the Executive Summary that "Over forty percent of the CERCLIS site actions reviewed were inaccurate or not adequately supported...As a result, CERCLIS users are misinformed about site activities." As outlined below, OECA believes that statement leaves out key facts and does not accurately portray the status of CERCLIS data for the Superfund enforcement program.

We also disagree with the report's statement that "...Congress, the public and EPA management, were being misinformed regarding the status and activities of many Superfund sites..." EPA routinely reports to Congress and the public on its Superfund accomplishments on an aggregate annual basis. Of the records sampled which were reported to Congress as accomplishments since 1995 only 1 of the discrepancies would have led to a change in our reported accomplishments. This change was limited only to the accomplishment having been reported in the incorrect fiscal year. To suggest that Congress, the public and EPA management were misinformed is a gross exaggeration.

As is reflected in our comments, we believe the report can be improved upon by adding the context necessary for readers to understand the CERCLIS database. While we share some of the same concerns about data quality, we believe some of the report's conclusions are overbroad and are themselves misleading.

Fundamentally, we do not think it is accurate to state that "Congress was misled" by such items as inputting of an incorrect code or the absence of a signature on a document. There is no doubt that the work on sites was done, in most cases by PRPs. This work was properly reported to Congress and the public. Although data improvements are possible, to conclude that the data is misleading as a result of data errors unfairly undermines the Superfund Program and gives a false impression regarding our knowledge of site activity. I would urge you to consider all of our comments again before this report is released.

Audit Methodology

OECA has concerns about the methodology used in conducting the audit and the extent to which it can be used to effect positive change. In its analysis, the OIG took a random sample of 221 records of the more than 30,000 records found in CERCLIS for accuracy of start and completion dates and adequacy of documentation for those dates. This random sampling technique did not focus primarily on that data which is key to the day-to-day management of the Superfund program: the core removal, remedial pipeline, and enforcement information. Eighty-three (37%) of the 221 records sampled were for the pre-remedial program while only 58 (26%) of the records examined were for enforcement at non-Federal Facilities. OECA believes that because the volume of CERCLIS records and types of data varies dramatically by program area (e.g., response, enforcement, site assessment), the OIG should have selected a stratified random sample. This method would have provided a statistically representative sample from each area and conclusions for each type of data based on the results of those samples. Results based on such a stratified sample for each program area would allow OECA and OSWER to focus on the areas with the most significant problems.

Enforcement Data Quality

Of the 221 actions sampled by the OIG, 61 are Federal and non-Federal Facility enforcement actions (Activity Codes: AV, JG, CA, DD, NS, RP, UA, AC, CD, AN, FN, IN, RN, NG, LI, SV, SX, and CL). Of those 61, 18 (29.5%) were identified as problematic: 13 having incorrect dates, and 5 as lacking documentation. But, the report's conclusions wildly overstate the problem. Of the 13 records with incorrect dates, 9 had discrepancies of fewer than 30 days. Only 4 of the 61 records audited had discrepancies greater than 30 days with none greater than 70 days.

CERCLIS Data Documentation

Of the 18 problematic enforcement records, the report identified five records for which documentation could not be located in the Regions. All but one of the records with missing documentation were for activities that occurred in 1994 or earlier.

Using CERCLIS Data to Manage Cost Recovery Statute of Limitations (SOLs)

The report concludes that OSRE's "...Cost Recovery Targeting Report uses the start and completion dates of actions to develop a list of sites that may have potential statute of limitations expiring for cost recovery at a site. Without accurate site action data, managers cannot rely on CERCLIS to effectively manage the Superfund program." This conclusion demonstrates a fundamental misunderstanding of how CERCLIS is used.

In managing potential Statutes of Limitations (SOLs), OECA works with Regions to ensure they target all cases with SOL expiring 6 months into the next fiscal year, e.g., in FY2002 the Regions were asked to target all sites with potential SOLs through March of 2003. In only 2 cases were the dates inaccurate by 180 days or more, and none of these were dates that would likely trigger the statute of limitations.

Ineffective Quality Assurance Process Led to Inadequate Data

We strongly disagree with the conclusion that an ineffective quality assurance process led to inadequate data as it pertains to Superfund enforcement data in CERCLIS. Over the past several years, OECA's data quality efforts have been extensive, particularly with respect to post-1990 data. However, given the large volume of enforcement action and PRP data, OECA focuses its efforts on areas which have been key to its program management and legislative analysis. As such, we have focused on PRP data, as well as administrative and judicial enforcement action data. These efforts include, for example, Regional visits and formal requests to obtain site source documents for all post-1990 enforcement actions and the names and addresses of PRPs associated with those actions. This represents more than 50,000 PRPs associated with more than 3,400 settlements at 1,846 sites. In addition, OECA reviews data on the status of judicial enforcement actions/litigation provided by the Department of Justice (DOJ) on a periodic basis and forwards information to the regions on the status of judicial settlements and litigation in order for them to update CERCLIS. There was only one data error identified in the audit sample in relation to a judicial enforcement action/litigation (Activity Code: CD, SX, SV, CL or LT) and the difference was only one day.

The serious over-generalization in the draft report, if allowed to proceed uncorrected in the final report, could lead CERCLA defendants to needlessly question the quality of EPA's CERCLIS and other information systems data. These defendants will be more likely to ask for original documents to examine during settlement discussions. The burden of assembling and transmitting those documents, and the delays involved, will slow down CERCLA enforcement and the cleanup dependent on that enforcement.

I hope that we can work together to resolve some of these issues and improve the data quality of CERCLIS. Should you have any questions or concerns, please give me a call. Thank you.

Attachment

cc: Marianne Horinko, OSWER
Barry Breen, OSRE
Mike Cook, OERR
Susan Bromm, OSRE
Elaine Davies, OERR
Michael Cullen, OERR
Paul Connor, OSRE
Neilima Senjalia, OSRE
Dela Ng, OERR
Monica Gardner, OSRE
Eric Burman, OSWER
Johnsie Webster, Audit Liaison, OSWER
Greg Marion, Audit Liaison, OECA

APPENDIX 8

REPORT DISTRIBUTION

Headquarters

Assistant Administrator for Solid Waste and Emergency Response
Assistant Administrator for Enforcement and Compliance Assurance
Director for Emergency and Remedial Response
Comptroller
Associate Administrator for Congressional and Intergovernmental Relations
Director, Office of Regional Operations
Agency Followup Official
Agency Audit Followup Coordinator
Audit Liaison, Office of Solid Waste and Emergency Response
Audit Liaison, Office of Enforcement and Compliance Assurance

Office of Inspector General

Inspector General
Assistant Inspector General for Planning, Analysis and Results
Assistant Inspector General for Audit
Assistant Inspector General for Program Evaluation
Media and Congressional Liaison
Director, Business Systems
Divisional Inspector General, Headquarters Audit
Divisional Inspector General, Washington Contracts
Divisional Inspector General, Information Technology Audits