Ref

United States
Environmental Protection
Agency

Office of Environmental
Information
Washington, DC 20460

LIBRARY COPY

EPA/240/R-02/004
November 2002

EPA

# Guidance on Environmental Data Verification and Data Validation

## EPA QA/G-8

DRAFT

9932

## FOREWORD

The U.S. Environmental Protection Agency (EPA) has developed an Agency-wide program of quality assurance for environmental data. Data verification and data validation are important steps in the project life cycle, supporting its ultimate goal of defensible products and decisions. This guidance document, *Guidance on Environmental Data Verification and Data Validation*, provides practical advice to individuals implementing these steps.

EPA works every day to produce quality information products. The information used in these products are based on Agency processes to produce quality data, such as the quality system described in this document. Therefore, implementation of the activities described in this document is consistent with EPA's Information Quality Guidelines and promotes the dissemination of quality technical, scientific, and policy information and decisions.

This document provides guidance to EPA program managers and planning teams. It does not impose legally binding requirements and may not apply to a particular situation based on the circumstances. EPA retains the discretion to adopt approaches on a case-by-case basis that differ from this guidance where appropriate. EPA may periodically revise this guidance without public notice.

This document is one of the *U.S. Environmental Protection Agency Quality System Series* documents. These documents describe the EPA policies and procedures for planning, implementing, and assessing the effectiveness of the Quality System. This document is valid for a period of up to five years from the official date of publication. After five years, this document will be reissued without change, revised, or withdrawn from the *U.S. Environmental Protection Agency Quality System Series* documents. Questions regarding this document or other *Quality System Series* documents should be directed to the Quality Staff at:

U.S. EPA
Quality Staff (2811R)
1200 Pennsylvania Avenue, NW
Washington, DC 20460
Phone: (202) 564-6830
Fax: (202) 565-2441
E-mail: quality@epa.gov

Copies of the *Quality System Series* documents may be obtained from the Quality Staff directly or by downloading them from its Home Page:

www.epa.gov/quality

## TABLE OF CONTENTS


Page

**1. INTRODUCTION ..... 1**
- 1.1 PURPOSE AND OVERVIEW ..... 1
- 1.2 DATA VERIFICATION/VALIDATION IN THE PROJECT LIFE CYCLE ..... 3
- 1.3 INTENDED AUDIENCE ..... 6
- 1.4 PERIOD OF APPLICABILITY ..... 6
- 1.5 ORGANIZATION OF THIS GUIDANCE ..... 6

**2. DATA VERIFICATION ..... 7**
- 2.1 INTRODUCTION TO THE DATA VERIFICATION PROCESS ..... 7
- 2.2 INPUTS TO DATA VERIFICATION ..... 9
- 2.3 IMPLEMENTATION OF DATA VERIFICATION ..... 11
- 2.4 OUTPUTS OF DATA VERIFICATION ..... 13

**3. DATA VALIDATION ..... 15**
- 3.1 INTRODUCTION TO THE DATA VALIDATION PROCESS ..... 15
- 3.2 INPUTS TO DATA VALIDATION ..... 17
  - 3.2.1 Project-Specific Planning Documents ..... 17
  - 3.2.2 Inputs from Field Activities ..... 18
  - 3.2.3 Inputs from the Analytical Laboratory ..... 19
- 3.3 IMPLEMENTATION OF DATA VALIDATION ..... 20
  - 3.3.1 Data Validation of Field Activities ..... 20
  - 3.3.2 Data Validation of Analytical Laboratory Activities ..... 21
  - 3.3.3 Focused Data Validation ..... 22
- 3.4 OUTPUTS OF DATA VALIDATION ..... 22

**4. DATA INTEGRITY ..... 25**
- 4.1 BACKGROUND ..... 25
- 4.2 IMPROPER LABORATORY PRACTICES ..... 26
  - 4.2.1 Examples of Improper Laboratory Practices ..... 26
  - 4.2.2 Warning Signs for Data Validators ..... 28
- 4.3 IMPROPER FIELD PRACTICES ..... 32
- 4.4 ETHICS CULTURE ..... 34

**5. TOOLS AND TECHNIQUES FOR DATA VERIFICATION AND VALIDATION ..... 35**
- 5.1 DATA VERIFICATION TOOLS AND TECHNIQUES ..... 35
  - 5.1.1 Identifying the Project Requirements ..... 35

5.1.2 Verifying Records Against the Method, Procedural, or Contractual Requirements . . . 36
5.2 DATA VALIDATION TOOLS AND TECHNIQUES . . . 52
5.2.1 Tools and Techniques for Data Validation of Field Activities . . . 52
5.2.2 Tools and Techniques for Data Validation of Analytical Laboratory Data . . . 59
5.2.3 Tools and Techniques for Focused Data Validation . . . 66

**6. DATA SUITABILITY . . . 71**
6.1 DETERMINING DATA SUITABILITY . . . 71
6.2 USING PROFESSIONAL JUDGMENT IN DATA VALIDATION . . . 72
6.3 FOCUSED DATA VALIDATION . . . 73
6.4 DATA QUALITY ASSESSMENT . . . 74
6.5 SUMMARY . . . 76

**7. REFERENCES . . . 77**

**APPENDIX A. OTHER DEFINITIONS OF DATA VERIFICATION AND DATA VALIDATION . . . A-1**

**APPENDIX B. GLOSSARY . . . B-1**

**APPENDIX C. EXAMPLES OF DATA QUALIFIERS USED BY SPECIFIC PROGRAMS . . . C-1**

## LIST OF FIGURES


| | | Page |
|---|---|---|
| Figure 1. | EPA Quality System Components and Tools | 2 |
| Figure 2. | Data Verification and Data Validation Components in the Project Life Cycle | 4 |
| Figure 3. | Data Verification Process | 12 |
| Figure 4. | Data Validation Process | 16 |
| Figure 5. | Example Data Verification Checklist for Sample Receipt | 36 |


## LIST OF TABLES


| | | Page |
|---|---|---|
| Table 1. | Records Commonly Used as Inputs to Data Verification | 10 |
| Table 2. | Examples of Documents and Records Generated during Field Activities | 19 |
| Table 3. | Examples of Improper Laboratory Practices and Warning Signs for Data Validators | 28 |
| Table 4. | Examples of Improper Field Sampling Practices and Warning Signs for Data Validators | 32 |
| Table 5. | Examples of Types of Field Records, Purpose of Each, and the Recorded Information | 53 |
| Table 6. | Examples of Items to Review for Consistency Checks for the Same Type of Information | 55 |
| Table 7. | Examples of Items to Review for Consistency Checks Between Types of Information | 56 |
| Table 8. | Examples of Data Validation Qualifiers and Definitions | 64 |
| Table 9. | Data Validation Versus Data Suitability | 73 |

## LIST OF ACRONYMS

| | |
|---|---|
| COC | chain of custody |
| DQA | data quality assessment |
| DQI | data quality indicator |
| GC | gas chromatography |
| LIMS | laboratory information management system |
| MS | mass spectrometry |
| MQO | measurement quality objective |
| PAH | polyaromatic hydrocarbon |
| PE | performance evaluation |
| QA | quality assurance |
| QC | quality control |
| SAP | sampling and analysis plan |
| SOP | standard operating procedure |
| SVOC | semivolatile organic compound |
| VOC | volatile organic compound |

# CHAPTER 1

# INTRODUCTION

## 1.1 PURPOSE AND OVERVIEW

A primary goal of the U.S. Environmental Protection Agency's (EPA's) Agency-Wide Quality System is "to ensure that environmental programs and decisions are supported by data of the type and quality needed and expected for their intended use...." (*EPA Quality Manual for Environmental Programs*, EPA Order 5360 A1) (EPA, 2000a). Accomplishment of this goal involves a set of activities conducted during the planning, implementation, and assessment phases of an environmental data collection project (Figure 1).

As used in this guidance, environmental data collection refers primarily to the sampling and analysis of environmental media. Though the main emphasis is on the collection of environmental samples and their analysis in a chemistry laboratory, many of the principles and practices described in this document are applicable to related measurement activities, such as bioassays, air monitoring, collection and use of geospatial data, and spatial data processing. The guidance does not address the collection or evaluation of other categories of data (economic, demographic, etc.) that play a role in environmental decision making, nor does it directly address the evaluation of secondary data (i.e., previously collected data compiled in EPA or other data sets).

Figure 1 shows that data verification and data validation are key steps in the assessment phase. The purpose of this guidance is to explain how to implement data verification and data validation in the context of EPA's Quality System, and to provide practical advice and references. This guidance describes an array of data verification and data validation practices in order to promote common understanding and effective communication among environmental laboratories, field samplers, data validators, and data users. This guidance also describes the related subjects of data integrity (how the data validator can help detect possible falsification of data) and data suitability [how the data validator can anticipate and support decisions about the usability of the data.

Although data verification and data validation are commonly-used terms, they are defined and applied differently in various organizations and quality systems. (See Appendix A for other definitions of data verification and data validation.) Without attempting to preempt other meanings or approaches, this guidance incorporates the following definitions:

> **Data Verification** is the process of evaluating the completeness, correctness, and conformance/compliance of a specific data set against the method, procedural, or contractual requirements.

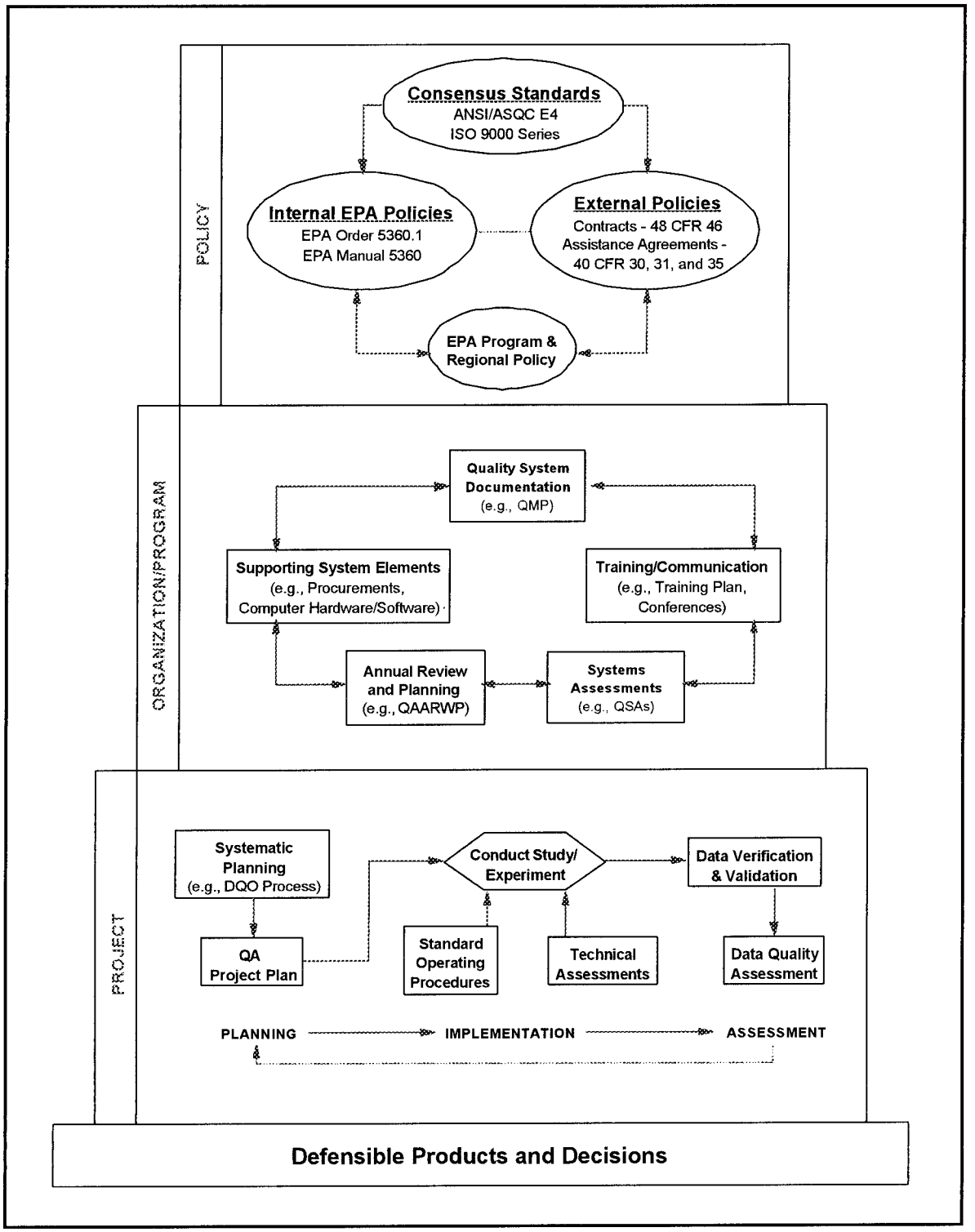


**Figure 1. EPA Quality System Components and Tools**

**Data Validation** is an analyte- and sample-specific process that extends the evaluation of data beyond method, procedural, or contractual compliance (i.e., data verification) to determine the analytical quality of a specific data set.

These definitions are parallel, and the processes that they describe are clearly related. Nevertheless, the terms data verification and data validation, as used in this guidance, reflect two separate processes with two separate functions. The fundamental difference between them is embedded in their respective emphases. Data verification is primarily an evaluation of performance against pre-determined (and often generic) requirements given in a document such as an analytical method procedure or a contract. Data validation, on the other hand, focuses on particular data needs for a project, as stated in a project-specific document such as a Quality Assurance (QA) Project Plan. Furthermore, data verification and data validation are typically sequential steps performed by different parties; data verification is performed during or at the culmination of field or laboratory data collection activities, whereas data validation is conducted subsequently, almost always by a party independent of both the data collector and the data user. Data validation begins with the outputs from data verification.

The definitions and approaches described in this guidance are not intended to be prescriptive or necessarily to be applied rigidly across all programs, organizations, and circumstances. Instead, this guidance will provide a clear overview of how data verification and data validation fit into EPA's Quality System, and will describe tools and techniques that can be employed to meet the goals that are common to all environmental data quality systems. Indeed, these verification, validation, and usability definitions and activities form a continuum and distinction between steps are somewhat artificial.

## 1.2 DATA VERIFICATION/VALIDATION IN THE PROJECT LIFE CYCLE

EPA's Quality System has been described in other documents issued by the EPA Quality Staff – see, for instance, *EPA Requirements for Quality Management Plans (QA/R-2)* (EPA, 2001a). This system provides an integrated set of policies, programs, and project-level tools, all with the common goal of producing defensible products and decisions. As shown in Figure 1, data verification and data validation fit into the category of project-level tools. This category of tools includes systematic project planning, project implementation in the field and analytical laboratory, and the assessment phase, where data are evaluated and prepared for use.

Figure 2 illustrates the overall framework and feedback loops that may be needed for data verification and data validation. Although data verification and data validation are both considered assessment tools, chronologically they occur prior to the formal data quality assessment (DQA) process. DQA is described in the *Guidance for Data Quality Assessment: Practical Methods for Data Analysis (QA/G-9)* (EPA, 200b). As discussed in subsequent chapters, the goal of data verification is to ensure and document that the data are what they purport to be, that is, that the reported results reflect what was actually done. Data validation is generally carried out (usually by an

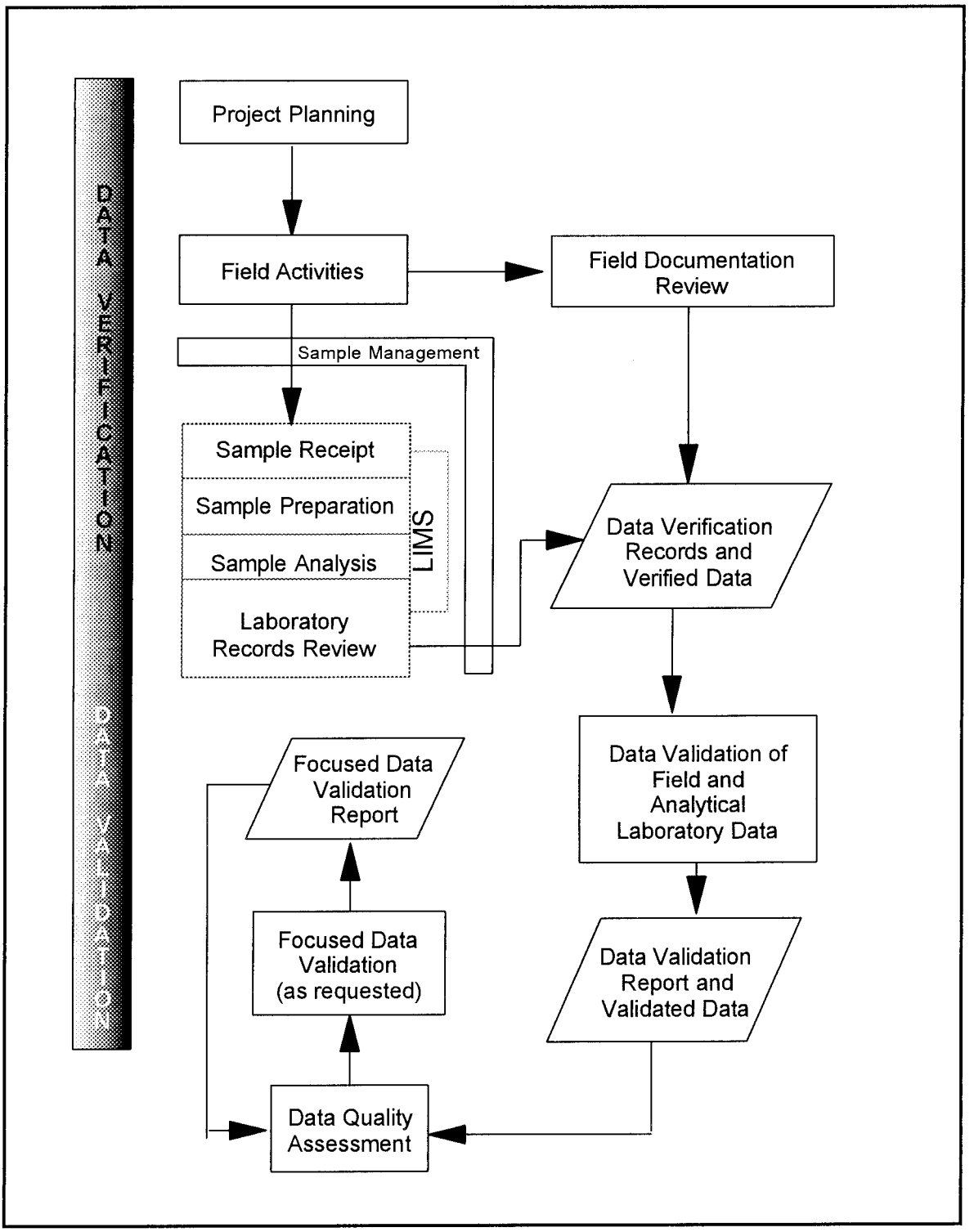


**Figure 2. Data Verification and Data Validation Components in the Project Life Cycle**

external party) as part of the assessment phase. The goal of data validation is to evaluate whether the data quality goals established during the planning phase have been achieved. As shown in Figure 2, data validation involves the outputs of the planning and implementation phases. The data validator may also be requested to perform a detailed investigation of particular data records that need special interpretation or review, referred to as a focused data validation (Section 3.3.3).

During the DQA process, the DQA analyst's focus is on environmental decision making, and whether the data sets that have been generated can effectively and credibly support those decisions. Data verification and data validation, on the other hand, do not concentrate on decisions, but on specific sampling and analysis processes and results. They may involve conclusions about whether project-specific measurement quality objectives (MQOs) for precision, bias, or other data quality indicators (DQIs) have been achieved. Note that MQOs are *inputs to* rather than the *culmination of* data quality assessment. For more information, see the peer review draft of *Guidance of Data Quality Indicators (QA/G-5i)* (EPA, 2001b).

To further clarify the respective roles of data verification, data validation, and DQA, consider the following example. As part of a site characterization soil sampling program for evaluating a potential remediation project, silver is a metal of interest. After samples have been collected, analyzed, and the results reported, the data set is submitted for data verification. The data verification process documents that silver recoveries for spiked samples fell below control limits. The data validation process traces the cause for the non-conformance to an elevated pre-spike sample concentration. The data validator notes that the laboratory control samples all have recoveries within criteria, that other spiked samples have recoveries within criteria, and that field duplicate results have significant variability. The data validation process determines that the low silver recovery is a result not of analytical bias, but of the heterogeneity of the matrix. The data quality assessment process considers the fact that all soil samples had silver concentrations below the action limit for the site by a factor of two or more, and therefore the data quality is adequate for the purpose of the site characterization. The matrix variability is noted and should be taken into account in planning future sample collection.

The EPA Quality System incorporates the principle of the graded approach. This principle recognizes that a "one size fits all" approach to quality will not be effective, given the wide variety of environmental programs. The graded approach applies to data verification and data validation on a project-specific basis, as established during project planning, and communicated in planning or implementation support documentation such as a QA Project Plan or a standard operating procedure (SOP). The level of detail and stringency of data verification and data validation efforts should depend on the needs of the project and program in question. Depending on the application of the graded approach, the individual data verifier or data validator may implement only a subset of the techniques offered in this document. For instance, while many data validation protocols "flag" data from a specific list of data qualifiers, other data validation protocols may use primarily narrative reports. In general,

exploratory studies do not need the same degree of rigor as would enforcement cases in which analytical results may be presented and defended in court.

In order to be useful to the widest audience possible, this guidance presents a broad array of data verification and data validation techniques and examples, not a prescription for how data verification and data validation is performed in all circumstances. Whenever program-specific terms or concepts are presented in this guidance, they are offered for illustrative purposes only.

## 1.3 INTENDED AUDIENCE

The primary audience for this guidance is practitioners directly involved in implementing or managing data verification or data validation efforts. This guidance should provide this audience with a conceptual overview, some "how-to" implementation details, and resources for additional information and exploration. A secondary audience for this guidance consists of DQA analysts (i.e., individuals responsible for conducting data quality assessments) as well as managers responsible for DQA or for the eventual use of verified and validated data; these groups will benefit from an understanding of the data verification and data validation processes and the potential uses and limitations of validated data.

Note that this guidance describes how to verify or validate field activities and results in addition to analytical laboratory activities and results. The concepts are equally applicable to both field and laboratory activities, and from the perspective of the data user, the validity of field results is at least as important as that of analytical data.

## 1.4 PERIOD OF APPLICABILITY

Based on the *EPA Quality Manual* (EPA, 2000a), this guidance will be valid for a period of five years from the official date of publication. After five years, this guidance will either be reissued without modification, revised, or removed from the EPA Quality System series.

## 1.5 ORGANIZATION OF THIS GUIDANCE

Chapters 2 and 3 introduce data verification and data validation, and describe their process inputs, activities, and outputs. Chapter 4 describes data integrity, primarily from the perspective of what the data validator can do to detect and counteract deliberate falsification of data. Chapter 5 presents "how-to" details for data verifiers and data validators. Chapter 6 completes this guidance with a look at data suitability, and how the data validator can support the needs of the DQA analyst.

# CHAPTER 2

# DATA VERIFICATION

## 2.1 INTRODUCTION TO THE DATA VERIFICATION PROCESS

For the purposes of this guidance, the term "data verification" is the process of evaluating the completeness, correctness, and conformance/compliance of a specific data set against the method, procedural, or contractual requirements. Again, the goal of data verification is to ensure and document that the data are what they purport to be, that is, that the reported results reflect what was actually done. When deficiencies in the data are identified, then those deficiencies should be documented for the data user's review and, where possible, resolved by corrective action. Data verification applies to activities in the field as well as in the laboratory.

Data verification may be performed by personnel involved with the collection of samples or data, generation of analytical data, and/or by an external data verifier. In general, the distinction can be made between the person producing the data to be verified (the sampler, surveyor, preparation technician, or bench analyst) and the person verifying the data (the sample custodian, lead chemist, or external data verifier). An external data verification may be performed by some agencies or programs upon receipt of data packages to confirm the completeness of the data package and to permit authorization of payment for the work. Personnel who may be involved in the collection of samples or the generation of the data, as well as individuals who may receive the final documentation and arrange for data verification include:

- sample collection personnel,
- surveyors/mappers,
- drillers,
- air monitoring personnel,
- sample custodians,
- preparation chemists,
- bench chemists,
- lead chemists,
- report preparers,
- data reviewers,
- project leaders,
- QA officers or managers,
- laboratory directors, and
- remediation project managers.

Any or all of these personnel may be involved in the data verification process. The functions performed by, not the titles assigned to, these personnel are what involves them in data verification. Each role might be filled by a separate person in larger laboratories or field operations, while in smaller organizations there may be fewer distinct job categories, with one person performing several functions.

Sampling protocols, analytical methods, and project-specific planning documents are examples of sources that can provide the specifications for the environmental data collection effort. Data verification evaluates how closely these documents and procedures were followed during data generation. Each person involved in data verification should understand the data generation procedures and should know project documentation requirements. Therefore, in order for data verification to be most effective, these planning documents and procedures should be readily available to all of the people involved in the process. The documents and procedures vary according to specific program requirements, but may include project-specific QA Project Plans, sampling and analysis plans (SAPs), reference methods from a variety of sources including EPA, as well as laboratory-specific SOPs and protocols. In some cases, a person or a facility involved with a portion of the data generation process may not have access to all, or any, of the project-specific planning documents. For example, a drilling subcontractor may be working from an internal SOP, or a subcontract laboratory may be provided only with method references from an analysis request form. If a project-specific document (e.g., a QA Project Plan) had additional specifications not known during data generation, this may hamper the achievement of the project objectives. In this example, data should be verified against the applicable standard (i.e., the internal SOP or reference method), and any deviations of these criteria from specifications provided in other, additional project-specific documents would be noted in the data verification documentation.

Not every project involving field or laboratory analyses will involve the same degree of planning. As noted in Section 1.3, EPA QA guidelines recognize that different programs for gathering environmental data will need different levels of detail through a graded approach. Similarly, different projects will have different needs regarding data verification. For some projects, data verification will be predominantly an internal function of the field or laboratory staff. For other projects, it may be more appropriate to have an external data verification.

Data verification is a part of what field and laboratory staff and managers routinely do to ensure that they are producing appropriate outputs. Using the bulleted list of personnel previously discussed, data verification in the field or within the laboratory should occur at each level (i.e., all personnel should verify their own work) and data verification should also occur as information is passed from one level to the next (i.e., the sample custodian should verify the information provided by the field personnel, and supervisors should verify the information produced by their staff).

Data verification by an external data verifier differs from that performed by the field or laboratory staff primarily in the timing. While field or laboratory staff verify data in "real time" or near

real time, external data verification is performed after receipt of field records or a complete data package. To the extent possible, records are reviewed for completeness, for factual content, and against project specifications.

## 2.2 INPUTS TO DATA VERIFICATION

Generating environmental data of any kind involves the production of documentation or records, from daily field logs regarding the collection of the samples to electronic records in a laboratory data system. All such records are potential inputs to the data verification process. Therefore, the first step in data verification is to identify the records that are produced, and to determine the criteria or specifications against which the records will be compared. Such criteria or specifications should be described in:

- project-specific planning documents for a given project;
- program-wide planning documents (e.g., Quality Management Plan);
- SOPs, including field and laboratory methods; or
- published, approved sampling or analytical methods [e.g., SW846 methods or American Society for Testing and Materials protocols].

Project-specific planning documents should include a QA Project Plan [see *Guidance for Quality Assurance Project Plans (QA/G-5)* (EPA, 1998)] or equivalent document.

As the data collection effort progresses from sample collection through sample analysis, the field and laboratory personnel produce a series of records that can be verified. These records may be verified at each sequential step and/or during the final record review process.

Table 1 presents information on a number of common operations in the process of environmental data generation, commonly-used records, and the likely source of the specifications for such records. The extent to which these records exist or apply will be a project-specific issue. The information in Table 1 should not be considered "requirements" for any particular project.

Records may be produced and maintained solely as hard copy, produced as hard copy and maintained electronically, or only produced and maintained electronically, depending on the project needs and the practices of the participants. Records that provide inputs to data verification may be in hard copy or electronic format. Field teams collecting samples may enter data in weatherproof, bound field notebooks, or they may use hand-held electronic devices to record field notes, log samples as they are collected, print labels for sample containers, etc. Other hand-held devices, such as global positioning system instruments, may also be used to record field information. A laboratory may employ an electronic data storage system, generically known as a laboratory information management system (LIMS), as a centralized repository for much of the information regarding analyses of samples. Newer

laboratory instrumentation is designed to be directly linked with a LIMS, thus eliminating much of the manual recording and transcription of data that has occurred in the past. Calculations once performed by hand are now made electronically in real time, or nearly real time, and automatically by the LIMS. Conversely, in a smaller laboratory or specialized analytical department, there may still be many hand-entered records that exist as hard copy only [e.g., multi-part manual chain-of-custody (COC) forms, pH results, or atomic absorption run logs]. Even a completely electronic sample collection and analysis process would still need data verification; the execution of the data verification process would change, not the goal or the inputs.

**Table 1. Records Commonly Used as Inputs to Data Verification**

| Operation | Common Records | Source for Record Specifications |
|---|---|---|
| Sample collection | Daily field logs, drilling logs, sample collection logs, COC forms, shipper's copy of air bill, surveys | QA Project Plan or SAP, SOPs for sample collection, pre-printed COC instructions |
| Sample receipt | COC forms from sampler, receiver's copy of air bill, internal laboratory receipt forms, internal laboratory COC forms, laboratory refrigerator or freezer logs | QA Project Plan or SAP, laboratory SOP for sample receipt, pre-printed COC instructions |
| Sample preparation | Analytical services requests, internal laboratory receipt forms, internal laboratory COC forms, laboratory refrigerator or freezer logs, preparation logs or bench notes, manufacturer's certificates for standards or solutions | QA Project Plan or SAP, reference method (EPA or other), laboratory SOP for preparation method, pre-printed instructions on internal forms |
| Sample analysis | Analytical services requests, internal laboratory receipt forms, internal laboratory COC forms, laboratory refrigerator or freezer logs, manufacturer's certificates for standards or solutions, instrument logs or bench notes, instrument readouts (raw data), calculation worksheets, quality control (QC) results | QA Project Plan or SAP, reference method (EPA or other), laboratory SOP for analysis method, pre-printed instructions on internal forms and worksheets |
| Records review | Internal laboratory checklists | QA Project Plan or SAP, laboratory SOP for analysis method or laboratory QA plan |

## 2.3 IMPLEMENTATION OF DATA VERIFICATION

This chapter provides an overview of data verification and outlines two steps in that process:

1. identifying the project needs for records, documentation, and technical specifications for data generation; and determining the location and source of these records.

2. verifying records that are produced or reported against the method, procedural, or contractual requirements, as per the field and analytical operations listed in Table 1, as applicable (specifically, sample collection, sample receipt, sample preparation, sample analysis, and data verification records review).

Figure 3 is a flow diagram depicting the organization of these steps. Chapter 5 provides a detailed discussion of how data verification may occur in a typical environmental data generation project.

The first part of step one, identifying the project needs, may begin by asking *"Why is this data collection project being conducted?"* Answering this question will generally lead the data verifier to review the various planning documents associated with the project. The data verifier should use these documents to determine the purpose of the data collection, and they should also specify the needs for the sample collection, data generation, and documentation of the analysis.

Planning document requirements will vary according to the purpose of the sample collection and anticipated end use of the analytical results. They will also vary with the nature of the analysis. For example, the requirements placed on a gas chromatography/mass spectrometry (GC/MS) analysis of semivolatile organic compounds (SVOCs) in a water sample would involve significantly more records than determining the pH of the same sample. However, even when using a relatively simple technique, such as pH determination, there may be differences between the project requirements, given different purposes. The determination of the pH of a sample relative to a regulatory requirement may involve more detailed record-keeping than a non-regulatory determination. Such differences should be reflected in the planning documents.

Project specifications may also include specifications for the analyses and for the resulting data reports. These specifications play an important role in verifying that what was done matches what was requested. For example, if the project needs a specific method employed, that should include a specification that the laboratory document what method was used for the analysis. In this example, data verification ensures that the method used by the laboratory was identified, and ensures that the specified method was used and that it met technical criteria that were established in the planning process.

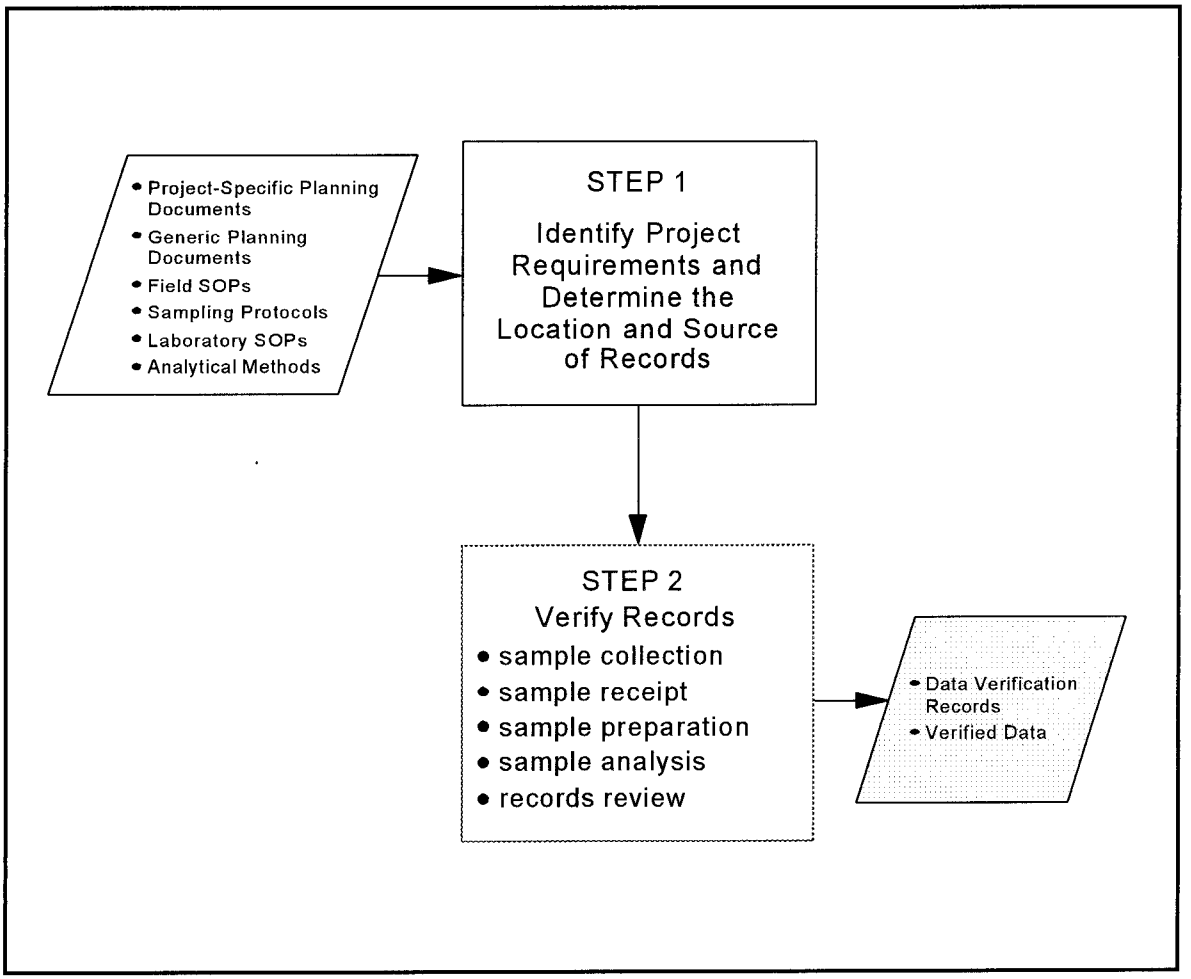


**Figure 3. Data Verification Process**

The second part of step one, determining the location and source of the records that are produced, is equally important. As noted earlier, the records may be produced by a number of personnel and maintained in a number of formats. All personnel should comply with the record-keeping procedures of the laboratory or the project. At any point in the data generation chain, the information needed for data verification should be available to the people responsible and the project requirements themselves should be clearly identified in the planning documents.

Many laboratory records may be maintained in a LIMS. The LIMS may also perform calculations using information (data) from those records. Therefore, identifying the source and location of the records also means identifying all the calculations performed on the input data. While the data verification process need not recheck the results of every automated calculation, the algorithms used for the calculations should be verified during the design of the LIMS. This is an example of records that

may or may not be needed by the project. However, whether a LIMS or manual system is used to process laboratory data and generate analytical reports, the data verification often includes a percentage of "raw data calculation verifications." The data verifier recalculates reported results using instrument outputs (e.g., absorbances) or recorded measurements (e.g., volume of titrant) for samples and standards, along with sample-specific preparation information (e.g., dilutions, percent moisture).

Step two of data verification compares the records that are produced against project needs. The project planning document that specifies the records to be reported should be used to determine what records to verify. In the absence of such an organizational specification, the determination of data to be verified may be left to the discretion of the project manager, lead person, or principal investigator. It is during this step of data verification that the results of the data collection activities are compared against the applicable standard, whether it is, for example, the SOP for sample collection, an EPA method for analysis, or the technical specifications provided in a detailed QA Project Plan for post-treatment soil sampling.

If electronic data are available to the data verifier, certain routine components of data verification are amenable to automation. These components may include interpreting the results of QC samples, holding times, and blank results. For example, EPA offers a Data Assessment Tool as a Contract Laboratory Program service.[1] Data Assessment Tool contains three separate programs: Contract Compliance Screening, Computer-Aided Data Review and Evaluation, and Data Assessment Rapid Transmittal to rapidly transfer analytical data into client databases. Computer-Aided Data Review and Evaluation examines the QC data for all analytical results and evaluates them against data review criteria which are appropriate for the corresponding analytical method/procedure and the intended use of the results. Computer-Aided Data Review and Evaluation uses both regional and national functional guidelines to review and evaluate the data. There is also commercial data verification software available that produces reports in common formats. These packages provide data qualification (flagging) and reports for precision, bias, detection limits, surrogates, and blank contamination. However, automated verification is not complete by itself for any data verification that may need visual, technical, inspection of chromatograms, mass spectra, and other instrument data. Data verification software may not be able to address all of the verification needs of a project. Any software package should be thoroughly evaluated before it is relied upon and used.

## 2.4 OUTPUTS OF DATA VERIFICATION

There are two general results or outputs of data verification, the verified data and the data verification records.

[1]For more information, see www.epa.gov/oerrpage./superfund/programs/clp/dat.htm.

The first output is verified data. Verified data are data that have been checked for a variety of factors during the data verification process, including transcription errors, correct application of dilution factors, appropriate reporting of dry weight versus wet weight, correct application of conversion factors, etc. Verified data may also include laboratory qualifiers, if assigned. Any changes to the results as originally reported by the laboratory should either be accompanied by a note of explanation from the data verifier or the laboratory, or reflected in a revised laboratory data report.

The second output from data verification is referred to as "data verification records" in this guidance. A main part of these records may be a "certification statement" certifying that the data have been verified. The statement should be signed by the responsible personnel, either within the organization or as part of external data verification. Data verification records may also include a narrative that identifies technical non-compliance issues or shortcomings of the data produced during the field or laboratory activities. If data verification identified any non-compliance issues, then the narrative should identify the records involved and indicate any corrective actions taken in response. The records routinely produced during the field activities and at the analytical laboratory (commonly referred to as a data package) and other documentation such as checklists, handwritten notes, or tables should also be included as part of the data verification records. Definitions and supporting documentation for any laboratory qualifiers assigned should also be included.

# CHAPTER 3

# DATA VALIDATION

## 3.1 INTRODUCTION TO THE DATA VALIDATION PROCESS

For the purposes of this guidance, the term "data validation" is an analyte- and sample-specific process that extends the evaluation of data beyond method, procedural, or contractual compliance (i.e., data verification) to determine the analytical quality of a specific data set. Data validation criteria are based upon the measurement quality objectives[2] developed in the QA Project Plan or similar planning document, or presented in the sampling or analytical method. Data validation includes a determination, where possible, of the reasons for any failure to meet method, procedural, or contractual requirements, and an evaluation of the impact of such failure on the overall data set. Data validation applies to activities in the field as well as in the analytical laboratory.

As shown in Figure 4, data validation includes inspection of the verified data and both field and analytical laboratory data verification records; a review of the verified data to determine the analytical quality of the data set; and the production of a data validation report and, where applicable, qualified data. A focused data validation may also be needed as a later step (see Section 3.3.3). The goals of data validation are to evaluate whether the data quality goals established during the planning phase have been achieved, to ensure that all project requirements are met, to determine the impact on data quality of those that were not met, and to document the results of the data validation and, if performed, the focused data validation. The main focus of data validation is determining data quality in terms of accomplishment of measurement quality objectives.

Data validation is typically performed by person(s) independent of the activity which is being validated. The appropriate degree of independence is an issue that can be determined on a program-specific basis. At a minimum, it is preferable that the validator does not belong to the same organizational unit with immediate responsibility for producing the data set.

As in the data verification process, all planning documents and procedures should be readily available to the data validators. A data validator's job cannot be completed properly without the knowledge of the specific project needs. In many cases, the field and analytical laboratory documents and records are validated by different personnel. Because the data validation process needs knowledge of the type of information to be validated, a person familiar with field activities is usually assigned to the data validation of the field documents and records. Similarly, a person with

---

[2]Measurement quality objectives are "acceptance criteria" for quality attributes measured by project DQIs. During project planning, MQOs are established as quantitative measures of performance against selected DQIs, such as precision, bias, representativeness, completeness, comparability, and sensitivity.

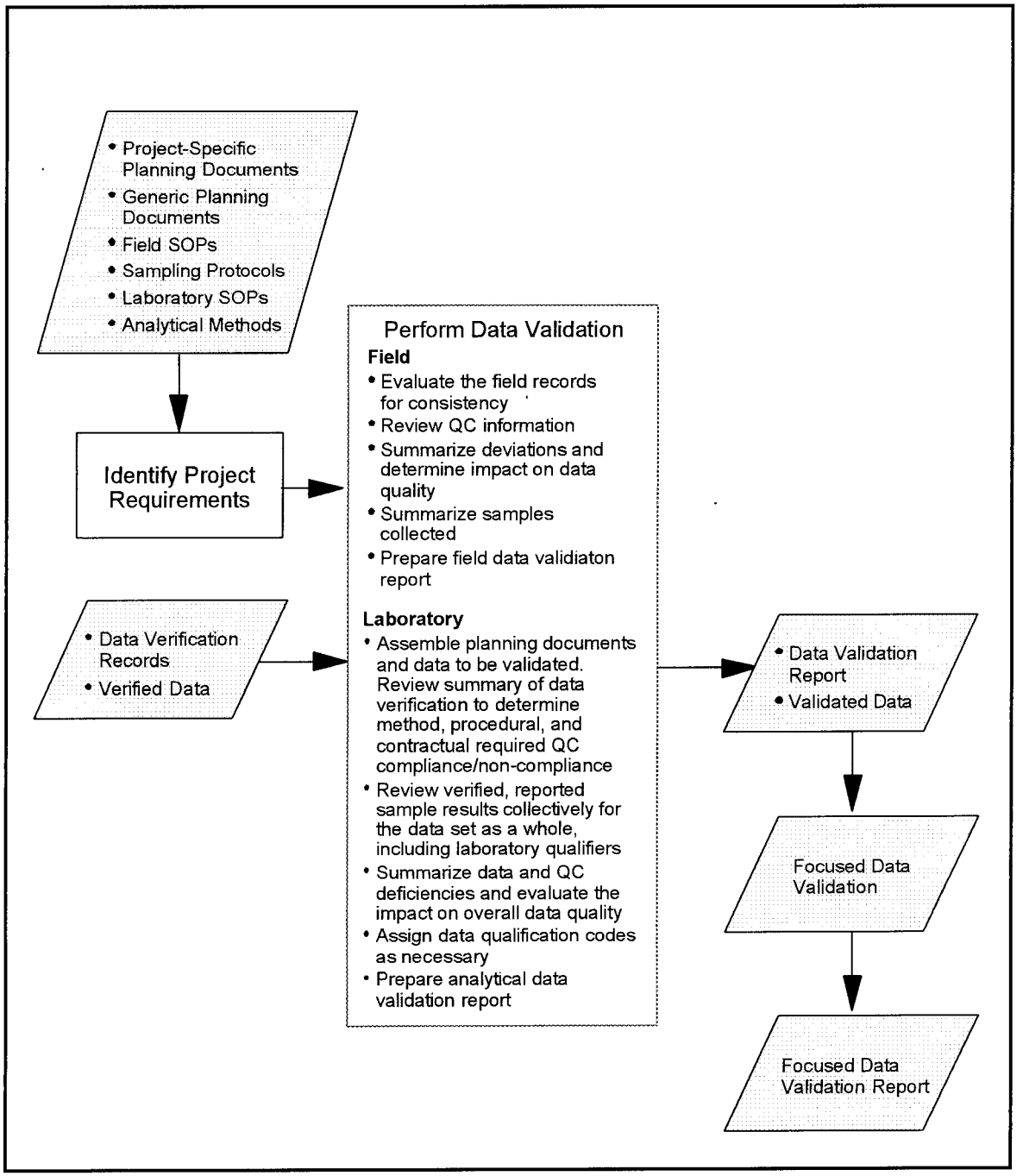


**Figure 4. Data Validation Process**

knowledge of analytical laboratory analysis, such as a chemist, aquatic biologist, or microbiologist (depending on the nature of the project, is usually assigned to the data validation of the analytical laboratory documents and records. In any case, the project needs should assist in defining the appropriate personnel to perform the data validation.

The personnel performing data validation should also be familiar with the project-specific DQIs and associated measurement quality objectives. One of the goals of the data validation process is to evaluate whether the data quality goals established during the planning phase have been achieved. In order to do so, certain data quality attributes are defined and measured. DQIs (such as precision, bias, comparability, sensitivity, representativeness, and completeness) are typically used as expressions of the quality of the data.

The inputs to data validation, the data validation process, focused data validation, and the outputs of data validation are described in this chapter. The level of data validation that is performed will be specific to each project. This chapter covers a wide range of records that may be involved in the data validation process. Because each project is unique, some topics discussed in this chapter may not be applicable to all projects, while a few projects may have more records than is discussed in this guidance.

## 3.2 INPUTS TO DATA VALIDATION

The planning stage of a project is vital to understanding what the expectations are for the project. Documents generated or reviewed during the planning stages of a project may include:

- project-specific planning documents (e.g., QA Project Plan or a SAP);
- program-wide planning documents (e.g., Quality Management Plan);
- SOPs including field and laboratory methods for any aspect of the data generation process; or
- published, approved sampling or analytical methods (e.g., SW846 methods or American Society for Testing and Materials protocols).

### 3.2.1 Project-Specific Planning Documents

The project-specific planning documents should state sampling objectives and identify project needs that should be met during the implementation of the project. Any products generated during the implementation of the project should be measured against specific needs from each of these planning documents.

The data validator should be familiar with planning document objectives and needs in order to identify those documents and records that should be included in data validation. Data validation begins

with the outputs from data verification discussed in Section 2.4. The verified data and data verification records, including a statement certifying that the data have been verified, are passed on to the data validator(s).

The verified data may be provided in hard copy or electronic format. A data validator may use electronic data, if available, to perform part of the data validation. When the verified data are available electronically, it is important to make sure that the data verification records and the electronic verified data present consistent information. If multiple sets of electronic data exist, these sets may be combined into a common database to facilitate the portion of the data validation process that can be done electronically. In this case, the database should be designed by the data user, so all electronic data will be available in a structured, usable format. The database may contain pre-defined fields to be populated with the analytical laboratory data as well as the field activities data. The data user should define electronic data needs in the appropriate planning documents to ensure that electronic data will easily upload to the database, that all necessary fields be reported by the field team and analytical laboratory, and that any other needs for electronic records are met.

### 3.2.2 Inputs from Field Activities

When samples are collected from environmental media for a project, the verified data and data verification records, including all field records generated from the sample collection activities, should be available for data validation. Field teams may have numerous members for some projects, while team members may have multiple roles for other projects. Field team members that may contribute to the data verification process include:

- field team leader,
- site safety officer,
- sampler,
- documenter,
- radiological technician,
- industrial hygienist,
- drilling team,
- heavy equipment operator, and
- decontamination team.

Most of the field team members contribute to the documentation of the field activities, some keeping records that provide information duplicated on another form. For example, the field team leader, the site safety officer, and the lead driller may each keep daily activity records, with each record focusing on a specific function. Although the records are for different purposes, they should be quite similar in content.

In a matter involving potential litigation, all of the records generated during field activities may become evidentiary documents and the needs of the project should be considered when these records are being validated. Table 2 contains a list of example records that may be generated during field activities and the purpose of each document. The data validator should note that the names of the records used here are typical, but each data validator will be working with field records specific to the project. In these cases, the data validator should identify the records that correspond to the tables here. A more detailed discussion of field records is presented in Chapter 5.

**Table 2. Examples of Documents and Records Generated during Field Activities**

| Type of Document or Record | Purpose of Document or Record |
|---|---|
| Instrument calibration records | Maintains accurate record of instrument calibration |
| Field notebook or daily activity log | Maintains accurate record of field activities by providing written notes of all activities |
| Sample collection logs | Maintains accurate record of samples collected |
| Chain-of-custody | Maintains proof that samples were not tampered with and that samples were under the appropriate possession at all times |

### 3.2.3 Inputs from the Analytical Laboratory

The data verification records should support the verified data that are reported. The data validator should already be aware of the needs from the planning documents so that the data validator knows what information the laboratory was to provide. Because each project is unique, the data validator should review the documentation that will allow determinations of the quality of the data to be made. For example, the data validator should ensure that the correct inorganic preparation method was followed (e.g., use of hydrofluoric acid for digestion).

In the process of receiving, preparing, and analyzing samples and reporting the results, the laboratory may generate numerous records. Not all of these records are generally included with the analytical data package normally provided by the laboratory but the validator should determine that all appropriate of records have been provided before initiating validation.

Electronic records that provide input to data validation may be referred to as electronic data deliverables. Data that can be entered into an electronic database may include sample results, units, dilution factors, sample numbers, and analytical methods. Items such as raw data, however, are usually available only in the hard-copy documentation unless a scanned version of the raw data is available electronically.

## 3.3 IMPLEMENTATION OF DATA VALIDATION

This chapter outlines the three basic steps of data validation, which include:

1. identifying the project needs for records;
2. obtaining the records that were produced during data verification; and
3. validating the appropriate records to determine the quality of data and whether or not project needs were met by performing data validation and focused data validation, as requested.

Figure 4 outlines the data validation process. Chapter 5 provides a detailed discussion of how data validation may occur in a typical environmental project.

The first step, identifying the project needs, begins with a review of the planning documents for the project. These documents should identify not only the objective of the analysis performed, but also the project-specific needs to be met. The data validator should outline all of the planning document needs in order to understand what documents and records should be reviewed during data validation.

The second step, obtaining verified data and the data verification records, including field records or an analytical data package, is important to ensure that the data validator has a complete set of information to perform the data validation. The data validator should account for all records that are needed by the planning documents. If the data validator does not possess all the documentation needed for the project, the data validation will be incomplete.

Once the project needs have been identified and all appropriate records have been obtained, the data validation begins. Through this process, the data validator should ensure that all samples collected and the data generated for those samples are fully supported by documentation that will assist in the defense of project decisions.

Some projects have the data validator assign qualifiers to the data records in order to identify potential deficiencies or concerns about the quality of the data. These qualifiers are referred to as "data validation qualifiers" for purposes of this guidance because they are assigned during data validation. Data validation qualifiers will be discussed in Chapter 5. Some projects may also have a focused data validation performed when the data user has a request for further information. Focused data validation is described in Section 3.3.3 as well as Chapters 5 and 6.

### 3.3.1 Data Validation of Field Activities

After reviewing the planning documents related to sample collection and field activities, the data validator should be aware of the sample collection needs. The data validator should be able to answer

questions such as: *Was a particular method needed for collecting any of the samples? Were field screening methods supposed to be used? Was pre- and post-measurement calibration and standardization completed and in control?* The data validation of the verified data, using the data verification records, and any other field records can be summarized in a series of steps as shown in Figure 4. Each of the steps for field activities data validation is outlined in Figure 4 and discussed in detail in Chapter 5. The five steps are:

1. evaluate the field records for consistency,
2. review QC information,
3. summarize deviations and determine impact on data quality,
4. summarize samples collected, and
5. prepare field data validation report.

If electronic verified data are available, the data validator may use these data for some steps of data validation, such as the sample summary table, in order to provide more efficiency in the overall data validation process.

### 3.3.2 Data Validation of Analytical Laboratory Activities

After reviewing the planning documents related to sample analysis, the data validator should be aware of the project requirements that the analytical laboratory was expected to meet. The data validator should be able to answer questions such as: *Was a particular analytical method specified for any analyses? Was a specific reporting limit specified for any particular chemical?* Planning document specifications, based on questions similar to these, help the data validator to focus on the appropriate information during the data validation of the verified data and associated records. The data validation of the analytical laboratory data can be summarized in a series of steps as shown in Figure 4. Each of the steps for data validation of analytical laboratory records is outlined in Figure 4 and discussed in Chapter 5. The five steps are:

1. assemble planning documents and data to be validated. Review data verification records to determine method, procedural, and contractual required QC compliance/non-compliance;
2. review verified, reported sample results collectively for the data set as a whole, including laboratory qualifiers;
3. summarize data and QC deficiencies and evaluate the impact on overall data quality;
4. assign data validation qualifiers as necessary; and
5. prepare analytical data validation report.

If electronic verified data are available, the data validator may use these data for some steps of data validation in order to provide more efficiency in the overall data validation process.

### 3.3.3 Focused Data Validation

A data validator's responsibility includes not only the evaluation of field and analytical data and the assignment of data validation qualifiers (if requested), but also communicating this information to the data user. The data validator should summarize the data validation in such a way that the data user can get a general overview of the data validation before using the data. A focused data validation is a detailed investigation of particular data records identified by the data validator or data user that need special interpretation or review by the data validator. In some cases, the data user may alert the data validator to anticipated problems before the data validation is performed. This may eliminate the need for further review later in the data validation process if the data validator can use this information during data validation. Otherwise, the data user may also identify the need for a focused data validation based on instances such as:

- errors or omissions in the data or data validation report,
- anomalies noted during review of the data and data validation report, and
- anomalies noted during the data quality assessment process.

Despite the best efforts of all data validators, errors and omissions may occur in the data validation process. If the data user identifies errors or omissions in the data or the data validation report, the data user may request a focused data validation by the data validator to correct the oversight. In some instances, the review of the data and data validation report may identify anomalies that the data user needs to resolve. In other instances, questions about the data or data validation report may not arise until during the DQA process. Any of these instances may need a focused data validation. A focused data validation involves communication between the data validator and the data user to resolve the issues that were raised. The data validator may be asked to further explain an aspect of the data validation report or the data validator may be requested to re-investigate some of the hard-copy documentation or the original electronic deliverable to provide additional information to the data user. Further details regarding focused data validation are discussed in Chapters 5 and 6.

## 3.4 OUTPUTS OF DATA VALIDATION

The three outputs that may result from data validation include validated data, a data validation report, and a focused validation report.

The first output is a set of data that has been validated and passed on to the project manager or data user. Validated data should be the same as the verified data with the addition of any data validation qualifiers that were assigned by the data validator. Any corrections or changes noted during the data validator's review of the verified data should be reflected in the validated data. Any specifications for reporting the validated data should be described in one of the planning documents.

The second output, the data validation report, documents the results of data validation for both the field data and analytical laboratory data. In some projects, the data validation report for the field data may be generated separately from the data validation report for the analytical laboratory data. This again illustrates the need to tailor this guidance for each project. The purpose of the data validation report is to provide a summary of data validation to the data user before the DQA process begins. In most cases, the data validator's report is the primary means of communication between the data validator and the data user, so it is important that the report reflects all details of data validation. A discussion of the objectives for sampling and analysis activities and a summary of the needs that the data validator gleaned from the planning documents should be included. Documentation from data validation of field data and analytical laboratory data should also be included in the report. The data validation report should emphasize any deficiencies encountered and clearly describe the impact of such deficiencies on overall data quality. If data validation qualifiers were a part of the data validation process, a summary of the data validation qualifier definitions, assignments, and reasons for the assignments should be included in the data validator's report. These data validation qualifiers should also be included in the validated data set. Any updates and/or corrections that were made to the validated data from the original verified data transfers should also be summarized and explained. The report(s) describing the data validation process should provide sufficient detail for the data user to have an overall idea of the quality of the data and how well the project needs were met.

The third output is a focused data validation report. As explained in Section 3.3.3, a focused data validation may or may not occur in a particular project, so this output is applicable only in certain instances.

If a data validator is asked to review specific information during data validation to clarify information in the data validation report, or review additional information in the hard-copy or electronic records, the data validator should provide a report of the additional clarification or review that was provided. This report should include details such as the question that was asked, how it was resolved, and the person who requested the information. The report may also include information such as a list of the samples collected, field information about how the samples were collected, the analysis performed on the samples, and the quality of the reported data depending on what question the data validator is trying to address. Any details that seem out of the ordinary during a data validator's review should also be documented. Specific formatting of this report should be determined by the content of the focused data validation. In any case, all focused data validation reports should be included with the data validation report to keep a complete record of all data validation efforts.

# CHAPTER 4

# DATA INTEGRITY

## 4.1 BACKGROUND

Traditionally, quality systems for environmental measurements have been based on the assumption that all those involved in implementing the system are adhering to the system specifications. Thus, the efficacy of the data verification and data validation processes discussed in the previous chapters depends (at least in part) on the integrity of all field, laboratory, and management personnel who contributed to the documents and records undergoing review.

Unfortunately, more than a decade's experience has demonstrated that integrity is not a safe assumption. A number of environmental testing laboratories have been subject to investigation, penalties, debarment, and successful criminal prosecution for improper practices that undermine the integrity and credibility of their data. These improper practices have prompted the need to build protective measures into quality systems. This is particularly so because many of these improper practices focus specifically on manipulating and falsifying the QC measurements that are the backbone of traditional QA programs. Although falsification may also be carried out by clients submitting the samples or results, this chapter is focused on the field, laboratory, and management personnel.

This chapter should help alert data validators and other reviewers/users of data to the possibility that a data package may have been tainted by improper field or laboratory practices. The express purpose of most improper field or laboratory practices is to manipulate and disguise the data set so that it looks "normal"; therefore, in many cases, the data validator will be unable to detect even flagrant abuse. Since the data validator may not have access to any analytical information beyond the contents of the field records or the data package, the data validator is often not in an advantageous position to detect falsification.

It should be noted that results of field and laboratory audits may prove useful in identifying potential problems with sample collection and analysis practices designed to provide misleading information. When project planning includes audits of both field and laboratory activities, much insight can be gained into whether there are sound ethical practices being implemented and documented. The data validator may be able to use audit results as a starting point for evaluating suspect data, but should keep in mind that, like the data validator, the auditor's primary purpose was probably not to detect falsification.

Data validators should watch for signs that may indicate improper field and laboratory practices. The following sections provide examples of abuse and warning signs that a data validator should recognize. This is not a complete list, as new methods of falsification are continually developed.

## 4.2 IMPROPER LABORATORY PRACTICES

### 4.2.1 Examples of Improper Laboratory Practices

To some degree, the detection of unethical and improper laboratory practices has proven to be a "moving target." As certain practices have been uncovered and appropriate safeguards built into the system, improper practices have developed in other components of the laboratory's processes. However, it is possible to detect patterns of improper conduct, and known examples of laboratory falsification can be arranged into the following categories. (Several commonly-used colloquial terms for laboratory practices are used throughout this chapter; the glossary in Appendix B includes definitions of these terms. Some terms may include multiple definitions because they are used in various ways.)

Improper practices include:

Failure to Analyze Samples

"Drylabbing" occurs when a laboratory reports analytical results without having actually performed the analyses. Results may be either invented from scratch, or previous legitimate results may be "borrowed" for inclusion in the present data package.

Failure to Conduct Specified Analytical Steps

Similar to "drylabbing," this practice occurs when a laboratory actually performs the analyses of the client's samples, but intentionally fails to conduct the associated QC analyses (such as batch-specific QC measurements); instead, the laboratory reports previously conducted successful QC results. As a result, all subsequent evaluations of the quality of the data become meaningless.

Manipulation of the Sample Prior to Analysis

It is possible to tamper with a sample prior to analysis in order to produce a desired analytical result. This technique is often employed on QC samples, including laboratory control samples, matrix spikes, standards, check standards, or known performance evaluation (PE) samples. Methods of tampering include:

- fortification of a sample with additional analyte (colloquially known as "juicing"),
- removal of small amounts of a known PE sample from an ampule and analyzing it directly before preparing the whole-volume sample that includes reagent water,
- over-dilution of the sample to create a false negative result or biased low recovery, and

- injection of an additional amount of continuing calibration verification solution when recoveries are poor.

In addition, techniques that are otherwise legitimate can be used for inappropriate purposes; for instance, QC samples such as matrix spikes can be excessively "blown down," or they can be "over spiked" with standards to increase the amount of analytes.

Manipulation of Results During Analysis

This category of improper laboratory practices attempts to disguise unacceptable results of QC measures in order to avoid the need to reject data and/or reanalyze samples. One approach is "peak shaving" or "peak enhancement" (i.e., manually adjusting the raw data by subtly reshaping a peak that is slightly out of specification). This practice, which is often referred to colloquially as shaving or juicing, may be the most prevalent, or at least the most frequently detected, form of laboratory falsification.

Another practice is artificially manipulating GC/MS tuning data to produce an ion abundance result that appears to meet specified QC criteria, when, in fact, the criteria were not met.

Another practice involves analysis of volatile organic compounds (VOCs) or other time-sensitive analytes. When a holding time has been exceeded, a laboratory may falsify the date of analysis in the laboratory's data system in order to conceal the exceedance. This practice is known informally as "time-traveling."

Post-Analysis Alteration of Results

This category of abuse involves the falsification or distortion of results following analysis but prior to transmittal of the data package. One practice is the transposition of figures to produce a desired result. For example, the matrix spike recovery was 58%, but was reported as 85%. Another practice is the suppression of particular laboratory qualifiers to conceal information about the analysis. For example, an "M" flag, which usually identifies manual integration of the analyses, may be suppressed to avoid further investigation of the extent of manual integration (see Section 5.2.2 for further discussion of flags). Another practice involves the selection of preferred data and suppression of the remainder (e.g., selectively cropping calibration points in a multi-point calibration curve without proper statistical or technical justification).

The common link in each of these categories is the misrepresentation of the laboratory's performance as it is reflected in the data package. This is usually done to enhance the laboratory's productivity and profitability at the expense of the integrity of the resulting data. Falsification may occur as a result of a systematic organization-wide policy, or it may be instigated by isolated individuals. Regardless, the consequences of this misbehavior can include major delays in the completion of

environmental projects, cost overruns due to the need to repeat sampling and analysis, and damage to the public credibility of the agencies and institutions involved. Perhaps most ominous is the possibility of continuing a threat to public health or the environment as a result of undetected falsification.

### 4.2.2 Warning Signs for Data Validators

External data validation is a good practice that helps maintain and improve data quality, and acts as a deterrent to falsification. But, it is often difficult for data validators to detect laboratory falsification based solely on examination of data packages. Data validation is not the only tool for detection and prevention of improper laboratory practices. A comprehensive approach should include other features, such as periodic on-site audits; analysis of PE samples; inspection/auditing of the laboratory's electronic data files; a systematic laboratory QA function led by an active QA Manager; providing proper training; and requiring sound organizational ethics, policies, and procedures.

The data validator is often the first line of defense against falsification. The data validator may detect the first indications of a problem, leading to further investigation and resolution of any problems. Therefore, the data validator needs to be alert to the various warning signs of potential falsification. Table 3 shows examples of improper laboratory practices and the data validator's warning signs.

**Table 3. Examples of Improper Laboratory Practices and Warning Signs for Data Validators**

| Category | Improper Practice | Data Validator's Warning Sign |
|---|---|---|
| Failure to analyze samples | "Drylabbing" – reporting results without analyzing samples | Overlapping analysis times on the same instrument |
| Failure to conduct specified analytical steps | Reporting previously conducted successful QC results instead of conducting specified QC analyses | QC measurements that are identical to those submitted with past projects. Inadequate run times for sample analysis (may suggest that specified QC checks were skipped) |
| Manipulation of sample prior to analysis | "Juicing" – fortification of a sample with additional analyte | A pattern of high responses for compounds that typically show a low response at that laboratory |
| | Overdilution of a sample | Differences in "background" from sample to sample (i.e., background chromatographic patterns are different for the matrix spike/matrix spike duplicate samples compared to the field samples) |

**Table 3. Examples of Improper Laboratory Practices and Warning Signs for Data Validators**

| Category | Improper Practice | Data Validator's Warning Sign |
|---|---|---|
| Manipulation of results during analysis | "Peak shaving" or "peak enhancement" – manually adjusting results to produce a desired outcome | Repeated manual integrations, especially on QC measurements |
| | Manipulation of GC/MS tuning data to produce a false ion abundance result | Raw data indicating numerous computer operations associated with tuning, tick marks suggesting possible "borrowing" from an adjacent peak |
| | "Time-traveling" – falsifying date of analysis to disguise exceedance of holding times | Inconsistencies in dates (e.g., analysis precedes extraction) |
| Post-analysis alteration of results | Transposition of figures to produce a desired result | Erasures or handwritten changes in the data package printed report from word processor or other software that allows editing, (absence of headers and footers) |
| | Suppression of all "M" flags | Absence of "M" flags even where they might be expected [e.g., polyaromatic hydrocarbons (PAHs) producing co-eluting peaks] |
| | Laboratory selection of preferred data from a larger data set (e.g., to demonstrate an acceptable method detection limit) | Raw data incompatible with calculated results |

The following is a series of questions that a data validator might ask while reviewing a data package. Note that these questions are based on a data validation that might be associated with a complex program (e.g., the references to "M" flags to indicate manual integrations); in practice, data validators may not have access to the information necessary to answer all of these questions. The answer to any of these questions by itself is not a sure indicator of falsification, but a series of disturbing responses suggests that further action may be beneficial. In the absence of previously defined

procedures, the data validator should report any concerns to the data validator's official contact, client, project manager, or project officer.[3]

*Are reported dates in the data package inconsistent (e.g., the date of analysis precedes the date of extraction)?* If so, this would suggest the possibility of "time-traveling" or some other improper manipulation of the analytical results.

*Are there repeated manual integrations or edits, especially related to QC measurements?* If so, this raises the suspicion of "peak shaving" or "peak enhancement," or some other improper manipulation.

*Have all "M" (manual integration) labels been removed, even where they might be expected? Is there an abnormal absence of laboratory qualifiers of any kind? Are the headers and footers that are a standard part of the report format missing from the printed reports?* If so, the laboratory may be suppressing all indicators of improper manual manipulation and editing. Reports that do not have standard headers and footers may have been printed from software that permits editing.

*Are there overlapping analysis times for the same instrument?* If so, this suggests the possibility of "drylabbing" or "time-traveling."

*Does the data package provide complete information on internal standard areas or similar QC-related measures?* If such information was expected, but not provided, in the laboratory data package, at a minimum this raises questions about the laboratory's performance and may suggest the use of improper practices.

*Is there a pattern of high response factors (i.e., sensitivity) for compounds where relatively low response factors are expected?* If so, this suggests the possibility of "juicing."

*Is there an indication that tuning or calibration data may have been manipulated? For example, do the raw data indicate numerous computer operations associated with tuning or calibration? Is there a possibility that an adjacent peak was "borrowed" in lieu of legitimate background subtraction procedures?* If so, this raises questions about the laboratory's performance and may suggest the use of improper practices.

---

[3]Data validators should report through official contacts only, in order to protect their own rights as well as those of the laboratory. Note that laboratories have legal rights to protect themselves against incorrect allegations. Especially in cases where there are only indications rather than compelling evidence of falsification, data validators should be sure to base such reports on demonstrated facts rather than speculation.

*Are there erasures, white-outs, and handwritten changes in the data package? Are all changes properly documented and dated?* Improperly documented changes may suggest improper manipulation of results.

*Are the QC data relevant and associated with the field sample data under review?* If not, the laboratory may be attempting to hide out-of-control performance.

*Is there any indication that the laboratory is selectively choosing desirable QC results while suppressing other data?* If so, the laboratory may be establishing improper calibration curves, method detection limits, etc., by performing more than the specified number of replicates, then selecting and using only the most beneficial results.

*If performance evaluation has been conducted, is there any indication that a PE sample was treated by the laboratory in an unusual fashion?* If so, this may raise questions about the laboratory's performance, but special treatment of a PE sample is not an automatic indicator of abuse.

*Has the laboratory experienced significant data validation problems in the past? Do current data packages look "too good to be true?"* Perhaps the laboratory has systematically addressed past quality problems and is now performing well. However, keep in mind that the laboratories that are tempted to falsify may be those that have experienced performance problems in the past.

*Does the case narrative include discussion of all failures or discrepancies detected during the data validation?* The data validator should consider why the laboratory might be neglecting to report failures or discrepancies.

*Were the operating conditions for QC samples and field samples different? For example was a fast GC ramp speed used for field samples and a slow GC ramp speed used for QC samples?* This could indicate preferential treatment of QC samples.

*Does the data validator have access to electronic data tapes or some other form of raw laboratory data?* Lack of access to raw data is not in itself improper, and in most cases the data validator should not expect to see it. However, when it is available, raw data is useful because it can pinpoint poor practices that would otherwise remain hidden.

This list is far from comprehensive and, as noted above, the patterns and techniques of environmental testing laboratory abuse continue to evolve over time. More important than any particular item is whether the data validator (and ultimately, the data user) can develop a sense of trust in the testing laboratory, based on the laboratory's performance, documentation, and history. In part, this depends on the existence of effective communication feedback mechanisms. It also depends on the fact that data validation is one part of a comprehensive approach to preventing falsification. Most

importantly, this depends on a meaningful and ongoing commitment to the highest ethical standards by all those involved in the collection, analysis, and use of environmental data.

## 4.3 IMPROPER FIELD PRACTICES

Analytical laboratories are not the only potential source of falsification. Field sampling personnel may engage in improper behavior that compromises the integrity of the resulting data. Unfortunately, the data validator can have a more difficult time detecting field activity abuses than laboratory abuses. Table 4 shows examples of improper field practices and warning signs for data validators.

Although improper field practices have not generated the headlines and notoriety that laboratory abuses have caused in recent years, that does not mean that the potential for field abuses is less important. Field work typically proceeds with less formality and automatic scrutiny than laboratory analyses; for instance, records are generally self-generated, often with pen and paper, rather than electronically captured as work proceeds. Unexpected field conditions such as adverse terrain or inclement weather can prompt the temptation to "cut corners" to get the job done. Most importantly, because the effectiveness of the sampling design is probably the single most significant driver of data quality, field abuses can dramatically and permanently compromise the utility of a data set.

**Table 4. Examples of Improper Field Sampling Practices and Warning Signs for Data Validators**

| Improper Practice | Description | Data Validator's Warning Sign |
|---|---|---|
| Mislabeling sample containers | Misrepresenting the sampling date, location, or other key parameter by putting false information on the sample container label | Crossed-out information, inconsistent information between the daily activity logs or the sample collection logs and the sample label |
| Documentation problems | Misrepresenting the sampling process by filling in log books improperly (i.e., to disguise the failure to sample in a location where sampling was specified) | Inconsistencies among daily activity logs, sample collection logs, sample labels, distances from sample locations, and times between samples |

**Table 4. Examples of Improper Field Sampling Practices and Warning Signs for Data Validators**

| Improper Practice | Description | Data Validator's Warning Sign |
|---|---|---|
| Problems with VOC sampling | Reducing the amount of VOCs in a sample prior to submitting the sample for analyses by collecting the sample properly, then leaving the cap off the container or collecting the VOC sample from a composite sample. | Air bubbles noted on laboratory receipt records. Leaving the cap off may result in air bubbles in the sample when the vials were capped |
| Problems with PAH sampling | Placing asphalt in a sample that is being analyzed for PAHs should result in high concentrations of PAHs | Sample description and site information indicates sample location close to a paved area |
| Improper sampling | Adding contamination to samples by collecting samples from an area of known contamination, mixing known contaminated material with material from the actual sample locations, or adding a contamination standard to the material | Inconsistencies among sample collection logs, field notebook, photos, and COC<br><br>Laboratory comments on heterogeneous material |
| | Biasing sampling locations or collecting improper samples by collecting samples from "clean" or "cleaner" areas or collecting samples from somewhere else entirely and forging location information | Records of a site visit made subsequent to sampling indicated that the sample location soil appears undisturbed. |
| | Improper purging of monitoring wells (i.e., samples from monitoring wells can appear "clean" and then suddenly appear "dirty") | Drastic change in sample results |
| | Collecting many samples from one location to avoid the time/cost of a sampling trip | Similar results for multiple samples |

## 4.4 ETHICS CULTURE

The establishment of a culture that promotes and sustains acceptable ethical behavior is a key management issue. An ethics culture should be a part of every organization that contributes to the collection and use of environmental data. This includes not just the testing laboratory, but also field personnel, data validators, and reviewers, and program managers in the client organization.

Chapter 5, Quality Systems Standard, of the 2000 National Environmental Laboratory Accreditation Conference Standard incorporates ethical standards for environmental laboratories (National Environmental Laboratory Accreditation Conference 2000). Highlighted practices include the following:

- laboratories should develop an ethics policy statement, with associated procedures for educating staff in their legal and ethical responsibilities;
- laboratories should maintain documentary evidence that each employee understands and acknowledges these legal and ethical responsibilities; and
- laboratories should develop a proactive program for prevention and detection of improper behavior, including internal testing, audits, reward programs, and SOPs identifying proper and improper practices.

# CHAPTER 5

# TOOLS AND TECHNIQUES FOR DATA VERIFICATION AND VALIDATION

## 5.1 DATA VERIFICATION TOOLS AND TECHNIQUES

As described in Chapter 2, the purpose of data verification is to ensure that the records associated with a specific data set actually reflect all of the processes and procedures used to generate them, and to evaluate the completeness, correctness, and compliance of the data set against the applicable needs or specifications. Chapter 2 also outlined, in general terms, the types of records that are commonly used as inputs to data verification, gave an overview of data verification, and gave the outputs generated as a result of data verification. This section describes the process of data verification in greater detail, focusing on the aspects of data verification that occur during field activities as well as in an environmental laboratory.

The analytical specifications and records needs will vary from project to project, depending to a large extent on the purpose of the sampling and analysis conducted. This section describes data verification using a relatively common project situation as an example—the analyses of samples to determine compliance with regulatory limits on specific constituents. When a project does not need the level of records or record-keeping described here, data verification will be less involved. The data verification process discussion and examples given can be applied to both an internal, real-time data verification as well as an external data verification. Hypothetical but realistic examples are interspersed throughout the chapter and are set off in italics in text boxes.

### 5.1.1 Identifying the Project Needs

The first step in data verification is identifying the project needs for records, documentation, and technical specifications, and determining the location and source of these records. These needs may be specified in a QA Project Plan, a SAP, a contract between the laboratory and the client, or a given regulation. Given a diverse group of potential needs, some organizations may decide to hold all activities to the most stringent record-keeping and documentation needs. This decision is made by each organization, based on their projects and clients.

Checklists are often inadequate for environmental analyses, because not every sample and not every analysis can be easily categorized. However, as records associated with a common analysis type are identified, it may be useful to develop a checklist of the records that will be verified. Figure 5 is an example of a checklist associated with sample receipt. It is intended *strictly* as an example of possible checklist content and format. Other formats may work as well or better, as long as the data verification process is in some way documented. For example, additional detail may be useful for some aspects of data verification or there may be no need for a formal checklist for other aspects.

| Records | Verified | Verified w/ Corrections | Not Verified | Not Applicable | Comments |
|---|---|---|---|---|---|
| Chain-of-custody form | | | | | |
| Shipper's airbill | | | | | |
| Lab log in sheets | | | | | |
| *Additional records as needed ...* | | | | | |

**Verified by: Name: Signature: Date:**

**Figure 5. Example Data Verification Checklist for Sample Receipt**

### 5.1.2 Verifying Records Against the Method, Procedural, or Contractual Requirements

Records are produced continually in the generation of sample data, both in the field and in the analytical laboratory. Chapter 2 lists five types of common operations that generate records which may be subject to data verification, beginning with sample collection and ending with records review. The following subsections describe the data verification process for each of these five types of operations. The first operation described, sample collection, may produce data verification records such as the records previously listed in Table 2. The four operations that may be performed at an analytical laboratory (sample receipt, sample preparation, sample analysis, and records review) produce various types of documentation, but the documentation from these steps may be compiled into what is commonly referred to as a data package.

A general hard-copy data package may include the following components: case narrative, COC documentation, summary of results for environmental samples (including quantitation limits), summary of QC results, and all associated raw data. The titles of these components might vary from one program to another or from one project to another, but the content should be similar. The following text describes these sections of a data package.

- The case narrative provides an overall summary of the verified data. The case narrative from the laboratory usually contains the signature of an authorized laboratory manager

for release of the data as well as the client's sample number, the corresponding laboratory sample number, analytical methods used for analysis, and information about holding times. A detailed description of any problems encountered with the analysis, a summary of QC samples outside of acceptance limits, and other observations that may affect sample integrity or data quality are also included in the case narrative. This overall summary should provide an immediate indication of any specific problems with the analysis.

- COC documentation may be included in a data package. Copies of the original COC forms as well as any internal laboratory tracking documents should be included to allow tracking of the sample through the entire process including sample collection, sample preparation, and sample analysis. Time and date of receipt as well as the condition of the sample may assist in checking consistency of information with other documentation.

- A summary of the results for the environmental samples is another important section of the data package. Not only are the sample results, units, and associated laboratory qualifiers usually reported in this section, but the specific information about the analysis for each individual sample may also be included here.

- A summary of QC results should also be included in the data package. This summary provides information about the QC samples that were run during the analysis of the environmental samples. Any QC samples outside of acceptance limits may be discussed here.

- The raw data may be included in the data package. The raw data will be presented in different forms depending on the type of analysis that was performed. In any case, the raw data provides the "back up" information to support the rest of the data package.

#### 5.1.2.1 Sample Collection

Samples are collected in the field in many different ways, depending upon the matrix, purpose, and analyte to be determined. Most sampling activities follow some sort of regulatory requirement including federal, state, tribal, or a combination of these. Sampling activities may be used in judicial proceedings and all records should follow appropriate guidelines. The following sequence describes typical sampling collection activities, the records generated during these efforts, and the data verification associated with the records.

A typical sampling day starts with trained and qualified team members gathering supplies for the sampling. At this time, the radiological technician, industrial hygienist, and/or site safety officer calibrates the field monitoring/field screening instruments that are needed for that day's activities. Each

instrument should be calibrated or standardized according to its own SOP. All calibrations should be recorded on an appropriate log sheet. Data verification should include review of the log sheets for calibration records. Calibration data recorded by the field staff should be compared to the criteria specified in the SOP.

Field log books or daily activity logs should be in the possession of the field team leader or designee at all times. All entries made should be signed by the person making the entries. If only one person is making entries in the log book, then that person may sign the bottom of the page. If custody is relinquished to someone else, both parties are responsible for signing the page. Usual entries may include:

- date;
- site name and location;
- weather conditions;
- team members present;
- time of field activities (i.e., the time of the tailgate safety meeting);
- sample numbers, locations, depths, and time of collection;
- sample matrix and volume of sample collected;
- name and signature of person making entries in the daily field log book;
- names of visitors to the site, their affiliation, and the time each person arrived and left;
- any deviations from established SOPs, the SAP, or the QA project plan, and the reasons for the deviations; and
- any unusual events or conditions.

Any incorrect information should be crossed out with a single line, initialed, and dated. The correct information should be added as close as possible to the incorrect information and should include a reason for the change. All information should be legible.

Sample collection should follow the approved QA Project Plan and SOPs. If not, any deviations should be documented. For example, a spade and scoop collection method would most likely be used to collect a surface soil sample. But if the soil is too hard, then a hand auger may be used. This change from one sampling method to another would be a deviation. In some cases, deviations may affect the comparability of the samples. The deviation should be noted in the daily field log book and on the sample collection log. Some sample collection logs are preprinted, so the sampler (or documenter) should draw a single line through the spade and scoop method, initial and date it, then write the method that was actually used. In the comment section of the sample collection log, the reason for the use of the alternate method should be given. The sample collection log should also include results of field screening and field monitoring. For example, if a soil sample is supposed to be screened for high explosives prior to collection, then the test should be performed and the results documented on the sample collection log. Data verification of the sample collection activities may

include an independent evaluation of the field log books to ensure the records are complete and properly signed. The data verifier should compare sample collection methods and locations to the specifications in the applicable planning documents (e.g., the QA Project Plan) to identify any deviations.

Once a sample is collected, it should be labeled and accompanied by a COC records. A label should be placed on the sample container to identify it and custody tape should be wrapped around the container lid to prevent tampering as soon as practical. The sample container and the sample collection logs are usually then placed in a cooler, which remains with the sampling team until they return to the field office. If the COC form was not completed in the field, then it should be completed when the team reaches the field office. The field team leader or sampler signs the COC when relinquishing custody of the sample to the shipping company or analytical laboratory. Data verification should include a comparison of the COC records against the field notebooks and the proposed samples specified in the planning documents against those collected. The data verifiers should confirm that any deviations are explained by entries in the field notebooks (i.e., notations regarding lack of borehole recovery or a well found damaged and unable to be sampled). Signatures on accompanying COCs should be verified, both upon release in the field and receipt in the laboratory (see Example 1).

**Example 1. Data Verification of Field Sample Collection Records**

*Emissions from the stack of a coal-fired power plant are collected to identify and measure levels of toxic air pollutants, including metals and dioxins. EPA standard methods are used for air emission sampling (i.e., EPA Method 29). Triplicate emission samples are collected from the stack in a three-day sampling period. Collected emission samples are transported to an off-site laboratory for analysis. The overall objective of the project is to conduct a comprehensive assessment of toxic emissions from two coal-fired electric utility power plants as part of an air toxics assessment of this source category. One of the project objectives is to collect a sufficient quantity of size-fractioned particulate flue gas emissions to permit evaluation of concentration of air toxic emissions as a function of particle size; as well as to collect a sufficient quantity of gas sample to establish comparable data for the particulate and vapor phases of air toxic emissions. As the data verifier begins reviewing the field notebooks and sample collection log, it is noted that there is no record of the acetone rinse sample specified in Method 29 when particulate emissions as well as gaseous metals are to be determined, as in this case. The procedure specifies that the probe nozzle, fitting and liner as well as the first half of the filter holder be brushed and rinsed with acetone, using 100 mls of solvent and collecting the rinsate as "Container 2." The data verifier includes in the verification documentation that this sample does not appear to have been collected as specified by the method.*

### 5.1.2.2 Sample Receipt

Samples are delivered to the laboratory most commonly by overnight air shipment or hand delivery. Samples may be accompanied by a COC form that is packed with the samples and delivered to the laboratory. Many types of samples are physically cooled (4 degrees C) or chemically "preserved" (e.g., addition of nitric or hydrochloric acid, sodium hydroxide, or sodium thiosulfate) to prevent or minimize degradation or other loss of the constituents of interest from the time that the sample is collected until analysis at the laboratory. The COC form will often indicate which samples have been preserved and with what preservative. Most COC forms will contain the following information at a minimum:

- sample numbers used by the field personnel for each sample;
- date and time that each sample was collected;
- client or project name and client address;
- sample matrix description;
- types of analyses requested for each sample;
- preservatives used, if any;
- number of containers for each sample;
- date and time of receipt; and
- most importantly, the signatures of all personnel who had custody of the samples.

Custody forms may also contain a section to use for comments about each sample, for example, to note the condition of the samples upon receipt, to record the temperature inside the cooler, or to document additional sample custody transfers within the laboratory (see Example 2).

Laboratories differ in the procedures used for receiving samples as well as in internal tracking mechanisms. Samples may be entered into a LIMS and/or manually into a log-in book. Project-specific planning documents may specify the sample receiving protocols or the procedures may be based upon the laboratory's SOPs. Data verification of the sample receipt information involves a review of all the pertinent records that were received with the samples as well as all the information generated by the laboratory during the receiving process.

The data verification process includes the following considerations:

*Completeness* - Are all the needed records present? Are the records filled out completely? Are the needed signatures present?

*Correctness* - Is the information in the records correct? For example, are the dates of sample collection, shipment, and receipt in the logical order? Does the count of samples match the

**Example 2. Typical Laboratory Receiving Procedures Including Evidentiary Chain of Custody**

*For projects involving regulatory compliance measurements or analyses that may be part of judicial proceedings, samples are often shipped in a manner that establishes and preserves an evidentiary COC between each successive person who handles the samples. Thus, samples may be shipped or delivered to the laboratory in a container (often a cooler) that is sealed with paper tape custody seals that break if the container is opened. The condition of the seals is checked to ensure that the container has been unopened during transfer from the field to the laboratory.*

*After the samples collected for regulatory compliance purposes are delivered to the laboratory, the person responsible for receiving them, usually known as the sample custodian, will follow the procedures established in the laboratory SOP for sample receipt. This will include inspecting the packaging and the samples to make sure the shipment is intact and not leaking. The sample custodian will note the presence and condition of custody seals on the packaging and record this information. The custodian will check the COC form for the name and signature of the sampler who relinquished the samples, and the date and time of the transfer. Samples listed on the COC will be compared to those received. If the samples arrived via an overnight delivery service, then there will be an airbill attached to the package. That airbill is removed from the package and placed in the laboratory's project files, since it provides documentation of the transfers of the package during shipping.*

*The sample custodian may check the temperature of the samples in the shipping container as needed for the specific project. Any problems will be documented and brought to the attention of the laboratory's project manager and resolved, if possible. The sample custodian will enter any necessary information on the COC form and sign and date the form as the individual receiving the samples. Internal laboratory identifiers may be assigned to each sample (if the laboratory uses this practice), and cross-referenced to the sample numbers used by the client or the samplers. The sample containers will then be stored under appropriate conditions, which may include refrigeration, freezing, or storage at ambient laboratory temperature, depending on the project specifications. The areas in which samples are stored may have written or electronic log-in sheets (e.g., refrigerator logs) that will be completed as the samples are placed in storage. Information from these steps may be recorded manually or entered into the laboratory's LIMS directly.*

number of containers received? Do the containers match what is generally needed for the analyses specified for the sample?

*Technical compliance* - Are the analytical methods referenced on the COC or analysis request the same as those given in the planning documents? Are samples properly preserved in accordance with the requested method? Were samples received in a timely manner to allow holding times to be met?

When data verification is taking place within the laboratory, the sample custodian or similar person should review the information to ensure it is factual, complete, and in compliance with established SOPs or the QA Project Plan. Errors or omissions may be identified and corrective action implemented. When data verification is done by an external data verifier, the process involves a similar review and non-compliance should be noted, although corrective action may not be possible. During this process, a checklist may be helpful, as was shown in Figure 5, with the data verifier marking the "verified" column for each record that was verified. If a record does not apply, then the data verifier should check the "not applicable" column. In addition, the data verifier should make a notation in the comment field to explain why the record did not apply (see Example 3).

**Example 3. Data Verification of Sample Receipt Records Using a Checklist**

*A data verifier is reviewing records associated with discharge samples using the example checklist shown in Figure 5. The data verifier checks the COC record and confirms all received samples were entered into the laboratory system. It is noted that the client collected the samples and hand delivered them to the laboratory; therefore, there is no shipper's airbill record. That record cannot be verified because it never existed. However, simply leaving the entry blank would not be adequate. The data verifier would check the "not applicable" column and add a note in the comment column to indicate "hand delivery."*

Despite the best efforts of everyone involved, errors and omissions will occur and will be identified during the data verification process. As with any systematic process, there should be feedback and corrective action procedures associated with such errors and omissions. However, it is critical that the data verification process address corrective actions in the appropriate context. This starts by recognizing that there are some errors and omissions that *cannot be corrected* (see Example 4).

**Example 4. Data Verification of Incomplete Record: No Corrective Action Possible**

*The sampler forgot to sign or date the COC form before it was shipped with the samples. The sample custodian discovered this error during sample receipt. The sample custodian notified the laboratory's project manager, the sampler, and the client. Those actions were recorded by the sample custodian and others, as appropriate to the situation; however, it was not possible to "correct" the missing signature after the fact. Data verification of the completed data package included a note as to the non-conformance, without corrective action possible.*

While the traditional practice of single-line cross-out, initialing, and dating the correction is an essential aspect of any correction, if the correction is made by someone other than the original producer of the record, there should be some formal notation in the records that explains the change. Data verifiers should never "assume" that they know the information and simply enter it into the records, even when they may consider the correction "obvious." The data verifier and the laboratory should never enter information into the sample records that they did not generate themselves, unless there is some form of documentation of the nature and resolution of the error (see Example 5). Equally important, there are situations where the apparent error or omission has no actual bearing on the results and therefore, need not be corrected as described in Example 6.

**Example 5. Data Verification of Incomplete Record: Documentation of Corrective Action Taken**

*Samples collected near the beginning of a new year sometimes suffer from the "obvious" mistake of having the wrong year listed in a date field. Most everyone has written a check or two in early January that is dated the year before, so it is easy to recognize the error on a COC form or in other laboratory records. However, simply changing the year to the correct entry without a formal notation of the problem could amount to falsification of the record.*

**Example 6. Data Verification of Incomplete Record: Corrective Action Not Needed**

*Using the same scenario presented in Example 4 (an unsigned COC), the sample custodian discovers the omission and calls the client, who informs the laboratory that the purpose of the analysis does not need custody to be maintained and therefore the COC form is not needed. As noted in Chapter 2, this becomes a situation where records were generated that are not needed. The sample custodian should make a formal notation that the COC form is not needed for the project. The form itself remains a part of the project records, as does the notation about the information from the client. It would not be appropriate to simply destroy the COC form after the fact. Using the example checklist, the data verifier would check the "not verified" column and add a note in the comment field that the COC form was not needed per the client.*

The least desirable outcome of the data verification process is the recognition that some records cannot be verified. The reasons will vary, but some records will simply be lost, damaged beyond recognition, etc. Often, an airbill arrives at the laboratory in such poor condition that it cannot be deciphered at all. Here again, the example checklist may be used by checking the "not verified" column and entering a note in the "comment" column.

Verifying hard-copy records is usually straightforward, based on the visual examination of the records themselves. When information is entered into a LIMS or other database directly, that

information is also subject to data verification. In designing the data verification process for a given laboratory, the first step is identification of the records that exist and that are needed. Once this has been accomplished, the laboratory staff can develop mechanisms for reviewing and verifying these records. Examples include reviewing a printout of every electronic record associated with the receipt of the samples, or developing an electronic checklist within the LIMS that displays the records in a clear format that lends itself to review. Again, the approach used in a given situation is a decision to be made in each laboratory.

External data verification (outside the laboratory) necessitates that these LIMS records be made available, usually by hard-copy printout. Since the use of external data verification is often known at the start of a project, project-specific planning documents should specify the availability of these LIMS records as part of the laboratory data package. When the need is not projected and the records are not available in the hard-copy data package, not all records of the sample receipt process may be verified. The impact of this would be assessed during the project's data validation phase.

The final step in data verification of the sample receipt records is to sign and date any records that data verification produced. The data verifier's name, signature, and date should be recorded at the end of the data verification.

#### 5.1.2.3 Sample Preparation

Following the sample collection field activities and after the samples are received at the laboratory, sample preparation for analysis begins. The process of preparing environmental samples for analysis includes a wide variety of procedures within the laboratory. The following discussion centers upon those procedures having a distinct preparation step, separate from actual analysis: e.g., the solvent extraction of a water sample prior to analysis for polychlorinated biphenyls or the acid digestion of a soil sample for metals analysis. In general, the following types of procedures may be employed during the preparation of samples for typical analyses:

- homogenizing the sample;
- removing a subsample (aliquot) from the original sample container or transferring the entire contents to another container, recording weight or volume;
- adjusting sample pH (generally only for aqueous samples);
- preparing new culture media;
- adding drying agents or other amendments to a solid sample prior to extraction;
- spiking surrogates, internal standards, or other analytes into sample aliquots;
- adding extraction solvents or digestion reagents to samples prior to extraction or digestion;
- separating the extract or digestate from the bulk sample by decanting, filtration, or other techniques;

- incubating pour plates at specified temperature and specified duration;
- sample clean-up by column chromatography, solid phase extraction, or other technique;
- drying or purifying a solvent extract;
- concentrating the extract or digestate to a smaller volume; and
- preparing the extract or digestate for storage prior to analysis.

The particulars will depend on the analyses to be conducted, the specific methods used, and the nature of the samples themselves.

Records are generated as a result of applying the procedures above. These records are typically in the form of "bench notes" from the chemist or technician performing the procedures. Such notes may be recorded in bound laboratory notebooks, on preprinted forms, or electronically in a LIMS. How these notes are recorded should be defined in the laboratory's QA manual, SOPs, or equivalent document. The documentation may be supplemented by other records including log-in sheets from refrigerators, internal COC or tracking forms, records of the preparation of standards and spiking solutions, etc. In addition to bench notes that describe the procedures used, there are a number of critical steps that may be performed by one staff member and witnessed by a second staff member in order to ensure that they were performed for each sample (see Example 7).

**Example 7. Data Verification of Process by a Witness**

*The spiking of surrogates or internal standards into samples prior to extraction or digestion is performed in some cases. Because the spiking process typically yields no visible change in the sample aliquot being spiked, the second person acts as an observer or witness to the spiking procedure. That witness will then record and verify the fact that the spiking was performed by the first person.*

The first step in data verification for sample preparation is to identify the project needs for records. Once those records are identified, they are verified in much the same way as the sample receipt records. The data verifier is someone other than the record producer. The records will be checked for completeness, consistency, and correctness against the project needs (see Example 8).

**Example 8. Consistency in Records**

*The records for preparation of 15 of 16 samples indicate that a 1000-milliliter aliquot was extracted. However, the record for the 16$^{th}$ sample lists the "volume" as "1000 gram." The data verifier needs to determine whether this discrepancy is real. There may be a problem with the units (gram versus milliliter), the entry may have been placed in the wrong field for the sample, or the final sample may have actually been weighed given certain circumstances.*

A data verification checklist for sample preparation might address questions about the following aspects of the sample preparation procedures:

- Is the sample identifier present?
- Is the amount (weight or volume) used in a preparation entered?
- Is the analyst's name and signature present?
- Are dates and times for all noted steps present?
- Is the method or SOP identified?
- Are initial weights/volumes and final weight/volumes for weighing or concentration steps listed?
- Is pH recorded as needed?
- Are QC samples identified?
- Are balance logs, refrigerator logs, etc., present?
- Can standards and spiking solutions be traced to their stocks and certificates of analysis?
- Are the additions of spikes recorded and witnessed?

The possible results of data verification for sample preparation records are similar to those described for sample receipt records. The records may be verified, verified with corrections, not verified, or not applicable. The latter three possibilities should generate some notation or comment on the data verification documentation.

Verifying electronic-only records is important for sample preparation, since much of the equipment used in sample preparation can be connected directly to a LIMS, thereby leaving fewer hard-copy records to be reviewed. In addition, the LIMS may also perform preprogrammed calculations using these data during the sample preparation stage (see Example 9).

**Example 9. Electronic, Preprogrammed Calculations**

*The determination of the dry-weight fraction of a solid sample or the solids content of an aqueous sample involves oven drying a subsample or a filter for a predetermined period or until constant weight is achieved. If the scale used to weigh the subsample or filter is connected to the LIMS, then the LIMS may perform the calculation automatically and the analyst will only see the final result displayed.*

The algorithms used by the LIMS to perform calculations need data verification as well. Since the purpose of automating data collection and calculation activities is to simplify and speed up the process, it is not realistic to expect that every automatic calculation be checked by hand for every sample, or even that a small percentage of such calculations be verified for each sample. Rather, the laboratory should verify *all* the calculations at some frequency (ideally, before implementation, and at

least annually thereafter for selected complete projects as defined in the laboratory's SOPs) and whenever new calculations are programmed into the LIMS. The frequency and manner of data verification should be defined in the laboratory's SOPs. External data verification of a completed data package would need hard-copy printouts of all raw data to allow calculations to be verified, at a project-specific frequency, by tracing all reported concentrations back to the original unit of measure including all preparation steps.

The same considerations about correcting records that were described in Section 5.1.2.2 apply to the sample preparation records. Indeed, there are even more possibilities for errors and omissions during preparation than during receipt (see Example 10).

**Example 10. Data Verification of Incomplete Record, Narrative Assessment Using Analysis Results**

*The record of spiking a given sample with surrogates may be blank. There may be no way to confirm or deny that the sample in question was ever spiked. Therefore, corrective action regarding the documentation of the spiking procedure may not be possible. However, when the analyses themselves are complete, it may become immediately apparent that the surrogates were present in the sample. Therefore, the sample results may be acceptable for their intended use, even without the ability to verify the record of the spiking. In that situation, it would not be appropriate to go back and "correct" the spiking records to show that the sample was spiked. Rather, the fact that the spiking records could not be verified might be noted in a report to the client, with an explanation of the logic used to demonstrate that the results indicate that the spike had been added.*

The final step in verifying the sample preparation records is the signing and dating of the data verification records themselves.

#### 5.1.2.4 Sample Analysis

Data verification associated with sample analysis varies based on the measurement to be made, the sample matrix, the project-specific QC measurements associated with the samples, and the purpose of the analysis (Example 11). Whether this data verification is performed in the laboratory by a lead chemist reviewing the work of the bench analyst or by an external data verifier reviewing the submitted data package, the process includes verification of the completeness, correctness, and technical compliance of the records and documentation associated with the analysis. The instrumental analysis procedures are among the most thoroughly automated aspects of the entire analytical process. The majority of the analyses of metals and organic chemicals are performed on instruments that can utilize autosamplers and similar devices designed around the "turn key" principle. In many cases, the analyst

**Example 11. Data Verification of Sample Analysis Records Using a Graded Approach**

*Consider the following records that may need data verification, using a GC/MS analysis for SVOCs including pesticides, in soil samples collected as part of a final remediation and site closure sampling effort. The following items might be checked by the data verifier in order to ensure the analytical process met technical compliance criteria and that documentation for the analysis is complete and factually correct:*

- *decafluoro-triphenyl-phosphine tuning results, summary and raw data;*
- *initial calibration or calibration verification response factors for critical analytes and review of sample chromatograms;*
- *dichloro-diphenyl-trichloroethane and endrin breakdown;*
- *method blank analyses, chromatograms, and spectra;*
- *internal standard areas and retention times;*
- *detector saturation;*
- *sample holding times;*
- *surrogate recovery compared to control limits;*
- *sample chromatograms and compound spectra;*
- *calculation of final concentration of positively identified compounds, including dry weight versus wet weight reporting as per project needs;*
- *verification of laboratory assigned data "flags;" and*
- *results of sample duplicate (field and/or laboratory) analysis or spiked sample analysis compared to laboratory control limits.*

*Slightly changing this sample set to encompass only PAH analyses in soil samples collected as part of a preliminary site characterization may significantly change the records and documentation needed to be verified. In this example, the records needing data verification may be reduced to:*

- *decafluoro-triphenyl-phosphine tuning results,*
- *tabulated summary of calibration results,*
- *summary of method blank results,*
- *sample surrogate recoveries,*
- *review of flagged data, and*
- *summary of sample duplicate and spiked results.*

*In either example, items on this list will be checked individually for correctness and collectively for completeness.*

simply sets up the samples and standards in the autosampler, pushes the "start" button, and evaluates the results of the QC samples after completion of the analysis batch. Many analytical results are maintained electronically in an instrument's data system and then transferred to the LIMS. There are also a number of other analyses used to assess various environmental samples, including spectrophotometric procedures that use an ultraviolet-visible spectrophotometer, titrations that depend on the visual differentiation of a color change endpoint, parameter-specific electrode methods, and gravimetric methods. Any or all of these analyses may be critical within a specific project to the overall assessment of environmental conditions or to delineate contamination at a site.

The first step in data verification is again to identify all the project needs for records that are produced during the actual analysis procedures. Of particular importance will be the records with the results of QC analyses and criteria associated with the analytical parameter, including calibration standards, method blanks, duplicate samples, spiked samples, spiked blanks, interference check standards, etc. Not all of these items will be needed for all analyses, nor will every project need that all of these be reported. Therefore, data verification of sample analysis results will be parameter specific and project specific. Techniques and analytical reference methods will have specific QC needs and there may be additional needs in the QA Project Plan, contract, or relevant rule or standard (e.g., the National Environmental Laboratory Accreditation Conference standard).

The comparison of the results of QC analyses against method needs may be largely automated by the instrumentation, in which case there may be records only when the instrument data system notes a problem and warns the analyst. Data verification of these and all method specified QC analyses should include confirmation that these analyses were indeed performed and that the results were technically compliant. The report should clearly identify any QC analysis that does not meet method criteria or project-specific specifications. If data are not available to perform this verification, the data verification records should state that these analyses and/or QC specifications could not be verified. The impact of not being able to verify QC analyses or specifications should be assessed in the data verification records and evaluated further by the data validation process and by the data user.

Where the dates and/or times of sample processing and analysis steps are reported, it is critical to verify that these dates and times match those recorded on the raw data or in the bench notes. Raw data from instruments such as inductively coupled plasma may include the date and time when raw instrument data were processed (sometimes shown as the "quantitation date" on the printout), which are not the same as the date and time of the analysis. The data verifier needs to make certain that the correct dates are provided in the record and that they match those reported elsewhere.

Other laboratory areas need evaluation when verifying sample analysis results. Records associated with an automated system include records produced by the analyst that are therefore subject to data verification. For example, in setting up standards and samples in an autosampler, the analyst has to make some record, either in hard copy or electronically, of the order of vials in the autosampler.

The exception would be in an automated laboratory in which each vial is labeled with a bar code and the autosampler is equipped with a bar code reader. Data verification of the sample analysis run logs would still be needed, usually by including a hard-copy printout of the electronic file and cross-referencing sample identifications. If needed, information should be available to either a data verifier within the laboratory setting, or the external data verifier, that allows the tracing of sample results to the original analytical result. That original analytical result may be an instrument response (e.g., absorbance), a titration volume, or a sample weight. Another possible area requiring special attention during data verification of sample analysis results is quantitation performed by instrumentation.

As discussed in Chapter 4, manual integration is one of the most commonly abused aspects of GC/MS analyses. Instances of falsification have begun with manipulations of the peak areas, often with practices known as "peak shaving" or " peak juicing" where integration points are moved to decrease (shaving) or increase (juicing) peak area to meet specification. Thus, it is *critical* that the laboratory have written procedures that describe how and when the analyst should perform manual integrations. These written procedures should also describe how to note in the laboratory records and data that manual integrations were performed. GC/MS data systems have the ability to "flag" the electronic and hard-copy records of manual integrations. Therefore, the data verifier should review procedures, records, and any bench notes from the analyst to make sure that when the electronic records indicate that a manual integration was performed, it was done in accordance with the laboratory's stated procedures, and that it is clearly evident to the data user. This is illustrated in Example 12.

A final example, usually applied to metals and organics results, is data verification of laboratory-assigned data qualifiers. The bench analyst is often responsible for assigning any laboratory qualifiers or

**Example 12. Data Verification of Chromatography Peak Integrations**

*A potentially crucial aspect of the data verification process for instrumental analysis (e.g., GC, GC/MS, high-pressure liquid chromatography) may be the review of peak integrations that were performed by the instrument software. Problems with low signal strength and interferences can cause automated algorithms in the software to integrate peaks in a manner that is less than ideal. These problems can cause some samples to fail to meet specifications, particularly for internal standard areas and surrogate recoveries. Thus, a legitimate aspect of the post-acquisition review of the results is to check on these peak integrations. Where the analyst can identify a specific and previously defined problem, documented in the laboratories written procedures, the peak in question may be manually integrated on the basis of the analyst's judgment of the proper integration points. The analyst should document in the raw data what was done and why it was done. In some cases, confirmation that peak integrations were conducted appropriately would necessitate a complete, i.e. 100 percent, review of all the original raw data associated with the analysis and an on-site laboratory audit.*

"flags" to the data to identify potential data quality problems for the data user. Some laboratory qualifiers may be applied by the instrument data system itself, based on preprogrammed rules. The bench analyst may review those qualifiers and overrule the data system, in which case there should be a record explaining why the qualifier was removed. When the bench analyst applies a qualifier, there should be some record explaining why the qualifier was applied. While there are several commonly-used sets of data qualifiers, there is no universal set that applies to all types of analyses, nor a universal specification for their use (Appendix C). If flags are being used, the data verifier should determine if their application was defined clearly in the data report, and whether the flags were appropriately assigned to sample results based on these definitions.

The data verifier may use a checklist or other means to record the results of the data verification process. Once the data verification is complete, the data verification records themselves are signed and dated, as discussed for the other aspects of sample analysis.

#### 5.1.2.5 Data Verification Records Review

The format and content of the data verification records sent to the client are as varied as the types of analyses performed and the end uses of the data. Data verification records may range from a one-page letter to a report that is several inches thick. The contract, the QA Project Plan, or the SAP may provide information about the content and format of the specified documentation. Thus, the data verification process may rely heavily on the specifications in those documents. All data verification records should be reviewed before they are delivered to the client.

The data verification records produced during field activities usually consists of documents and records such as the ones described in Section 5.1.2.1. These documents, including sample collection logs, field screening results, and daily activity logs, should be reviewed by field personnel to ensure that all information was recorded in accordance with the appropriate procedures. Any deviations may be noted either in the standard field documentation or provided in a separate summary that is included as part of the data verification records.

The laboratory usually produces a data package that includes the documentation from sample receipt to sample analysis. This documentation was described in Sections 5.1.2.2 to 5.1.2.4. Laboratory personnel should review the data package to ensure that all information, including any deviations, was recorded appropriately.

Data verification within the laboratory will make extensive use of any internal verification results already generated. Data verification by an external data verifier of the completed laboratory data package will be performed as discussed in the previous three steps from sample receipt to sample analysis.

Many laboratories use a formal checklist to guide the assembly of the data package. The most comprehensive records will be those where the laboratory has been instructed to include copies of all of the raw data, including the bench notes, internal tracking forms, and COC records. The data verification records should ensure that the sample identifiers used by the client can be clearly associated with any internal sample identifiers assigned by the laboratory. This is particularly important on the printouts of the raw data from the instrumentation, since it often displays only the internal laboratory identifier.

The samples sent to and analyzed by the laboratory are associated with a variety of QC samples. This could include various blanks (field and laboratory), spiked samples, laboratory control samples, etc. The associations between the field samples and the QC samples vary widely. In addition to the previous data verification steps that evaluated technical compliance, the records review should ensure that the QC samples can be associated with the field samples.

By the time data verification records have been assembled and reviewed, it is often too late for any corrective action of technical problems with the data. The records should reflect what was done and describe any corrective actions that may have been applied to the sample analysis results. The data verification records should demonstrate the chain of events involved in the analysis of a given sample and describe what was done, how it was done, and whether what was done fulfilled the project needs.

## 5.2 DATA VALIDATION TOOLS AND TECHNIQUES

Chapter 3 introduced the inputs to data validation including verified data, data verification records, and associated records such as a data package or field records. The following sections describe the step-by-step process that a data validator may follow. It is important to note that not all steps may be needed for a particular project.

### 5.2.1 Tools and Techniques for Data Validation of Field Activities

The data validator should have access to a complete set of verified data and data verification records, including field records. The typical field records identified in Chapters 2 and 3 are described in more detail in Table 5. Not all records are needed for every field sampling campaign, nor are they all called by the same name in all projects. Using the description of each record described in Table 5, the data validator can determine if the field records that are being reviewed contain a similar document. Table 5 summarizes common records that may be generated by the field team. There may also be records generated that are not usually available to the data validator. Examples of these include sample labels and field notebooks although this information may be available if necessary.

The five steps outlined in Section 3.3.1 are presented here in more detail. These steps lead the data validator through a logical sequence to review the field records.

**Step 1. Evaluate the field records for consistency**. The first thing that the data validator should check in the field records is the consistency of the recorded information. Similar information may be recorded on multiple forms and may provide a means for consistency checks. Consistency may be reviewed by comparing the same field of information from different records or it may involve checking the agreement between different fields that are expected to be related. For example, the time that each sample was collected should be consistent in all the records generated in the field. The time that a sample was taken may be recorded in records such as the field notebook, the sample collection log, and the COC. The data validator should review the field records that are available and the contents of each document in order to determine which information may be needed to perform a consistency check. The suggestions in Table 6 give examples of how to start looking for consistency within the field records. Any inconsistencies found in the field records should be compared to the verified data and the data verification records for further explanation.

**Table 5. Examples of Types of Field Records, Purpose of Each, and the Recorded Information**

| Document Type | Document Purpose | Summary of Document Information |
|---|---|---|
| Sample location survey | Records all sample locations so they can be accurately plotted on a map | Should indicate that sample locations are based on either global positioning system or a fixed marker. Survey information may be used for a computer-generated map. |
| Instrument calibration records | Maintains accurate record of instrument calibration | May include instrument name, model number, date and time of calibration, and calibration results. |
| Field notebook/ daily activity log | Maintains accurate record of field activities by providing written notes of all activities | Information may include personnel in the field, weather conditions, health and safety briefing, location and name of job, zone set-up, time of sample collection and sample descriptions, visitors to the site including arrival time and departure time, any unusual occurrences or events, field instrument surveys, decontamination procedures, any sampling deviations, etc. Each page is signed by the person making the entry. |

**Table 5. Examples of Types of Field Records, Purpose of Each, and the Recorded Information**

| Document Type | Document Purpose | Summary of Document Information |
|---|---|---|
| Sample collection logs | Maintains accurate record of samples collected | Information may include sample number, date/time of sample collection, sample type/description, sampler identification, collection method, sample location, depth of the sample, QC type, compositing details, sample matrix, analyses requested, bottle type and volume for each requested analyses, preservation method, the COC number, any field measurements, photo number, etc. |
| Photo logs | Maintains accurate sampling activities photo record | Photo number and what sample or activity it corresponds to, the date, and the direction of the picture. |
| Driller's/ heavy equipment operator's daily activity log | Maintains accurate record of field activities with emphasis on drilling or heavy equipment operation | Maintained by the driller, may include drill rig type, type of drilling (air rotary, split spoon, etc.), sample location, depth, problems encountered, material drilled, down time, the names of the driller/driller's assistants, the angle of the drill hole, etc. Heavy equipment operator's log may include type of equipment, the name of the operator, procedures used, etc. |
| Field monitoring results | Maintains record of potential contaminant hazards to the field team | Field monitoring results should include date, type of field instrument, and monitoring results, as well as the type of personal protective equipment worn by the field team. |
| Field screening results | May support characterization or clean-up of a site | Field screening results should include date, location, type of field instrument, and screening results with any QC information that is available. |
| Chain-of-custody | Maintains proof that samples were not tampered with and that samples were under appropriate possession at all times | Includes COC number, sample collection information (sample ID, collection date and time, preservative, matrix, etc.), analysis request (method reference, QC requested, etc.), and signatures of persons relinquishing and receiving samples to document custody transfer. |

**Table 6. Examples of Items to Review for Consistency Checks for the Same Type of Information**

| Type of Information | Documentation to Check | Reason for Checking Consistency |
|---|---|---|
| Sample matrix | Sample collection log<br>Photo log<br>COC | To check the type (soil, water, sediment) of material that was sampled |
| Sample number | Sample collection log<br>COC | To check the list of sample numbers |
| Location identification | Sample collection log<br>COC<br>Field notebook | To check the list of location identifications |
| Date and time of sample collection | Sample collection log<br>COC<br>Field notebook | To check the date and time of sample collection |
| Depth of sample | Sample collection log<br>Field notebook<br>Driller's log | To review sample depths and consistency of units for each depth |
| Sampling method | Field notebook<br>Sample collection log<br>Photo log (if available) | To check that the intended sampling method was used and that it was used appropriately |

Similarly, the data validator should also review the field records to ensure that there is agreement between different fields that are expected to be related. In order to review consistency between fields, the data validator should be knowledgeable about field activities in order to identify appropriate fields to review. Table 7 gives examples to consider when reviewing the field records.

**Step 2. Review quality control information**. The planning documents should indicate any type of field measurements that were part of field implementation. If field screening or field monitoring was performed during the course of the field activities, the data validator should review the QC documentation recorded for this work as well as any other data verification records. All instruments used in the field should be calibrated. Instruments and techniques that may be used include photo ionization detectors, radiation detectors (alpha, beta/gamma, and sodium iodide detectors), x-ray fluorescence instruments, pH meters, immunoassay techniques, and Laser Induced Breakdown Spectroscopy instruments. The data validator should ensure that calibrations were performed at the necessary intervals and the instruments were calibrated correctly. If background readings were

recorded for a field instrument, the data validator should also review the data recorded for the background readings and check any calculations that were done to determine site background values. The data validator may also review data from any samples collected for field quality control, such as trip blanks (see Example 13).

**Table 7. Examples of Items to Review for Consistency Checks Between Types of Information**

| Type of Information for Comparison | Examples of Questions to Check Consistency |
|---|---|
| Sample collection method<br>SOP used for sample collection<br>Depth of sample | Do the sample collection method and the SOP that is referenced agree? Is the depth of the sample appropriate for this type of sampling? |
| Sample location<br>Local area information such as buildings, utilities, or roads<br>Sample collection method | Were samples collected in an area that may have needed special sampling (e.g., angled borehole)? Do the sampling locations appear to be in the correct area based on where sampling was supposed to occur? |

**Example 13. Qualification of Sample Data Based on Field Blanks**

*Sampling activities during the base realignment and closure program included the preparation of field QC samples, concurrent with collection of the environmental samples. The field QC samples included:*

- *field blanks, designed to determine if samples were contaminated by ambient conditions in the field such as wind-blown dust;*
- *rinsate blanks, designed to determine if samples were contaminated by improperly decontaminated sampling equipment; and*
- *field blanks, designed to determine if empty sample containers were contaminated during transport to the field, or if samples were contaminated during shipment from the field to the laboratory.*

*Both the environmental samples and field QC samples were analyzed for an extensive suite of target analytes, including metals, pesticides, herbicides, VOCs and SVOCs, and analyses intended to detect spilled fuel (gasoline range organics and diesel range organics). Samples were analyzed for gasoline range organics and diesel range organics because the base facilities included underground storage tanks, as well as vehicle fueling and maintenance facilities, all potential sources of environmental contamination with petroleum products.*

**Example 13. Qualification of Sample Data Based on Field Blanks (continued)**

*The results of the groundwater analyses showed well-defined plumes of petroleum contamination in areas that could logically be attributed to an intact underground storage tank documented to have leaked in the past, as well as from a site where underground storage tanks had previously been removed. However, other samples unassociated with these plumes also showed contamination with petroleum compounds, primarily widely varying concentrations of gasoline range organics reported. The contamination in these samples appeared random, forming no directional or concentration pattern. Review of the associated field QC data showed similar contamination in the field blanks collected with many, although not all, of those samples. It was apparent that the sample results were the result of environmental contamination, but the question remained, "How had the contamination occurred?" The data validator's review of the field logs showed that all of the samples were collected by the same field crew. When the field crew was interviewed, it was determined that it was cold when the samples were collected (verified with the ambient temperature notation in the field log), that the field crew had kept their vehicle running while they were collecting samples so that they could warm themselves in the truck cab, and that is was their practice to keep the sample coolers in the bed of the truck near the exhaust. It was determined that the field blanks were probably contaminated from the truck exhaust, and that it was likely that the sample were similarly contaminated. All gasoline range organics collected by this crew were disregarded based upon the suspected source of the contamination. Furthermore, given the nature of the analysis and the inability to fully delineate the source of the contamination (e.g., exhaust sample could not be collected from comparison) it was recommended that diesel range organics results from these samples be used with caution by the client, and that this qualification be considered when planning subsequent sampling activities.*

**Step 3. Summarize deviations and determine impact on data quality.** In some cases, it may not have been possible to carry out all elements of the field activities according to the original specifications in the sampling plan. The data verification records should include a summary of deviations encountered during sampling activities. Depending on the data validator's familiarity with the sampling plan, the data validator may also identify additional deviations from the original plan based on the review of all of the field records. In the data validator's summary of the deviations, the reason for each deviation should be discussed if it is clear from the field records. Deviations may include changes in sample locations, changes in samples collected, changes in the sample analyses, change in length of time for field activities to occur, or any unusual readings from the field instruments that resulted in either additional sampling or fewer samples. As the data validator reviews the deviations, their effect on the overall quality of the data should also be considered. Examples 14 and 15 illustrate how deviations can have a significant impact on an overall project and any deviation from the original plan should be documented.

**Example 14. Impact of Sample Collection Method on Data Quality**

*The SAP for site characterization soil sampling specified the use of a shovel for sample collection. A hand auger was used instead and this deviation was recorded. In this case, the impact on data quality is probably minimal or has no impact.*

*However, if the SAP for site characterization soil sampling specified cone and quartering homogenization and this was not performed, the effect on the overall data quality may be quite significant.*

**Example 15. Evaluating Documentation of Field Sample Matrix**

*An industrial waste stream is chemically characterized based on known engineering design specifications and historical data provided by the facility in order to make a hazardous waste listing determination. Since each listing determination includes extensive investigations, requiring literature and database searches, industry surveys, engineering site visits, sampling and analysis, and risk assessment, it is critical that any deviations from the project planning documents be carefully recorded and documented. In particular, deviations in the anticipated sample waste streams listed in the SAP should be recorded by the sampling personnel and the receiving laboratory.*

*For this example, a sample identified as a wastewater in the SAP only needs laboratory analyses for an aqueous matrix, i.e., total solids and leachate analyses are not specified. However, at the time of sample collection it is noted the wastewater actually contains 30% solids. Upon sample receipt, the laboratory sample custodian also notes that the sample composition is more representative of a sludge matrix rather than the expected wastewater.*

*During the data validation of the laboratory data package, the data validator investigates duplicate results that were non-compliant. Reviewing sample identification records, including laboratory log-in records, uncovers the notation on the sludge-like matrix. Obtaining and reviewing the sample collection log confirms the unexpected matrix. The data validator then reviews the planning documents again and ascertains that analyzing the sample as though it is a wastewater matrix may not have been in keeping with the intent of the SAP.*

*The data validation report includes a note that the project team should decide whether to modify the SAP and whether it would be more appropriate to treat the sample in question as a solid matrix in order to measure the mobility of constituents entrained in the solid particles using conventional leaching methodology.*

**Step 4. Summarize samples collected.** After reviewing the verified data and the field records, the data validator should summarize the sample data and field data that were collected during field activities. The sample data for each individual sample may include information such as sample identification numbers, date and time of collection, sample location, depth of sample, sample matrix, and any duplicate or split sample information. The field data may include measurements such as pH, conductivity, or field immunoassay. Based on the data verification of the field data, these data may be qualified as necessary based on associated field QC samples such as immunoassay control solutions or pH check standards. If the field information was provided in the electronic format, the data validator may easily summarize all of these data. If an electronic version of the field data is not available, the data validator should choose the most important information about each sample and include this information in a summary table for the data validation report. Similarly, the data validator should also summarize any of the field data that are relevant for making project decisions according to the planning documents.

**Step 5. Prepare field data validation report**. The data validator should document the information from each step as outlined above. The content and format of the data validation report will depend on the project, and may be specified in one of the planning documents. The data validator may check some of the same information that the field personnel verified during their work, and should consider the field information in the context of the overall project needs. For example, the deviations from the sampling plan and the reasons the deviations occurred should have been included in the data verification records, and the impact discussed in the field data validation report. If field screening methodology was used for a particular project, the data validator should include a review of the field screening results in the data validation report. Any QC data that were produced with the field screening results should be presented with a discussion of the confidence in the field screening data to assist in making project decisions in the field. The data validation report should provide the data user with an overall picture of the quality of the field data and how well it supports the project needs that were initially defined.

### 5.2.2 Tools and Techniques for Data Validation of Analytical Laboratory Data

In order to understand the needs for data validation of analytical data, the data validator should have a list of the applicable specifications from the planning documents. The data validator uses all data verification records, including the verified data, to perform the steps outlined in Section 3.3.2. These steps, which are presented in more detail below, lead the data validator through a logical sequence to review the analytical laboratory data. (Each project will have a unique set of needs, and all steps identified below may not be applicable to all projects.)

**Step 1. Assemble planning documents and data to be validated. Review summary of data verification to determine method, procedural, and contractual required QC compliance/non-compliance.** As the data validator begins the data validation process, a complete set of records from the laboratory analysis should be available. The data validator should also have the planning documents available in order to ensure that the data verification records and verified data are complete. Based on

the planning document and the results provided in the data verification records, the methods, analytical results, and QC results can be examined to attempt to determine why certain non-compliances were encountered, as illustrated in Example 16.

**Example 16. Using Analytical and QC Results in Data Validation**

*A data validator is evaluating the results of an analysis of volatile organic compounds by GC/MS. d5-chlorobenzene is used as an internal standard for this analysis. The quantification ion for d5-chlorobenzene is 117 m/z. There are a number of alkyl benzene compounds, commonly found in gasoline, that have fragments ions equal to 117. It is quite possible that a sample that contains very high levels of these gasoline components would result in an internal standard recovery for d5-chlorobenzene that exceeded the limits (greater than 200%). The data validator inspects the chromatograms, and finds that target plus tentatively identified compounds point to an interfering (co-eluting) peak with the internal standard. If the internal standard is biased high, the result would be to underestimate the concentration of target analytes that use that internal standard for quantification. Since absolute confirmation of an interference necessitates inspection of the ion chromatogram for peak shape and retention time irregularities, the data validator contacts the client to obtain access to this information.*

The QC data should be compared to any specifications in the planning documents including the type and frequency of each QC sample (Example 17). Whenever possible, a determination should be made as to the cause of the non-conformance. QC data may include results from analysis of blanks, matrix spikes, laboratory duplicates, laboratory control samples, etc.

**Example 17. Using the Data Verification Documentation in Initiating Data Validation**

*To initiate data validation efforts for ambient air samples collected via Method TO-4A on polyurethane foam filters, the data validator reviews the data verification documentation and project planning documents for the data report. It was noted in the data verification documentation that the lab blank contained trace polychlorinated biphenyl levels. The data validator then looks through the assembled data for the certification results of the polyurethane foam cartridge assembly analyses for the batch used in the field and notes that the data are missing. Contacting the laboratory indicates that this check was not performed. As the data validation process continues, the data validator should ascertain the impact of this on project objectives, taking into consideration blank levels, sample concentrations, and end-use of the data.*

**Step 2. Review verified, reported sample results collectively for the data set as a whole, including laboratory qualifiers**. The data validator can confirm that the reported sample results make

sense by checking the calculations that were used. Inputs to the calculation, such as dilution factors, should be checked for accuracy as well. In some cases, data reduction may be performed by an instrument or a computer at the laboratory. If there is concern about the data reduction performed by the laboratory, the data validator may have to request further information from the laboratory in order to validate the data to the detail of each calculation.

Some projects specify that the laboratory add qualifiers to the data to serve as an indication of the quality of the data. The laboratory should provide a list of the laboratory qualifiers that were used and a definition for each one. This information will assist the data validator in determining the data validation qualifiers that may be assigned to the data during the data validation process. The definition and use of these laboratory qualifiers should be checked for consistency and correctness in the data package. If data are reported in an electronic format, sample results and laboratory qualifiers, if assigned, would most likely be fields included both in the electronic data and in the data package. This is illustrated in Example 18.

**Example 18. Impact of Method Blank Contamination**

*A set of samples was analyzed for SVOCs. A majority of the sample results for bis(2-ethyl hexyl)phthalate were qualified "B" indicating that bis(2-ethyl hexyl)phthalate was detected in the method blank. The data validator would consider not only the concentrations of bis(2-ethyl hexyl)phthalate in the method blank and the samples, but also whether or not the bis(2-ethyl hexyl)phthalate may have been a contaminant of concern for the particular project. By putting this information into the context of the project, the data validator can make recommendations about the quality of the data for the intended use in the project.*

**Step 3. Summarize data and QC deficiencies and evaluate the impact on overall data quality**. In some cases, the verified data may not meet the needs that were stated in the planning documents. The data validator may discover the non-compliance during data validation or it may have been noted and documented during the data verification process. The reasons for any deficiency encountered may vary, and one of the goals of the data validation process is to try to determine the reason for the non-compliance, and to evaluate the impact of the deficiency on the overall quality of the data set. QC deficiencies may include a particular type QC sample that should have been run but was not, low matrix spikes, or laboratory control samples that were not within laboratory control limits. Any QC deficiency may bring particular sample results into question. The data validator should consider the deficiency and make a determination as to whether a particular analytical batch is adversely affected, whether the non-conformance indicates a widespread bias in the analysis that affects all samples, or whether the deficiency has no significant impact on data quality and the sample results can be used as reported. As noted earlier, the purpose of the sampling and analysis effort should be taken into account during the data validation process in order to understand the end-use of the data. Discussions with the project

manager or lead technical person may also clarify the intended end-use of the data. This is illustrated in Example 19.

**Example 19. Impact of Holding Time Non-Compliances**

*Regulatory holding times are different for different sample parameters. Non-compliance of the holding time for a sample or sample set may be a result of laboratory oversight, delayed sample shipment, need for reanalysis, or poor planning. The data validator should evaluate the impact of the non-compliance, taking into account the nature of the analysis (Was it a critical parameter in the determination of project objectives?), the extent of the non-compliance (Was holding time missed by 1 day or 1 week? Is the regulatory limit 48 hours or 40 days?), the sample matrix, any supporting data (Was there a diluted analysis performed within holding times?), and the purpose and goals of the sampling and analysis program. Consider the following comparisons. Samples for nitrite have a holding time of 48 hours. Extracted samples for SVOC analysis should be analyzed within 40 days. A holding time violation of two days for a nitrite sample will have a bigger impact on data quality than the same two day lapse for SVOCs, based upon the differences in the regulatory limit as well as the nature and stability of the parameter. On the other hand, a two-day holding time violation for SVOC analysis of samples collected as part of an industrial discharge permit litigation effort may result in rejecting the affected samples. The same two-day holding time violation for SVOC analysis of samples collected for a preliminary site characterization effort may simply indicate that the non-compliance be noted, without limiting the use of the data.*

In some cases, data validation may mean determining whether an analysis met the intended method or technical specifications even if there was no obvious non-conformance or deficiency. Spiked samples, for example, are analyzed for many parameters to provide an indication of method accuracy for the matrix in question. Recovery results are expected to be within method control limits, laboratory-derived statistical limits, or limits established in the planning document (e.g., the QA Project Plan) for spiked samples. The matrix spike and matrix spike duplicate results provide an indication of possible matrix interferences that may impact the analysis. Surrogate standards added to each standard, blank, and sample are analyzed to assess analytical performance for each analysis. Surrogate recovery results may also indicate matrix effects. The procedures used for spiking, the samples selected for spiking (e.g., a field sample or a trip blank), and the levels spiked should all be considered (see Example 20).

Samples are often collected in duplicate during the field sampling effort and sent to the laboratory as "blind" duplicates. For soil samples in particular, the data validator can examine the analytical results of these samples to evaluate the combined variability of both sampling and analytical techniques (see Example 21).

**Example 20. Evaluating if Data Meet Intended QC Needs**

*A set of samples is analyzed for metals by inductively coupled plasma atomic emissions spectroscopy. The data verification noted that the interference check sample was outside of the control limits and all samples were reanalyzed at a dilution due to high aluminum levels. All matrix spike and matrix spike duplicate results and laboratory control samples had spike recoveries within limits. Upon closer examination, it was noted that all matrix spike and matrix spike duplicate results were performed on the samples that were diluted 10-fold. Thus, although all spiked recovery results were compliant, the data validator reports that the way the spikes were performed precludes an evaluation of accuracy for any samples analyzed without dilution. The interference check sample indicates the potential for interference; the true impact on sample results is unknown. Therefore, any undiluted sample analysis results may be qualified as estimated.*

**Example 21. Inherent Field and Analytical Variability of Field Duplicates**

*Soil samples are collected in an effort to determine the baseline contamination at a Brownfields site being considered for an evaluation of an emerging, in-situ remediation technology. Results for field duplicate samples fall outside of the precision objectives established in the SAP. During the data validation process, it was noted that the non-conformances associated with the field duplicates appeared to be concentrated in samples collected from a specific area of the site. The data validator, therefore, may look into differences in that area of the site (e.g., known dump site? different particle size distribution indicating more rock, pebbles? etc.). Samples from that area may be qualified as estimated values. If non-compliant field duplicate results were random, but all laboratory duplicate results were within control limits, sampling techniques may be investigated. If rocks and aggregate soil clusters were indiscriminately discarded from some samples but not from others without any consistent rationale, all results may be considered suspect. Again, the non-conformance should be evaluated in the context of the project goals and objectives to determine the impact on overall data quality.*

**Step 4. Assign data validation qualifiers as necessary**. The data validator reviews the analytical data to provide an overall assessment of the quality of the data. Some data may need a data validation qualifier to give an indication of potential bias of the data. Data validation qualifiers may be assigned to particular sample results based on information such as laboratory qualifiers, QC summaries, and data summaries. Any data validation qualifiers that the data validator assigns should be documented in a report. This report will be used to support the assignment of the data validation qualifiers as well as providing the data validation qualifier information for entry into an electronic database. Data validation qualifiers are not mandated by all projects, but when a qualifier is assigned to a sample result, it gives the data user some indication about the data quality. Examples of data validation qualifiers and typical

definitions are given in Table 8. Appendix C also provides additional examples of data validation qualifiers used by specific programs.

**Table 8. Examples of Data Validation Qualifiers and Definitions**

| Data Validation Qualifier | Typical Definition |
|---|---|
| U | The analyte was analyzed for, but was not detected above the reported sample quantitation limit. |
| UJ | The analyte was not detected above the reported sample quantitation limit. However, the reported quantitation limit is approximate and may or may not represent the actual limit of quantitation necessary to accurately and precisely measure the analyte in the sample. |
| J | The analyte was positively identified; the associated numerical value is the approximate concentration of the analyte in the sample. |
| R | The sample results are rejected due to serious deficiencies in the ability to analyze the sample and meet QC criteria. The presence or absence of the analyte cannot be confirmed. |

Source: EPA, 1999.

For projects that do not mandate any form of data validation qualifiers, recommendations for data qualification may be summarized in text format in a narrative.

**Step 5. Prepare analytical data validation report.** The purpose of preparing a data validation report is to summarize all of the information about the analytical data that was reviewed during the data validation and to detail how the project needs were met. The data validator should document each step outlined above and assemble this documentation into an analytical data validation report. The report should outline the data that were reported as well as any deficiencies in the sample data or QC data and the data validation qualifiers assigned. The information in the data validation report should also support any additional information that is reported as part of the validated data (see Example 22).

As the data validation process is completed, the analytical data validation report should include:

- a summary of project objectives and needs,
- a summary of the quality of the data,
- a summary of the fulfillment of the project objectives and needs, and
- the validated data.

**Example 22. Data Validation Report Documentation**

*A regulation lists wastes that have been determined to be hazardous. Industrial waste streams are evaluated for potential inclusion as entirely new "listed" waste streams. Evaluation entails chemically characterizing the waste stream based on known engineering design specifications and historical data provided by the facility. Listing determinations entail extensive investigations, generally including literature and database searches, industry surveys, engineering site visits, and sampling and analysis. Generally, sampling and analysis in conjunction with a hazardous waste determination is performed in three stages – engineering site visit, familiarization sampling and analysis, and record sampling and analysis. This is due in part to the chemical uncertainty and lack of process knowledge for many industrial waste streams.*

*During the initial engineering site visit phase, a number of facilities that are unique to the industrial category are selected to obtain information on their current waste management practices. This is followed by a familiarization phase in which a select number of samples are collected and analyzed in order to allow the laboratories to become familiar with the anticipated sample matrices and the potential analytical problems they may pose. For the final record sampling phase, samples are collected from points within the process that are representative of the waste as managed prior to disposal. At least one record sample is collected for each waste stream under consideration.*

*The constituents that are to be measured are determined from the starting material composition and the suspected byproducts obtained from the industrial process. Samples of solid waste streams are evaluated for potential to leach target analytes into the environment using leaching tests, such as the toxicity characteristic leaching procedure (TCLP) (Method 1311) and synthetic precipitation leaching procedure (SPLP) (Method 1312).*

*In one listing determination, thallium was identified in the planning documents as a target analyte. After the respective TCLP and SPLP procedures were completed, the leachates were prepared according to Method 6010B, followed by analysis for thallium with an ICAP-61E Trace Level Analyzer. During data validation, the reviewer questioned seemingly inconsistent detection limits reported for thallium in the two leachate matrices. He resolved the questions by reviewing the analytical methods and the specifications of the SAP. He then reconstructed the calculations, determining that:*

- *The TCLP leachate was diluted by a factor of 20 prior to extraction to compensate for the high level of sodium contained in the acetate buffer leaching solution. Without*

**Example 22. Data Validation Report Documentation (continued)**

*consideration of the dilution factor, the typical laboratory reporting limit for thallium in an aqueous matrix is 5 parts per billion (ppb), with a calculated instrument detection limit of 2.2 ppb.*

- *After considering the leachate dilution factor of 20 multiplied by 5 ppb, the TCLP thallium actual reporting limit should have been 100 ppb.*
- *The laboratory correctly reported all SPLP thallium concentrations down to 5 ppb but arbitrarily set the TCLP thallium reporting limit at 2000 ppb.*
- *After further discussions with the laboratory, it was confirmed that the actual TCLP thallium reporting limit is 100 ppb based on the dilution factor correction.*

*Therefore, the laboratory corrected the initial TCLP thallium result <2000 ppb. The corrected sample value of 280 ppb was further substantiated, based on a duplicate analysis yielding 270 ppb and matrix spike and matrix spike recoveries of 94% and 92%.*
*In addition, all TCLP and SPLP leachate preparation and method blank analyses contained no thallium concentrations above the laboratory reporting limits of 100 ppb and 5 ppb, respectively.*

*These findings and the corrected values were documented in the analytical data validation reports.*

Documentation of the data validation process is needed for the DQA. Therefore, it is vital that the data validator compiles all possible information from the data validation process into a usable format. In some cases, the field and analytical data validation reports may be combined into one report. Similarly, the validated field data and validated analytical laboratory data may also be combined into one database in order to facilitate the review of validated data by the data user. These options are dependent upon the needs specified in the planning documents and the resources available to carry out these options.

### 5.2.3 Tools and Techniques for Focused Data Validation

As defined in Chapter 3, a focused data validation is a detailed investigation of particular data records that need special interpretation or review. These data records may be related to the field activities, the analytical laboratory data, or the assignment of data validation qualifiers. However, not all projects need a focused data validation be performed. Three instances were identified in Section 3.3.3

to illustrate when a focused data validation may be requested for a project. Examples of these instances are discussed below.

As the data user reviews the data and data validation report for a project, the data user may identify an error or omission in these documents or records, as shown in Example 23. In some cases, as shown in Example 24, the data user may not note any errors or omissions during the review of the data or the data validation report, but the review may identify anomalies or inconsistencies in the information.

**Example 23. Further Investigation into Data Validation Assigned Data Qualifiers**

*One of the project needs may be that the data validator should apply data validation qualifiers to the data records based on review of the laboratory qualifiers and the QC data. Upon review of the data validation report, the data user notes that a subset of mercury data are qualified with a "UJ." The "UJ" indicates that the results were not detected above the reported sample quantitation limit, which is approximate and may or may not represent the actual limit of quantitation necessary to accurately measure the analyte in the sample. The report contains no further documentation to support the "UJ" qualification. Because the mercury data are important to support project decisions and the estimated quantitation limits were higher than what was specified by the planning documents, the data user may request that the data validator perform a focused data validation to supply information about these "UJ" qualifiers. The focused data validation would be directed at issues such as:*

*Why were these records qualified "UJ?"*
*Why were the quantitation limits higher than the specified reporting limits?*

*After the data validator has provided the requested information to the data user to resolve this issue, the data validator should also document how the issue was resolved for the project records.*

**Example 24. Further Investigation into Analytical Method Comparability**

*Numerous samples were collected for a project, but particular analytical methods were not specified in the planning documents for the analysis of the samples. The data user noted that the samples were sent to two different laboratories for analysis. One laboratory analyzed for uranium by kinetic phosphorescence analysis and the other laboratory analyzed for uranium by inductively coupled plasma - mass spectroscopy. Although both methods are acceptable, the data user may request a focused data validation to look closer at the laboratory procedures for analyzing the samples to determine the comparability of the analytical methods.*

The most common instance that may call for a focused data validation occurs when anomalies are identified during the DQA process. As the data user begins to perform exploratory data and statistical analysis, the data user may notice anomalies in the data set as a whole. Examples 25 and 26 illustrate instances where the data user began to look at the whole data set and noted an anomaly in the field or analytical results, and so instigated a focused data validation to find the error. In any case where a focused data validation is performed, even for the smallest detail, the data validator should document all of the efforts that were put forth to reconcile the question.

**Example 25. Use of Historic Site Records in Field Data Validation**

*An extensive sampling and analysis program was conducted in support of a base realignment and closure effort at a large military installation. This program included collection and analysis of soil, groundwater, surface water, and vegetation samples. Review of the analytical data from these samples indicated that a relatively small area of the site was contaminated with high concentrations of a herbicide that is used only in agricultural applications. This result was confusing, since there was no known or logical use of this compound on this, or any other military installation. The data were verified and validated from the analytical standpoint. The data appeared to be valid, but remained illogical. The project team requested a focused data validation to review the field documentation, including the extensive site background records. These records included files kept by the military documenting the activities on the base, periodic aerial photographs and maps, and permitting files maintained by the state's department of environmental protection. Review of the aerial photographs of the base spanning 60 years led to the ultimate solution to the question and validation of the data. Photographs taken sporadically over a decade showed little change in the area of suspected contamination – a heavily vegetated area with no obvious activity. A photo taken five years later showed vegetation was growing back, strongly implying that something had occurred to adversely impact the plant life in the area – a finding consistent with application or disposal of a large quantity of herbicide. Searches of records from that five-year period yielded a memo requesting permission to dispose of large quantities of off-spec material on the base, and a map with cryptic notes indicating disposal in the precise location where the contaminated samples had been collected. These findings resulted in further field activities to confirm and delineate the disposal site.*

**Example 26. Further Investigation into Validated Analytical Results**

*In the course of developing a new analytical method for dioxins and furans, EPA solicited the voluntary participation of 22 laboratories in 5 countries. Because of the concern over this class of analytes and the regulatory implications, the data from the study were subjected to exceptional scrutiny, including thorough data validation. All of the valid data from the study were then used to develop statistical QC specifications for use in the final method. During the course of those statistical evaluations, all of the results from the study were plotted and the data distributions were examined. As expected, most of the data were distributed either normally, or log-normally. However, data for one of the spiked compounds were clearly bimodal in their distribution, a completely unexpected result.*

*Based on the distributions, all of the results from the minor mode in the distribution were re-examined and found to come from a single laboratory. All of these results were within a range of reasonable recoveries for a spiked compound, and although they might have been set aside as statistical outliers, the data were too consistent to be ignored.*

*All of the results had already passed the data validation. Therefore, the laboratory was contacted about the situation. Based on their examination of the computerized calculations, it became apparent that the problem was caused at the laboratory. The method called for the compound in question to be spiked into the samples at twice the concentration of all of the other compounds in the same spiking solution. The laboratory had used the spiking solution provided to them for the purposes of the study, but had failed to take into account the higher concentration of this one compound, which causes all of their results for this compound to be off by a factor of 2. They subsequently corrected the computerized calculation for this compound, revised their data reports, and verified all the other automated calculations. All of the revised results fell within the distribution from the other laboratories in the study.*

*The cause of the problem was the lack of verification of the automated calculations by the laboratory. The results had passed the subsequent data validation efforts because the data reported for the study fell within the range of recoveries that were acceptable for the study. Had the results been discarded as outliers based on a statistical test, the power of the study would have been needlessly reduced.*

# CHAPTER 6

# DATA SUITABILITY

## 6.1 DETERMINING DATA SUITABILITY

Data verification and data validation are two key steps in the project life cycle (Figure 2). They are important because they determine whether sampling and analytical activities were performed in accordance with the planned approach, and because they document the known quality of the data and specific concerns or vulnerabilities associated with data points and data sets.

However, the outputs of data verification and data validation by themselves are not sufficient to answer the fundamental question: can these data be used for their intended purpose in environmental decision-making? While data verification and data validation are essential precursors to answering this question, the data user should also take other considerations into account when evaluating the utility of the data. This is true for a number of reasons:

- More than one laboratory and more than one data validator may be involved in producing or reviewing project data. Therefore only the data user may have access to the complete set of data which will be used to make decisions.

- Even if they have full access to all planning documentation such as QA Project Plans and SAPs, neither data verifiers nor data validators are knowledgeable about the full range of goals and constraints that shape the data user's actions and perspective. For example, the data user may have to address the risk management tradeoff between taking immediate action to resolve a pressing problem on one hand, versus taking additional time to resolve uncertainty in data on the other.

- Analysis of the utility of data sets needs more than a knowledge of how individual data points have been qualified during data verification or data validation. In most cases, data will be combined into mathematical results or models, and statistical tests may be applied in order to determine whether and how the data can be used.

The process for determining the utility of data sets is known as data quality assessment, which has been defined by EPA in *Guidance for Data Quality Assessment: Practical Methods for Data Analysis (QA/G-9)* (EPA, 2000b) as "the scientific and statistical evaluation of data to determine if data obtained from environmental data operations are of the right type, quality, and quantity to support their intended use." That guidance provides extensive information about DQA and the statistical tools that it employs.

The focus of the present chapter is what the data validator can do to facilitate the transition from data validation to data quality assessment. As used here, the term "data suitability" refers to efforts of the data validator to foresee and support the needs of the DQA analyst and ultimate data user. Since this role is feasible only to the degree that the data validator has been informed about the intended use of the data, it is vital for the data user to share this information with the data validator to the extent possible.

Section 6.2 describes how the data validator can employ professional judgment to anticipate and document in the data validation report any concerns that the data validator anticipates might become important to the DQA analyst. Section 6.3 discusses the concept of focused data validation, in which the data validator may answer specific questions raised by the data user after review of the data validation report. Section 6.4 is a brief overview of DQA, highlighting how it is influenced by data validation outputs.

## 6.2 USING PROFESSIONAL JUDGMENT IN DATA VALIDATION

As described in previous chapters, the data validator typically follows project-specific protocols that guide the review process and shape the content and format of the data validation report. The data validation process is constrained by a number of factors, including contract requirements, client and management expectations, and competing demands on the data validator's time.

However, in most cases there remains an opportunity for the data validator to exercise professional judgment in order to maximize the benefits of the data validation process. For instance, *USEPA Contract Laboratory Program, National Functional Guidelines for Organic Data Review* (EPA, 1999) includes a section titled "Overall Assessment," which is described as "a brief narrative in which the data reviewer expresses concerns and comments on the quality and, if possible, the usability of the data." To develop this narrative, the data validator uses "professional judgment to determine if there is any need to qualify data which were not qualified based on the QC criteria previously discussed."

Data validators may be able to examine the issues associated with this "Overall Assessment" step. To do so, they would need access to project planning documentation such as the QA Project Plan or SAP, and they would need sufficient communication with the data user to develop a clear understanding of the intended use and desired quality of the data. It would also be useful to obtain a more complete record of the laboratory's activities, including logs for sample preparation, calibration, and instrument performance; instrument printouts; and raw data. The extent to which data validators have access to these types of information depends on the graded approach to data validation discussed in Chapter 1 (Section 1.3).

There is an opportunity for the data validator to play a proactive role on behalf of the data user. Ideally, there should be a two-way dialogue between the data validator and data user. In some cases, the data validator's input may make clear that the data package would benefit from additional review by someone with professional expertise that would not otherwise be called for (e.g., hydrogeology, radiological chemistry, or engineering).

Table 9 lists typical data validation questions in the left column, and in the other columns demonstrates how those questions could be expanded to incorporate data suitability concerns.

## 6.3 FOCUSED DATA VALIDATION

Focused data validation is a detailed investigation of particular data records that need special interpretation or review. The purpose of focused data validation is to answer questions about the data

**Table 9. Data Validation Versus Data Suitability**

| Data Validation Question | Data Suitability Question | Examples |
|---|---|---|
| Have the analytical methods been followed properly? | Now that data are available, do we still think that these were the appropriate analytical methods? | Were there extreme matrix interferences? Were matrix spike/matrix spike duplicate recoveries unusually low using these methods? |
| Have the detection limits been calculated properly? | Are these detection limits adequate for the goals of this project? | Were detection limits appropriate (do they cover the threshold of concern for each compound)? Were the technical basis for calculation of detection limits documented correctly? |
| Have MQO goals, such as precision and bias, been achieved? | Based on the available data, do these MQO goals still seem reasonable? | Were the initial calibration criteria (response factors, precision, correlation coefficient) appropriate for these analytes? |
| Are the appropriate data points flagged with qualifiers? | What do patterns in the qualified data suggest about the overall data set? | For data that fall between the detection limit and the quantitation limit, has the laboratory provided numeric values rather than flags only? How do you interpret flags indicating contaminated blanks when the real samples have the same contaminants? |

that arise as a result of the data user's review of the validated data and data validation report. The inputs to focused data validation may include the planning documents, data validation report, hard-copy data package, the validated data set, and a general knowledge of the environmental problem and its history.

A focused data validation may be requested by the data user during the initial review of the data validation report, or it may occur later during the DQA process. As the information is reviewed, the data user is looking at whether the data appear to be appropriate to support decision making based on the original project needs. The data user may also identify errors or omissions in the data or data validation report that need to be corrected. The report should include items such as a list of the samples collected, field information about how the samples were collected, the analysis performed on the samples, and the quality of the reported data. The data validator should attempt to document anything out of the ordinary that is noticed about the data during their review.

If the data user has questions about the data validator's report, the data user may go back to the data validator and request further explanation or information. For example, the data user may notice that a majority of the data were rejected for a particular analyte in the data set. Although the data validator provided an explanation for the rejection in the report, the data user may request additional information from the data validator to determine if the data may be useful in some context to meet project objectives. The data validator would then go back and review the data in the context of the data user's question and provide additional input.

Often the data user will initially accept the data validator's report directly, but as the DQA process unfolds (Section 6.4), the data user may observe that some information appears anomalous. This situation may also motivate the data user to request a focused data validation. Additional effort from the data validator may be needed in this situation because the data user may be seeing possible anomalies that could be caused by any number of various sources. In either case, the focused data validation should provide the data user with additional information so that the data user can make decisions about the suitability of project data.

## 6.4 DATA QUALITY ASSESSMENT

Once the data validation process, including any focused data validation steps, has been completed, it is time for the DQA process. Data quality assessment, like data validation, can be more or less rigorous depending on how the graded approach has been applied to the project. EPA's *Guidance for Data Quality Assessment: Practical Methods for Data Analysis (QA/G-9)* (EPA, 2000b) describes it as a five-step process:

Step 1: Review the Data Quality Objectives and Sampling Design
Step 2: Conduct a Preliminary Data Review

Step 3: Select the Statistical Test
Step 4: Verify the Assumptions of the Statistical Test
Step 5: Draw Conclusions from the Data

Although the process is presented as a series of steps, it can be iterative to allow for steps to be repeated as necessary. The outputs of data validation are important to accomplishing the DQA process steps. For example:

**Step 1** includes a review of the implementation of the sampling design. If the data validator has determined that the sampling and analysis process deviated in significant ways from that envisioned during the planning phase, that determination should be included in the narrative section of the data validation report.

**Step 2** involves a preliminary evaluation of the data set. This step makes extensive use of the data validation report, especially with respect to QC measures. The DQA analyst looks to the data validation report not only to examine flagged data, but also to note "anomalies in recorded data, missing values, deviations from SOPs, and the use of nonstandard data collection methodologies" (EPA, 2000b).

In **Steps 3 and 4**, the DQA analyst uses the collected data to determine whether they are consistent with the assumptions underlying the statistical test(s) to be employed. For some assumptions, the analyst may rely on the data validator's conclusions. For instance, a key assumption for many statistical tests is an absence of bias in the data set. If the data validation report's analysis of QC measurements indicates that the data are biased, the DQA analyst may be compelled either to develop a technique to adjust for the bias, or to select an alternative suite of statistical tests.

Other points at which the data validation report can be used during Steps 3 and 4 include the evaluation of potential outliers and development of a strategy for handling values reported as being below the detection limit. The data validation qualifiers and the data validator's narrative report constitute the most important source of evidence as the DQA analyst attempts to determine whether apparent outliers or non-detects are in fact suspect results, and whether and how they can be used.

In **Step 5**, the DQA analyst draws conclusions from the validated data and the statistical tests performed on it. In doing so, the analyst may rely on the data validator's professional judgment. For instance, if outliers have proved to be a problem with the data set, the analyst may perform calculations both with and without the questionable data in order to make comparisons in order to ascertain the influence of these anomalies on decision making.

## 6.5 SUMMARY

As reflected in Figure 1 at the beginning of this guidance, data quality assessment marks the culmination of the "assessment" phase of the project life cycle. In the broadest sense, the assessment phase commences with data verification activities that are conducted in conjunction with field sampling and laboratory analysis. The primary goal of data verification is to document that applicable method, procedural, or contractual requirements have been met.

Once the data packages and related documentation have been transmitted, the next step in the assessment phase belongs to the data validator. Data validation determines whether a data set has met the specifications for a project-specific intended use. It provides the data user and DQA analyst with crucial inputs that will enable them to evaluate whether and how the data can be used for decision making.

From data verification to data validation to DQA, each step in the assessment phase of the project life cycle benefits from and builds on the previous one. Together, they assure achievement of the ultimate goal of environmental data collection: credible products and sound and defensible decisions.

# CHAPTER 7

## REFERENCES


National Environmental Laboratory Accreditation Conference Standards, 2000. *Chapter 5: Quality Systems Standard.* www.epa.gov/ttnnela1/standard.html.

LANL (Los Alamos National Laboratory) Environmental Restoration Project, 1999. *Baseline Analytical Data Validation,* ER-SOP-15.17, ER Catalog Number ER19990078, Revision 0.

Poppiti, James. Environmental Science and Technology, 28.6 (1994).

U.S. Army Corp of Engineers, 1997. *Environmental Quality - Chemical Quality Assurance for Hazardous, Toxic, and Radioactive Waste (HTRW) Projects,* EM 200-1-6.

U.S. Department of Defense, 2000. *Department of Defense Quality Systems Manual for Environmental Laboratories,* DoD Environmental Data Quality Workgroup.

U.S. Department of Energy, 1999. *Oak Ridge Reservation Annual Site Environmental Report 1999.*

U.S. Environmental Protection Agency, 1992. *USEPA SW-846 Test Methods for Evaluating Solid Waste, Physical/Chemical Methods*, Office of Solid Waste.

U.S. Environmental Protection Agency, 1994. *USEPA Contract Laboratory Program National Functional Guidelines for Inorganic Data Review*, EPA 540/R-94/013, Office of Emergency and Remedial Response.

U.S. Environmental Protection Agency, 1996. *Region I, EPA-New England Data Validation Functional Guidelines for Evaluating Environmental Analyses, U.S. EPA-New England*, Region I, Quality Assurance Unit Staff, Office of Environmental Measurement and Evaluation.

U.S. Environmental Protection Agency, 1998. *Guidance for Quality Assurance Project Plans (QA/G-5)*, EPA/600/R-98/018, Office of Environmental Information.

U.S. Environmental Protection Agency, 1999. *USEPA Contract Laboratory Program National Functional Guidelines for Organic Data Review*, EPA 540/R-99/008, Office of Emergency and Remedial Response.

U.S. Environmental Protection Agency, 2000a. *EPA Quality Manual for Environmental Programs*, EPA Manual 5360 A1, Office of Environmental Information.

U.S. Environmental Protection Agency, 2000b. *Guidance for Data Quality Assessment: Practical Methods for Data Analysis (QA/G-9)*, EPA/600/R-96/084, Office of Environmental Information.

U.S. Environmental Protection Agency, 2001a. *EPA Requirements for Quality Management Plans (QA/R-2)*, EPA/240/B-01/002, Office of Environmental Information.

U.S. Environmental Protection Agency, 2001b. *EPA Guidance on Data Quality Indicators (QA/G-5i)*, peer review draft.

## APPENDIX A

## OTHER DEFINITIONS OF DATA VERIFICATION AND DATA VALIDATION

| | |
|---|---|
| U.S. Army Corps of Engineers Environmental Quality - Chemical Quality Assurance for Hazardous, Toxic, and Radioactive Waste (HTRW) Projects (1997)<br><br>http://www.usace.army.mil/inet/usace-docs/eng-manuals/em200-1-6/basdoc.pdf | Data verification is the most basic assessment of data. Data verification is a process for evaluating the completeness, correctness, consistency, and compliance of a data package against a standard or contract. In this context, "completeness" means all required hard-copy and electronic deliverables are present. Data verification should be performed by the government or independent entity for QA laboratory deliverables, and by the laboratory contract holder for primary laboratory deliverables.<br>Validation: Process of data assessment in accordance with EPA regional or national functional guidelines, or project-specific guidelines.<br>Assessment of the whole raw data package from the lab. They break the process down into data verification, data review, data evaluation, and data validation. |
| James Poppiti<br>*Environmental Science and Technology*<br>Vol. 28, No. 6, 1994 | Validation is more complicated than verification, it attempts to assess the impacts of data use, especially when requirements are not met.<br>Data that do not meet all the measurement requirements (verification) do not have to be rejected or considered useless (validation). |
| Department of Energy/Oak Ridge Reservation Annual Site Environmental Report 1999<br><br>http://www.ornl.gov/Env_Rpt/aser99/aser99.htm | Validation of field and analytical data is a technical review performed to compare data with established quality criteria to ensure that data are adequate for intended use. |

| | |
|---|---|
| Department of Defense (DoD) Quality Systems Manual (2000)<br><br>https://www.denix.osd.mil | Validation: the process of substantiating specified performance criteria. (EPA-QAD)<br>Verification: confirmation by examination and provision of evidence that specified requirements have been met. (National Environmental Laboratory Accreditation Conference) |
| US EPA Region 1, New England 1996<br><br>Data Validation Functional Guidelines for Evaluating Environmental Analyses<br><br>http://www.epa.gov/region01/oeme/DVMANUAL.pdf | Data Validation, the first step in assessing data quality; is a standardized review process for judging the analytical quality and usefulness of a discrete set of chemical data. Thus, data validation identifies the analytical error associated with a data set. Data validation can also identify some (e.g., incorrect preservation techniques), but not all of the sampling error associated with a data set. |
| USEPA SW-846 Test Methods for Evaluating Solid Waste, Physical/Chemical Methods. Third Edition<br><br>http://www.epa.gov/epaoswer/hazwaste/test/sw846.htm | Data Validation: The process of evaluating the available data against the project data quality objectives to make sure that the objectives are met. Data validation may be very rigorous, or cursory, depending on project data quality objectives. The available data review will include analytical results, field QC data and lab QC data, and may also include field records. |

## APPENDIX B

## GLOSSARY

**calibration** – comparison of a measurement standard, instrument, or item with a standard or instrument of higher accuracy to detect and quantify inaccuracies and to report or eliminate those inaccuracies by adjustments.

**chain-of-custody** – an unbroken trail of accountability that ensures the physical security of samples, data, and records.

**data quality assessment** – a statistical and scientific evaluation of the data set to determine the validity and performance of the data collection design and statistical test, and to determine the adequacy of the data set for its intended use.

**data quality indicators** – quantitative and qualitative measures of principal quality attributes, including precision, accuracy, representativeness, comparability, completeness, and sensitivity

**data quality objectives** – qualitative and quantitative statements that clarify study objectives, define the appropriate type of data, and specify tolerable levels of potential decision errors that will be used as the basis for establishing the quality and quantity of data needed to support decisions.

**data validation** – an analyte- and sample-specific process that extends the evaluation of data beyond method, procedural, or contractual compliance (i.e., data verification) to determine the analytical quality of a specific data set.

**data validation qualifier** – code applied to the data by a data validator to indicate a verifiable or potential data deficiency or bias.

**data validator** – an individual (typically an independent third party) responsible for conducting data validation activities.

**data verification** – the process of evaluating the completeness, correctness, and conformance/compliance of a specific data set against the method, procedural, or contractual requirements.

**data verifier** – an individual (typically an employee of the field or laboratory organization whose operations are being verified) responsible for conducting data verification activities.

**drylabbing** – a laboratory may report analytical results without having actually performed the analyses. Results may be either invented from scratch, or previous legitimate results may be "borrowed" for inclusion in the present data package.

**environmental data** – any measurements or information that describe environmental processes, location, or conditions; ecological or health effects and consequences; or the performance of environmental technology. For EPA, environmental data include information collected directly from measurements, produced from models, and compiled from other sources such as data bases or the literature.

**focused data validation** – a detailed investigation of particular data records identified by the data user that need interpretation or review.

**graded approach** – the process of basing the level of application of managerial controls applied to an item or work according to the intended use of the results and the degree of confidence needed in the quality of the results.

**juicing** – fortification of a sample with additional analyte such as re-spiking a spiked sample or adding peak area. See also peak enhancement and peak juicing.

**laboratory qualifier** – code applied to the data by the contract analytical laboratory to indicate a verifiable or potential data deficiency or bias.

**measurement quality objectives** – "acceptance criteria" for the quality attributes measured by project data quality indicators. During project planning, measurement quality objectives are established as quantitative measures of performance against selected data quality indicators, such as precision, bias, representativeness, completeness, comparability, and sensitivity.

**peak shaving** – manually adjusting the raw data by reducing a peak area that is out of specification.

**peak enhancement** – manually adjusting the raw data by increasing a peak area that is out of specification. See also juicing and peak juicing.

**peak juicing** – manually adjusting the raw data by increasing a peak area that is out of specification. See also juicing and peak enhancement.

**performance evaluation** – a type of audit in which the quantitative data generated in a measurement system are obtained independently and compared with routinely obtained data to evaluate the proficiency of an analyst or laboratory.

**quality** – the totality of features and characteristics of a product or service that bear on its ability to meet the stated or implied needs and expectations of the user.

**quality assurance** – an integrated system of management activities involving planning, implementation, documentation, assessment, reporting, and quality improvement to ensure that a process, item, or service is of the type and quality needed and expected by the customer.

**quality assurance project plan** – a document describing in comprehensive detail the necessary QA, QC, and other technical activities that should be implemented to ensure that the results of the work performed will satisfy the stated performance criteria.

**quality control** – the overall system of technical activities that measures the attributes and performance of a process, item, or service against defined standards to verify that they meet the stated needs established by the customer; operational techniques and activities that are used to fulfill needs for quality.

**quality system** – a structured and documented management system describing the policies, objectives, principles, organizational authority, responsibilities, accountability, and implementation plan of an organization for ensuring quality in its work processes, products (items), and services. The quality system provides the framework for planning, implementing, documenting, and assessing work performed by the organization and for carrying out needed QA and QC activities.

**record** – a completed document that provides objective evidence of an item or process. Records may include photographs, drawings, magnetic tape, and other data recording media.

**time-traveling** – falsification of the date of analysis in the laboratory's data system in order to conceal such things as exceeding a holding time.

**validation** – confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use are fulfilled. In design and development, validation concerns the process of examining a product or result to determine conformance to user needs.

**verification** – confirmation by examination and provision of objective evidence that specified requirements have been fulfilled. In design and development, verification concerns the process of examining a result of a given activity to determine conformance to the stated requirements for that activity.

## APPENDIX C

## EXAMPLES OF DATA QUALIFIERS USED BY SPECIFIC PROGRAMS

The following examples are quoted from the programs referenced.

### EXAMPLE 1: USEPA CONTRACT LABORATORY PROGRAM NATIONAL FUNCTIONAL GUIDELINES FOR INORGANIC DATA REVIEW (EPA, 1994)

*"The following definitions provide brief explanations of the national qualifiers assigned to results in the data review process. If the Regions choose to use additional qualifiers, a complete explanation of those qualifiers should accompany the data review.*

*U* *The material was analyzed for, but was not detected above the level of the associated value. The associated value is either the sample quantitation limit or the sample detection limit.*

*J* *The associated value is an estimated quantity.*

*R* *The data are unusable. (Note: Analyte may or may not be present.)*

*UJ* *The material was analyzed for, but was not detected. The associated value is an estimate and may be inaccurate or imprecise."*

### EXAMPLE 2: USEPA CONTRACT LABORATORY PROGRAM NATIONAL FUNCTIONAL GUIDELINES FOR ORGANIC DATA REVIEW (EPA, 1999)

*"The following definitions provide brief explanations of the national qualifiers assigned to results in the data review process. If the Regions choose to use additional qualifiers, a complete explanation of those qualifiers should accompany the data review.*

*U* *The analyte was analyzed for, but was not detected above the reported sample quantitation limit.*

*J* *The analyte was positively identified; the associated numerical value is the approximate concentration of the analyte in the sample.*

*N* *The analysis indicates the present of an analyte for which there is presumptive evidence to make a "tentative identification."*

*NJ* *The analysis indicates the presence of an analyte that has been "tentatively identified" and the associated numerical value represents its approximate concentration.*

*UJ* *The analyte was not detected above the reported sample quantitation limit. However, the reported quantitation limit is approximate and may or may not represent the actual limit of quantitation necessary to accurately and precisely measure the analyte in the sample.*

*R* *The sample results are rejected due to serious deficiencies in the ability to analyze the sample and meet quality control criteria. The presence or absence of the analyte cannot be verified."*

**EXAMPLE 3: REGION I, EPA-NEW ENGLAND DATA VALIDATION FUNCTIONAL GUIDELINES FOR EVALUATING ENVIRONMENTAL ANALYSES, U.S. EPA-NEW ENGLAND (EPA, 1996)**

*"Only codes defined by this document are permitted to qualify data. Should it be necessary to include other codes, prior approval must be obtained from the EPA-NE CLP-TPO. If approval is given, complete definitions must be supplied in the key for the Data Summary Table. The standard data validation codes used in qualifying data in accordance with this guidance are:*

*U* *The analyte was analyzed for, but was not detected. The associated numerical value is the sample quantitation limit. The sample quantitation limit accounts for sample specific dilution factors and percent solids corrections or sample sizes that deviate from those required by the method.*

*J* *The associated numerical value is an estimated quantity.*

*R* *The data are unusable (analyte may or may not be present). Resampling and reanalysis is necessary for verification. The R replaces the numerical value or sample quantitation limit.*

*UJ* *The analyte was analyzed for, but was not detected. The sample quantitation limit is an estimated quantity.*

*EB, TB, BB* *An analyte that was identified in an aqueous equipment blank, trip blank, or bottle blank that was used to assess field contamination associated with soil/sediment samples.* ***These qualifiers are to be applied to soil/sediment sample results only.*** *(For additional guidance refer to Blank Section V of Parts II, III or IV)"*

**EXAMPLE 4: LOS ALAMOS NATIONAL LABORATORY ENVIRONMENTAL RESTORATION PROJECT (LANL, 1999)**

*"The following are definitions of laboratory qualifiers and laboratory reason codes for radiochemistry analysis:*

*U* *The analyte was analyzed for but not detected above the reported estimated quantitation limit.*

*J* *The analyte was positively identified, the associated numerical value is the approximate concentration of the analyte in the sample:*

*J+ = likely to have a high bias,*
*J- = likely to have a low bias.*

*UJ* *The analyte was analyzed for but not detected. The associated value is an estimate.*

*R* *The sample results are rejected due to serious deficiencies in the ability to analyze the sample and meet quality-control criteria. Presence or absence cannot be verified.* ***Note:*** *Any results qualified as "R" should be looked at for relevance for data use. Thus, "R" implies "PM" also, and must not be used alone.*

*P* *Use professional judgment based on data use. It usually has an "M" with it, which indicates that a manual check should be made if the data that are qualified with the "P" are important to the data user. In addition, "PM" also means that a decision must be made by the project manager or a delegate with regard to the need for further review of the data. This review should include some consideration of potential impact that could result from using the "P" qualified data. (For example, in the case of holding-time exceedance, the project manager or delegate can decide to use the data with no qualification when analytes of interest are known to not be adversely affected by holding-time exceedances. Another example*

*is the case where soil sample duplicate analyses for metals exceed the precision criteria. Because this is likely due to sample nonhomogeneity rather than contract laboratory error, the manager or delegate must decide how to use the data.)*

*PM* *Manual review of raw data is recommended in order to determine if the defect impacts data use, as in "R" above."*