EPA GUIDANCE FOR QUALITY ASSURANCE PROJECT PLANS

EPA QA/G-5

Quality Assurance Division
United States Environmental Protection Agency

Washington, DC 20460

EXTERNAL WORKING DRAFT

November 1996

PREFACE ADDRESSED TO EXTERNAL REVIEWERS

This document is an external review draft of EPA QA/G-5, *Guidance on Quality Assurance Project Plans*. The Quality Assurance Division of the U.S. Environmental Protection Agency (EPA) is developing this guidance to help the environmental community comply with Agency policies and Quality System requirements regarding environmental data collection planning, implementation, and assessment (the EPA Quality System is described briefly in the Foreword and elsewhere in this draft document). In particular, this guidance is intended to help project managers, environmental scientists and engineers, and quality assurance professionals understand and comply with the companion quality assurance policy document EPA QA/R-5, *Requirements for Quality Assurance Project Plans*.

This external review draft is being distributed broadly to the environmental community for review and comment. Reviewers are encouraged to submit constructive suggestions for changes, additions, and editorial improvements to this draft. Broad participation in the review of this document will benefit the environmental community as a whole, so please share your hard-won experience, new ideas, and diverse perspectives.

The Quality Assurance Division recognizes that many organizations already have developed or are in the process of developing their own guidance on how to prepare quality assurance project plans that comply with EPA QA/R-5. Those organizations may find that their internal documentation provides more detailed recommendations or considerations than are discussed in various sections of this draft. The Quality Assurance Division encourages those organizations to share their knowledge and experience by submitting examples of how they have addressed the QAPP elements in their guidance.

This draft also contains technical appendices that address certain topics in greater depth. Your comments on the usefulness of these materials also are solicited. In particular, it would be very helpful to know which topics could be added or which areas addressed in more depth to improve the utility of the document.

Please submit your comments by 31 January 1997 to:

John Warren
Quality Assurance Division (8724)
U. S. Environmental Protection Agency
401 M Street, S.W.
Washington, DC 20460

E-mail: warren.john@epamail.epa.gov
Fax: 202-401-7002

FOREWORD

The U.S. Environmental Protection Agency (EPA) has developed the Quality Assurance Project Plan (QAPP) as an important tool for project managers and planners to document the type and quality of data needed for environmental decisions and to use as the blueprint for collecting and assessing those data from environmental programs. The development, review, approval, and implementation of the QAPP is part of the mandatory Agency-wide Quality System that requires all organizations performing work for EPA to develop and operate management processes and structures for ensuring that data or information collected are of the needed and expected quality for their desired use. The QAPP is an integral part of the fundamental principles of Quality Management that form the foundation of the Agency's Quality System.

This document is one of the *U.S. Environmental Protection Agency Quality System Series* requirements and guidance documents. These documents describe the EPA policies and procedures for planning, implementing, and assessing the effectiveness of the Quality System. Requirements documents (identified as EPA/R-x) establish criteria and mandatory specifications for quality assurance (QA) and quality control (QC) activities. Guidance documents (identified as EPA QA/G-x) provide suggestions and recommendations of a nonmandatory nature for using the various components of the Quality System.

Other guidance documents related to EPA QA/G-5 include:

EPA QA/G-4	*Guidance for the Data Quality Objectives Process*
EPA QA/G-4D	*Data Quality Objectives Decision Error Feasibility Trials (DQO/DEFT)*
EPA QA/G-4R	*Guidance for the Data Quality Objectives Process for Researchers* (in preparation)
EPA QA/G-4HW	*Data Quality Objectives Process for Hazardous Waste Site Testing*
EPA QA/G-5S	*Guidance on Sampling Designs to Support QAPPs* (in preparation)
EPA QA/G-6	*Guidance for the Preparation of Standard Operating Procedures (SOPs) for Quality-Related Documents*
EPA QA/G-9	*Guidance for Data Quality Assessment*
EPA QA/G-9D	*Data Quality Evaluation Statistical Tools (DataQUEST)*

Effective use of this document assumes that appropriate management systems for QA and QC have been established by the implementing organization and are operational.

Questions regarding this document or other Quality System Series documents may be directed to:

U.S. EPA
Quality Assurance Division (8724)
Office of Research and Development
401 M Street, SW
Washington, DC 20460

Phone: (202) 260-5763
Fax: (202) 401-7002

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EXHIBITS

APPENDICES

CHAPTER I

INTRODUCTION

OVERVIEW

This document presents detailed guidance on how to develop a Quality Assurance Project Plan (QAPP) for environmental data operations performed by or for the U.S. Environmental Protection Agency (EPA). This guidance discusses how to address and implement the specifications in *Requirements for QA Project Plans for Environmental Data Operations* (EPA QA/R-5).

The QAPP is the critical planning document for any environmental data collection operation. It documents how quality assurance (QA) and quality control (QC) activities will be implemented during the life cycle of a program, project, or task. The QAPP is the blueprint for how a particular project (and associated technical goals) is integrated into the quality system of the organization performing the work. QA is a system of management activities designed to ensure that the data produced by the operation will be of the type and quality needed and expected by the data user. QA is performed at the management level, with emphasis on systems and policies, and it aids the collection of data of known quality appropriate to support management decisions in a resource-efficient manner.

A project may be viewed as a series of three phases: Planning, Implementation, and Assessment. The QAPP development may be viewed as the transition between the first two phases, Planning and Implementation (see Figure 1). The first phase is the development of Data Quality Objectives (DQOs) using the DQO Process or a similar systematic planning process. The DQOs provide statements about the expectations and requirements of the *data user* (such as the decision maker). In the QAPP, these requirements are translated into measurement performance specifications and QA/QC procedures for the *data suppliers* to provide the information needed to satisfy the data user's needs. See Appendix A for a crosswalk between the outputs of the DQO Process and the inputs of the QAPP. This guidance links the results of the DQO Process with the QAPP in order to complete documentation of the planning process. Once the data have been collected and validated in accordance with the elements of the QAPP, the data should be evaluated to determine whether the DQOs have been satisfied. The final phase, Data Quality Assessment (DQA), involves the application of statistical tools to determine whether the data meet the assumptions made during planning and whether the total error in the data is small enough to support a decision within tolerable decision error rates expressed by the decision maker. Plans for data validation and Data Quality Assessment are discussed in the final sections of the QAPP. Thus, the activities addressed in the QAPP cover the entire project life-cycle, integrating elements of the planning, implementation, and assessment phases.

A QAPP is made up of four sections called "classes," which are further broken into divisions called "elements." The QAPP for a particular project may not require every element to be included. It is expected that some projects may require additional information that is not contained in the elements. This document provides a discussion and background of the elements of a QAPP that will typically be necessary. The final decision on the use of any or all of these elements for project-specific QAPPs will be made by the overseeing or sponsoring EPA organization(s). The Agency encourages the specific tailoring of implementation documents within the EPA's general QA framework.

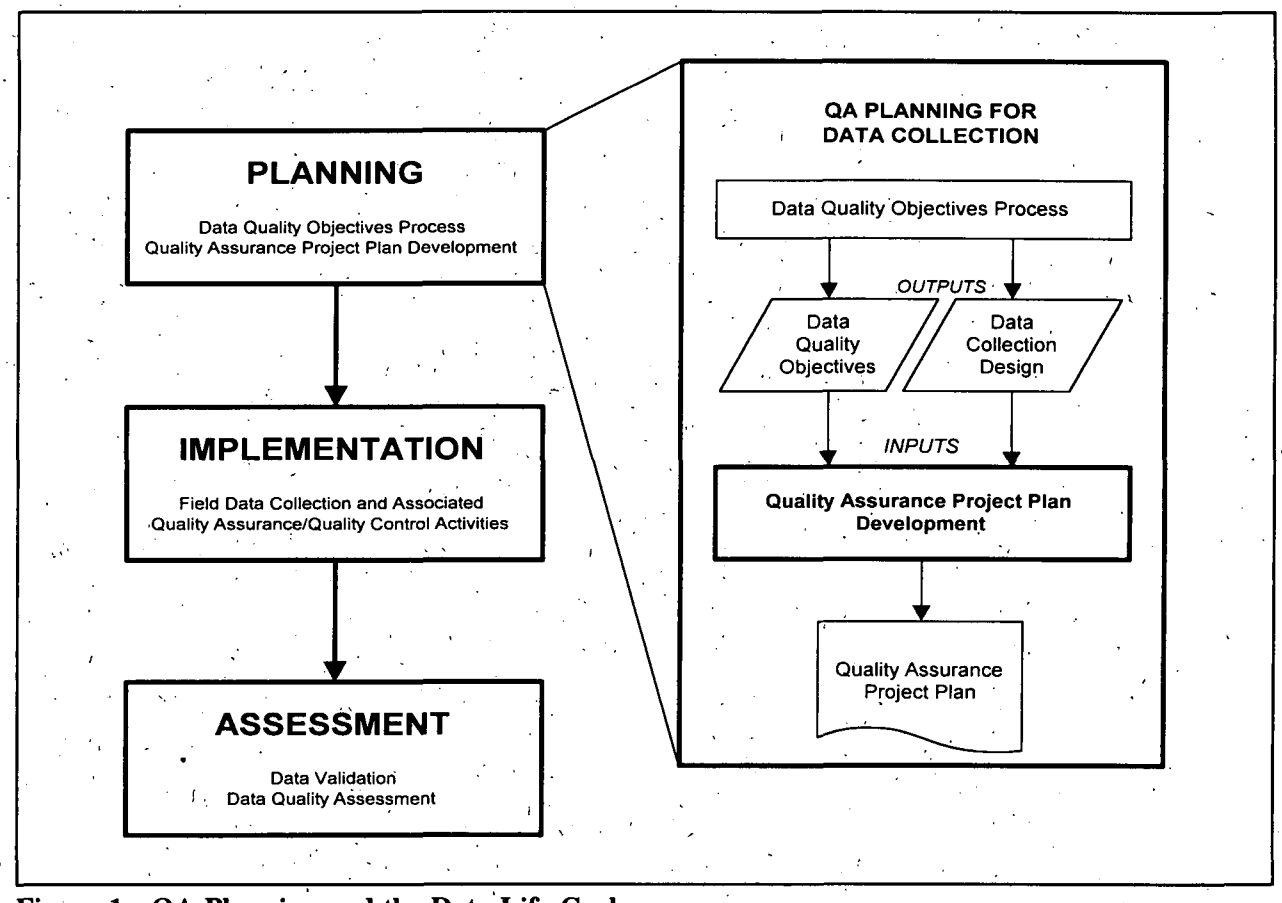


Figure 1. QA Planning and the Data Life Cycle.

PURPOSE OF QA PLANNING

The EPA Quality System is a structured and documented management system describing the policies, objectives, principles, organization, authority, responsibilities, accountability, and implementation plan of an organization for ensuring quality in its work processes, products, and services.

One requirement of the EPA Quality System is that all projects involving the generation, acquisition, and use of environmental data shall be planned and documented and require an Agency-approved QAPP. The primary purpose of the QAPP is to provide an overview of the project, the measurements, and QA/QC system to be applied to the project within a single document. It is detailed enough to provide a clear description of every aspect of the project and include information for every member of the staff including samplers, lab staff, and data reviewers. It facilitates communication among clients, data users, project staff, management, and external reviewers, and assists project management by keeping the projects on schedule and within the resource budget. Because procedural changes may occur at any time during the course of a project, it may be necessary to modify or append the QAPP. A QAPP should be treated as a "living document" throughout the life of the project.

Materials from one QAPP may be used in other QAPPs and applicable materials can be copied as needed into other project documentation. Documents prepared prior to the QAPP (e.g., standard operating procedures, test plans, and sampling plans) can be appended or, in some cases, incorporated by reference. Procedures for revising an approved QAPP are discussed in Chapter IV of this document.

CHAPTER II

QUALITY ASSURANCE PROJECT PLAN REQUIREMENTS

EPA POLICY ON QAPPS

It is EPA policy[1] that the collection of environmental data by the Agency be supported by a mandatory QA program, or Quality System. This requirement also applies to work done for EPA through extramural agreements including 48 CFR, Chapter 15, Part 1546 for contractors, and 40 CFR, Chapter 1, Parts 30 and 31, for financial assistance recipients, negotiated interagency agreements, and consent agreements in enforcement actions.

One part of this mandatory Quality System is the development, review, approval, and implementation of the QAPP. QAPPs are required for all environmental data collection operations involving direct measurements performed by or for the EPA. A QAPP must address all of the elements contained in QA/R-5 unless otherwise specified by the EPA QA Manager responsible for the data collection.

The QAPP is the logical product of the planning process for any data collection. It documents how the QA and QC activities will be planned and implemented to the technical activities of the project. In order to be complete, the QAPP must meet certain specifications for detail and coverage, but the extent of detail is dependant on the type of project, the data to be collected, and the decisions that need to be made. Overall, the QAPP must provide sufficient detail to demonstrate that:

- the project's technical and quality objectives are identified and agreed upon;
- the intended measurements or data acquisition methods are appropriate for achieving project objectives;
- assessment procedures are sufficient for confirming that data of the type and quality needed and expected are obtained; and
- any limitations on the use of the data can be identified and documented.

QAPP CLASSES AND ELEMENTS

The elements of QAPPs are grouped into "classes" according to their function:

Class A: Project Management

This group of QAPP elements covers the general areas of project management, project history and objectives, and roles and responsibilities of the participants. The following ten elements ensure that the project's goals are clearly stated, that all participants understand the goals and the approach to be used, and that project planning is documented:

---

[1]EPA Order 5360.1, "Policy and Program Requirements to Implement the Quality Assurance Program," was issued in April 1984 and established the policy and program requirements for QA at EPA.

A1	Title and Approval Sheet
A2	Table of Contents
A3	Distribution List
A4	Project/Task Organization
A5	Problem Definition/Background
A6	Project/Task Description
A7	Quality Objectives and Criteria for Measurement Data
A8	Project Narrative (ORD Only)
A9	Special Training Requirements/Certification
A10	Documentation and Records

Class B: Measurement/Data Acquisition

This group of QAPP elements covers all of the aspects of measurement system design and implementation, ensuring that appropriate methods for sampling, analysis, data handling, and QC are employed and will be thoroughly documented:

B1	Sampling Process Design (Experimental Design)
B2	Sampling Methods Requirements
B3	Sample Custody Requirements
B4	Analytical Methods Requirements
B5	Quality Control Requirements
B6	Instrument/Equipment Testing, Inspection, and Maintenance Requirements
B7	Instrument Calibration and Frequency
B8	Inspection/Acceptance Requirements for Supplies and Consumables
B9	Data Acquisition Requirements (Non-direct Measurements)
B10	Data Management

Class C: Assessment/Oversight

The purpose of assessment is to ensure that the QAPP is implemented as prescribed. This group of QAPP elements addresses the activities for assessing the effectiveness of the implementation of the project and associated QA/QC:

C1	Assessments and Response Actions
C2	Reports to Management

Class D: Data Validation and Usability

Implementation of Class D elements ensures that the individual data elements conform to the specified criteria, thus enabling reconciliation with the project objectives. This group of QAPP elements covers the QA activities that occur after the data collection phase of the project is completed:

D1	Data Review, Validation, and Verification Requirements
D2	Validation and Verification Methods
D3	Reconciliation with Data Quality Objectives

The specifications for each element are to be found in *Requirements for QA Project Plans for Environmental Data Operations* (EPA QA/R-5). Quotes from that document are contained in a box at the beginning of each specific element.

QAPP RESPONSIBILITIES

QAPPs may be prepared by different groups outside EPA such as contractors, assistance agreement holders, or other Federal agencies under interagency agreements. Generally, all QAPPs prepared by non-EPA organizations should be approved by EPA for implementation. Writing QAPPs is often a collaborative effort within an organization, depending on the technical expertise, writing skills, knowledge of the project, and availability of staff. Organizations are encouraged to involve technical project staff and the QA Manager in this effort to ensure that the QAPP has adequate detail and coverage.

None of the environmental data collection work addressed by the QAPP should be started until the initial QAPP has been approved by the EPA Project Officer and the EPA QA Manager and then distributed to project personnel. In some cases, EPA may grant conditional approval to a QAPP to permit some work to begin while noncritical deficiencies in it are being resolved. However, the QA Manager should be consulted to determine the length of time that work may continue under a conditional QAPP.

The group performing the work is responsible for implementing the approved QAPP. This responsibility includes ensuring that all personnel involved in the work have copies of or access to the approved QAPP along with all other necessary planning documents. In addition, the group must ensure that these personnel understand their requirements prior to the start of data generation activities. Communication among responsible managers is essential to the accurate fulfilment of a QAPP.

CHAPTER III

QAPP ELEMENTS

A PROJECT MANAGEMENT

The following ten Project Management elements provide guidance on the procedural aspects of QAPP development and what to include in the QAPP project background, task description, and quality objectives elements.

A1 TITLE AND APPROVAL SHEET

Organizations and Approving Officials

The title and approval sheet includes the title of the QAPP; the name(s) of the organization(s) implementing the project; and the names, titles, and signatures of the appropriate approving officials, and the signature date. The approving officials typically include:

- the organization's Technical Project Manager,
- the organization's Quality Assurance Officer or Manager,
- the EPA (or other funding agency) Technical Project Manager,
- the EPA (or other funding agency) Quality Assurance Officer or Manager, and
- other key staff, such as QA Officer of prime contractor when a QAPP is prepared by a subcontractor organization.

The purpose of the approval sheet is to have an area where officials can note their approval and commitment to implementing the QAPP. The title and approval sheet should also indicate the date of the revision and a document number, if appropriate.

A2 TABLE OF CONTENTS

List the sections, figures, tables, references and appendices.

The Table of Contents lists all the elements, references, and appendices contained in a QAPP, including a List of Tables and a List of Figures that are used in the text. The major headings for most QAPPs should closely follow the list of required elements; an example is shown in Figure 2.

The Table of Contents of the QAPP should include a document control component in the upper right-hand corner. This information should appear in the upper right corner of each page of the QAPP. For example:

Project No. ____
Element No. ___
Revision No. __
Date: _____
Page ___ of ___

This component, together with the distribution list (see element A3), facilitates control of the document to help ensure that the most current QAPP is in use by all project participants. Each revision of the QAPP should have a different revision number and date.

A3 DISTRIBUTION LIST

List all individuals designated to receive the QAPP.

The Table of Contents should be followed by a Distribution List of all persons designated to receive copies of the QAPP and any future revisions. This list, together with the document control information, will help the project manager ensure that all key personnel have up-to-date copies of the QAPP.

A well planned QA program can best be implemented if those responsible for and engaged in the project work know the contents of the approved QAPP. A typical distribution list appears in Figure 2.

A4 PROJECT/TASK ORGANIZATION

Identify the individuals or organizations participating in the project and discuss their roles and responsibilities.

A4.1 Purpose/Background

The purpose of the project organization is to provide EPA and other involved parties with a clear understanding of the role that each party plays in the investigation or study.

A4.2 Roles and Responsibilities

The specific roles and responsibilities of participants as well as the internal lines of authority and communication within and between organizations should be detailed. The information for this element is best presented graphically as well as in writing. A short narrative about each individual and organization should be included and their involvement in the investigation should be outlined.

A concise chart showing the project organization, the lines of responsibility, and the lines of communication should be presented; an example is given in Figure 3. For complex projects, it may be useful to include more than one chart—one for the overall project with at least the primary contact and others for each organization.

Where direct contact between project managers and data users does not occur, such as between a project consultant for a potentially responsible party and the EPA risk assessment staff, the organization chart should show the route by which information is exchanged.

A5 PROBLEM DEFINITION/BACKGROUND

State the specific problem to be resolved or decision to be made. Include sufficient background information to provide a historical perspective for this particular project.

A5.1 Purpose/Background

The background information provided in this element will place the problem in historical perspective, giving readers and users of the QAPP an appreciation for the project's value and its place relative to other project and program initiatives.

A5.2 Problem Statement and Background

The discussion must include enough information about the problem, the past history, any previous work or data, and any other regulatory or legal context to allow a technically trained reader to make sense of the project objectives and activities. This discussion should include:

- a description of the problem as currently understood, indicating its seriousness and programmatic, regulatory, or research context;
- a summary of existing information on the problem, including any conflicts or uncertainties that are to be resolved by the project;
- a discussion of initial ideas or approaches for resolving the problem that were considered before selecting the approach described in element A6, "Project/Task Description"; and
- the identification of the principal data user or decision maker (if known).

Note that Problem Statement is the first step of the Data Quality Objectives (DQO) Process. This step is discussed in QA Publication EPA QA/G-4.

CONTENTS

Section

Distribution List

N. Wentworth, EPA/ORD (Work Assignment Manager)
B. Waldron, EPA/ORD (QA Manager)
J. Warren, State University (Principal Investigator)
T. Dixon, State University (QA Officer)
G. Johnson, State University (Field Activities)
F. Haeberer, State University (Laboratory Activities)
B. Odom, State University (Data Management)
E. Renard, ABC Laboratories (Subcontractor Laboratory)
P. Lafornara, ABC Laboratories (QA Manager Subcontractor Laboratory)

Figure 2. An Example of a Table of Contents and a Distribution List.

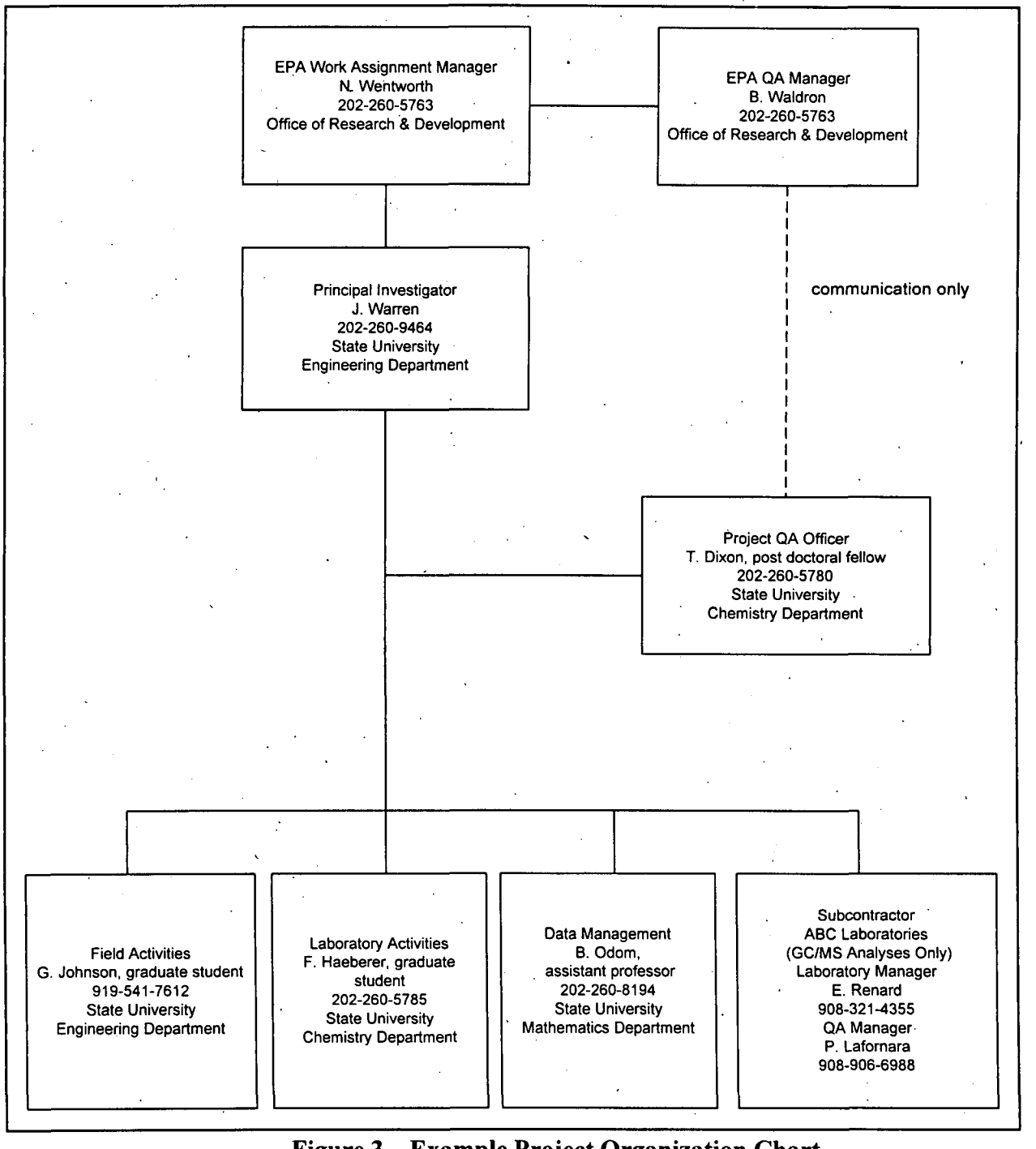


Figure 3. Example Project Organization Chart

A6 PROJECT/TASK DESCRIPTION

Provide a description of the work to be performed. This discussion may not be lengthy or overly detailed, but it should give an overall picture of how the project will resolve the problem or question described in A5.

A6.1 Purpose/Background

The purpose of the Project/Task Description is to provide the participants with a background understanding of the types of project activities to be conducted, the measurements that will be taken, and the associated QA/QC goals, procedures, and timetables for collecting the measurements.

A6.2 Description of the Work to be Performed

(1) **Measurements that are expected during the course of the project.** Describe the characteristic or property to be studied and the measurement processes and the gathering techniques that will be used to collect data. Determine the most appropriate or effective sampling, measurement, and analytical techniques for acquiring the data. Define which measurements are "critical" (ones that will be used to meet the limits on decision errors) and "noncritical" (generally peripheral samples that provide background information).

(2) **Applicable technical quality standards or criteria.** Describe the relevant regulatory standard, criteria, or objectives. If environmental data are collected to test for compliance with a standard, the standard should be cited and the numerical limits should be given in the QAPP. The DQO Process refers to these limits as "action levels," because the type of action taken by the decision maker will depend on whether the measured levels exceed the limit.

(3) **Any special personnel and equipment requirements that may indicate the complexity of the project.** Describe any special personnel or equipment required for the specific type of work being planned or measurements being taken. For example, because of the Occupational Safety and Health Act (OSHA) requirements and depending on the conditions, sometimes, personnel entering a hazardous waste site for field sampling will need safety suits and breathing apparatus.

(4) **The assessment techniques needed for the project.** The degree of quality assessment activity for a project will depend on the project's complexity, duration, and objectives. In general, projects that involve subcontracting for environmental measurement activities and projects that produce data for regulatory or programmatic decision making will be the subject of audits and reviews coordinated with the EPA QA Manager. Examples of assessment techniques include program technical review, peer review, surveillance, technical audits, readiness reviews, and performance evaluation studies. (Refer to Appendix I for definitions of these terms.) Discuss the timing of each planned assessment and briefly outline the roles of the different parties to be involved.

(5) **A schedule for the work performed.** Anticipated start and completion dates for the project should be given. In addition, the discussion should include an approximate schedule of important project milestones, such as the start of environmental measurement activities. Dates for the start of environmental measurement activities should follow the QAPP approval date.

(6) **Project and quality records required, including the types of reports needed.** Environmental studies generate numerous records, including field notebooks, logbooks, custody records, laboratory sample logs, files, and reports. Most of these records are considered routine, and many are detailed in other, more targeted, elements of the QAPP (e.g., B10, "Data Management," discusses the maintenance of electronic records, data entry forms, and checklists). This element of the QAPP should list, in condensed form, the records discussed in elements A10 ("Documentation and Records"), B10 ("Data Management"), and C2 ("Reports to Management"). When applicable, these records will include periodic progress reports, QA audit reports, the project final report, and any planned publications.

A7 QUALITY OBJECTIVES AND CRITERIA FOR MEASUREMENT DATA

The QAPP must include a statement of the project quality objectives.

A7.1 Purpose/Background

The purpose of this element is to document the quality objectives for the measurement data that will be generated under the project and establish performance criteria for the measurement system that will be employed in generating the data.

A fundamental principle underlying the EPA Quality System is that data quality must be defined, specified, and documented by the data user. By clarifying the intended use of the data, and specifying qualitative and quantitative criteria for how well the measurement system and the data set as a whole should perform, this element establishes the critical link between the needs of the data user and the performance requirements to be placed on the data generator.

A7.2 Specifying Quality Objectives

This element of the QAPP should define what quality the final results of the study should have in order to satisfy the needs of the data user. The Agency strongly recommends the DQO Process, a systematic procedure for planning data collection activities, to ensure that the right type, quality, and quantity of data are collected to satisfy the data user's needs. DQOs are qualitative and quantitative statements that:

- clarify the intended use of the data,
- define the type of data needed to support the decision,
- identify the conditions under which the data should be collected, and
- specify tolerable limits on the probability of making a decision error due to uncertainty in the data.

Figure 4 shows the seven steps of the DQO Process, which is explained in detail in EPA QA/G-4, QA/G-4D, QA/G-4R, and QA/G-4S. In addition, Appendix A.4 provides a crosswalk between the requirements of the QAPP and the DQO outputs.

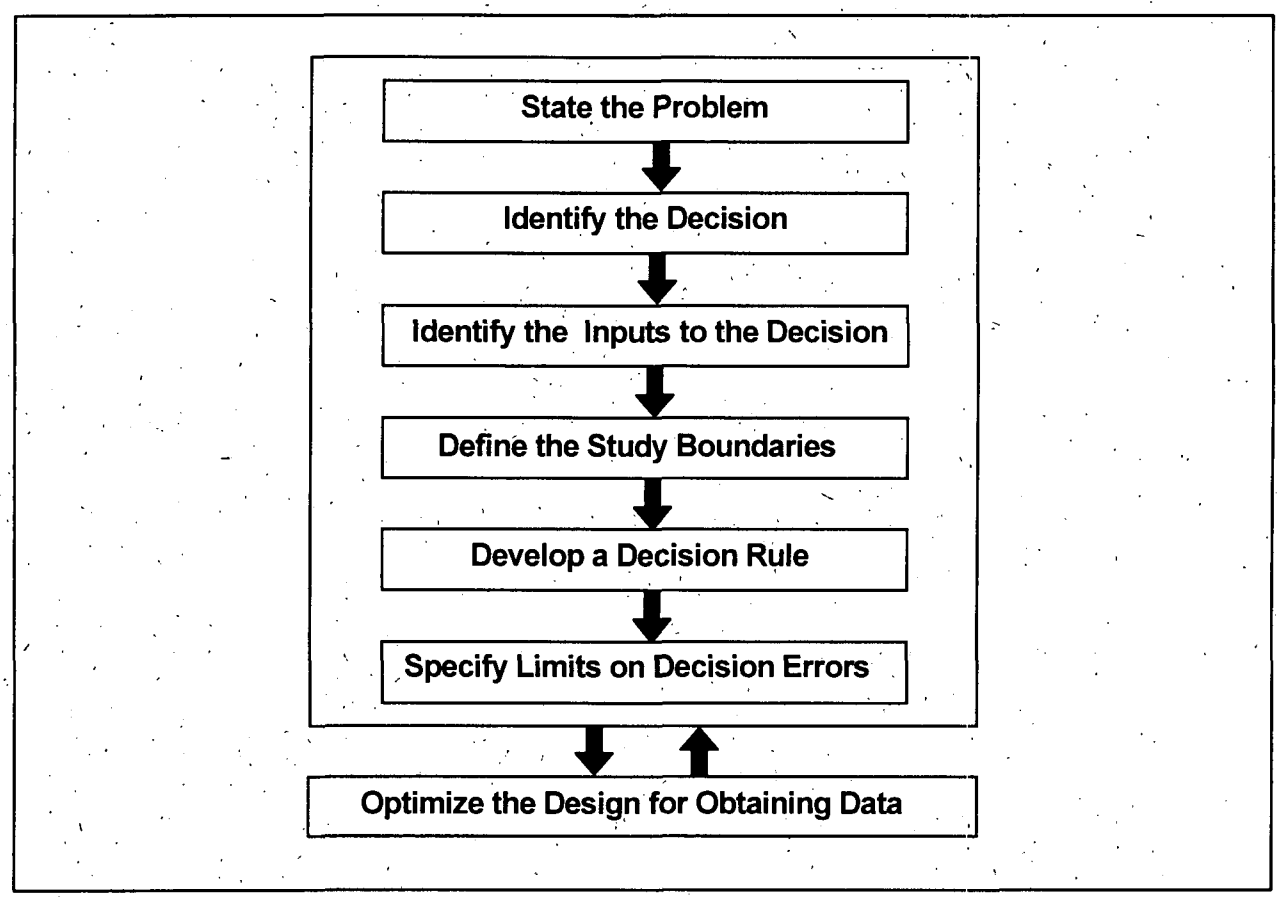


Figure 4. The Data Quality Objectives Process.

For exploratory research, sometimes the goal is to develop questions that may be answered by subsequent work. Therefore, researchers may modify activities advocated in QA/G-4 to define decision errors (see EPA QA/G-4R *Data Quality Objectives for Researchers).*

A7.3 Specifying Measurement Performance Criteria

While the quality objectives state what the data user's needs are, they do not provide sufficient information about how these needs can be satisfied. The specialists who will participate in generating the data need to know the measurement performance criteria that must be satisfied to achieve the overall quality objectives. One of the most important features of the QAPP is that it links the data user's quality objectives to verifiable measurement performance criteria. Although the level of rigor with which this is done and documented will vary widely, this linkage represents an important advancement in the implementation of QA; Appendices F and G discuss this topic further. Once the measurement performance criteria have been established, sampling and analytical methods criteria can be specified in Part B.

A8 PROJECT NARRATIVE

The narrative should allow technical or QA readers to relate the project or task to the DQOs and to the Problem Definition given earlier in the QAPP.

A8.1 Purpose/Background

Some areas within the Agency prefer to categorize their projects I through IV (see Appendix B). Category IV projects involve environmental data operations to study basic phenomena or issues, including proof of concept, feasibility studies, and qualitative screening for a particular analytical species. For example, extramural work funded under the Office of Research and Development's (ORD's) Research Grants Program are often Category IV projects. This element may be omitted if deemed non-applicable.

A8.2 Project Narrative

EPA recognizes that Category IV projects may require more flexibility in QA requirements, due to the exploratory nature of this type of project. The only recommended elements for inclusion in a Category IV QAPP are the title and approval sheet, the distribution list, and the project narrative. The project narrative covers many of the QAPP elements, but in a level of detail that is more appropriate for the nature of these projects. The project narrative is not needed for Category I, II, or III QAPPs since it overlaps with the other elements included in greater detail in those QAPPs.

The following issues are appropriate for inclusion in most Category IV project narratives and should be discussed in narrative form, if relevant to the project:

- project/task organization (A4),
- work to be performed or hypothesis to be tested (A5, A6),
- anticipated use of the data (A5, A6),
- how the success of the project will be determined (A7),
- survey design requirements and description (B1),
- sample type and sampling location requirements (B2),
- sample handling and custody requirements (B3),
- selection of analytical methods (B4),
- calibration and performance evaluation samples for sampling and analytical methods used (B5, B7),
- sampling or analytical instrumentation requirements (B6),
- plans for peer or readiness reviews prior to data collection (C1),
- any on-going assessments during actual operation (oversight) (C1), and
- reconciliation with DQOs or other objectives (D3).

References to other elements of this guidance that provide details on these issues are given in parentheses. As always, topics should be addressed appropriately for the goals and intended data use for the particular project. The EPA QA Officer can offer assistance to clarify which topics need to be addressed on any particular study.

If the project employs methods that have not been validated for the intended application, the QAPP should include information about the intended procedure, how it will be validated, and what criteria must be met before it is accepted for the application.

A9 SPECIAL TRAINING REQUIREMENTS/CERTIFICATION

Identify and describe any specialized training or certification requirements for personnel in order to successfully complete the project or task. Discuss how such training will be available.

A9.1 Purpose/Background

The purpose of this element is to ensure that the training requirements necessary to complete the projects are known and furnished and the procedures are described to ensure that proper training skills can be verified, documented, and updated as necessary. This element should define any specialized training or certification requirements for personnel in order to successfully complete the project or task. The discussion should also show how specialized knowledge and skills acquired through the training program will be retained should changes to personnel occur. This element of the QAPP should cover both voluntary training programs set up through organizational management and specialized training mandated through project-specific requirements.

A9.2 Training

Training of employees may be accomplished by the following:

(a) *A system of training where the aspects of training, including quality standards, are defined and included in a training checklist.* The employee's work is immediately rechecked for errors or defects and the information is fed back instantaneously for corrective action. Personnel who have an impact on quality (e.g., calibration, maintenance, and analytical staff) are trained in the reasons for and benefits of standards of quality and the methods by which high quality is to be achieved. Personnel training accomplishments are documented in written records and periodically reviewed by management. Personnel proficiency is evaluated on a continuing basis and the results used by management to establish the need for and type of special training.

(b) *On-the-job training by the supervisor who gives an overview of quality standards.* Details of quality standards are learned as normal results are fed back to the employee. Personnel who have an impact on quality are told about quality only when work falls below certain levels. Personnel training accomplishments are documented in written records periodically reviewed by management.

(c) *On-the-job learning with training on the rudiments of the job by senior coworkers.* Personnel who have an impact on quality are approached when quality deficiencies are directly traceable to their work. Proficiency is based on observation of performance by management.

Depending on the nature of the environmental data operation, the QAPP may need to address compliance with specifically mandated training requirements. For example, contractors or employees working at Resource Conservation and Recovery Act (RCRA) regulated facilities or a Superfund site need specialized training as mandated by the OSHA regulations. If hazardous materials are moved offsite after samples have been taken, a project team may need to comply with the training requirements for shipping hazardous materials as mandated by the Department of Transportation in association with the International Air Transportation Association. This element of the QAPP should show that the management and project teams are aware of specific health and safety needs as well as any other organizational safety plans, such as the plans developed by contractors.

A9.3 Certification

Usually, the organizations participating in the project are responsible for conducting training and health and safety programs and ensuring certification. Various commercial training courses are available that meet some government regulations. Training and certification should be planned well in advance for necessary personnel prior to the implementation of the project.

All certificates or documentation representing completion of specialized training should be maintained in personnel files and copies incorporated into the project data reporting package.

A10 DOCUMENTATION AND RECORDS

Itemize the information and records which must be included in a data report package for the project or task, and specify the reporting format, if desired.

A10.1 Purpose/Background

The purpose of this element is to define what records are critical to the project and what information needs to be included in reports, as well as the report format and the document control procedures for both data and reports. The proper reporting format will facilitate clear, direct communication of the investigation and its conclusions and be a resource document for the design of future studies.

A10.2 Information Included in the Reporting Packages

The selection of which records to include in a data reporting package must be determined based how the data will be used. Different "levels of effort" require different supporting QA/QC documentation. For example, organizations conducting basic research need different reporting requirements from organizations collecting data in support of litigation. Information such as blank forms and custody labels should be included as figures and appendices in the QAPP.

A10.2.1 Field Operation Records

The information contained in these records will document overall field operations and generally consist of:

- *Sample collection records*. These records show that proper sampling protocol was performed in the field. At a minimum, this documentation should include the names of the persons conducting the activity, sample number, sample collection points, maps and diagrams, equipment/method used, climatic conditions, and unusual observations. Bound field notebooks, pre-printed forms, or computerized notebooks can serve as the recording media. Bound field notebooks are generally used to record raw data and make references to prescribed procedures and changes in planned activities. They should be formatted to include pre-numbered pages with date and signature lines.

- *Chain-of-custody records*. Chain-of-custody records document the progression of samples as they travel from the original sampling location to the laboratory and finally to their disposal area. These records should contain the project name, the signature of the sample collector, the sample number, the date and time of collection, the nature of the sample, and the signatures of individuals involved in the transfer of samples from one project event to another. (See Appendix C.1 for an example of a chain-of-custody checklist.)

- *Quality control sample records*. Quality control sample records document the generation of quality control samples, such as field, trip, and equipment rinsate blanks and duplicate samples and include documentation on sample integrity and preservation. These records should also include calibration and standards' traceability documentation that will be used to provide a reproducible reference point to which all sample measurements can be correlated. Quality control sample records should contain information on the frequency, conditions, level of standards, and instrument calibration history. This quality control information could be provided in a chart format.

- *Personnel files*. Personnel files record the names and training certifications of the staff that collected data.

- *General field procedures*. General field procedures record the procedures that were used in the field to collect data.

- *Corrective action reports*. Corrective action reports show what methods were used in cases where general field practices or other standard procedures were violated for any reasons.

If applicable, to show regulatory compliance in disposing of waste generated during the data operation, the procedures manifest and testing contracts should also be included in the field record section of the data reporting package.

A10.2.2 Laboratory Records

The QAPP must document all laboratory activities that may affect data quality. In addition to continuing the documentation of records initiated during field operations (e.g., sample custody), laboratory personnel must document activities unique to their responsibilities. The following list describes some of the laboratory-specific documentation that should be included in the data reporting package if available and appropriate.

- *Data Reporting Turnaround Time*. These records note the time that samples were analyzed to verify that they met the time deadlines.

- *Sample Management Records.* Sample management records will document sample receipt, handling and storage, and scheduling of analysis. The records will verify that the chain-of-custody and proper preservation has been maintained, reflect any anomalies in the samples (such as receipt of damaged samples), note proper log-in of samples into the laboratory, and address procedures used to ensure that holding time requirements were met.

- *Test Methods.* Unless analyses are performed exactly as prescribed by standard methods, this documentation will show how the analyses were carried out in the laboratory and note any deviations. This documentation includes sample preparation and analysis, instrument standardization, detection and reporting limits, and test-specific quality control criteria. Documentation demonstrating laboratory proficiency with each method used should also be a part of the data reporting package.

- *Data Handling Record.* This record documents a prescribed protocol for reducing field measurement data or the method of measurement appropriate for the data operation. When computer programs are used for data reduction, documentation should show validation of the program before use and on a regular basis. Hard copies of information downloaded from computerized field notebooks and backup diskettes would also be included in the data reporting package.

A10.3 Data Reporting Package Format and Documentation Control

All individual records that represent actions taken to achieve the objective of the data operation and the performance of specific quality assurance functions are potential components of the final data reporting package. This element of the QAPP should discuss how these various components will be assembled to represent a concise and accurate record of all activities impacting data quality. The discussion should detail the recording medium for the project, guidelines for hand-recorded data (e.g., using indelible ink), procedures for correcting data (e.g., single line drawn through errors and initial by the responsible person), and documentation control. Procedures for making revisions to technical documents should be clearly specified and the line of authority indicated.

A10.4 Data Reporting Package Archiving and Retrieval

The length of storage for the data reporting package may be governed by regulatory requirements, organizational policy, or contractual project requirements. This element of the QAPP should note the governing authority for both storage and final disposal. In describing how the records will be maintained, the discussion should address how to store hard copy records to minimize deterioration over the expected period, how to use computer systems for expedient information retrieval, how to maintain a system that offers security by limiting access to records, and how to document access to the records.

A10.5 Reference

Kanare, Howard M. 1985. *Writing the Laboratory Notebook.* Washington, DC: American Chemical Society.

B MEASUREMENT/DATA ACQUISITION

B1 SAMPLING PROCESS DESIGN (EXPERIMENTAL DESIGN)

Outline in general terms the experimental design of the project and the anticipated project activities, including the types of samples required, sampling network design, sampling frequencies, sample matrices, measurement parameters of interest, and the rationale for the design.

Describe techniques or guidelines to be followed in selecting sampling points and frequencies, well installation design (when applicable), field decontamination procedures and materials needed, and sampling equipment.

B1.1 Purpose/Background

The purpose of this element is to describe all the relevant components of the experimental design and indicate the number of samples expected and where, when, and how samples are to be taken. For example, the characterization of a wastewater-treatment effluent stream on a particular day might be accomplished by collecting one-liter samples from mid-stream at one-hour intervals throughout the day. Strategies such as stratification, compositing, and clustering should be discussed and diagrams or maps showing sampling points should be included.

In addition to describing the design, this element of the QAPP should include discussions on the following subjects:

- scheduled project activities,
- rationale for the design (in terms of meeting DQOs),
- design assumptions,
- procedures for locating and selecting environmental samples,
- classification of measurements as critical or noncritical, and
- validation of any nonstandard methods.

The sub-elements that follow (B1.2 through B1.8) address these subjects.

B1.2 Scheduled Project Activities, Including Measurement Activities

This element of the QAPP should give anticipated start and completion dates for the project as well as anticipated dates for major milestones, such as:

- schedule of sampling and analysis events,
- schedule for phases of sequential sampling (or testing) if applicable,
- schedule of test runs, and
- schedule for peer review activities.

B1.3 Rationale for the Design

The objectives for an environmental study should be formulated prior to designing the study and include:

- a definition of the characteristic of the population of interest,
- a discussion of whether the population parameter can change over time,
- the relationship of the parameter to relevant thresholds,
- a discussion of the potential range of the parameter, and
- an evaluation of the potential effects of uncontrollable factors.

The rationale of the design should directly address the objectives of the study. For example, when estimating a mean, it is important that samples are chosen in such a way as to be representative of the entire population. This is often best accomplished when samples are chosen in some random design with all parts of the population having some chance of being selected (see Appendix H and EPA QA/G-5S).

It is always useful to test whether the population of interest has been completely and unambiguously specified. If this population is open to different interpretations, there will be a good chance of a very common error—producing the right answer to the wrong question. Agreement on the boundaries and constituents of the population of interest is essential.

When the intended use of the experimental data is hypothesis testing (e.g., to test whether some threshold concentration has been exceeded), quantitative project objectives are usually expressed in terms of the design's ability to achieve prescribed false positive and false negative error rates. When the intended use of experimental data is estimating some characteristic of the environment, the quantitative project objectives are often expressed in terms of a desired confidence or probability interval width. For either of these cases, investigators should give evidence or references that the proposed design is expected to satisfy the DQOs, provided that design assumptions are valid. (See also B1.4 "Design Assumptions.")

B1.4 Design Assumptions

This element of the QAPP should discuss assumptions about the magnitude and structure of measurement error and the population variability that are an inherent part of the sampling design. This element should answer the following questions:

- Are the data expected to be relatively free of bias (sampling and analytical) and representative of the medium being investigated?

- Is the random component of measurement error constant or some other function of the measured value (e.g., characterized by *relative* standard deviation)?

- Where does information on bias and variance come from (e.g., from the data alone, from prior information alone, or from some combination of the two)?

- Are measurement error and sampling error expected to be normal (Gaussian) or log normal, or will a mathematical transformation be required?

- What are the largest components of total variability? (If a pilot study is planned to validate the assumptions about bias and variability, or to provide preliminary estimates of bias and variance components, then the study should be described.)

- To what extent will correlation issues influence the data?

EPA QA/G-5S provides nonmandatory guidance on the practicality of constructing sampling plans to meet the guidelines outlined in the DQO Process. Refer to Appendix D for a detailed discussion of bias, Appendix E for a discussion of error in the measurement process, and EPA QA/G-9 for a discussion on the effects of violations of assumptions on decision making.

B1.5 Procedures for Locating and Selecting Environmental Samples

The best plan for a particular sampling application will depend on issues of practicality and feasibility (e.g., determining specific sampling locations), the population characteristic to be estimated (e.g., with respect to the contaminant and physical matrix, can the samples be composited?) and implementation costs (e.g., the costs of sample collection, transportation, and analysis).

Depending on the population matrix, this element of the QAPP should also describe sample port locations and traverses (for emissions source testing), well installation designs (for ground water investigations), field decontamination procedures, and sampling materials. Sometimes decisions on the number and location of samples will be made in the field; therefore, the QAPP should describe how these decisions will be driven by observations or by field screening data. When locational data are to be collected, stored, and transmitted, the methodology used must be specified and described (or referenced) and include:

- procedures for finding prescribed sample locations,
- contingencies for cases where prescribed locations are inaccessible, and
- location bias and its assessment.

EPA QA/G-5S provides nonmandatory guidance on the practicality of constructing sampling plans and references to alternative sampling procedures.

B1.6 Classification of Measurements as Critical or Noncritical

All measurements should be classified as critical (i.e., required to achieve project objectives) or noncritical (informational purposes only). Critical measurements will undergo closer scrutiny during the review and the data gathering process, and will have first claim on limited budget resources. A simple way to identify which measurements are critical is to annotate a table listing analytical procedures and measurement objectives. It is also possible to include the expected number of samples to be tested by each procedure and the acceptance criteria for QC checks (as described in element B5, "Quality Control Requirements").

B1.7 Validation of Any Non-Standard Methods

For nonstandard sampling methods or unusual sample matrices and situations, appropriate method validation study information is needed to confirm the performance of the method for the

particular matrix. Such validation studies may include round-robin studies performed by EPA or other organizations. If previous validation studies are not available, some level of single-user validation study or ruggedness study should be performed during the project and included as part of the project's final report. This element of the QAPP should clearly reference any available validation study information.

B2 SAMPLING METHODS REQUIREMENTS

Describe the procedures for collecting samples. Identify the required sampling methods (and/or equipment, if automated), including any implementation requirements, decontamination procedures and methods needed, and any specific performance requirements for the method.

B2.1 Purpose/Background

Environmental samples should reflect the population and parameters of interest (see the discussion on representativeness, Appendix H). As with all other considerations involved with environmental measurements, sampling methods should be chosen with respect to the intended application of the data. Just as methods of analysis vary in accordance with project needs, sampling methods can also vary according to these requirements. Different sampling methods have different operational characteristics, such as cost, difficulty, and necessary equipment. In addition, the sampling method can materially affect the representativeness, comparability, bias, and precision of the final analytical result.

In the area of environmental sampling, there exists a great variety of sample types and it is beyond the scope of this document to provide detailed advice for each sampling situation and sample type. Nevertheless, it is possible to define certain common elements that are pertinent to many sampling situations with discrete samples.

If a separate sampling and analysis plan has been created for the project, it should be included as an appendix to the QAPP. The QAPP should simply refer to the appropriate portions of the sampling and analysis plan for the pertinent information and not reiterate information.

B2.2 Describe the Sample Collection, Preparation, and Decontamination Procedures

(1) *Identify appropriate sampling methods from the EPA compendia of methods relating to the most important media and sampling scenarios.* When EPA-sanctioned procedures are not available, standard procedures from other organizations and disciplines may be used. A complete description of non-EPA methods should be provided in (or attached to) the QAPP because reviewers and project personnel may not have ready access to them.

(2) *Identify sampling methods' requirements.* Having identified appropriate and applicable methods, it is necessary to determine the requirements of each method. If there is more than one acceptable sampling method applicable to a particular situation, it is necessary to choose among them. DQOs should be considered in choosing these methods to ensure that a) the sample accurately represents the portion of the environment to characterized; b) the sample is of

sufficient volume to support the planned chemical analysis; and c) the sample remains stable during shipping and handling.

(3) *Identify sampling methods' preparation procedures including equipment set-up, calibrations, removal of extraneous overburden, etc.* The investigator may find that other preparations may also be necessary when setting up remote sensing or noninvasive measurement procedures.

(4) *Identify sampling methods, decontamination procedures, and decontamination materials.* Decontamination is primarily applicable in the situation of sample acquisition from solid, semi-solid, or liquid media, but should be addressed, if applicable, for continuous monitors as well. Existing EPA documents provide guidance for decontamination and related procedures for the various media. The investigator must consider the appropriateness of the decontamination procedures to the project at hand. For example, if contaminants are present in the environmental matrix at the 1% level, it is probably unnecessary to clean sampling equipment to parts-per-billion (ppb) levels. Conversely, if ppb-level detection is required, rigorous decontamination or the use of disposable equipment is called for.

(5) *Describe procedures for disposal of decontamination byproducts, if applicable.* Disposal of the rinsates and other byproducts of decontamination can be trivial or very complex depending on the situation. For example, sampling of radioactive mixed wastes is a case in which the decontamination byproducts may themselves be hazardous. Good scientific or engineering judgment should be applied in the disposal of wastes. There may also be a variety of applicable rules and regulations that would pertain to a particular situation, such as the regulations of OSHA, the Nuclear Regulatory Commission (NRC), and state and local governments.

(6) *Define sampling methods' performance requirements.* Aggregate error in a measurement consists of several components, one of the most important of which is sampling method performance. Investigators should examine the feasibility of the proposed sampling method's ability to achieve the level of performance demanded by the DQOs.

Each medium or contaminant matrix has its own characteristics that define the method performance and the type of material to be sampled. Investigators should address:

- actual sampling locations,
- choice of sampling method/collection,
- delineation of a properly shaped sample,
- inclusion of all particles within the volume sampled, and
- correct subsampling to reduce the representative field sample into a representative laboratory aliquot.

A full theoretical discussion of these issues is to be found in *Pierre Gy's Sampling Theory and Sampling Practice* (see references in section B2.6).

B2.3 Identify Support Facilities for Sampling Methods

Support facilities vary widely in their capabilities, from percentage-level analyses capability, to others oriented toward ppb levels. The investigator must determine the required capabilities of the support facilities with respect to the DQOs established in the planning phase.

B2.4 Define Sampling/Measurement System Corrective Action.

This section should address issues of responsibility for the quality of the data, methods for making changes and corrections, the criteria for the decision of a new sample location, and how this change will be documented.

B2.5 Describe Sampling Equipment, Preservation, and Holding Time Requirements.

Characteristics of appropriate sampling equipment include:

(1) *Appropriate material of construction to prevent sample contamination.* The sampler material must either be easy to decontaminate or should be partially or completely disposable.

(2) *The volume of material collected must be sufficient to minimize field fluctuations and errors.* Often this will mean that a large sample will be taken in the field, only to be subsampled to a much smaller amount for analysis by documented compositing techniques.

(3) *Sample preservation methods must be reviewed to ensure that none of the other target analytes would experience interferences.*

(4) *Holding times for extraction and analysis must also be defined to ensure the integrity of the samples against degradation.* Failing to achieve the holding times can potentially affect the variability of measurements.

B2.6 References

Solid and Hazardous Waste Sampling

U.S. Environmental Protection Agency. 1986. *Test Methods for Evaluating Solid Waste (SW-846).* 3rd Ed., Chapter 9.

U.S. Environmental Protection Agency. 1985. *Characterization of Hazardous Waste Sites - A Methods Manual.* Vol. I, *Site Investigations.* EPA-600/4-84-075. Environmental Monitoring Systems Laboratory. Las Vegas, NV.

U.S. Environmental Protection Agency. 1984. *Characterization of Hazardous Waste Sites - A Methods Manual.* Vol. II, *Available Sampling Methods.* EPA-600/4-84-076. Environmental Monitoring Systems Laboratory. Las Vegas, NV.

U.S. Environmental Protection Agency. 1987. *A Compendium of Superfund Field Operations Methods.* NTIS PB88-181557. EPA/540/P-87/001. Washington, DC.

Ambient Air Sampling

U.S. Environmental Protection Agency. *Quality Assurance Handbook for Air Pollution Measurement Systems.* Vol. I, *Principles.* EPA 600/9-76-005. Section 1.4.8 and Appendix M.5.6.

U.S. Environmental Protection Agency. *Quality Assurance Handbook for Air Pollution Measurement Systems.* Vol. II, EPA 600/4-77-27a. Sections 2.0.1 and 2.0.2 and individual methods.

U.S. Environmental Protection Agency. 1984. *Compendium of Methods for the Determination of Toxic Organic Compounds in Ambient Air*. EPA/600-4-84-41. Environmental Monitoring Systems Laboratory. Research Triangle Park, NC. *Supplement:* EPA-600-4-87-006. September 1986.

Source Testing (Air)

U.S. Environmental Protection Agency. *Quality Assurance Handbook for Air Pollution Measurement Systems*. Vol. III, EPA 600/4-77-27b. Section 3.0 and individual methods.

Acid Precipitation

U.S. Environmental Protection Agency. *Quality Assurance Handbook for Air Pollution Measurement Systems*. Vol. V, EPA 600.

Meteorological Measurements

U.S. Environmental Protection Agency. *Quality Assurance Handbook for Air Pollution Measurement Systems*. Vol. IV, EPA 600.

Radioactive Materials and Mixed Waste

U.S. Department of Energy. 1989. *Radioactive-Hazardous Mixed Waste Sampling and Analysis: Addendum to SW-846*. September.

Soils and Sediments

U.S. Environmental Protection Agency. 1985. *Sediment Sampling Quality Assurance User's Guide*. NTIS PB85-233542. EPA/600/4-85/048. Environmental Monitoring Systems Laboratory. Las Vegas, NV.

U.S. Environmental Protection Agency. 1989. *Soil Sampling Quality Assurance User's Guide*. EPA/600/8-89/046. Environmental Monitoring Systems Laboratory. Las Vegas, NV.

Barth, D.S., and T.H. Starks. 1985. *Sediment Sampling Quality Assurance User's Guide*. EPA/600-4-85/048. Prepared for Environmental Monitoring and Support Laboratory. Las Vegas, NV. July.

Statistics, Geostatistics, and Sampling Theory

Ingamells, C.O., and F.F. Pitard. 1986. *Applied Geochemical Analysis*. New York: Wiley-Interscience,

Pitard, F.F. 1989. *Pierre Gy's Sampling Theory and Sampling Practice*. Vol I and II. Boca Raton, FL: CRC Press.

Miscellaneous

American Chemical Society Joint Board/Council Committee on Environmental Improvement. 1990. *Practical Guide for Environmental Sampling and Analysis, Section II. Environmental Analysis*.

ASTM Committee D-34. 1986. *Standard Practices for Sampling Wastes from Pipes and Other Point Discharges*. Document No. D34.01-001R7. October.

Keith, L. 1990. *EPA's Sampling and Analysis Methods Database Manual.* Austin, TX: Radian Corp.

B3 SAMPLE CUSTODY REQUIREMENTS

Describe the provisions for sample handling and shipment, taking into account the nature of the samples and the maximum allowable sample holding times before extraction or analysis.

B3.1 Purpose/Background

This element of the QAPP should clearly describe all procedures that are necessary for ensuring that:

- samples are collected, transferred, stored, and analyzed by authorized personnel;
- sample integrity is maintained during all phases of sample handling and analyses; and
- an accurate written record is maintained of sample handling and treatment from the time of its collection through laboratory procedures to disposal.

Proper sample custody minimizes accidents by assigning responsibility for all stages of sample handling and ensures that problems will be detected and documented if they occur. A sample is in custody if it is in actual physical possession or in a secured area that is restricted to authorized personnel.

B3.2 Sample Custody Procedure

The QAPP should discuss the sample custody procedure having the following steps at a level commensurate with the intended use of the data:

1) List the names and responsibilities of all sample custodians in the field and laboratories.

2) Give a description and example of the sample numbering system.

3) Define acceptable conditions and plans for maintaining sample integrity in the field prior to and during shipment to the laboratory (e.g., proper temperature and preservatives).

4) Give examples of forms and labels used to maintain sample custody and document sample handling in the field and during shipping. An example of a sample log sheet is given in Figure 5; an example sample label is given in Figure 6.

5) Describe the method of sealing of shipping containers with chain-of-custody seals. An example of a seal is given in Figure 7.

6) Describe procedures that will be used to maintain chain-of-custody and document sample handling during transfer from the field to the laboratory, within the laboratory, and among contractors. An example of chain-of-custody record is given in Figure 8.

7) Provide for the archiving of all shipping documents and associated paperwork.

8) Discuss procedures that will ensure sample security at all times.

9) Describe procedures for within-laboratory chain-of-custody together with verification of printed name, signature, and initials of the personnel responsible for custody of samples, extracts, or digests during analysis at the laboratory. Finally, document disposal or consumption of samples. A chain-of-custody checklist is included in Appendix C.3 to aid in managing this element.

Minor documentation of chain-of-custody procedures is generally applicable when:

1) samples are generated and immediately tested within a facility; and

2) continuous rather than discrete or integrated samples are subjected to real- or near-real-time analysis (e.g., continuous monitoring).

B4 ANALYTICAL METHODS REQUIREMENTS

Identify the analytical methods and/or equipment required, including any extraction methods needed, laboratory decontamination procedures and materials needed (such as in the case of hazardous or radioactive samples), waste disposal requirements (if any), and any specific performance requirements for the method. The QAPP should also address what to do if there is a failure in the analytical system and who is responsible for corrective action.

B4.1 Purpose/Background

The choice of methods will be influenced by performance criteria (as defined by project data quality indicator goals for bias, precision, and limits of detection), Data Quality Objectives, and possible regulatory or document-driven criteria. Qualification requirements may range from functional group identification only to complete individual specification. Quantification needs may range from only order-of-magnitude quantities to parts-per-trillion concentrations.

The matrix containing the subject analytes often dictates the sampling and analytical methods. Gaseous analytes often must be concentrated on a trap in order to collect a measurable quantity. If the matrix is a liquid or a solid, the analytes usually must be separated from it using various methods of extraction. Sometimes the analyte is firmly linked by chemical bonds to other elements and must be subjected to digestion methods to be freed for analysis.

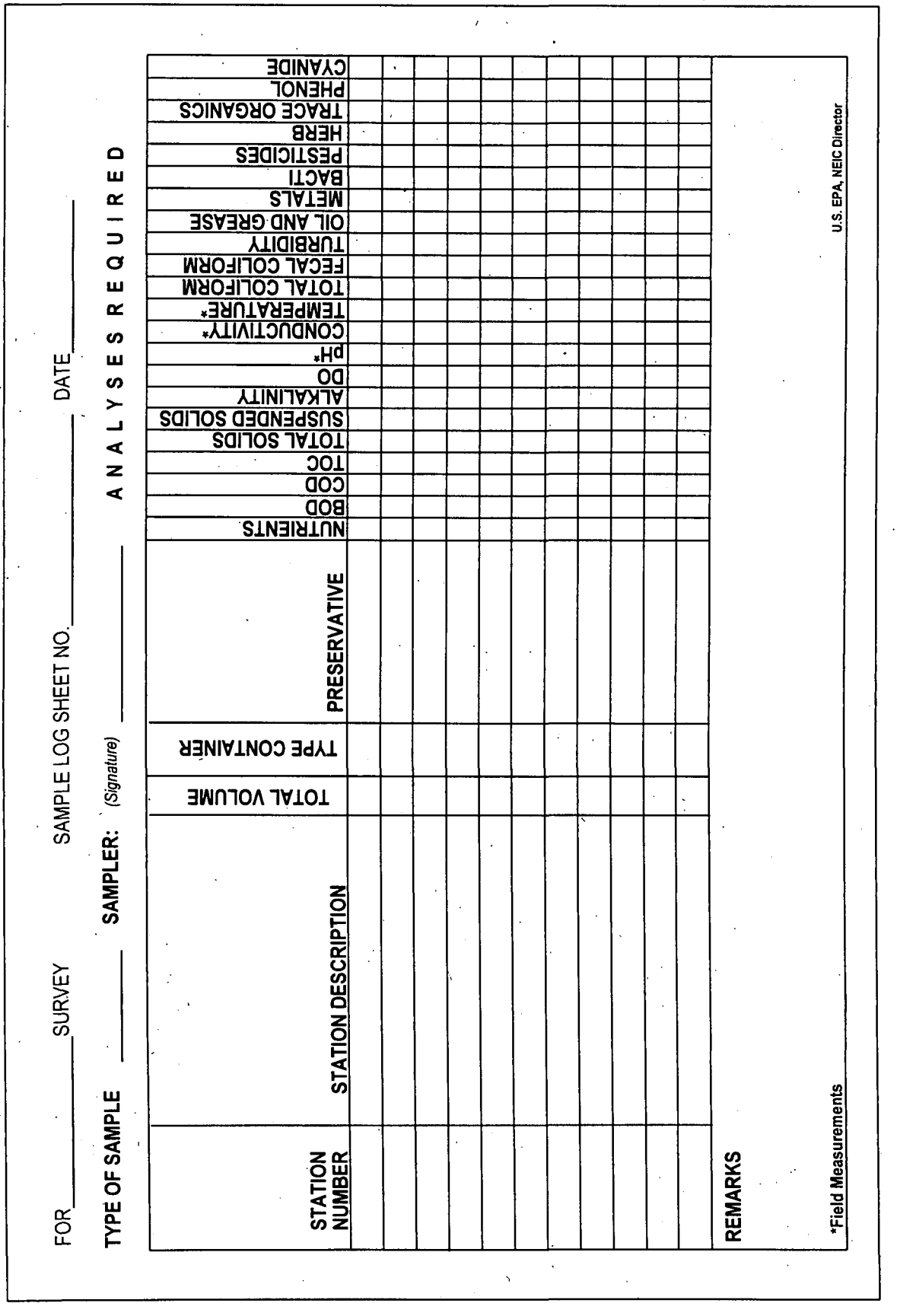


Figure 5. An Example of a Sample Log Sheet

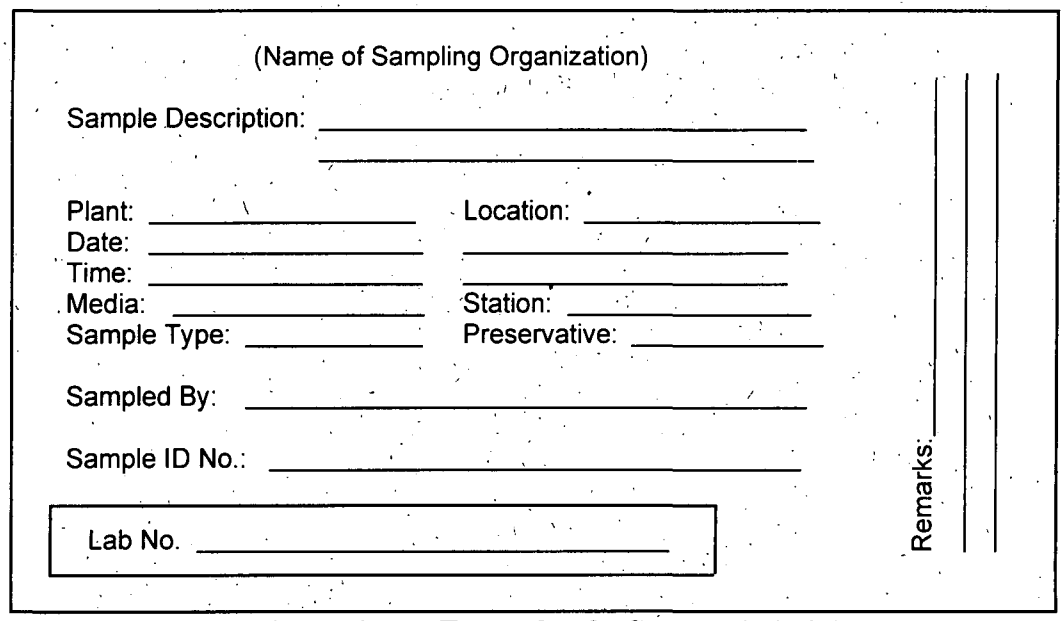


Figure 6. An Example of a Sample Label

Figure 7. An Example of a Custody Seal

CHAIN OF CUSTODY RECORD

				SAMPLERS *(Signature)*							
STATION NUMBER	STATION LOCATION	DATE	TIME	SAMPLE TYPE			SEQ NO.	NO. OF CONTAINERS	ANALYSIS REQUIRED		
				WATER		AIR					
				Comp	Grabx						
Relinquished by: *(Signature)*				Received by: *(Signature)*						DATE/TIME	
Relinquished by: *(Signature)*				Received by: *(Signature)*						DATE/TIME	
Relinquished by: *(Signature)*				Received by: *(Signature)*						DATE/TIME	
Received by: *(Signature)*				Received by Mobile Laboratory for field analysis: *(Signature)*						DATE/TIME	
Received by: *(Signature)*		DATE/TIME		Received for Laboratory by:						DATE/TIME	
Method of Shipment:											
Distribution: Original - Accompany Shipment 1 Copy - Survey Coordinator Field Files											

Figure 8. An Example of a Chain-of-Custody Record

Laboratory contamination from the processing of hazardous materials such as toxic or radioactive samples for analysis and their ultimate disposal should be a consideration during the planning stages for selection of methods for analysis. The safe handling of project samples in the laboratory with appropriate decontamination and waste disposal procedures should also be defined.

Often the selected analytical methods can be best given in a table or several tables describing the matrix, the analytes to be measured, and the analysis methods. The sampling containers, methods of preservation, holding times, conditions of holding, the number and types of all QA/QC samples to be collected, and the names of the laboratories who will be performing the analyses should be referenced. Appendix C1 contains a checklist of many important components to consider when selecting analytical methods.

B4.2 Subsampling

If subsampling is required, the procedures should be described in this QAPP element, and the full text of the subsampling operating procedures should be appended to the QAPP. Subsampling methods are generally combined with compositing in order to reduce the variance of samples in an effort to determine the mean concentration. Because subsampling may involve more than one stage, it is imperative that the procedures be documented fully so that the results of the analysis can be evaluated properly.

B4.3 Preparation of the Samples

Preparation procedures should be described and standard methods cited and used where possible. Step-by-step operating procedures for the preparation of the project samples should be listed in an appendix.

B4.4 Analysis Methods

The simple citing of a method usually is not sufficient because often the analysis of a project's samples will require some deviations from a standard method and some selection from the range of options in the method. The step-by-step operating procedures should explicitly state those nuances, and all deviations from the QAPP should be listed through an amendment.

B4.5 References

U.S. Environmental Protection Agency. *Test Methods for Evaluating Solid Waste.* SW-846. Chapter 2, "Choosing the Correct Procedure."

Greenberg, A. E., L. s. Clescer, and A. D. Eaton, eds. 1992. *Standard Methods for the Examination of Water and Wastewater. 18th ed.* American Public Health Association. Water Environment Federation.

Simes, Guy F. 1996. *Quality Control: Variability in Protocols.* EPA/600/9-91/034. Risk Reduction Engineering Laboratory. U.S. EPA. Cincinnati, OH. September.

B5 QUALITY CONTROL REQUIREMENTS

Discuss QC procedures that should be associated with each sampling, analysis, or measurement technique. For projects at or beyond the "proof-of-concept" stage, or for projects employing well-characterized methods, this section should list each required QC procedure, along with the associated acceptance criteria and corrective action.

B5.1 Purpose/Background

Quality Control (QC) is "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 requirements established by the customer." (See Appendix I, "Additional Terms and Definitions.") This element will rely on information developed in A7, "Quality Objectives and Criteria for Measurement Data," which established measurement performance criteria.

B5.2 QC Procedures

This element will need to furnish information on any QA checks not defined in other QAPP elements and should reference other elements that contain this information where possible.

Most of the QC acceptance limits of EPA methods are based on the results of interlaboratory studies. Because of improvements in measurement methodology and continual improvement efforts in individual laboratories, these acceptance limits may not be stringent enough for some projects; therefore, consultation with expert analysts may be necessary. Other elements of the QAPP that contain related sampling and analytical QC requirements include:

- **Sampling Process Design** (B1), which identifies the planned field QC samples as well as procedures for QC sample preparation and handling;
- **Sampling Methods Requirements** (B2), which includes requirements for determining if the collected samples accurately represent the population of interest;
- **Sample Handling and Custody Requirements** (B3), which discusses any QC devices employed to ensure samples are not tampered with (e.g., custody seals) or subject to other unacceptable conditions during transport;
- **Analytical Methods Requirements** (B4), which includes information on the subsampling methods and information on the preparation of QC samples in the sample matrix (e.g., splits, spikes, and duplicates); and
- **Instrument Calibration and Frequency** (B7), which defines prescribed criteria for triggering recalibration (e.g., failed calibration checks).

Table 1 lists QC checks often included in QAPPs.

Table 1. QC Checks That Should Be Included in the QAPP

Type of QC Check	Information Provided by Check
Blanks	
trip blank, field blank	transport and field handling bias
reagent blank	contaminated reagent
reinstate blank	contaminated reinstate
Spikes	
field matrix spike	handling + preparation + analysis bias
matrix spike	analytical (preparation + analysis) bias
matrix spike duplicate	analytical bias and precision
analysis matrix spike	instrumental bias
surrogate spike	analytical bias
(internal standard)	(non-QC, used for quantitation, but indicates instrument performance)
Calibration Check Samples	
zero check	calibration drift and memory effects
span check	calibration drift and memory effects
mid-range check	calibration drift and memory effects
Duplicates, splits, etc.	
collocated samples	sampling + measurement precision
field duplicates/replicates	precision of all steps after acquisition
field splits	shipping + interlaboratory precision
laboratory splits	interlaboratory precision
laboratory duplicates	analytical precision
analysis duplicates	instrument precision

Many of these QC checks result in measurement data that are used to compute statistical indicators of data quality. For example, a series of dilute solutions may be measured repeatedly to produce an estimate of instrument detection limit. The formulae for calculating such data quality indicators should be provided or referenced in the text. This element should also prescribe any limits that define acceptable data quality for these indicators (see also Appendix D, "Data Quality Indicators," and Appendix K, "Calculation of Statistical Quantities"). A QC checklist should be used to discuss the relation of QC to the overall project objectives with respect to:

- the frequency of the check and the point in the measurement process in which the check sample is introduced,
- the traceability of standards,
- the matrix of the check sample,
- the level or concentration of analyte of interest,
- actions to be taken in the event that a QC check identifies a failed or changed measurement system,

- formulae for estimating data quality indicators, and
- procedures for documenting QC results, including control charts.

Refer to Appendix G, "Quality Control," has a more detailed discussion of instrument calibration, aspects of quality control checks, and quality assurance samples.

Finally, this element should describe how the QC check data will be used to determine that measurement performance is acceptable. This step can be accomplished by establishing QC "warning" and "control" limits for the statistical data generated by the QC checks; see Appendix G and standard quality control textbooks for operational details.

B6 INSTRUMENT/EQUIPMENT TESTING, INSPECTION, AND MAINTENANCE REQUIREMENTS

Discuss how inspections and acceptance testing, including the use of QC standards and reference materials, of environmental sampling and measurement systems and their components must be performed and documented to ensure their intended use as specified by the design.

Discuss how the periodic preventive and corrective maintenance of measurement or test equipment shall be performed to ensure availability and satisfactory performance of the systems.

B6.1 Purpose/Background

The purpose of this element of the QAPP is to discuss the procedures used to verify that all instruments and equipment are maintained in sound operating condition and are capable of operating at acceptable performance levels.

B6.2 Testing, Inspection, and Maintenance

The procedures described should (1) reflect consideration of the possible effect equipment failure will have on overall data quality, including timely delivery of project results, (2) address any relevant site- specific effects (e.g., environmental conditions), and, (3) include procedures for assessing equipment status. This element of the QAPP should address the scheduling of routine calibration and maintenance activities, the steps that will be taken to minimize instrument down-time, and the prescribed corrective action procedures for addressing unacceptable inspection or assessment results. This element should also include periodic maintenance procedures and describe the availability of spare parts and how an inventory of these parts is monitored and maintained. The discussion in this element should be in-depth enough to allow for reviewing the adequacy of the instrument/equipment management program.

Inspection and testing procedures may employ reference materials, such as the National Institute of Standards and Technology's (NIST's) Standard Reference Materials (SRMs), as well as quality control standards or an equipment certification program. The accuracy of calibration standards is

important because all data will be in reference to the standard used. The types of standards or special programs should be noted in this element, including the inspection and acceptance testing criteria for all components. The acceptance limits for verifying the accuracy of all working standards against primary grade standards should also be provided.

B7 INSTRUMENT CALIBRATION AND FREQUENCY

Identify all tools, gauges, instruments, and other sampling, measuring, and test equipment used for data collection activities affecting quality that must be controlled and, at specified periods, calibrated to maintain bias within specified limits. Discuss how calibration shall be conducted using certified equipment and/or standards with known valid relationships to nationally recognized performance standards.

B7.1 Purpose/Background

This element of the QAPP concerns the calibration procedures that will be used for instrumental analytical methods and other measurement methods that are used in environmental measurements. The development of these procedures must be coordinated with the development of quality control requirements under B5, "Quality Control Requirements." Refer to Appendix G, "Quality Control," for additional discussion on calibration.

B7.2 Identify the Instrumentation Requiring Calibration

The QAPP should identify any instrumentation that requires calibration to maintain acceptable performance. While the primary focus of this element is on instruments of the measurement system (sampling and measurement equipment), other instrumentation should be included if its improper calibration could impact data quality.

B7.3 Document the Calibration Method That Will Be Used for Each Instrument

The QAPP must describe the calibration method for each instrument in enough detail for another researcher to duplicate the calibration method. It may reference external documents such as EPA-designated calibration procedures or standard operating procedures, providing that these documents can be easily obtained. Nonstandard calibration methods or modified standard calibration methods should be fully documented and justified.

Some instrumentation may be calibrated against other instrumentation or apparatus (e.g., NIST thermometer), while other instrumentations are calibrated using standard materials traceable to national reference standards. QAPP documentation for calibration apparatus and calibration standards are addressed in B7.4 and B7.5.

Calibrations normally involve challenging the measurement system or a component of the measurement system at a number of different levels over its operating range. The calibration may cover a narrower range if accuracy in that range is critical, given the end use of the data. Single-point calibrations are of limited use and two-point calibrations do not provide information on nonlinearity. If single- or two-point calibrations are used for critical measurements, the potential shortcomings should

be carefully considered and discussed in the QAPP. Most EPA-approved analytical methods require multipoint (three or more) calibration that include zeros, or blanks, and higher levels so that unknowns fall within the calibration range and are "bracketed" by calibration points. The number of calibration points, the calibration range, and any replication (repeated measures at each level) should be given in the QAPP.

The QAPP should describe how calibration data will be analyzed. Any goodness-of-fit tests (e.g., calculation of the correlation coefficient) should be described together with acceptance criteria (e.g., "correlation coefficient must exceed 0.99"). The use of statistical quality control techniques to process data across multiple calibrations to detect gradual degradations in the measurement system should be described. The QAPP should describe any corrective action that will be taken if calibration (or calibration check) data fail to meet the acceptance criteria including recalibration.

B7.4 Document the Calibration Apparatus

Some instruments and equipment are calibrated using calibration apparatus, rather than calibration standards. For example, an ozone generator is part of a system used to calibrate continuous ozone monitors. Commercially available calibration apparatus should be listed together with its make (the manufacturer's name), the model number, and the specific variable control settings that will be used during the calibrations. A calibration apparatus that is not commercially available should be described in enough detail for another researcher to duplicate the apparatus and follow the calibration procedure.

B7.5 Document the Calibration Standards

Most measurement systems are calibrated by processing materials that are of known and stable composition; these calibration standards must be described in the QAPP. Calibration standards are normally traceable to national reference standards, and the traceability protocol should be discussed. If the standards are not traceable, the QAPP must include a detailed description of how the standards will be prepared. Any method used to verify the certified value of the standard independently should be described.

B7.6 Document Calibration Frequency

The QAPP must describe how often each measurement method will be calibrated. It is desirable that the calibration frequency be related to any known temporal variability (i.e., drift) of the measurement system. The calibration procedure may involve less-frequent comprehensive calibrations and more-frequent simple drift checks.

B7.7 References

Dieck, R.H. 1992. *Measurement Uncertainty Methods and Applications*. Research Triangle Park, NC. Instrument Society of America.

Dux, J. P. 1986. *Handbook of Quality Assurance for the Analytical Chemistry Laboratory*. New York, NY. Van Nostrand Reinhold.

ILAC Task Force E. 1984. *Guidelines for the Determination of Recalibration Intervals of Testing Equipment Used in Testing Laboratories.* International Organization for Legal Metrology (OIML). International Document No. 10. 11 Rue Twigot, Paris 95009, France.

Ku, H. H., ed. 1969. *Precision Measurement and Calibration. Selected NBS Papers on Statistical Concepts and Procedures.* Special Publication 300. Vol. 1. Gaithersburg, MD: National Bureau of Standards.

Liggett, W. 1986. "Tests of the Recalibration Period of a Drifting Instrument." In *Oceans '86 Conference Record.* Vol. 3. Monitoring Strategies Symposium. The Institute of Electrical and Electronics Engineers, Inc. Service Center. Piscataway, NJ.

Pontius, P. E. 1974. *Notes on the Fundamentals of Measurement as a Production Process.* Publication No. NBSIR 74-545. Gaithersburg, MD: National Bureau of Standards.

Taylor, J. T. 1987. *Quality Assurance of Chemical Measurements.* Boca Raton, FL: Lewis Publishers, Inc.

B8 INSPECTION/ACCEPTANCE REQUIREMENTS FOR SUPPLIES AND CONSUMABLES

Discuss how and by whom supplies and consumables shall be inspected and accepted for use in the project.

B8.1 Purpose

The purpose of this element is to establish and document a system for inspecting and accepting all supplies and consumables that may directly or indirectly affect the quality of the project or task.

B8.2 Identification of Critical Supplies and Consumables

Clearly identify and document all supplies and consumables that may directly or indirectly affect the quality of the project or task. In particular, list all items that, if inferior or deficient, could have a significant or adverse effect on the quality of the project or task. (See Exhibits 1 and 2 for example documentation of inspection/acceptance testing requirements.)

For each item identified, document the inspection or acceptance testing requirements or specifications (e.g., concentration, purity, cell viability, activity, or source of procurement) in addition to any requirements for certificates of purity or analysis.

B8.3 Establishing Acceptance Criteria

Acceptance criteria must be consistent with overall project technical and quality criteria (e.g., concentration must be within $\pm$ 2.5%, cell viability must be $>$90%). If special requirements are needed for particular supplies or consumables, a clear agreement should be established with the supplier, including methods used for evaluation and provisions for settling disparities.

B8.4 Inspection or Acceptance Testing Requirements and Procedures

Inspections or acceptance testing should be documented, including procedures to be followed, responsible individuals, and the frequency of evaluation. In addition, handling and storage conditions for supplies and consumables should be documented.

B8.5 Tracking and Quality Verification of Supplies and Consumables

Procedures should be established to ensure that inspections or acceptance testing of supplies and consumables are adequately documented by permanent, dated, and signed records or logs that uniquely identify the critical supplies or consumables, the date received, the date tested, the date to be retested (if applicable), and the expiration date. These records should be kept by the responsible individual(s). (See Exhibit 3 for an example log).

In order to track supplies and consumables, labels with the information on receipt and testing should be used.

These or similar procedures should be established to enable project personnel to (1) verify, prior to use, that critical supplies and consumables meet specified project or task quality objectives, and (2) ensure that supplies and consumables that have not been tested, have expired, or do not meet acceptance criteria are not used for the project or task.

Unique identification No. (if not clearly shown)_____
Date received_____
Date opened_____
Date tested (if performed)_____
Date to be retested (if applicable)_____
Expiration date_____

Exhibit 1. Example of a Record for Consumables

Critical Supplies and Consumables	Inspection/ Acceptance Testing Requirements	Acceptance Criteria	Testing Method	Frequency	Responsible Individual	Handling/Storage Conditions

Exhibit 2. Example Inspection/Acceptance Testing Requirements

Critical Supplies and Consumable (Type, ID. No.)	Date Received	Meets Inspection/ Acceptance Criteria (Y/N, Include Date)	Requires Retesting (Y/N, If Yes, Include Date)	Expiration Date	Comments	Initials /Date

Exhibit 3. Example Log for Tracking Supplies and Consumables

B9 DATA ACQUISITION REQUIREMENTS (NON-DIRECT MEASUREMENTS)

Identify the type of data acquired from non-measurement sources such as computer data bases, spreadsheets, and programs, and literature files. Define acceptance criteria for the use of the data in this project. Discuss any limitations on the use of the data based on uncertainty in the quality of the data and discuss the nature of that uncertainty.

B9.1 Purpose/Background

This element of the QAPP should clearly identify the intended sources of previously collected data and other non-measurement data that will be used in this project. Information that is nonrepresentative and possibly biased and is used uncritically may lead to decision errors. The care and skepticism applied to the generation of new data is also appropriate to the use of previously compiled data (for example, data sources such as handbooks and computerized databases).

B9.2 Acquisition of Non-Direct Measurement Data

This element's criteria should be developed to support the objectives of element A7. Acceptance criteria for each collection of data being considered for use in this project should be explicitly stated especially with respect to:

- **Representativeness.** Were the data collected from a population that is sufficiently similar to the population of interest and the population boundaries? How will potentially confounding effects (for example, season, time of day, and cell type) be addressed so that these effects do not unduly alter the summary information? This issue is discussed at length in Appendix H.

- **Bias.** Are there characteristics of the data set that would shift the conclusions (for example, has bias in analysis results been documented?) Is there sufficient information to estimate and correct bias?

- **Precision.** How is the spread in the results estimated? Does the estimate of variability indicate that it is sufficiently small to meet the objectives of this project as stated in element A7? See also Appendix D.

- **Qualifiers.** Are the data evaluated in a manner that permits logical decisions on whether or not the data are applicable to the current project? Is the system of qualifying or flagging data adequately documented to allow combination of data sets?

- **Summarization.** Are the available data a summary with a summarization process clear and sufficiently consistent with the goals of this project? (See element D2 for further discussion.) Ideally, observations and the transformation equations are available so that their assumptions can be evaluated against the objectives of the current project.

B10 DATA MANAGEMENT

B10.1 Purpose/Background

Outline the project data management scheme, tracing the path of the data, beginning from receipt from the field or laboratory, to the use or storage of the final reported form. Describe the standard record keeping procedures, document control system, and the approach used for data storage and retrieval on electronic media. Discuss the control mechanism for detecting and correcting paperwork errors and for preventing loss of data during data reduction (i.e., calculations), data reporting, and data entry to forms, reports, and data bases. Provide examples of any forms or checklists to be used.

This element of the QAPP should present an overview of all manipulations performed on raw ("as-collected") data to change their form of expression, location, quantity, or dimensionality. These manipulations include data recording, validation, transformation, transmittal, reduction, analysis, management, storage, and retrieval. A diagram that illustrates the source(s) of data, processing steps, intermediate and final data files, and reports produced may be helpful, particularly when there are multiple data sources and data files.

B10.2 Data Recording

Any internal checks (including verification and validation checks) that will be used to ensure data quality during the data collection process should be identified. Examples of data entry forms and checklists should be included.

B10.3 Data Validation

The details of the process of data validation and prespecified criteria should be documented in this element of the QAPP. This element of the QAPP should address the validation to be performed as data are generated. Part D of this document addresses the overall project data validation, which is performed after the project is completed. Refer to Appendix F, "Verification and Validation," for more detailed discussion.

B10.4 Data Transformation

Data transformation is the conversion of individual data point values into related values or possibly symbols using conversion formulae (e.g., units conversion or logarithmic conversion) or a system for replacement. The transformations can be reversible (e.g., as in the conversion of data points using a formulae) or irreversible (e.g., when a symbol replaces actual values and the value is lost). The procedures for all data transformations should be described and recorded in this element. The procedure for converting calibration readings to an equation that will be applied to measurement readings should be documented in the QAPP.

B10.5 Data Transmittal

Data transmittal occurs when data are transferred from one person or location to another or when data are copied from one form to another. Some examples of data transmittal are copying raw data from a notebook onto a data entry form for keying into a computer file and electronic transfer of data over a telephone or computer network. The QAPP should describe each data transfer step and the procedures that will be used to characterize data transmittal error rates and to minimize information loss in the transmittal.

B10.6 Data Reduction

Data reduction includes all processes that change the number of the data items. This process is distinct from data transformation in that it entails an irreversible reduction in the size of the data set and an associated loss of detail. For manual calculations, the QAPP should include an example in which typical raw data are reduced. For automated data processing, the QAPP should clearly indicate how the raw data are to be reduced with a well-defined audit trail, and reference to the specific software documentation should be provided.

B10.7 Data Analysis

Data analysis sometimes involves comparing suitably reduced data with a conceptual model (e.g., a dispersion model or an infectivity model). It frequently includes computation of summary statistics, standard errors, confidence intervals, tests of hypotheses relative to model parameters, and goodness-of-fit tests. This element should briefly outline the proposed methodology for data analysis and a more detailed discussion should be included in the final report.

B10.8 Data Tracking

Data management includes tracking the status of data as they are collected, transmitted, and processed. Projects should have established procedures for tracking the flow of data through the data processing system.

B10.9 Data Storage and Retrieval

The QAPP should discuss data storage and retrieval including security and time of retention. The QAPP should also discuss the performance requirements of the data processing system including provisions for batch processing schedule and data storage facilities.

C ASSESSMENT/OVERSIGHT

C1 ASSESSMENTS AND RESPONSE ACTIONS

Identify the number, frequency, and type of assessment activities needed for this project. Assessments include, but are not limited to, the following:

- **surveillance,**
- **peer reviews,**
- **management systems reviews,**
- **readiness review,**
- **technical systems audits,**
- **performance evaluations,**
- **audits of data quality, and**
- **Data Quality Assessment.**

C1.1 Purpose/Background

During the planning process, many options for sampling design (see EPA QA/G-5S), sample handling, sample cleanup and analysis, and data reduction are evaluated and chosen for the project. In order to ensure that the data collection is conducted as planned, a process of evaluation and validation is necessary. This process will ensure that:

- all elements of the QAPP are correctly implemented as prescribed,
- the quality of the data generated by the QAPP is adequate, and
- a corrective action plan is in place if unforeseen circumstances force a deviation from the plan.

Although any external assessments that are planned should be described in the QAPP, the most important part of this element is documenting all planned internal assessments. Generally, internal assessments are initiated or performed by the internal QA Officer so the activities described in this element of the QAPP should be related to the responsibilities of the QA Officer as discussed in A4.

C1.2 Assessment Activities and Project Planning

The following is a description of various types of assessment activities available to managers in evaluating the effectiveness of QA programs.

C1.2.1 Assessment of the Subsidiary Organizations

A. *Management Systems Review (MSR).* This review consists of a qualitative assessment of a data collection operation or organization to establish whether the prevailing quality management structure, policies, practices, and procedures are adequate for ensuring that the type and quality of data needed are obtained. The MSR is used to ensure that sufficient management controls are in place and carried out by the organization to adequately plan, implement, and assess the results of the project. See also *Guidance for the Management Systems Review Process* (EPA QA/G-3).

B. *Readiness Review*. A readiness review is a technical check to determine if all components of the project are in place so that work can commence on a specific phase of a project.

C1.2.2 Assessment of Project Activities

A. *Surveillance*. Surveillance is continual or frequent monitoring of the status of a project and the analysis of records to ensure that specified requirements are being fulfilled.

B. *Technical Systems Audit (TSA)*. A TSA is a thorough and systematic onsite qualitative audit, where facilities, equipment, personnel, training, procedures, and record keeping are examined for conformance to the QAPP. The TSA is a powerful audit tool with broad coverage that may reveal weaknesses in the management structure, policy, practices, or procedures. The TSA is ideally conducted after work has commenced, but before it has progressed very far, thus giving opportunity for corrective action.

C. *Performance Evaluation (PE)*. The performance evaluation is a type of audit in which the quantitative data generated by the measurement system are obtained independently and compared with routinely obtained data to evaluate the proficiency of an analyst or laboratory. "Blind" PE samples are those whose identity is unknown to those operating the measurement system. Blind PEs often produce better performance assessments because they are handled routinely and are not given the special treatment that undisguised PEs sometimes receive.

The QAPP should list the performance evaluations that are planned, identifying:

- constituents to be measured,
- target concentration ranges,
- timing/schedule for PE sample analysis, and
- aspect of measurement quality to be assessed (e.g., bias, precision, and detection limit).

PE materials are now available from commercial sources and a number of EPA program offices coordinate various interlaboratory studies and laboratory proficiency programs. Participation in these or in the National Voluntary Laboratory Accreditation Program (run by the National Institute of Standards and Technology) should be mentioned in the QAPP.

D. *Audit of Data Quality (ADQ)*. An ADQ will reveal how the data were handled, what judgement calls were made, and whether uncorrected mistakes were made. Performed prior to producing a project's final report, ADQs can often identify means to correct systematic data reduction errors.

During the data reduction phase of a project, there are many decisions that can arise concerning the evaluation of results from blanks, surrogates, and spike recoveries. It may be necessary to decide whether to subtract target analyte concentrations that appear in blanks from the project sample results. Referring to the guidance for performing the specific analytical method may aid in making this decision.

E. *Peer Review*. Whether a planning team will choose audits of data quality or peer reviews depends upon the nature of the project, the intended use of the data, and the policies established by the sponsor of the project. Reviewers are chosen who have technical expertise comparable to the project performers but who are independent of the project. They ensure that the project activities:

- were technically adequate,
- were competently performed,
- were properly documented,
- satisfied established technical requirements, and
- satisfied established quality assurance requirements.

Peer reviewers assess the assumptions, calculations, extrapolations, alternative interpretations, methods, acceptance criteria, and conclusions documented in the project's report. The names, titles, and positions of the peer reviewers should be included in the final QAPP.

F. *Data Quality Assessment (DQA).* DQA involves the application of statistical tools to determine whether the data meet the assumptions that the DQOs and data collection design were developed under and whether the total error in the data is tolerable. *Guidance for the Data Quality Assessment Process* (EPA QA/G-9) provides nonmandatory guidance for planning, implementing, and evaluating retrospective assessments of the quality of the results from environmental data operations.

C1.3 Documentation of Assessments

The following material describes what should be documented in a QAPP after consideration of the above issues and types of assessments.

C1.3.1 Number, Frequency, and Types of Assessments

Depending upon the nature of the project, there may be more than one audit. A schedule of the number, frequencies, and types of assessments required should be given.

C1.3.2 Assessment Personnel

Internal audits are usually performed by QA personnel who work for the contractor performing the project work but who are organizationally independent of the management of the project. External audits are performed by personnel of organizations not connected with the project, but who are technically qualified and who understand the quality assurance requirements of the project.

C1.3.3 Schedule of Assessment Activities

A schedule of audit activities, together with relevant criteria for assessment, should be given to the extent it is known in advance of project activities.

C1.3.4 Reporting and Resolution of Issues

Audits and other assessments often reveal findings of practice or procedure that do not conform to the written QAPP. Because these issues must be addressed in a timely manner, the protocol for resolving them should be given here. The person to whom the concerns should be addressed is given,

and the decision making hierarchy is delineated. The schedule and format for oral and written reports are given in this element, and responsibility for corrective action is assigned. This element should explicitly define the unsatisfactory conditions upon which the assessors are authorized to act and list the project personnel who should receive assessment reports.

C2 REPORTS TO MANAGEMENT

Identify the frequency, content, and distribution of reports issued to inform management of the following:

- **status of the project;**
- **results of performance evaluations and systems audits;**
- **results of periodic data quality assessments; and**
- **significant quality assurance problems and recommended solutions.**

Identify the responsible organization(s) that will prepare the reports and the recipients of the reports. Identify any other status reports to management as well as their content and frequency.

C2.1 Purpose/Background

Effective communication is an integral part of a quality system. Planned reports provide a structure for apprising management of the project schedule, the deviations from approved quality assurance and test plans, the impact of these deviations on data quality, and the potential uncertainties in decisions based on the data.

C2.2 Frequency, Content, and Distribution of Reports

The QAPP should indicate the frequency, content, and distribution of the reports so that management may anticipate events and move to ameliorate potentially adverse results. An important benefit of the status reports is the opportunity to alert the management of data quality problems, propose viable solutions, and procure additional resources. If program assessment (including the evaluation of the technical systems, the measurement of performance, and the assessment of data) is not conducted on a continual basis, the integrity of the data generated in the program may not meet the quality requirements. These audit reports, submitted in a timely manner, will provide an opportunity to implement corrective actions when most appropriate.

C2.3 Identify Responsible Organizations

It is important that the QAPP identify the personnel responsible for preparing the reports, evaluating their impact, and implementing follow-up actions. It is necessary to understand how any changes made in one area or procedure may affect another part of the project. Furthermore, the documentation for all changes should be maintained and included in the reports to management.

At the end of the project, a Data Quality Assessment and a reporting of the findings to management makes for a formal conclusion to the life cycle of a project.

D DATA VALIDATION AND USABILITY

D1 DATA REVIEW, VALIDATION, AND VERIFICATION REQUIREMENTS

State the criteria used to review and validate—that is, accept, reject, or qualify—data, in an objective and consistent manner. Provide examples of any forms or checklists to be used.

D1.1 Purpose/Background

The purpose of this element is to state the criteria for deciding the degree to which each data item has met its quality specifications as described in element B. Investigators should estimate the potential effect that each deviation from a QAPP may have on the usability of the associated data item, its contribution to the quality of the reduced and analyzed data, and its effect on the decision.

The following discussion applies to situations in which a sample is separated from its native environment and transported to a laboratory for analysis and data generation. However, these principles can be adapted to other situations (for example, in-situ analysis or laboratory research).

D1.2 Sampling Design

How correctly a measurement at a given time and location represents the actual environment is a complex issue that is considered during development of element B1. See *Guidance on Sampling Designs to Support QAPPs* (EPA QA/G-5S). Acceptable tolerances on each critical sample coordinate should be specified in element B1, along with the action to be taken if the tolerances are exceeded.

Each sample should be checked for conformity to the specifications, including type and location (spatial and temporal). By noting the deviations in sufficient detail, subsequent data users will be able to determine the data's usability under scenarios different from those included in project planning. The strength of conclusions that can be drawn from data (see *Guidance Document for Data Quality Assessment*, EPA QA/G9) has a direct connection to the sampling design and deviations from that design. Where auxiliary variables are included in the overall data collection effort (for example, microbiological nutrient characteristics or process conditions), they should be included in this evaluation.

D1.3 Sample Collection Procedures

Details of how a sample is separated from its native time/space location is important for properly interpreting the measurement results. Element B2 provides these details, which include sampling and ancillary equipment and procedures (including equipment decontamination). Acceptable departures (for example, alternate equipment) from the QAPP, and the action to be taken if the requirements cannot be satisfied, should be specified for each critical aspect. Validation activities should note potentially unacceptable departures from the QAPP.

D1.4 Sample Handling

Details of how a sample is physically treated and handled during relocation from its original site to the actual measurement site are extremely important. Correct interpretation of the subsequent

measurement results requires that deviations from element B3 of the QAPP and the actions taken to minimize or control the changes be detailed. Data validation activities should indicate out-of-tolerance events.

At a minimum, investigators should evaluate the sample containers and the preservation methods used, and ensure they are appropriate to the nature of the sample and the type of data generated from the sample. Checks on the identity of the sample (e.g., proper labeling and chain-of-custody records) as well as proper physical/chemical storage conditions (e.g., chain-of-custody and storage records) should be made to ensure that the sample continues to be representative of its native environment as it moves through the analytical process.

D1.5 Analytical Procedures

Each sample should be verified to ensure that the procedures used to generate the data (as identified in element B4 of the QAPP) were as specified. Acceptance criteria should be developed for important components of the procedures, along with suitable codes for characterizing each sample's deviation from the procedure. Data validation activities should determine how seriously a sample deviated beyond the acceptable limit so that the potential effects of the deviation can be evaluated during DQA.

D1.6 Quality Control

Element B5 of the QAPP specifies the quality control (QC) checks that are to be performed during sample collection, handling, and analysis. These checks include information on blanks, spikes, and duplication, and assist in estimating of the quality of data being produced by specified components of the measurement process. For each specified QC check, the procedure, acceptance criteria, and corrective action (and changes) should be specified. Data validation should document the corrective actions that were taken, which samples were effected, and the actions' potential effect on the validity of the data.

D1.7 Calibration

Element B7 addresses the calibration of instruments and equipment and the information that should be presented to ensure that the calibrations:

- were performed within an acceptable time prior to generation of measurement data;
- included the proper number of calibration points;
- were performed using standards that "bracketed" the range of reported measurement results (otherwise, results falling outside the calibration range should be flagged as such); and
- had acceptable linearity checks and other checks to ensure that the measurement system was stable when the calibration was performed.

When calibration problems are identified, any data produced between the suspect calibration event and any subsequent recalibration should be flagged to alert data users.

D1.8 Data Reduction and Processing

Checks on data integrity evaluate the accuracy of "raw" data and include the comparison of important events and the duplicate rekeying of data to identify data entry errors.

Data transformations include relatively simple scaling changes to the raw data such as unit conversions (for example, centimeters from inches), coordinate transformations (for example, rectangular to polar coordinates), or use of calibration equations (for example, concentrations from voltages). How transformation equations are checked, the requirements for the outcome, and how deviations are handled should be documented in this element.

Data Reduction is an irreversible process that involves a loss of detail in the data and may involve averaging across time (for example, hourly or daily averages) or space (for example, compositing results from samples thought to be physically equivalent). Since this summarizing process produces few values to represent a group (population) of many data points, its validity should be well-documented in the QAPP. Potential data anomalies can be investigated by simple statistical analyses (see *Guidance for Data Quality Assessment*, EPA QA/G-9).

The information generation step involves the synthesis of the results of the previous operations and the construction of tables and charts suitable for use in reports. Operations at this level involve correlation of different variables, model fitting, and three-dimensional visualization presentations. This is a difficult process to evaluate due to frequently massive amounts of sequentially processed data, with little or no access to the detailed processing logic. How information generation is checked, the requirements for the outcome, and how deviations from the requirements will be treated, should be addressed in this element.

Additional checks may be developed that require an understanding of the project that extends to the fundamental manner in which interactions occur in the specific environmental system. For example, in evaluating a process, mass balance calculations can be useful in determining that part of the process has been overlooked. Similarly, in many complex systems, reactions are coupled such that an increase in one will quantitatively trigger a decrease (or increase) in another. The basis for these additional checks and the requirements for the outcomes, and how deviations from the requirements will be treated, should be addressed in this element. Inconsistencies discovered here, with the use of properly validated raw data and subsequent data management processes, require explanation.

D2 VALIDATION AND VERIFICATION METHODS

Describe the process to be used for validating and verifying data, including the chain-of-custody for data throughout the life cycle of the project or task.

D2.1 Purpose/Background

The purpose of this element is to describe, in detail, the process for validating (determining if data satisfy QAPP-defined user requirements) and verifying (ensuring that conclusions can be correctly drawn) project data. Diagrams should be developed showing the various roles and responsibilities with

respect to the flow of data as the project progresses. The QAPP should have a clear definition of what is implied by "verification," and what is implied by "validation." (Refer to Appendix F, "Verification and Validation" for a more detailed discussion.)

D2.2 Describe the Process for Validating and Verifying Data

The individuals responsible for data validation together with lines of authority should be shown on an organizational chart and may be indicated in the chart in element A7. A diagram, similar to the one developed in element B10, depicting the flow of data from its generation through its use in reports, should be included. The chart should indicate who is responsible for each activity of the overall validation and verification.

The data to be validated should be compared to "actual" events using the criteria documented in the QAPP. The criteria for comparison may be physically contained in the QAPP itself, or the QAPP may reference other documents such as contract statements of work, SOPs, work plans, or facility manuals.

D3 RECONCILIATION WITH DATA QUALITY OBJECTIVES

Describe how the results obtained from the project or task will be reconciled with the results of the DQO Process. Describe how issues will be resolved. Discuss how limitations on the use of the data will be reported to decision makers. Identify the procedures used to assess precision, bias, and completeness of the project data.

D3.1 Purpose/Background

The purpose of element D3 is to outline and specify, if possible, the acceptable methods for evaluating the results obtained from the project. It includes scientific and statistical evaluations of data to determine if the data are of the right type, quantity, and quality to support their intended use. This element should apply to all projects, regardless of whether formal DQOs were developed.

The Data Quality Assessment (DQA) process has been developed for cases where formal DQOs have been established. *Guidance for Data Quality Assessment* (EPA QA/G-9) focuses on evaluating data for fitness in decision making and also provides many graphical and statistical tools.

D3.2 Reconciling Results with DQOs

DQA is a key part of the assessment phase of the data life cycle, as shown in Figure 9. During the planning phase, DQOs are developed and a sampling and analysis design is chosen and together with plans for QA and QC, these are documented in the QAPP. In the assessment phase, following data validation and verification, DQA determines how well the validated data can support their intended use.

From the EPA QA/G-9 guidance document, the 5-step DQA Process is presented as follows:

> The DQA Process involves 5 steps that begin with a review of the planning documentation and end with an answer to the question posed during the planning phase of the study. The 5 steps, which are described in EPA QA/G-9, are briefly summarized as follows:

1. *Review the Data Quality Objectives (DQOs) and Sampling Design:* Review the DQO outputs to assure that they are still applicable. If DQOs have not been developed, specify DQOs before evaluating the data (for environmental decision, define the statistical hypothesis and specify tolerable limits on decision errors; for estimation problems, define an acceptable confidence or probability interval width). Review the sampling design and data collection documentation for consistency with the DQOs.

2. *Conduct a Preliminary Data Review:* Review quality assurance (QA) reports, calculate basic statistical quantities and generate graphs of the data. Use this information to learn about the structure of the data and identify patterns, relationships, or potential anomalies.

3. *Select the Statistical Test:* Select the most appropriate procedure for summarizing and analyzing data, based on the preliminary data review. Identify the key underlying assumptions that must hold for the statistical procedures to be valid.

4. *Verify the Assumptions of the Statistical Test:* Evaluate whether the underlying assumptions hold, or whether departures are acceptable, given the actual data and other information about the study.

5. *Draw Conclusions from the Data:* Perform the calculations required for the statistical test and document the inferences drawn as a result of these calculations. If the design is to be used again, evaluate the performance of the sampling design.

These 5 steps are presented in a linear sequence, but the process is by its very nature iterative. For example, if the preliminary data review reveals patterns or anomalies in the data set that are inconsistent with the DQOs, then some aspects of the study planning may have to be reconsidered in Step 1. Likewise, if the underlying assumptions of the statistical test are not supported by the data, then previous steps of the DQA Process may have to revisited. The strength of the process is that it is designed to promote an understanding of how well the data satisfy their intended use by progressing is a logical and efficient manner. Nevertheless, it should be emphasized that the DQA Process cannot absolutely prove that one has or has not achieved the DQOs set forth during the planned phase of a study. This situation occurs because DQOs depend on the true parameter(s) inherent to a site or process (e.g., the true mean concentration). While more information is available after the data collection-namely, an estimate of the parameter-the true value of the parameter is still unknown.

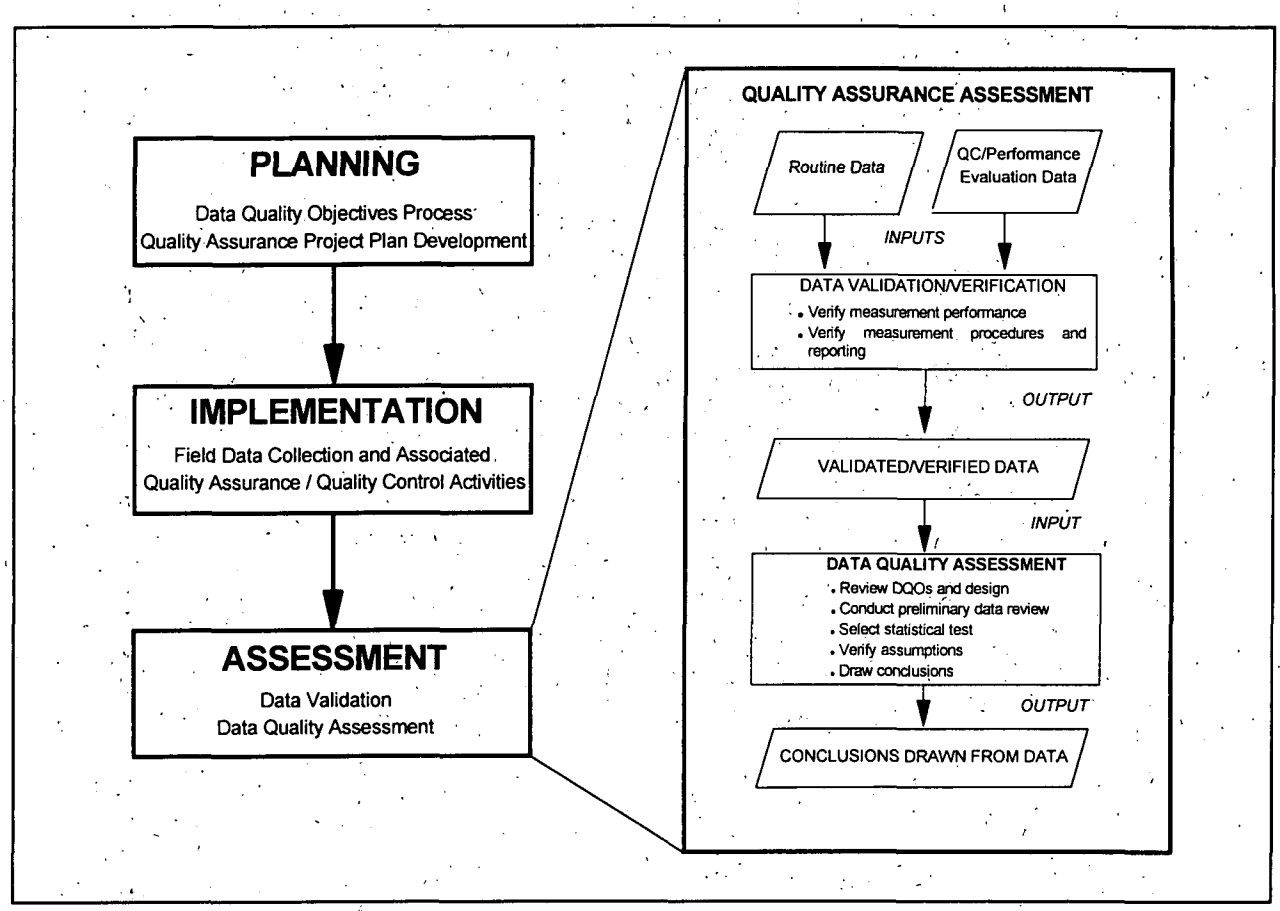


Figure 9. Data Quality Assessment in the Data Life Cycle

CHAPTER IV

QAPP REVISIONS AND RELATED GUIDANCE

QAPP REVISIONS

During the course of environmental data collections, it is probable that changes will occur and revisions to the QAPP will have to be made. Any changes to the technical procedures should be evaluated by the EPA QA Officer and Project Officer to determine if they significantly affect the technical and quality objectives of the project. If so, the QAPP should be revised and reapproved, and a revised copy should be sent to all personnel on the distribution list. For projects of long duration, the QAPP should be reviewed at least annually and revised as appropriate.

COMPARISON WITH PREVIOUS GUIDANCE (QAMS-005/80)

EPA's previous guidance for preparing QAPPs, *Interim Guidelines and Specifications for Preparing Quality Assurance Project Plans* (QAMS-005/80) was released in December 1980. The evolution of the EPA programs, changing needs, and changes to quality management practices have mandated the preparation of a new guidance. The QAPPs that will be generated based on this guidance will be slightly different from those in the past because:

1) Additional guidance documents from the agency including *Guidance for the Data Quality Objectives Process* (EPA QA/G-4), and *Guidance for Data Quality Assessment* (EPA QA/G-9), are available on important quality management practices. The G-4 guidance was released in June 1994. The G-9 document was issued in September 1996. The QAPP guidance (EPA QA/G-5) incorporates the concepts addressed in these other two guidance documents for a more complete guidance on planning. These guidance documents show how the DQO Process, the QAPP, and the DQA Process link together in a coherent way. (See Appendix A.3 for a crosswalk between the DQOs and the QAPP.)

2) The new guidance includes flexibility in requirements. However, if an element of the QAPP is not applicable to a particular project, rationale for not addressing the element should be included.

3) The elements of the QAPP are now organized in an order that corresponds to the customary planning, implementation, and assessment phases of a project and have been grouped into four classes:

   - Project Management,
   - Measurement/Data Acquisition,
   - Assessment/Oversight, and
   - Data Validation and Usability.

4) There are more elements identified than in the previous QAMS-005/80 guidance and this encourages flexibility in construction of defensible QAPPs.

A comparison between the requirements of QAMS-005/80 and this document is presented in Appendix A.2, "Crosswalk Between EPA QA/R-5 and QAMS-005/80." A description of the relationship of this document with the Agency's quality system, national consensus standards, and the ISO 9000 series is presented in Appendix A.1.

APPENDIX A

CROSSWALKS BETWEEN QA DOCUMENTS

This appendix consists of five sections. The first section describes the relationship between the systems requirements developed by the American National Standards Institute (ANSI) and the EPA Quality System requirements. The second section provides a crosswalk between the requirements document for Quality Assurance Project Plans (QAPPs), EPA QA/R-5, *EPA Requirements For Quality Assurance Project Plans For Environmental Data Operations*, and its predecessor document QAMS 005/80, *Interim Guidelines And Specifications For Preparing Quality Assurance Project Plans*. The third section provides a crosswalk between QA/R-5 and the elements of ISO 9000. The fourth section is a crosswalk between the requirements of the QAPP and the steps of the Data Quality Objectives (DQOs) Process. The fifth section lists and discusses the relationship among the different EPA Quality System requirements and guidance documents.

A1. Relationship Between E4 and EPA Quality System

The Environmental Protection Agency (EPA) has developed a mandatory Agency-wide Quality System that applies to all organizations performing work for EPA. These organizations must ensure that data collected for the characterization of environmental processes and conditions are of the appropriate type and quality for their intended use, and environmental technologies are designed, constructed, and operated according to defined expectations. All Quality Systems established in accordance with these requirements shall comply with ANSI/ASQC E4-1994, *Quality Systems Requirements for Environmental Programs* (E4 document), which is in compliance with ISO 9000. In addition, EPA has developed two documents, EPA QA/R-1, *EPA Quality Systems Requirements for Environmental Programs* (R-1 document) and EPA QA/R-2, *EPA Requirements for Quality Management Plans* (R-2 document), that specify the requirements for developing, documenting, implementing, and assessing a Quality System. This appendix describes these three Agency documents in order to show their relationship and role in laying the foundation for EPA's Quality System.

The E4 Document provides the basis for the preparation of a quality system for an organization's environmental programs. The document provides the requisite management and technical area elements necessary for developing and implementing a quality system. The document first describes the quality management elements that are generally common to environmental problems regardless of their technical scope. The document then discusses the specifications and guidelines that apply to project-specific environmental activities involving the generation, collection, analysis, evaluation, and reporting of environmental data. Finally, the document contains the minimum specifications and guidelines that apply to the design, construction, and operation of environmental technology.

The R-1 document provides the details on EPA quality management requirements to organizations conducting environmental programs. The R-1 document states that "... all EPA organizations and all organizations performing work for EPA shall develop and establish Quality Systems, as appropriate, that are compliant with the American National Standard ANSI/ASQC E4-1994 *Quality Systems Requirements for Environmental Programs*, and its additions and supplements from the American National Standards Institute (ANSI) and the American Society for Quality Control (ASQC)." The R-1 applies to all EPA programs and organizations, unless explicitly exempted, that produce, acquire, or use environmental data depending upon the purposes for which the data will be used. The R-1 also applies to systems, facilities, processes, and methods for pollution control, waste treatment, waste remediation, and waste packaging and storage.

EPA Requirements for Quality Management Plans, EPA QA/R-2 discusses the development, review, approval, and implementation of the Quality Management Plan (QMP). The QMP is a means of documenting how an organization will plan, implement, and assess the effectiveness of the management processes and structures (required under R-1) that relate to the Quality System. The R-2 document describes the program elements that should be part of a QMP. These requirements match the quality management elements described in the E4 document that are generally common to environmental projects. These elements include the following: (1) management and organization, (2) quality system and description, (3) personnel qualification and training, (4) procurement of items and services, (5) documents and records, (6) computer hardware and software, (7) planning, (8) implementation of work processes, (9) assessment and response, and (10) quality improvement.

Quality Assurance Project Plans normally will be addressed as part of an organization's QMP. In essence, the QMP will establish the nature of the requirements for QAPPs for work done by or for that organization.

The *International Organization for Standardization (ISO) 9000 Series* is a set of five international standards developed by the ISO Technical Committee 176 on quality systems. Published in 1987 and adopted by over 70 countries, conformance with these standards is being demanded in purchasing specifications with increasing frequency. The standards are:

- ISO 9000: (ANSI/ASQC Q90), Quality Management and Quality Assurance Standards—Guidelines for selection and use;
- ISO 9001: (ANSI/ASQC Q91), Quality Systems—Model for quality assurance in design/development, production, installation, and servicing;
- ISO 9002: (ANSI/ASQC Q92), Quality Systems—Model for quality assurance in production and installation;
- ISO 9003: (ANSI/ASQC Q93), Quality Systems—Model for quality assurance in final inspection and test;
- ISO 9004: (ANSI/ASQC Q94), Quality Management and Quality System Elements—Guidelines.

The objectives of the ISO quality system are to:

- Achieve and sustain the quality of the product or service produced so as to meet purchaser's needs;
- Provide confidence to management that the intended quality is being achieved;
- Provide confidence to the purchaser that the intended quality will be achieved in the delivered product or service.

The ISO 9000 series may be regarded as generic systems standards that facilitate demonstration of conformance. They do not, however, apply to the quality, useability, or applicability of specific products or services. Figure A1. illustrates the relationships among the ISO standards and other elements of quality systems.

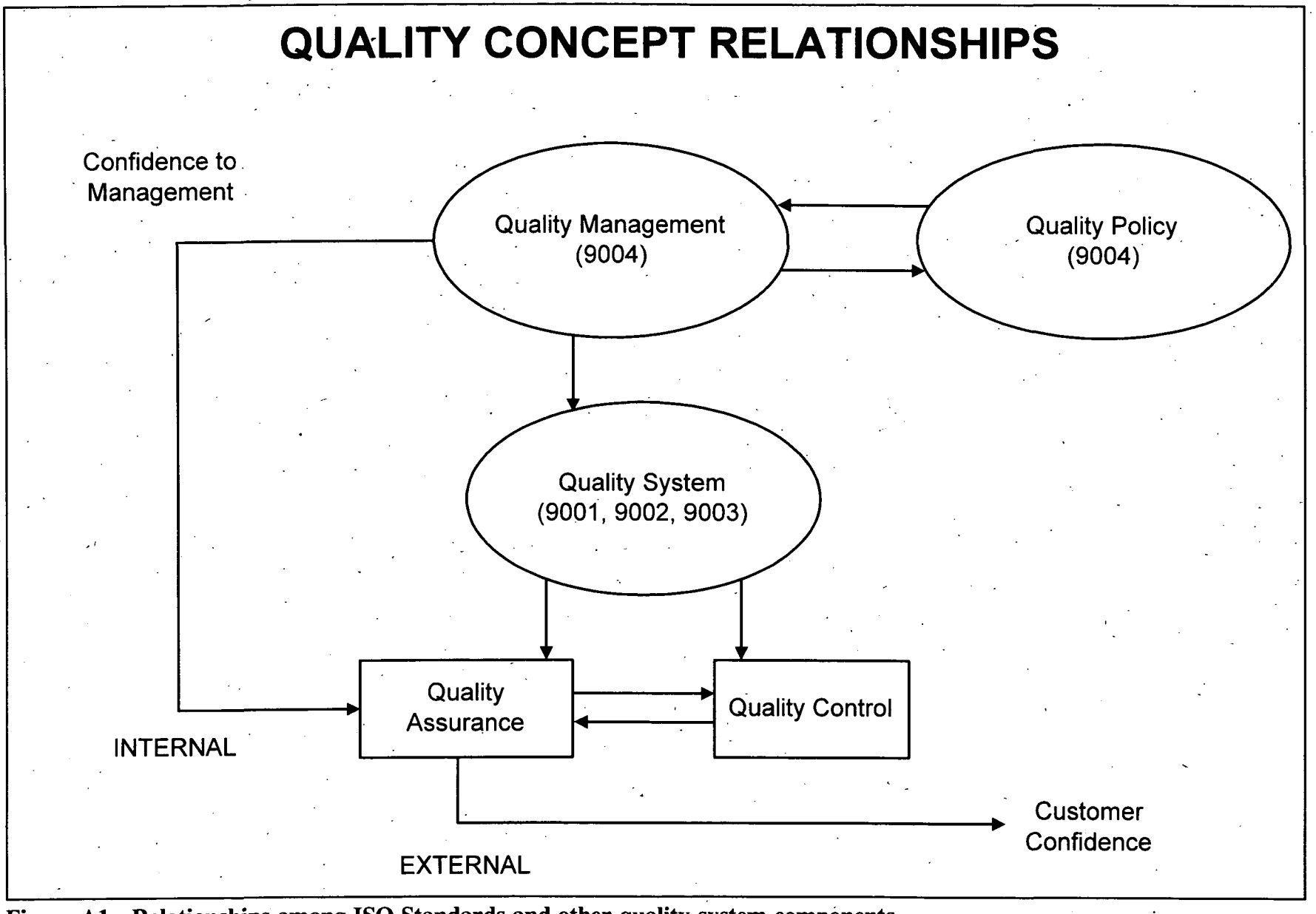


Figure A1. Relationships among ISO Standards and other quality system components

A2. Crosswalk Between EPA QA/R-5 and QAMS-005/80

QAMS-005/80 ELEMENTS		QA/R-5 ELEMENTS	
1.0	Title Page with Provision for Approval Signatures	A1	Title and Approval Sheet
2.0	Table of Contents	A2	Table of Contents
3.0	Project Description	A5 A6	Problem Definition/Background Project/Task Description
4.0	Project Organization and Responsibility	A4	Project/Task Organization
5.0	QA Objectives for Measurement Data (PARCC)	A7	Quality Objectives and Criteria for Measurement Data
6.0	Sampling Procedures	B1 B2	Sampling Process Design Sampling Methods Requirements
7.0	Sample Custody	B3	Sample Handling and Custody Requirements
8.0	Calibration Procedures and Frequency	B7	Instrument Calibration and Frequency
9.0	Analytical Procedures	B4	Analytical Methods Requirements
10.0	Data Reduction, Validation, and Reporting	D1 D2 B10	Data Review, Validation, and Verification Requirements Validation and Verification Methods Data Quality Management
11.0	Internal Quality Control Checks and Frequency	B5	Quality Control Requirements
12.0	Performance and Systems	C1	Assessments and Response Audits Actions

QAMS-005/80 ELEMENTS		QA/R-5 ELEMENTS	
13.0	Preventive Maintenance	B6	Instrument/Equipment Testing, Procedures and Schedules Inspection, and Maintenance Requirements
14.0	Specific Routine Procedures Measurement Parameters Involved	D3	Reconciliation with Data Used to Assess PARCC for Quality Objectives
15.0	Corrective Action	C1	Assessments and Response Actions
16.0	QA Reports to Management	C2	Reports to Management
(No Corresponding QAMS-005/80 Elements)		A8	Project Narrative
		A9	Special Training Requirements or Certification
		A10	Documentation and Records
		B8	Inspection/Acceptance Requirements for Supplies and Consumables
		B9	Data Acquisition Requirements (Non-direct Measurements)
		B10	Data Quality Management

A3. Crosswalk between EPA QA/R-5 and ISO 9000

EPA/R-5 Elements		ISO 9000 Elements	
A1	Title and Approval Sheet		
A2	Table of Contents		
A3	Distribution List		
A4	Project/Task Organization	4	Management Responsibility
A5	Problem Definition/Background		
A6	Project/Task Description		
A7	Quality Objectives and Criteria for Measurement Data	5 5.2	Quality System Principles Structure of the Quality System
A8	Project Narrative		
A9	Special Training Requirements/Certification		
A10	Documentation and Records		
B1	Sampling Process Design	8	Quality in Specification and Design
B2	Sampling Methods Requirements	10	Quality of Production
B3	Sample Handling and Custody Requirements	16	Handling and Post Production Functions
B4	Analytical Methods Requirements	10	Quality of Production
B5	Quality Control Requirements	11	Control of Production
B6	Instrument/Equipment Testing, Inspection, and Maintenance Requirements	13	Control of Measuring and Test Equipment
B7	Instrument Calibration and Frequency		
B8	Inspection/Acceptance Requirements for Supplies and Consumables	9 11.2	Quality in Procurement Material Control and Traceability
B9	Data Acquisition Requirements		
B10	Data Quality Management		
C1	Assessments and Response Actions	5.4 14 15	Auditing the Quality System Nonconformity Corrective Action
C2	Reports to Management	5.3 6	Documentation of the Quality System Economics - Quality Related Costs
D1	Data Review, Validation, and Verification Requirements	11.7	Control of Verification Status
D2	Validation and Verification Methods	12	Verification Status
D3	Reconciliation with User Requirements		
		7	Quality in Marketing

A4. Crosswalk Between DQOs and the QAPP
(Roman Numerals refer to QAPP use categories, see Appendix B)

	Elements	Requirements	DQO Overlap
	PROJECT MANAGEMENT		
A1	Title and Approval Sheet I, II, III, IV	Title and approval sheet.	None
A2	Table of Contents I, II, III	Document control format.	None
A3	Distribution List I, II, III, IV	Distribution list for the QAPP revisions and final guidance.	List the members of the scoping team. Step 1: State the Problem.
A4	Project/Task Organization I, II, III, IV	Identify individuals or organizations participating in the project and discuss their roles, responsibilities and organization.	Step 1: State the Problem requires definition of the DQO scoping or planning team, which includes the decision maker, technical staff, data users, etc. This step also requires the specification of each member's role and responsibilities.
A5	Problem Definition/Background I, II, III, IV	1) State the specific problem to be solved or decision to be made. 2) Identification of the decision maker and the principal customer for the results.	Step 1: State the Problem requires a description of the problem. It also identifies the decision maker and decision makers who could use the data.
A6	Project/Task Description I, II, III	1) Hypothesis test, 2) expected measurements, 3) ARARs or other appropriate standards, 4) assessment tools (technical audits), 5) work schedule and required reports.	Step 1: State the Problem requires work schedule. Step 3: Identify the Inputs requires the ARARs or standards and expected measurements. Step 6: Specify Limits on Decision Errors.
A7	Data Quality Objectives for Measurement Data I, II, III	Decision(s), population parameter of interest, action level, summary statistics and acceptable limits on decision errors. Also scope of the project (domain or geographical locale).	Steps 1: State the Problem, Step 4 Define the Boundaries, Step 5: Develop a Decision Rule, Step 6: Specify Limits on Decision Errors.
A8	Project Narrative (ORD Only) IV	Anticipated data use, definition of project success, survey design requirements and description, sample type and location, COC, PE samples for total measurement process, sampling and analytical instrumentation requirements, and audit and review plans.	Steps 5: Develop a Decision Rule and Step 7: Optimize the Design for Obtaining Data.
A9	Special Training Requirements/Certification I	Identify special training that personnel will need.	None

A4. Crosswalk Between DQOs and the QAPP
(Roman Numerals refer to QAPP use categories, see Appendix B)

Elements		Requirements	DQO Overlap
A10	Documentation and Record I, II, III	Itemize the information and records which must be included in a data report package including report format and requirements for storage etc.	None
MEASUREMENT/DATA ACQUISITION			
B1	Sampling Process Designs (Experimental Design) I, II, III	Outline the experimental design, including sampling design and rationale, sampling frequencies, matrices, and measurement parameter of interest.	Step 7: Optimize the Design for Obtaining Data Step 5: Develop a Decision Rule.
B2	Sampling Methods Requirements I, II, III	Sample collection method and approach	Step 7: Optimize the Design for Obtaining Data
B3	Sample Handling and Custody Requirements I, II, III	Describe the provisions for sample labeling, shipment, chain-of- custody forms, procedures for transferring and maintaining custody of samples.	None
B4	Analytical Methods Requirements I, II, III	Identify analytical method(s) and equipment for the study include method performance requirements.	Step 3: Identify Inputs to the Decision, Step 7: Optimize the Design for Obtaining Data
B5	Quality Control Requirements I, II, III	Describe routine (real time) QC procedures that should be associated with each sampling and measurement technique. List required QC checks and corrective action procedures.	None
B6	Instrument/Equipment Testing Inspection and Maintenance Requirements I, II	Discuss how inspection and acceptance testing, including the use of QC samples, must be performed to ensure their intended use as specified by the design.	None
B7	Instrument Calibration and Frequency I, II, III	Identify tools, gauges and instruments, and other sampling or measurement devices that need calibration. Describe how the calibration should be done.	None
B8	Inspection/Acceptance Requirements for Supplies and Consumables I	Define how and by whom the sampling supplies and other consumables will accepted for use in the project.	None
B9	Data Acquisition Requirements (Non-direct Measurements) I, II,III	Define criteria for the use of non-measurement data such as data that come from databases or literature.	None

A4. Crosswalk Between DQOs and the QAPP
(Roman Numerals refer to QAPP use categories, see Appendix B)

Elements		Requirements	DQO Overlap
B10	Data Management I, II	Outline of data management scheme including path of data, use of storage and the record keeping system. Identify all data handling equipment and procedures that will be used to process, compile and analyze the data.	None
		ASSESSMENT/OVERSIGHT	
C1	Assessments and Response Actions I, II, III	Describe the assessment activities needed for this project. These may include DQA, PE, TSA, MSR/PR/RR,	None
C2	Reports to Management I, II, III	Identify the frequency, content and distribution of reports issued to keep management informed.	None
		DATA VALIDATION AND USABILITY	
D1	Data Review, Validation, and Verification Requirements I, II, III	State the criteria used to accept or reject the data based on quality.	None
D2	Validation and Verification Methods I, II	Describe the process to be used for validating and verifying data, including chain of custody for data throughout the lifetime of the project.	None
D3	Reconciliation With Data Quality Objectives I, II, III	Describe how results will be evaluated to determine if DQOs are satisfied.	None

A5. U.S. EPA Quality System Requirements and Guidance Documents

Summary of Documents

This section presents a brief discussion of the EPA Quality System requirements series (EPA QA/R-x) and guidance series (EPA QA/G-x) documents. Figure A2 illustrates the relationship among the various documents.

EPA QA/R-1: EPA Quality Systems Requirements for Environmental Programs. QA/R-1 is the external policy document by which EPA will announce its implementation of the American National Standard ANSI/ASQC E4-1994, *Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs*. An internal preliminary draft has been completed and is awaiting formal adoption of the standard by EPA. The same information will be part of the EPA *Quality Manual for Environmental Programs*, an internal policy manual. When E4 has been formally adopted by EPA, the draft will be distributed for comment. Target Availability: External Draft, Spring 1997.

EPA QA/G-1: Guidance for Developing Quality Systems for Environmental Data Operations. QA/G-1 provides non-mandatory guidance to help organizations develop a QA program that will meet EPA expectations and requirements. There is no draft currently available. Target Availability: Draft, Summer 1997.

EPA QA/R-2: EPA Requirements for Quality Management Plans. QA/R-2 is the policy document containing the specifications and requirements for Quality Management Plans (QMPs) for organizations with which EPA has extramural agreements. An Interim Final version is awaiting Agency approval for release and is expected to be available for public comment and use shortly. QA/R-2 is the intended replacement for QAMS-004/80. The same information contained in this document is found in the EPA *Quality Manual for Environmental Programs*, an internal policy manual. Current Draft Version: August 1994. Target Availability: Final, Spring 1997.

EPA QA/R-2A: EPA Requirements for Quality Management Plans for Analytical Laboratories and Facilities. QA/R-2A will provide detailed requirements for environmental analytical labs. Since there may be a national consensus standard for labs, the content of this document is unclear at present. This is still a planning item. Target Availability: Undetermined.

EPA QA/G-2: Guidance for Preparing Quality Management Plans. QA/G-2 provides non-mandatory guidance to help organizations develop a Quality Management Plans (QMPs) that will meet EPA expectations and requirements. The document will contain tips, advice, and case studies to help users develop improved QMPs. There is no draft currently available. Target Availability: Draft, Spring 1997.

EPA QA/G-3: Guidance for the Management Systems Review Process. QA/G-3 provides non-mandatory guidance to help organizations plan, implement, and evaluate management assessments of their quality systems. The guidance will present a step-by-step description of the MSR process. The revised third draft will be issued for internal EPA comments in early 1997. Current Draft Version: January 1994. Target Availability: Spring 1997.

EPA QA/G-4: Guidance for the Data Quality Objectives Process. QA/G-4 provides non-mandatory guidance to help organizations plan, implement, and evaluate the Data Quality Objectives (DQO) process, with a focus on environmental decision-making for regulatory and enforcement decisions. The guidance presents a step-by-step description of the DQO process. This document is available now. Final Version: EPA/600/R-96/055, September 1994.

EPA QA/G-4D: DEFT Software for the Data Quality Objectives Process. QA/G-4D provides non-mandatory guidance for using the Decision Error Feasibility Trials (DEFT) software to help organizations plan, implement, and evaluate the Data Quality Objectives (DQO) process. The guidance presents a step-by-step description of the use of the PC-based DEFT software DQO process. This document is available now. Final Version: EPA/600/R-96/056, September 1994.

EPA QA/G-4R: Guidance for the Data Quality Objectives Process for Researchers. QA/G-4R provides non-mandatory guidance on the application of the Data Quality Objectives (DQO) Process for researchers and experimenters. The guidance integrates the DQO Process with statistical design of experiments. There is no draft currently available. Target Availability: August 1997.

EPA QA/G-4HW: Guidance for the Data Quality Objectives Process for Hazardous Waste Site Testing. QA/G-4HW provides non-mandatory guidance to help organizations plan, implement, and evaluate the statistics-based Data Quality Objectives (DQO) process as applied to hazardous waste sampling activities. The guidance will present a step-by-step description of the DQO process and its application to sampling designs for environmental remediation and waste management activities. There is no draft currently available, although a predecessor document, *Data Quality Objectives Process for Superfund: Interim Final Guidance* (EPA540-R-93-071, September 1993), was developed by OERR with support from QAD and has been available since early 1994 through NTIS (PB94-963203). Availability of G-4HW: Final, January 1997.

EPA QA/R-5: EPA Requirements for Quality Assurance Project Plans. QA/R-5 is the intended replacement for QAMS-005/80. This external policy document will establish the requirements for QA Project Plans prepared for activities conducted by or funded by EPA. It is intended for use by organizations having contracts or extramural agreements with EPA. Current Draft Version: August 1994. Availability: Final, Spring 1997.

EPA QA/G-5: Guidance on Quality Assurance Project Plans. QA/G-5 provides non-mandatory guidance to help organizations develop a Quality Assurance Project Plans (QAPPs) that will meet EPA expectations and requirements. The document will provide a linkage between the DQO process and the QAPP. It will contain tips, advice, and case studies to help users develop improved QAPPs. Target Availability: External Draft, January 1997.

EPA QA/G-5S: Guidance on Sampling Designs to Support QA Project Plans. QA/G-5S provides non-mandatory guidance on practical methods for developing sampling plans to satisfy the guidelines outlined in the statistics-based DQO Process (QA/G-4) and the QAPP (QA/G-5). Different sampling schemes are discussed and the relative strengths and weaknesses outlined. There is no draft currently available. Target Availability: External Draft, August 1997.

EPA QA/G-6: Guidance for the Preparation of Operating Procedures for Quality-Related Operations. QA/G-6 provides nonmandatory guidance to help organizations develop and document Standard Operating Procedures (SOPs). The document contains tips, advice, and case studies to help users develop improved SOPs. This document is available now. Final Version: EPA/600/R-96/027, November 1995.

EPA QA/G-7: Guidance for Determining Quality Training Requirements for Environmental Data Operations. QA/G-7 will provide non-mandatory guidance to help organizations determine and develop training requirements for their programs. The document will contain tips, advice, and case studies to help users develop improved processes for making training determinations and estimates. This is currently a planning item. QAD expects to use a Work Group process to develop this guidance. Target Availability: Undetermined.

EPA QA/G-8: Guidance on Technical Assessments for Environmental Data Operations. QA/G-8 will provide non-mandatory guidance to help organizations plan, conduct, evaluate, and document technical assessments for their programs. Such technical assessments include Technical Systems Audits (TSAs), surveillance, readiness reviews, and Performance Evaluations (PEs). The document will contain tips, advice, and case studies to help users develop improved processes for conducting technical assessments. This is currently a planning item. QAD expects to use a Work Group process to develop this guidance. Target Availability: Draft, Fall 1997.

EPA QA/G-9: Guidance for the Data Quality Assessment Process. QA/G-9 provides non-mandatory guidance for planning, implementing, and evaluating retrospective assessments of the quality of the results from environmental data operations. DQA is a statistically-based, quantitative evaluation of the extent to which a data set satisfies the user's needs (or DQOs). This particular document is aimed at the project managers who are responsible for conducting the environmental data operations and assessing the usability of the results. This document is available now. Final Version: EPA/600/R-96-084, July 1996.

EPA QA/G-9D: Guidance for DataQUEST, the Data Quality Assessment Process Software. QA/G-9D provides non-mandatory guidance for planning, implementing, and evaluating retrospective assessments of the quality of the results from environmental data operations using the PC-based software, DataQUEST. Availability: External Working Draft, August 1996, is currently available.

EPA QA/R-10: EPA Quality Assurance Requirements for Computer Hardware and Software Systems for Environmental Programs. QA/R-10 will establish requirements for quality in the use of computer hardware and software. This is a planning item. There is no draft currently available. The document will be developed jointly by QAD and the Office of Information Resources Management (OIRM). Target Availability: Undetermined.

EPA QA/G-10: Guidance for Implementing Quality Assurance Requirements for Computer Hardware and Software Systems for Environmental Programs. QA/G-10 will provide non-mandatory guidance for assuring quality in the use of computer hardware and software. This is a planning item. There is no draft currently available. The document will be developed jointly by QAD and the Office of Information Resources Management (OIRM). Target Availability: Undetermined.

EPA QA/G-11: Guidance on Decision Quality Planning for Project Managers. QA/G-11 will provide non-mandatory guidance for assuring quality in the planning of environmental programs and projects. Its intention is to help project managers integrate quality management principles and practices into their project activities. This is a planning item. There is no draft currently available. Target Availability: Undetermined.

Notes on the Quality System Series Documents

(1) Requirements Documents (identified as QA/R-x) will also be the subject of chapters in the EPA *Quality Manual for Environmental Programs*. The *Quality Manual* requirements will apply to EPA Program Offices, Regions, and ORD laboratories. The QA/R-x versions will apply to EPA contractors and organizations receiving financial assistance from EPA (e.g., grants, cooperative agreements, and inter-agency agreements). They also will be issued as policy documents under the signature of the AA/ORD.

(2) Guidance Documents (identified as QA/G-x) will be published as ORD reports after the appropriate peer and policy reviews and issued under the signature of the AA/ORD.

Availability of Documents as of October 1996

Documents that are in final form are as follows:

- QA/G-4: *Guidance for the Data Quality Objectives Process* (EPA/600/R-96/055, September 1994)
- QA/G-4D: *DEFT Software for the Data Quality Objectives Process*, V. 4.0 (EPA/600/R-96/056, September 1994)
- QA/G-6: *Guidance for the Preparation of Operating Procedures for Quality-related Operations* (EPA/600/R-96/027, November 1995)
- QA/G-9: *Guidance for the Data Quality Assessment Process* (EPA/600/R-96/084, July 1996)

Draft reports that are available for distribution are as follows:

- QA/R-2: *EPA Requirements for Quality Management Plans* (August 1994)
- QA/G-3: *Guidance for the Management Systems Review Process* (January 1994)
- QA/R-5: *EPA Requirements for Quality Assurance Project Plans* (August 1994)
- QA/G-9D: *Guidance for DATAQUEST - the Data Quality Assessment Process Software* (August 1996)

Documents that are in progress or for which drafts are not available are as follows:

- QA/R-1: *EPA Quality Systems Requirements for Environmental Programs*
- QA/G-4HW: *Guidance for the Data Quality Objectives Process for Hazardous Waste Site Testing*

Documents that are planned for future development are as follows:

- QA/G-1: *Guidance for Developing Quality Systems for Environmental Data Operations*
- QA/G-2: *Guidance for Preparing Quality Management Plans*
- QA/R-2A: *EPA Requirements for Quality Management Plans for Analytical Laboratories and Facilities*
- QA/G-4R: *Guidance for the Data Quality Objectives Process for Researchers*
- QA/G-5S: *Guidance on Sampling Plans*
- QA/G-7: *Guidance for Determining Quality Training Requirements for Environmental Data Operations*
- QA/G-8: *Guidance on Technical Assessments for Environmental Data Operations*
- QA/R-10: *EPA Quality Assurance Requirements for Computer Hardware and Software Systems for Environmental Programs*
- QA/G-10: *Guidance for Implementing Quality Assurance Requirements for Computer Hardware and Software Systems for Environmental Programs*
- QA/G-11: *Guidance on Decision Quality Planning for Project Managers*

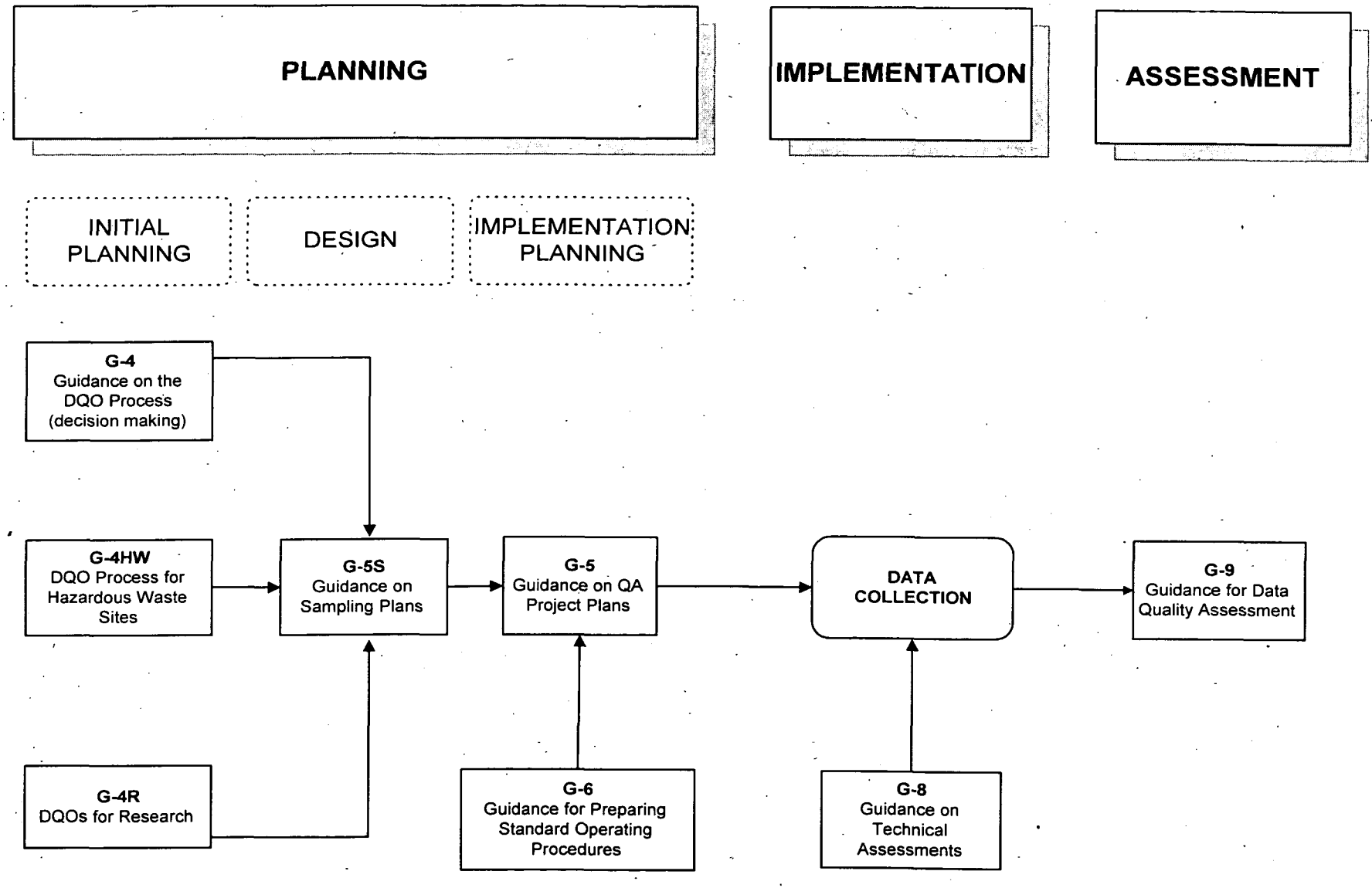


Figure A2. Relationship among EPA Quality System documents

APPENDIX B

QAPP-USE CATEGORIES

The diversity and variability in the mission requirements of the organizations that make up EPA (e.g., program offices, regions, research laboratories) may not allow the user to define a single checklist of elements and details needed for all QAPPs. To provide flexibility, several EPA organizations have used an optional approach that categorizes QAPP requirements according to the type of work being performed and the intended use of the data. These nonmandatory QAPP-use categories vary the level of detail and rigor prescribed for a particular QAPP. While this approach is being used most frequently by the Office of Research and Development (ORD), it may have applicability to other programs.

These categories may be an aid to determining the level of detail that may be needed in a QAPP for a particular type of work. This approach recognizes that not all environmental data operations require QAPPs with the same level of detail. For example, data collected for compliance or enforcement decisions in a Region will require a more comprehensive QAPP than an exploratory research project conducted for an ORD R&D laboratory. The categories are:

- **Category I: Direct Support to Rulemaking, Enforcement, Regulatory, or Policy Decisions.** The projects include environmental data operations that directly support rulemaking, enforcement, regulatory, or policy decisions. They also include research projects of significant national interest, such as those typically monitored by the Administrator. Category I projects require the most detailed and rigorous QA and QC for legal and scientific defensibility. Category I projects are typically stand-alone; that is, the results from such projects are sufficient to make the needed decision without input from other projects.

- **Category II: Complementary support to Rulemaking, Regulatory, or Policy Decisions.** These projects include environmental data operations that complement other projects in support of rulemaking, regulatory, or policy decisions. Such projects are of sufficient scope and substance that their results could be combined with those from other projects of similar scope to provide the necessary information for decisions. Category II projects may also include certain high-visibility projects as defined by EPA management.

- **Category III: Interim Studies.** These projects include environmental data operations performed as interim steps in a larger group of operations. Such projects include testing research hypotheses, estimating effects, developing methods, and other work producing results that are used to evaluate and select options for interim decisions or to perform feasibility studies or preliminary assessments or unexplored areas for possible future work.

- **Category IV: Basic Studies.** These are projects involving environmental data operations to study basic phenomena or issues, including proof of concept and qualitative screening for particular analytical species.

The determination of a project's category is made by the EPA QA Manager in consultation with the EPA project manager in the organization[1] responsible for the work. It should be noted that projects

[1]Organization refers to the EPA Program Office, Region, or ORD Laboratory having an approved Quality Management Plan that describes its quality system for planning, implementing, and assessing environmental programs.

may contain specific tasks or subtasks that vary in the level of QA/QC requirements and these conditions should be considered when deciding on the use category for a particular project. A chart identifying the categories assigned to each QAPP Element follows.

CATEGORY	ELEMENT	DESCRIPTION
PROJECT MANAGEMENT		
I, II, III, IV	A1	Title and Approval Sheet
I, II, III	A2	Table of Contents
I, II, III, IV	A3	Distribution List
I, II, III	A4	Project/Task Organization
I, II, III	A5	Problem Definition/Background
I, II, III	A6	Project/Task Description
I, II, III	A7	Quality Objectives and Criteria for Measurement Data
IV	A8	Project Narrative (ORD Only)
I	A9	Special Training Requirements/Certification
I, II, III	A10	Documentation and Records
MEASUREMENT/DATA ACQUISITION		
I, II, III	B1	Sampling Process Design (Experimental Design)
I, II, III	B2	Sampling Methods Requirements
I, II, III	B3	Sample Handling and Custody Requirements
I, II, III	B4	Analytical Methods Requirements
I, II, III	B5	Quality Control Requirements
I, II	B6	Instrument/Equipment Testing, Inspection, and Maintenance Requirements
I, II, III	B7	Instrument Calibration and Frequency
I	B8	Inspection/Acceptance Requirements for Supplies and Consumables
I, II, III	B9	Data Acquisition Requirements (Non-direct Measurements)
I, II	B10	Data Management
ASSESSMENT/OVERSIGHT		
I, II, III	C1	Assessments and Response Actions
I, II, III	C2	Reports to Management
DATA VALIDATION AND USABILITY		
I, II, III	D1	Data Review, Validation, and Verification Requirements
I, II	D2	Validation and Verification Methods
I, II, III	D3	Reconciliation with Data Quality Objectives

References

Johnson, Gary L., and Judith S. Ford. 1985. *AEERL Quality Assurance Procedures Manual.* U.S. Environmental Protection Agency. April.

Simes, Guy F. 1991. *Preparation Aids for the Development of Category I Quality Assurance Project Plans.* U.S. Environmental Protection Agency. EPA/600/8-91/003. February.

Simes, Guy F. 1991. *Preparation Aids for the Development of Category II Quality Assurance Project Plans.* U.S. Environmental Protection Agency. EPA/600/8-91/004. February.

Simes, Guy F. 1991. *Preparation Aids for the Development of Category III Quality Assurance Project Plans.* U.S. Environmental Protection Agency. EPA/600/8-91/005. February.

Simes, Guy F. 1991. *Preparation Aids for the Development of Category IV Quality Assurance Project Plans.* U.S. Environmental Protection Agency. EPA/600/8-91/006. February.

APPENDIX C

CHECKLISTS USEFUL IN QA REVIEW

C1. Sample Handling, Preparation, and Analysis Checklist

This checklist covers most of the appropriate elements performed during the analysis of environmental samples. Functions not appropriate for a specific analysis should be annotated.

Information on the collection and handling of samples should be completely documented to allow the details of sample collection and handling to be recreated. All information should be entered in ink at the time the information was being generated in a permanently bound logbook. Errors should not be erased or clocked-out but corrected by putting a line through the erroneous information and by entering, initializing, and dating the correct information. Blank spaces should have an obliterating line drawn through to prevent addition of information. Each set of information should have an identifying printed name, signature, and initials.

Sample Handling

• Field Logs	The documentation of events occurring field sampling and to identify individual field samples
• Sample Labels	Used to link individual samples with field log and Chain-of-Custody record
• Chain-of-Custody	Documentation of exchange and transportation of samples form the field to final analysis
• Sample Receipt Log	Documentation of receipt of the laboratory or organization of entire set of individual samples for analysis

Sample Preparation and Analysis

• Sample Preparation Log	Documents the preparation of samples for a specific method or procedure
• Sample Analysis Log	Records information on the analysis and calculation of analytical results
• Instrument Run Log	Records analyses of calibration standards, field samples, and quality control samples

Chemical Standards

• Chemical Standard Receipt Log	Records of receipt analytical standards and chemicals
• Standards/Reagent Preparation Log	Records of preparation of internal standards, reagents, spiking solutions, surrogate solutions, and reference materials

Field Logs

ELEMENT	COMMENT
Project name/ID and location	
Sampling personnel	
Geological observations including map	
Atmospheric conditions	
Field measurements	
Sample dates, times, and locations	
Sample identifications present	
Sample matrix identified	
Sample descriptions (e.g., odors and colors)	
Number of samples taken per location	
Description of any QC samples	
Any deviations from the sampling plan	
Difficulties in sampling or unusual circumstances	

Sample Labels

Sample ID	
Date and time of collection	
Sampler's signature	
Characteristic or parameter investigated	
Preservative used	

Chain of Custody Records

ELEMENT	COMMENT
Project name/ID and location	
Sample custodian signatures verified and on file	
Date and time of each transfer	
Carrier ID number	
Integrity of shipping container and seals verified	
Standard Operating Procedures for receipt on file	
Samples stored in same area	
Holding time protocol verified	
Standard Operating Procedure for sample preservation on file	
Identification of proposed analytical method verified	
Proposed analytical method documentation verified	
QA Plan for proposed analytical method on file	

Sample Receipt Log

Date and time of receipt	
Sample collection date	
Client sample ID	
Number of samples	
Sample matrices	
Requested analysis, including method number(s)	
Signature of the sample custodian or designee	
Sampling kit code (if applicable)	
Sampling condition	
Chain-of-Custody violations and identities	

SAMPLE PREPARATION AND ANALYSIS

Sample Preparation Logs

ELEMENT	COMMENT
Parameter/analyte of investigation	
Method number	
Date and time of preparation	
Analyst's initials or signature	
Initial sample volume or weight	
Final sample volume	
Concentration and amount or spiking solutions used	
Quality control samples included with the sample batch	
ID for reagents, standards and spiking solutions used	

Sample Analysis Logs

ELEMENT	COMMENT
Parameter analyte of investigation	
Method number/reference	
Date and time of analysis	
Analyst's initials or signatures	
Laboratory sample ID	
Sample aliquot	
Dilution factors and final sample volumes (if applicable)	
Absorbance values, peak heights, or initial concentrations reading	
Final analyte concentration	
Calibration data (if applicable)	
Correlation coefficient (including parameters)	
Calculations of key quantities available	
Comments on interferences or unusual observations	
Quality control information, including percent recovery	

Instrument Run Logs

ELEMENT	COMMENT
Name/type of instrument	
Instrument manufacturer and model number	
Serial number	
Date received and date placed in service	
Instrument ID assigned by the laboratory (if used)	
Service contract information, including service representative details	
Description of each maintenance or repair activity performed	
Date and time when of each maintenance or repair activity	
Initials of maintenance or repair technicians	

CHEMICAL STANDARDS

Chemical/Standard Receipt Logs

ELEMENT	COMMENT
Laboratory control number	
Date of receipt	
Initials or signature of person receiving chemical	
Chemical name and catalog number	
Vendor name and log number	
Concentration or purity of standard	
Expiration date	

Standards/Reagent Preparation Log

ELEMENT	COMMENT
Date of preparation	
Initials of the analyst preparing the standard solution or reagent	
Concentration or parity of standard or reagent	
Volume or weight of the stock solution or neat materials	
Final volume of the solution being prepared	
Laboratory ID/control number assigned to the new solution	
Name of standard reagent	
Standardization of reagents, titrants, etc. (If applicable)	
Expiration date	

Reference

1. Roserance, A. and L. Kibler. 1994. Generating Defensible Data, *Environmental Testing and Analysis.* May/June.

2. Roserance, A. and L. Kibler. 1996. "Documentation and Record Keeing Guidelines." In *Proceedings of the 12th Annual Waste Testing and Quality Assurance Symposium.* July.

C2. QAPP Review Checklist

QAPP REVIEW CHECKLIST

	COMMENTS
1. Title & Approval Sheet	
Title	
Organization's name	
Dated signature of project manager	
Dated signature of quality assurance officer	
Other signatures, as needed	
2. Table of Contents	
3. Distribution List	
4. Project/Task Organization	
Identifies key individuals, with their responsibilities (data users, decision-makers, project QA manager, subcontractors, etc.)	
Organization chart shows lines of authority & reporting responsibilities	
5. Problem Definition/Background	
Clearly states problem or decision to be resolved	
Provides historical & background information	
6. Project/Task Description	
Lists measurements to be made	
Cites applicable technical, regulatory, or program-specific quality standards, criteria, or objectives	
Notes special personnel or equipment requirements	
Provides work schedule	
Notes required project & QA records/reports	
7. Quality Objectives & Criteria for Measurement Data	
States project objectives and limits, both qualitatively & quantitatively	
States & characterizes measurement quality objectives as to applicable action levels or criteria	
8. Project Narrative (ORD projects only)	

QAPP REVIEW CHECKLIST

	COMMENTS
9. Special Training Requirements/Certification Listed	
States how provided, documented, & assured	
10. Documentation & Records	
Lists information & records to be included in data report (e.g. raw data, field logs, results of QC checks, problems encountered)	
States requested lab turnaround time	
Gives retention time and location for records & reports	
11. Sampling Process Design (Experimental Design) States the following:	
Samples required as to type & number	
Sampling network design & rationale	
Sampling locations & frequency of sampling	
Sample matrices	
Classification of each measurement parameter as either critical or needed for information only	
Appropriate validation study information, for non-standard situations	
12. Sampling Methods Requirements	
Identifies sample collection procedures & methods	
Lists equipment needs	
Identifies support facilities	
Identifies individuals responsible for corrective action	
13. Sample Handling & Custody Requirements	
Notes sample handling requirements	
Notes chain of custody procedures, if required	
14. Analytical Methods Requirements	
Identifies analytical methods to be followed (with all options) & required equipment	
Provides validation information for non-standard methods	

QAPP REVIEW CHECKLIST

	COMMENTS
Identifies individuals responsible for corrective action	
15. Quality Control Requirements	
Identifies QC procedures & frequency for each sampling, analysis, or measurement technique, as well as associated acceptance criteria & corrective action	
References procedures used to calculate QC statistics (precision & bias or accuracy)	
16. Instrument/Equipment Testing, Inspection, & Maintenance Requirements	
Identifies acceptance testing of sampling & measurement systems	
Describes equipment preventive & corrective maintenance	
Notes availability & location of spare parts	
17. Instrument Calibration & Frequency	
Identifies equipment needing calibration & frequency for such calibration	
Notes required calibration standards and/or equipment	
Cites calibration records & manner traceable to equipment	
18. Inspection/Acceptance Requirements for Supplies & Consumables	
States acceptance criteria for supplies & consumables	
Notes responsible individuals	
19. Data Acquisition Requirements for Non-direct Measurements	
Identifies type of data needed from non-measurement sources (e.g. computer data bases and literature files), along with acceptance criteria for their use.	
Describes any limitations of such data	
20. Data Management	
Describes standard record keeping & data storage & retrieval requirements	
Checklists or standard forms attached to QAPP	

QAPP REVIEW CHECKLIST

	COMMENTS
Describes data handling equipment & procedures used to process, compile, and analyze data (e.g. required computer hardware and software)	
21. Assessments & Response Actions	
Lists required number, frequency & type of assessments, with approximate dates & names of responsible personnel (Assessments include but are not limited to peer review, management systems review, technical systems audits, performance evaluations, and audits of data quality)	
Identifies individuals responsible for corrective actions	
22. Reports to Management Identifies frequency and distribution of reports for:	
Project status	
Results of performance evaluations and audits	
Results of periodic data quality assessments	
Any significant QA problems	
Preparers and recipients of reports	
23. Data Review, Validation, & Verification	
States criteria for accepting, rejecting, or qualifying data	
Includes project-specific calculations or algorithms	
24. Validation & Verification Methods	
Describes process for data validation & verification	
Identifies issue resolution procedure & responsible individuals	
Identifies method for conveying these results to data users	
25. Reconciliation with User Requirements	
Describes process for reconciling project results with DQOs & reporting limitations on use of data	

Reference

Personal Communication, Margo Hunt, EPA Region II, February, 1996.

C3. Chain-of-Custody Checklist

Item	Y	N	Comment
1. Is a sample custodian designated? If yes, name of sample custodian.			
2. Are the sample custodian's procedures and responsibilities documented? If yes, where are these documented?			
3. Are written Standard Operating Procedures (SPO) developed for receipt of samples? If yes, where are the SOP documented (laboratory manual, written instructions, etc.)?			
4. Is the receipt of chain-of-custody record(s) with samples being documented? If yes, where is this documented?			
5. Is the non-receipt of chain-of-custody record(s) with samples being documented? If yes, where is this documented?			
6. Is the integrity of the shipping container(s) being documented (custody seal(s) intact, container locked, or sealed properly, etc.)? If yes, where is security documented?			
7. Is the lack of integrity of the shipping container(s) being documented (i.e., evidence of tampering, custody seals broken or damaged, locks unlocked or missing, etc.)? If yes, where is non-security documented?			
8. Is agreement between chain-of-custody records, and sample tags being verified? If yes, state source of information.			
9. Is the agreement or non-agreement verification being documented? If yes, where, is this documented?			
10. Are sample tag numbers recorded by the Sample Custodian? If yes, where are they recorded?			
11. Are written Standard Operating Procedures (SOP) developed for sample storage? If yes, where are the SOP documented (laboratory manual, written instructions, etc.)?			
12. Are samples stored in a secure area? If yes, where and how are they stored?			

Item	Y	N	Comment
13. Is sample identification maintained? If yes, how?			
14. Is sample extract (or inorganics concentrate) identification maintained? If yes, how?			
15. Are samples that require preservation stored in such a way as to maintain their preservation? If yes, how are the samples stored?			
16. Based upon sample records examined to determine holding-times, are sample holding-times limitations being satisfied? Sample records used to determine holding-times:			
17. Are written Standard Operating Procedures (SOP) developed for sampling handling and tracking? If yes, where are the SOP documented (laboratory manual; written instructions, etc.)?			
18. Do laboratory records indicate personnel receiving and transferring samples in the laboratory? If yes, what laboratory records document this?			
19. Does each instrument used of sample analysis (GC,GC/MS, AA, etc.) have an instrument log? If no, which instruments do not?			
20. Are analytical methods documented and available to the analysts? If yes, where are these documented?			
21. Are quality assurance procedures documented and available to the analysts? If yes, where are these documented?			
22. Are written Standard Operating Procedures (SOP) developed for compiling and maintaining sample document files? If yes, where are the SOP documented (laboratory manual, written instructions, etc.)?			
23. Are sample documents filed by case number? If no, how are documents filed?			
24. Are sample document file inventoried?			
25. Are documents in the case files consecutively numbered according to the file inventories?			

Item	Y	N	Comment
26. Are documents in the case files stored in a secure area? If yes, where and how are they stored?			
27. Has the laboratory received any confidential documents?			
28. Are confidential documents segregated from other laboratory documents? If no, how are they filed?			
29. Are confidential documents stored in a secure manner? If yes, where an how are they stored?			
30. Was a debriefing held with laboratory personnel after the audit was completed?			
31. Were any recommendations made to laboratory personnel during the debriefing?			

APPENDIX D

DATA QUALITY INDICATORS

INTRODUCTION

Data Quality Indicators (DQIs) are qualitative and quantitative descriptors used to interpret the degree of acceptability or utility of data to the user. The principal DQIs are precision, bias, representativeness, comparability, and completeness. Establishing acceptance criteria for the DQIs sets quantitative goals for the quality of data generated in the analytical measurement process. DQIs may be expressed for entire measurement systems, but it is customary to allow DQIs to be applied to only to laboratory measurement processes. The issues of design and sampling errors, the most influential components of variability, are discussed separately in EPA QA/G-5S, *Guidance on Sampling Designs to Support QAPPs.*

Of the five principal DQIs, precision and bias are quantitative measures that can be controlled; representativeness, comparability, and completeness are more qualitative. Less detailed definitions are provided for other DQIs.

The five principal DQIs are also referred to by the acronym PARCC, with the "A" in PARCC referring to accuracy instead of bias. This inconsistency is because some analysts believe accuracy and bias are synonymous, and PARCC is a more convenient acronym than PBRCC. Accuracy is comprised of random error (precision) and systematic error (bias), and these indicators are discussed separately.

D1. PRINCIPAL DQIS: PARCC

Precision

Precision is a measure of agreement among individual measurements of the same property, under prescribed similar conditions. Precision is determined by measuring the agreement among a number of individual measurements of the same sample or concentration. This agreement is calculated as either the range (R) (for duplicate measurements) or as the standard deviation (s). It may also be expressed as a percentage of the mean of the measurements, relative range (RR) (for duplicates) or relative standard deviation (RSD). Appendix K, "Calculation of Statistical Quantities," contains formulae and examples of these quantities.

For analytical procedures, precision may be specified as either intralaboratory (within a laboratory) or interlaboratory (between laboratories) precision. Intralaboratory precision estimates represent the agreement expected when a single laboratory uses the method to make repeated measurements of the same sample. Interlaboratory precision refers to the agreement expected when two or more laboratories analyze the same or identical samples with the same method. Intralaboratory precision is more commonly reported; however, where available, both intralaboratory and interlaboratory precision are listed in the data compilation.

The Measurement of Precision

A sample subdivided in the field and preserved separately is used, where possible, to assess the variability of sample handling, preservation, and storage along with the variability of the analysis process. The subsection of field instrument measurement discusses this further.

Collocated samples when collected, processed, and analyzed by the same organization provide intralaboratory precision information on sample acquisition, handling, shipping, storage, preparation and analysis. Both samples can be carried through the steps in the measurement process together providing an estimate of short-term precision. Likewise, the two samples, if separated and processed at different times or by different people, and/or analyzed using different instruments, provide an estimate of long-term precision. This subject is discussed further in the subsection on laboratory measurement.

Calculation of the Summary Precision Statistics

The summary statistics are developed from the basic statistics gathered throughout the project or time period represented. Because the precision of environmental measurement systems is often a function of concentration (e.g., as concentration increases, standard deviation increases), this relationship should be evaluated before selecting the most appropriate form of the summary statistic. An evaluation of the basic precision statistics as a function of concentration will usually lead to one of three conclusions: (1) standard deviation (or range) is independent of concentration (i.e., constant); (2) standard deviation (or range) is directly proportional to concentration, and coefficient of variation (or relative range) is constant; or (3) both standard deviation (or range) and coefficient of variation (or relative range) vary with concentration.

For simplicity of use and interpretation, the relationship most easily described should be selected for use; i.e., for case (1) the standard deviation (or range) is simplest to work with, whereas, for case (2), the coefficient of variation (or relative range) is simplest. If the relationship of precision to concentration falls into case (3), regression analysis can be used to estimate the relationship between standard deviation (or range) and concentration.

The decision as to which case is applicable can be based on plots of precision versus concentration or by regressions of s (or R) or CV (or RR) versus concentration. The preferred measure of precision is standard deviation as this provides the maximum amount of information.

Reporting Precision

Procedures for presenting precision estimates in order of preference are as follows:

1. Precision as a function of the measured value across the applicable range (Ideally, the presentation could be a graph containing the actual data points, the mathematical relationship providing the best-fit curve, and confidence intervals about the best-fit curve.);

2. A table showing data quality assessment data points derived from a linear regression equation and the regression equation coefficients when appropriate; and

3. Calculated values of the standard deviation (or relative standard deviation) at discrete measured values that cover the applicable range.

Some components of precision include field instrument measurement variation, laboratory measurement variation, temporal variation, seasonality, spatial variation, physical support, nonresponse (or non-analyzed) component of variation, and data preparation variation.

Field Instrument Measurement Variation

Field instrument measurement variation is the lack of precision in the repetition of measurements taken by equipment in the field under the same conditions. This variation is a combination of potentially three different sources: (1) variation among different instruments used for the same type of measurement; (2) variation between repeated measurements taken by the same field instrument on the same sample; and (3) variation among different field technicians collecting field measurements using the same instrument and the same sample.

It is important to regularly calibrate field instruments against a common standard to minimize variation between instruments. This also allows one to estimate the amount of variation among instruments for measurements taken between calibrations. (It is important not to make too many adjustments based on frequent calibrations. This could result in increasing the variability in instrument measurements.) To measure the amount of variation between repeated measurements taken by the same field instrument on the same sample it is necessary to take multiple readings. The average of these readings can be used to improve estimation, and the variability about this average will estimate the instrument measurement error. To minimize variation between field technicians, it is necessary to establish Standard Operating Procedures, train all personnel in their use, and conduct quality audits to check on their implementation.

It is important that the analysis of field instrument measurement variation be conducted at multiple concentration levels. This is because field instrument precision is frequently a function of the concentration level being analyzed.

Laboratory Measurement Variation

Laboratory measurement variation is the lack of precision in measurements taken by equipment in the laboratory. This variation is the combination of potentially four different sources: (1) variation among measurements of the same sample taken in different laboratories; (2) variation among different laboratory instruments used for the same of type of measurement; (3) variation among repeated measurements taken by the same laboratory instrument on the same sample; and (4) variation among different laboratory technicians preparing and taking measurements using the same instrument and the same sample. For example, variation in laboratory measurements of dioxin contamination may result from differences among laboratories, among instruments used in the same laboratory to take the measurements, among repeated measurements taken by the same instrument of the same sample, and among different technicians using the same instrument to measure the same surface.

It is important to regularly calibrate laboratory instruments against a common standard to minimize variation between instruments. (It is important not to make too many adjustments based on frequent calibrations. This could result in increasing the variability in laboratory measurements.) This also allows one to estimate the amount of variation between instruments for measurements between calibrations. To measure the amount of variation between repeated measurements taken by the same laboratory instrument on the same sample it is necessary to take multiple readings. The average of these readings can be used to improve estimation, and the variability about this average will estimate the laboratory measurement error.

It is important that the analysis of laboratory measurement variation be conducted at multiple concentration levels. This is because laboratory precision is frequently a function of the concentration level being analyzed.

Temporal Variation

Temporal variation results when the true value that is being estimated fluctuates during the data collection time period. Among the possible causes of this fluctuation are an overall trend in the data, changes in water or other atmospheric levels, introduction of new sources of contamination during the data collection period, or cyclical patterns that are regularly repeated. Temporal variability results in an increase in the total variability of sample estimates.

A second form of temporal variability is temporal correlation. This occurs when samples taken across time are not independent, but rather are correlated. That is, the value of one reading is (at least partially) a function of previous values. The formula for the standard error of the mean assumes that the n sampled measurements are independent. If they are positively correlated (for example, air emissions taken every minute), there is less new information provided by each data point than would be expected. This results in an "effective sample size" less than n and understates the true variability. Corresponding confidence intervals based on this standard error are too small. A common procedure for describing temporally correlated data is the correlogram or auto correlation function, see guidance document EPA QA/G-9, *Guidance for Data Quality Assessment* for details.

Seasonality

Seasonality is a special type of cyclical temporal variation. Typical cycles might be quarterly, semi-annual, or annual. This may result , for example, from regular atmospheric/weather related patterns. Other sources of seasonality are a function of such patterns; for example, fertilizer application is timed to certain weather patterns. The term seasonality is not restricted to patterns that perfectly match with spring, summer, fall, and winter.

Depending upon the frequency of data collection relative to the seasonal cycle, seasonality may cause the data to be positively or negatively correlated. If data collection is very frequent it will be positively correlated. If the collection frequency is equal to half the seasonal cycle time and its timed with peaks and valleys, it will be negatively correlated (for example, if fertilizer use follows an annual cycle and sampling is conducted twice a year). If the data collection frequency equals the cycle time the data will not appear to be temporally correlated. However, in such a situation it is highly unlikely that the sample average will reflect the overall average. If data collection is frequent relative to cycle time, all seasonal patterns will show up clearly on correlograms; see EPA QA/G-9, *Guidance for Data Quality Assessment*, Section 2.3 for details.

Spatial Variation

Spatial variation results when the true value that is being estimated is not constant throughout the location being sampled. In many environmental data collection efforts this location is three-dimensional. For example, when sampling from a landfill or water from a lake, the contamination is likely to vary with both surface location (two dimensional) and depth.

Levels of contamination almost always vary spatially. In general, two samples taken (spatially) close together are more likely to have similar levels of the contaminant than those taken far apart. The variability of samples therefore typically increases monotonically as the physical distance between the samples increases. Unlike the situation with temporal correlograms, spatial variograms are not likely to show cyclical patterns.

A balance between spatial and measurement variation must be achieved when designing a sampling plan for characterizing a physical location. If measurement variability is thought to be the larger problem, it is advantageous to take composite samples from many sampling locations, and analyze aliquots from each composite. If spatial variation is large, it is important to keep samples from each location separate. Given the high cost of environmental measurements, it may be impossible to accurately assess both measurement and spatial variation. Guidance document EPA QA/G-5S, *Guidance on Sampling Designs to Support QAPPs*, discusses this further.

Physical Support

The physical support of a physical sample is the volume from which an individual sample is extracted. This volume is defined by its shape, size, and location. All three of these characteristics can influence the quality of the data and the inferences that can reliably be drawn.

For composite samples, the size of the physical support of the physical sample affects the variability of the estimates. (When using grab samples, the physical support is exactly equal to the size of the physical sample.) In general, the larger the support the smaller the variance of estimates. The actual numerical relationship between size and variability is complex and depends upon the spatial correlation within the support.

Many laboratory analyses do not report such that as to the original support is apparent. The size of the support has a significant effect on the modeling process and there is a difference in estimating the average value over a sample of large volume and in estimating the average value over a sample of small volume. Appendix H, "Representativeness of Environmental Data," also discusses this problem.

Nonresponse (or Nonanalyzed) Component of Variation

Precision decreases when the data are missing because of nonresponse or not being analyzed. See also Appendix G, "Representativeness of Environmental Data."

Typical examples of this situation include:

- physical samples were collected incorrectly for some sites so that all analyses for these sites are invalidated (e.g., water samples were collected using a metal beaker);
- the maximum allowable holding time was exceeded for a particular analysis, invalidating this one procedure but not others from the same physical sample; or
- a laboratory technician did not follow established good laboratory practices invalidating a series of analyses.

Nonresponse is often separated into unit and item nonresponse. Unit nonresponse (or nonanalysis) is the failure to obtain any valid measurements or answers to the questionnaire for a given case (e.g., the sample record will contain missing values for all variables). Item nonresponse (or nonanalysis) is the failure to obtain a valid answer for a particular question or a valid analysis for a particular analyte for a given case (e.g., the sample record will contain missing values for some variables).

Missing data decrease the effective sample size on which the analyses are based. Since standard errors are inversely proportional to the square root of the sample size, reducing the responding or analyzed sample size will increase the standard error, thus decreasing the sample's precision. Unit nonresponse (or

nonanalysis) will decrease the precision of all variables, while item nonresponse (or nonanalysis) will decrease the precision of only those variables whose data are missing. (In addition to decreasing precision, the above causes for nonresponse or nonanalysis also introduce a potential source of bias.)

Data Preparation Variation

Data preparation can introduce variation into the data by inconsistent coding, editing, or data entry. Frequently, revised coding and editing instructions will be developed during data preparation activities. This compounds the variation resulting from differences among or inconsistent practices being followed by individual coders and editors.

There are a number of ways in which data preparation activities can introduce variability into the data. When coding procedures are revised during the processing of the data (e.g., it is decided to code two different pesticide applications as equivalent) and it is decided not to go back through already processed forms to make the same correction, the resulting database will be more variable than need be.

If Standard Operating Procedures (SOPs) have not been established, the procedures used by each coder or editor are likely to be quite different. The same response may be coded into two different categories, or edited differently, depending upon who does the coding. It is therefore important to establish SOPs and automate as much of the data preparation processes as possible. Quality audits should be conducted to ensure that established SOPs are being followed.

Continual Precision Assessments

For organizations in which sample lots are routinely analyzed and data are reported on a frequent basis, the basic precision statistics from multiple lots of a given sample matrix may be combined to provide an estimate of long-term precision and an improved estimate of short-term precision. This assessment can also be extended to include subsequent lots, unless test results for these new lots indicate that method precision is significantly different. This combining of assessment results permits the laboratory to provide a precision assessment derived from a substantial amount of background data rather than from limited precision data produced in a small study.

This procedure also provides the basis for the use of control charts to monitor the performance of the measurement system. The procedure is based upon the availability of a precision assessment (normally developed from prior performance of the system), the use of control chart limits, and routine replicate pairs. This is discussed further in Appendix G, "Quality Control."

Bias

Bias is the systematic or persistent distortion of a measurement process that causes errors in one direction (i.e., the expected sample measurement is different from the sample's true value).

The Measurement of Bias

Bias assessments for environmental measurements are made using personnel, equipment, and spiking materials or reference materials as independent as possible from those used in the calibration of the measurement system. Where possible, bias assessments should be based on analysis of spiked samples rather than reference materials so that the effect of the matrix on recovery is incorporated into the assessment. A documented spiking protocol and consistency in following that protocol is an important element in obtaining meaningful data quality estimates. Spikes should be added at different concentration levels to

cover the range of expected sample concentrations. For some measurement systems (e.g., continuous analyzers used to measure pollutants in ambient air), the spiking of samples is not practical, and assessments are made using appropriate blind reference materials.

For certain multianalyte methods, bias assessments are complicated by mutual interference between certain analytes that prevent all of the analytes from being spiked into a single sample. For such methods, lower spiking frequencies can be employed for analytes that are seldom, or never, found. The use of spiked surrogate compounds for multianalyte GC/MS procedures, while not ideal, may be the best available procedure for assessment of bias. An added attraction is the ability to obtain recovery data on every field sample at relatively low costs. It is used, for example, to evaluate the applicability of methodology and, indirectly, data quality assessments to individual members of a sample lot. Such practices do not preclude the need to assess bias by spiking with the analytes being measured or reported.

Calculation of Bias Statistics

The most widely used summary of bias is by linear regression of bias on T; or, equivalently, regression of assessment results ($\overline{X}$ or $\overline{X}_l$) on T. For the important special case of spiked samples, the following approach may also be useful.

An estimate of the bias (B) is the difference between the average value $\overline{X}$ of a set of measurements of a standard and the reference value of the standard T given by:

$$B=\overline{X}-T$$

An alternative estimate of bias is percent bias:

$$\%B=100(\overline{X}-T)/T$$

Bias can also be expressed in terms of percent recovery (P), where P is defined as follows:

$$P_i=100(A_i-B_i)/T$$

where A_i = the analytical result from the spiked sample, and B_i = the analytical result from separate analysis of the unspiked sample. From this equation, the average percent recovery can be derived:

$$\overline{P}=\sum_{1}^{n}P_i$$

The relationship between percent bias and percent recovery is:

$$\%B=\overline{P}-100$$

If reference materials instead of spiked samples are analyzed to assess bias, percent recovery is calculated by the equation above with B_i equal to zero.

Reporting Bias

The preferred measure of bias is the difference between the average measured value and the true value. Percent recovery (100% + bias) is also frequently used in environmental measurement programs.

Procedures for presenting bias information are as follows, in order of preference:

1. Bias as a function of the true value over the applicable range (This could be a graph containing the actual data points, the best-fit curve, confidence intervals about the best-fit curve, and the regression equation for the best-fit curve when appropriate.);

2. A table of data points derived from a linear regression equation and the regression equation coefficients; and

3. Bias values calculated at discrete measured values covering the applicable range.

Bias data should be accompanied with certain supporting information. This information should include (but not be limited to) a description of how and under what conditions the bias data were collected, the number of data points involved, the applicable range of the data, and an equation of the best-fit curve.

Some components of bias include average percent recovery, measurement (equipment) bias, nonresponse (or nonanalyzed) bias, data preparation bias, and statistical biases.

Average Percent Recovery. Average percent recovery is a measure of how well laboratory equipment, protocols, and technicians can detect known concentrations of a contaminant. This measure is the ratio of the average detected concentration to the average known concentration, in known-concentration samples. Ideally, multiple samples are examined at multiple concentrations at each laboratory. Consistency in percent recovery can then be examined across concentrations and laboratories. Lack of consistency can be used to suggest improvements in quality assurance procedures.

Average percent recovery is almost always less than 100 percent. If no adjustment is made in the data analyses, this condition will result in a downward bias in both average concentrations and percent detections. Dividing measured concentrations by the average recovery will adjust for most of this bias in estimating average concentrations. It will not, however, adjust for the underestimate in either the average concentration or the percent detected resulting from samples estimated below the detection limit whose true concentrations are above the detection limit. Thus, it is important to develop protocols and use laboratory equipment that can achieve an average percent recovery as close to 100 percent as possible.

Adjustments for average percent recovery should not be based upon a single known-concentration sample. A single sample can be subject to enough laboratory measurement error that the reduction in bias will be counteracted by a decrease in precision. It is therefore necessary to accurately estimate the average percent recovery from averaging multiple know-concentration samples. The number of samples required to reduce the overall mean square error is a function of the laboratory measurement error and the average percent recovery. For smaller average percent recoveries (larger biases) the less accurate it needs to be estimated for the adjustment of sample estimates of concentrations to improve overall accuracy. For example, if percent recovery is more than 90 percent, the laboratory measurement error (associated with this estimate of 90 percent) will have to be made quite small in order for dividing all estimated concentrations by .90 to improve overall accuracy. Thus it may not be cost-effective to analyze all of the necessary known-concentration samples that would be required. If, however, the average percent recovery is under 50 percent, it is much more likely that the measurement error associated with estimating the average percent recovery can be reduced to the point that adjusting for this bias (dividing estimated concentrations by .50) will reduce the overall mean square error.

Determination of the average percent recovery should not be conducted at only one concentration. It is necessary to examine a variety of concentrations to determine if the average percent recovery is a

function of the concentration level. If it is, it will be necessary to interpolate bias adjustments between those levels that are actually analyzed.

If multiple laboratories test for the same contaminant it is important to determine if percent recovery is consistent across laboratories. By testing each laboratory on the same set of known-concentration samples it is possible to use the statistical technique of analysis of variance with randomized blocks to compare laboratories.

Measurement (Equipment) Bias. There are two ways in which measurement bias can result from field instruments and laboratory equipment. First, analogous to average percent recovery in a laboratory, field instruments may on average detect an amount not equal to the true amount being measured. This bias may be adjusted for by recalibration of the equipment or the development of mathematical adjustments to the raw data. Second, estimates are sometimes based on the maximum of a series of field instrument or laboratory equipment measurements. The calculations for the adjustment of bias are the same as those for the average percent recovery.

Nonresponse (or Nonanalyzed) Bias. Bias may result when the data are missing due to nonresponse or not being analyzed. For example: physical samples may have been collected incorrectly so that all analyses are invalidated (e.g., water samples were collected using a metal beaker), the maximum allowable holding time may have been exceeded for a particular analysis, or a laboratory technician may not have followed established SOPs and invalidated a series of analyses.

Nonresponse is often separated into unit and item nonresponse. Unit nonresponse (or nonanalysis) is the failure to obtain any of the measurements (it will contain missing values for all variables). Item nonresponse (or nonanalysis) is the failure to analyze the data for any particular analyte for a given case (it will contain missing values for some variables).

These missing data can potentially bias the analyses unless the probability of being missing is random; i.e., if the chance the data are missing is not correlated with the variables being analyzed. For example, assume metal beakers were used to collect all samples from a specific type of well, making the samples nonanalyzable. If this type of well is more (or less) likely to contain contamination than other wells, then the exclusion of these wells from the analyses will bias estimates of overall contamination. If not all of this type of well are nonanalyzable, it may be possible to minimize the extent of this bias through post-stratification or other weighting procedures.

The extent of bias resulting from nonresponse or nonanalysis is the product of two factors: the percent of cases whose responses are missing and the difference in the average values between those who responded (were analyzed) and those who did (were) not. If much of the data are missing and there are no differences between respondents and nonrespondents, then there will not be any bias. If only a small percentage of the data are missing, then even relatively large differences between respondents and nonrespondents will result in small bias in the estimates.

Data Preparation Biases. Data preparation can introduce biases into the data by how the coding or editing is performed. Decisions are made by supervisory staff as to how to code open-ended questions, how to treat multiple responses, and how to handle similar situations. If these actions cause the database to understate (or overstate) the incidence of certain responses, the data may be biased.

When editing data, it is sometimes decided to categorize responses to a continuous variable. For example, the volume of waste water that is produced may be categorized into three of four categories,

rather than retaining the actual reported volume. This will bias downward potential correlations with other continuous variables, such as concentration levels.

Statistical Biases. Statistical procedures can also introduce biases into the analyses. For example, ratio estimators are frequently used to reduce the sampling error when estimating population parameters. However, ratio estimators are not unbiased for the parameters; it is hoped that the resulting mean square error of the ratio estimator is smaller than that for the alternative biased estimator.

The technical discussion of statistical bias is beyond the scope of this document and a statistician should be consulted whenever ratio estimators are used (e.g., when the estimate of interest is a function of one measurement having error being divided by another measurement having error).

Accuracy

Accuracy is a measure of the closeness of an individual measurement or the average of a number of measurements to the true value. Accuracy includes a combination of random error (precision) and systematic error (bias) components that result from sampling and analytical operations.

Accuracy is determined by analyzing a reference material of known pollutant concentration or by reanalyzing a sample to which a material of known concentration or amount of pollutant has been added. Accuracy is usually expressed either as a percent recovery (P) or as a percent bias (P - 100). Determination of accuracy always includes the effects of variability (precision); therefore, accuracy is used as a combination of bias and precision. The combination is known statistically as mean square error.

Mean Square Error

Mean square error (MSE) is the quantitative term for overall quality of individual measurements or estimators. To be accurate, data must be both precise and unbiased. Using the analogy of archery, to be accurate, one must have one's arrows land close together and on average at the spot where they are aimed. That is, the arrows must all land near the bull's-eye.

Mean square error is the sum of the variance plus the square of the bias. (The bias is squared to eliminate concern over whether the bias is positive or negative.) Frequently, it is impossible to quantify all of the components of the mean square error—especially the biases—but it is important to attempt to quantify the magnitude of such potential biases, often by comparison with auxiliary data.

Representativeness

A measure of the degree to which data accurately and precisely represents a characteristic of a population parameter variations at a sampling point, a process condition, or an environmental condition. Representativeness is the qualitative term that should be evaluated to determine that in situ and other measurements are made, and physical samples collected, at such locations and in such a manner as to result in data reflecting the media and phenomenon measured or studied. Refer to Appendix H for a more detailed definition.

Comparability

Comparability is the qualitative term that expresses the measure of confidence that two data sets can contribute to a common analysis and interpolation. Comparability must be carefully evaluated in order

to establish whether two data sets can be considered equivalent in regard to the measurement of a specific variable or groups of variables. In a laboratory analysis, the term comparability is directed to method type comparison, holding times, stability issues, and aspects of overall analytical quantitation.

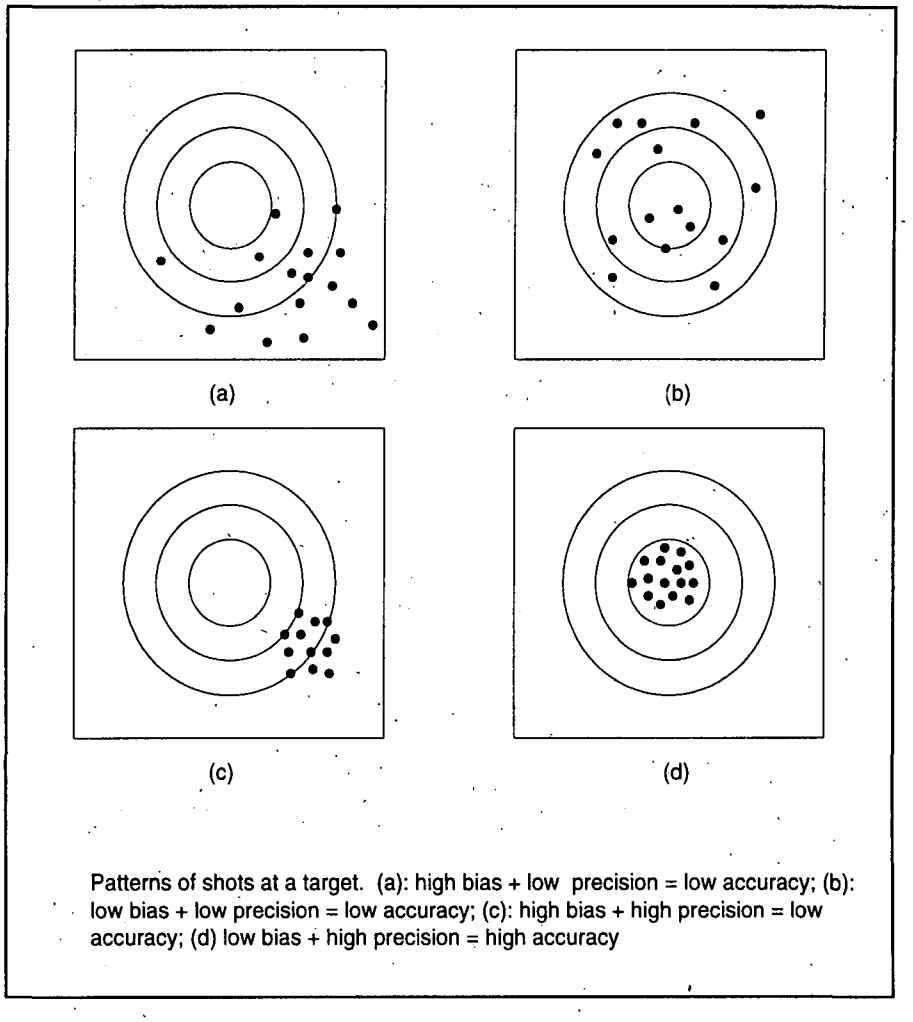


Figure D1. Measurement Bias and Random Measurement Uncertainties. *Adapted from Gilbert (1987), Figure 2.4.*

There are a number of issues that can make two data sets comparable, and the presence of each of the following items enhances their comparability:

- two data sets should contain the same set of variables of interest;
- the units in which these variables were measured should be convertible to a common metric;
- similar analytic procedures and quality assurance should be used to collect data for both data sets;

- the time of measurements of certain characteristics (variables) should be similar for both data sets;
- the measuring devices used for both data sets should have approximately similar detection levels;
- the rules for excluding certain types of observations from both samples should be similar
- samples within data sets should be selected in a similar manner;
- the sampling frames from which the samples were selected should be similar; and
- the number of observations in both data sets should be of the same order or magnitude.

These characteristics vary in importance depending on the final use of the data. The closer two data sets are with regards to these characteristics, the more appropriate it will be to compare them. Large differences between characteristics may be of only minor importance depending on the decision that is to be made from the data.

Comparability is very important when conducting meta-analysis, an attempt to combine the results from numerous studies to identify commonalities which are then hypothesized to hold over a range of experimental conditions. To the extent that the studies being evaluated are not truly comparable, the meta-analysis can be very misleading. The hypothesized findings of the meta-analysis may be an artifact of the differences among the studies rather than the experimental conditions. Expert opinion to classify the importance of differences in characteristics is invaluable.

Completeness

Completeness is a measure of the amount of valid data obtained from a measurement system, expressed as a percentage of the number of valid measurements that should have been collected (i.e., measurements that were planned to be collected).

Completeness is not intended to be a measure of representativeness; that is it does not describe, how closely the measured results reflect the actual concentration or distribution of the pollutant in the media sampled. A project could produce 100% data completeness (i.e., all samples planned were actually collected and found to be valid), but the results may not be representative of the pollutant concentration actually present.

Alternatively, there could only be 70% data completeness (30% lost or found invalid) but, due to the nature of the sample design, the results could still be representative of the target population and so yield valid estimates. Where lack of completeness is of vital concern is with stratified sampling. Substantial incomplete sampling of one or more strata can seriously compromise the validity of conclusions from the study. In other situations (for example, simple random sampling of a relatively homogenous medium), lack of completeness only results in a loss of statistical power. The degree to which lack of completeness affects the outcome of the study is a function of many variables, ranging from deficiencies in the number of field samples acquired, to failure to analyze as many replications as deemed necessary by the QAPP and Data Quality Objectives. The intensity of effect of lack of completeness is sometimes best expressed as a qualitative measure and not just as a quantitative percentage.

Completeness can have an effect on the DQO parameters. Lack of completeness may require reconsideration of the limits for the false negative and positive error rates because insufficient completeness will decrease the power of the curve.

The following four situations demonstrate the importance of considering the planned usage of the data when determining the completeness of a study. The purpose of the study is to test the hypothesis that

the average concentration of dioxin in surface soil is no more than 1.0 ppb. The established DQO specified simple random sampling with 30 samples being drawn and that the sample average should estimate the true average concentration to within ±0.30 ppb with 95 percent confidence.

	Study result	Completeness	Outcome
1)	1.5 ppb ± 0.28 ppb	97%	satisfies DQO and study purpose,
2)	500 ppb ± 0.28 ppb	87%	satisfies DQO and study purpose,
3)	1.5 ppb ± 0.60 ppb	93%	doesn't satisfy either,
4)	500 ppb ± 0.60 ppb	67%	fails DQO but meets study purpose.

For all but the third situation, the data that were collected completely achieved their purpose, meeting data quality requirements originally set out, or achieved the purpose of the study. The degree of incompleteness did not affect some situations (numbers 2 and 4) but may have been a prime cause for situation 3 to fail the DQO requirements. Expert opinion would then be required to ascertain if further samples for situation 3 would be necessary in order to meet the established DQO.

When a study is found to lack completeness, the reasons for this shortcoming should be investigated. It may be a result of poor assumptions on which the DQOs were established, poor implementation of the survey design, or that the design proved impossible to carry out given resource limitations. Lack of completeness should always be investigated and the lessons learned from conducting the study incorporated into the planning of future studies.

D2. OTHER DATA QUALITY INDICATORS

Recovery

Recovery refers to whether or not the methodology measures all of the analyte that is contained in the sample. This is best evaluated by the measurement of reference materials or other samples of known composition. In their absence, spikes or surrogates may be added to the sample matrix. The recovery is often stated as the percentage measured with respect to what was added. Complete recovery (100%) is the ultimate goal. At the minimum, recoveries should be constant (only varying within acceptable limits), and should not differ significantly from an acceptable value. This means that control charts or some other means should be used for verification. Significantly low recoveries should be pointed out, and any corrections made for recovery should be stated explicitly.

Blunder

Blunders are simply mistakes that occur on occasion and produce erroneous results. Measuring the wrong sample, errors of transcription or transposition of measured values, misreading a scale, and mechanical losses, are examples of blunders. They produce outlying results that may be recognized as such by statistical procedures, but they cannot be treated by statistics. Appropriate quality control procedures can minimize the occurrence of some kinds of blunders but may not eliminate carelessness which often is their principal cause.

Memory effects

The effect that a relatively high concentration sample has on the measurement of a lower concentration sample of the same analyte when the higher concentration sample precedes the lower concentration sample in the same analytical instrument.

Identification

Misidentification of an analyte. Results in the contaminant of concern not being identified and the measured concentration being incorrectly assigned to another contaminant.

Sensitivity

The capability of a method or instrument to discriminate between measurement responses representing different levels of a variable of interest. Sensitivity is evaluated from the value of the standard deviation at the concentration level of interest. It represents the minimum difference in two samples of approximately equal concentration that can be distinguished with a 95% confidence.

Limit of quantitation

The minimum concentration of an analyte or category of analytes in a specific matrix that can be identified and quantified above the method detection limit and within specified limits of precision and bias during routine analytical operating conditions.

Repeatability

The degree of agreement between independent test results produced by the same analyst, using the same test method and equipment on random aliquots of the same sample within a short time period.

Reproducibility

The precision, usually expressed as a variance, that measures the variability among the results of measurements of the same sample at different laboratories.

DQIs and the QAPP

At a minimum, the following DQIs should be addressed in the QAPP: accuracy and/or bias, precision, completeness, comparability and representativeness. Accuracy (or bias), precision, completeness, and comparability should be addressed in Section A7.3, Specifying Measurement Performance Criteria. Refer to that section of the text for a discussion of the information to present and a suggested format. Representativeness should be discussed in Section B4.2, Sub-Sampling and in Section D1.2, Sampling Design.

D3. LINKING QUANTITATIVE DATA QUALITY INDICATORS TO DATA QUALITY OBJECTIVES

Introduction

One of the barriers to EPA's institutionalization of the Data Quality Objectives (DQOs) Process is the confusion that exists between the definitions and relationship of DQOs and Data Quality Indicators (DQIs). Early EPA guidance (QAMS-005/80, *Interim Guidance and Specifications for Preparing Quality Assurance Project Plans*, December, 1980) used the two terms interchangeably to represent the specific statistical parameters of precision, accuracy, representativeness, completeness, and comparability. (These DQIs are referred to as the PARCC terms, which are discussed in Appendix D1.) Later, EPA adopted the term "Data Quality Objectives" to refer to the new and more encompassing process of establishing criteria for overall data quality and for developing data collection designs. (Refer to Section A7.2 for a description of the DQO Process.) However, many in the environmental community mistook the term DQOs to

represent the specific objectives set for the DQIs. The difference between the two terms is that DQOs include performance measures and goals for the entire project while DQIs represent measures and goals for the project's sample measurement process. More specifically, the DQIs quantify the amount of error in the data collection process and the analytical measurement system.

This portion of the appendix describes the relationship between DQOs and DQIs. The description first entails a general discussion of the common types of error that occur while measuring environmental properties. The errors that can propagate throughout the measurement process are then discussed in more detail. The affect that errors (as measured by DQIs) have on the DQOs are then discussed, followed by a description of establishing DQIs.

1. Types of Error Possible in the Measurement Process

The purpose of an environmental measurement is to characterize a portion of the environment with respect to a specific property such as its temperature, pH, or contaminant concentration. Unfortunately, errors can occur throughout the measurement process and DQIs are measures of this error. While not exhaustive, the list below covers the more common types of error in environmental measurements and their causes. The common error types listed include those measured by the PARCC terms and two additional types, misidentification and blunder. The fifth PARCC term, comparability, is not addressed. Comparability is a qualitative measure of the confidence with which one data set can be compared to another. This DQI is not translated to the DQOs directly but is a comparison between different data sets.

TYPES OF ERROR	SOURCES OF ERROR
Random or Imprecision (lack of precision, P in PARCC)	Natural variability in the population from which the sample is taken. Measurement system variability, introduced at each step of sample handling and measurement processes.
Systematic or Bias (part of accuracy, A in PARCC)	Interferences that are present in sample matrix. Loss (or addition) of contaminants during sample collection and handling. Loss (or addition) of contaminants during sample preparation and analysis. Calibration error or drift in the response function estimated by the calibration curve.
Representativeness (R in PARCC)	Sample isn't representative of the population, which often occurs in judgmental sampling because not all the units of the population have equal or known selection probability (also caused by bias). Sample collection method does not extract the material from its natural setting in a way that accurately captures the desired qualities to be measured. Subsample (taken from a sample for chemical analysis) isn't representative of the sample, which occurs because the sample is not homogenous and the subsample is taken from the most readily available portion of the sample. Consequently, other parts of the sample had less chance of being selected for analysis.
Incompleteness (lack of completeness, C in PARCC)	Lack of completeness sometimes caused by loss of a sample, loss of data, or inability to collect the planned number of samples. Incompleteness also occurs when data are discarded because they are of unknown or unacceptable quality.
Mis-identification	Failing to detect a contaminant that is present. Giving a detected compound the wrong name (e.g., error in computer-matched mass spectra).
Blunder	Generally huge error caused by human or natural phenomenon, usually unpredictable and often undetected.

2. Propagation of Error

The first two types of error listed above, precision and bias, will be discussed in further detail. These error types are of particular concern because they can be introduced at every step of the measurement process and can propagate throughout this process. For example:

- volatiles can be lost during sample acquisition or during subsequent sample storage and handling, causing negative bias (and, because the bias is different for different samples, this contributes to imprecision);
- subsampling very small portions of non-homogenous samples can introduce imprecision (and, if some portions are less "available" than others, bias is also introduced); and
- analytical instruments, when challenged repeatedly by the same reference material will produce variable results and may show increasing or decreasing trends due to calibration drift.

It is convenient (though not strictly correct) to think of both kinds of error as being additive at every step of the measurement process. Biases added at different steps combine to produce a net bias that is the sum of the individual bias errors. Variances (squares of standard deviations that characterize imprecision) also combine to produce a total variance for the measured value. This idea can be illustrated mathematically as follows.

1. Let **b** and s^2 represent bias and variance.
2. Let the subscripts **s, h, b, e,** and **a** represent the sampling, sample handling, subsampling, extraction, and instrumental analysis steps of a measurement process.
3. The subscript **p** represents population variance (variance between the different population units that may be chosen to form the "samples")
4. The subscript **t** represents total, for total bias and total variance. Biases and variances combine to produce total bias and variances for the environmental datum in the additive form:

$$b_t = b_s + b_h + b_b + b_e + b_a \quad (1)$$

$$s_t^2 = s_p^2 + s_s^2 + s_h^2 + s_b^2 + s_e^2 + s_a^2 \quad (2)$$

The above equations are an over-simplification because, in reality, there will be more error components, and the errors will not necessarily be additive. In some circumstances, the relationship between the errors will result in a multiplicative effect. However, in these instances, a logarithmic transformation could be used to derive additive relationships as shown in equations (3) and (4).

$$\ln(1+b_t) = \ln(1+b_s) + \ln(1+b_h) + \ln(1+b_b) + \ln(1+b_e) + \ln(1+b_a) \quad (3)$$

$$CV_t^2 = CV_p^2 + CV_s^2 + CV_h^2 + CV_b^2 + CV_e^2 + CV_a^2 \quad (4)$$

(1+b, above is also known as recovery and CV is the coefficient of variation, also known as the relative standard deviation.)

To describe precisely the relationship between total MSE (variance plus the square of bias) often takes a mixed model of linear and multiplicative components. Estimation for such mixed models is usually impossible and a simple linear model chosen as an approximation of the true relationship. The use of MSE often results in an overestimate of total error as occasionally biases may cancel out at various stages of the

analytical procedure; as these biases are difficult to estimate in magnitude and direction, using the squared value enables cross-comparisons to be made.

Error propagation can be illustrated using equation (1) as a basis for discussing control of bias and equation (2) as a basis for discussing approaches to controlling variance. Bias check samples, such as spiked samples or use of standard reference materials, can be introduced at various points in the measurement process. A bias check introduced at the instrumental analysis step will only provide insight into that part of total bias that is due to the instrument: b_a. A bias check introduced earlier in the measurement process, for example, prior to extraction will provide an estimate of $b_e + b_a$. In general, the best check of total bias is a spiked sample or standard material that is introduced at the beginning and carried through the entire measurement process. This is the best procedure even though matrix effects, that cause measurement interference, may arise from the use of standard materials or from spiking a field sample.

In equation (2), there will almost always be one or two types of error that dwarf all the others in contributing to total variance. A good rule of thumb is offered by W.J. Youden:

> "Once the analytical uncertainty [$s_t^2 - s_p^2$] has been reduced to a third or less of the sampling uncertainty [s_p^2], further reduction in the analytical uncertainty is of little importance."

Often, the population variance is several times larger than any of the other variance components. It follows that, the most efficient approach to estimating the average contaminant level is to collect a number of field samples (each consisting of individual grab samples or composited samples) and to subject each one of them to a single chemical analysis. For cases such as this, the link between measurement performance criteria and DQOs is not very important and planners should consider other factors (e.g., cost, appropriateness) when selecting measurement methods. However, if error in the analytical measurement process is the main problem, samples can be split at various points to produce a greater number of analytical results for each field sample. Planners should consider the relative magnitudes of the different variance components of error. Those components that contribute most significantly to total error should be the prime candidates for replication. For example, if instrumental error is huge, then planners should consider producing repeated instrumental analyses of each laboratory sample as a means to reduce the contribution of that error component.

Guidance Document EPA QA-G-4D, *DEFT Software for the Data Quality Objectives Process* is useful in assisting data collectors to efficiently allocating resources for sample collection design. The user supplies information obtained while completing the DQO Process, sampling and analytical costs, and variances. DEFT then generates estimates of resource effective sample designs that achieve the DQOs. From this set of designs the most appropriate can be selected.

References

Calculation of Precision, Bias, and Method Detection Limit for Chemical and Physical Measurements. Chapter 5. 1984.

Definitions of Environmental Quality Assurance Terms. American Society for Quality Control. Milwaukee, WI: ASQC Press, 1996.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring.* New York: Van Nostrand.

Taylor, J.K. and T.W. Stanley. eds. 1985. *Quality Assurance for Environmental Measurements.* Philadelphia, PA: American Society for Testing and Materials.

Taylor, J.K. 1987. *Quality Assurance of Chemical Measurements.* Chelsea, MI: Lewis Publishers Inc.

U.S. Environmental Protection Agency. 1992. *Tools for Determining Data Quality.* Office of Policy, Planning, and Evaluation. July.

U.S. Environmental Protection Agency. 1994. *AEERL Quality Assurance Procedures Manual for Contractors and Financial Assistance Recipients.*

U.S. Environmental Protection Agency. 1994. *EPA Requirements for Quality Management Plans.* EPA QA/R-2, Draft Interim Final. August.

Youden, W.J. 1967. *Journal of the Association of Official Analytical Chemists.* Vol. 50. 1007.

APPENDIX E

DETECTION LIMITS

E1. INTRODUCTION

Variability in replicate analytical determinations (also referred to as analytical uncertainty), especially variability associated with the measurement of low concentrations, may impose limitations on EPA in setting regulatory standards. EPA has utilized the general concept of a detection limit (DL) to quantify this variability. Various alternative DLs appear in EPA regulatory literature including method detection limits (MDLs), practical quantitation limits (PQLs), limits of detection (LODs), and limits of quantification (LOQs). This appendix describes the alternative DL definitions and computational methods that appear in EPA literature.

This appendix has four further sections: Section E2 contains a discussion of the DL concept and a review of three different approaches used to define and compute DLs; Section E3 presents definitions and computational formulas for alternative DLs found in EPA literature; Section E4 contains comparisons of alternative DL definitions; and finally, Section E5 is a bibliography.

E2. BACKGROUND

E2.1 Detection Limit Concept

The DL is a concept concerning the capability of an analytical method to distinguish samples that do not contain a specific analyte from samples that contain low concentrations of the analyte. DLs are intended to transmit information about the general efficacy of analytical methods in the analysis of low concentrations. DLs are analyte and matrix specific, and may be laboratory dependent.

An analytical method may produce a non-zero signal even when the target analyte is not present. Conceptually, the DL is the minimum true concentration of target analyte producing a non-zero signal that can be distinguished with an appropriate degree of certainty from non-zero signals produced when that analyte is not present.

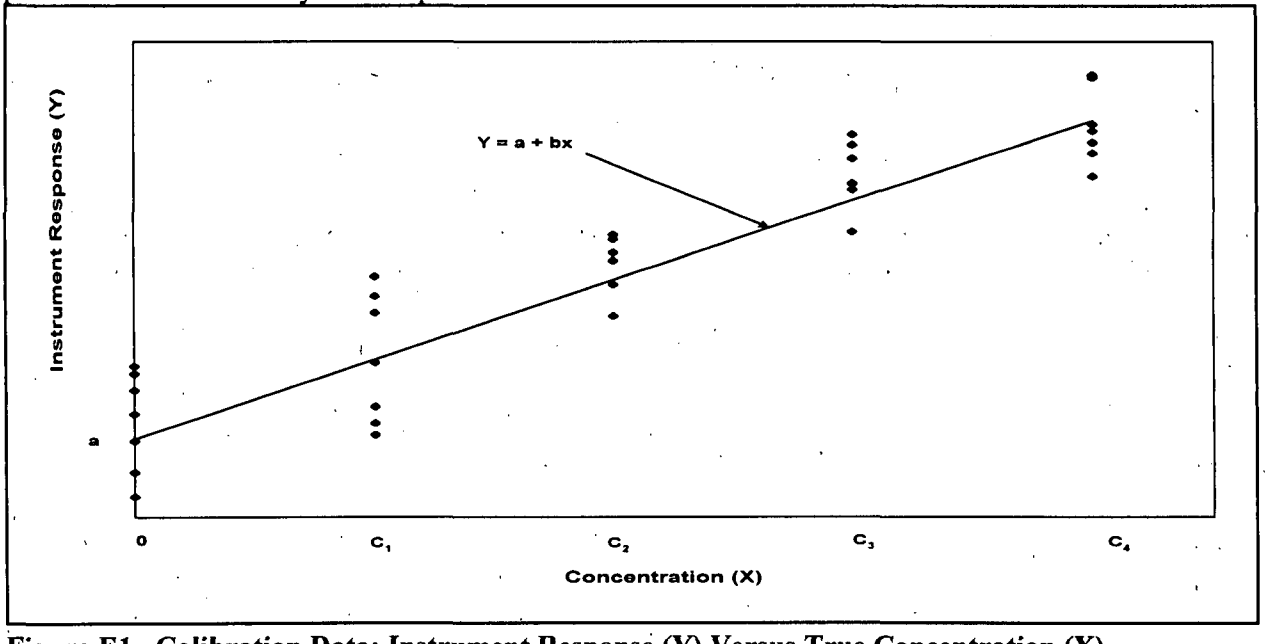


Figure E1. Calibration Data: Instrument Response (Y) Versus True Concentration (X).

Figure E1 presents the DL concept graphically and shows why the DL is not necessarily near zero. The figure shows the spread of values associated with replicate measurements corresponding to samples with true concentrations of 0, C_1, C_2, C_3, and C_4. The spread is due to inherent analytical variability. The measurements corresponding to the concentration at C_1 overlap considerably with the measurements when the concentration is zero. Therefore, differentiating a true concentration of C1 from a true concentration of zero on the basis of the measurements alone would be subject to a high degree of uncertainty. Differentiating C_4 from zero would involve less uncertainty because the overlap in measurements corresponding to those concentrations is less than the overlap in measurements between C_1 and zero. C_4, therefore, is a better candidate for the DL than C_1.

The DL value, under almost all definitions found in EPA literature, is a multiple of the analytical standard deviation (σ). The analytical standard deviation is assumed to be constant over a relatively short range of low concentrations. The structure underlying these DL definitions, whether or not explicitly stated, is the statistical decision problem, choosing between the Null Hypothesis: $C_T = 0$ and the Alternative Hypothesis: $C_T > 0$, where C_T is the true concentration of the target analyte in the sample. Using observations to make a decision, deciding in favor of "$C_T > 0$" means "detection."

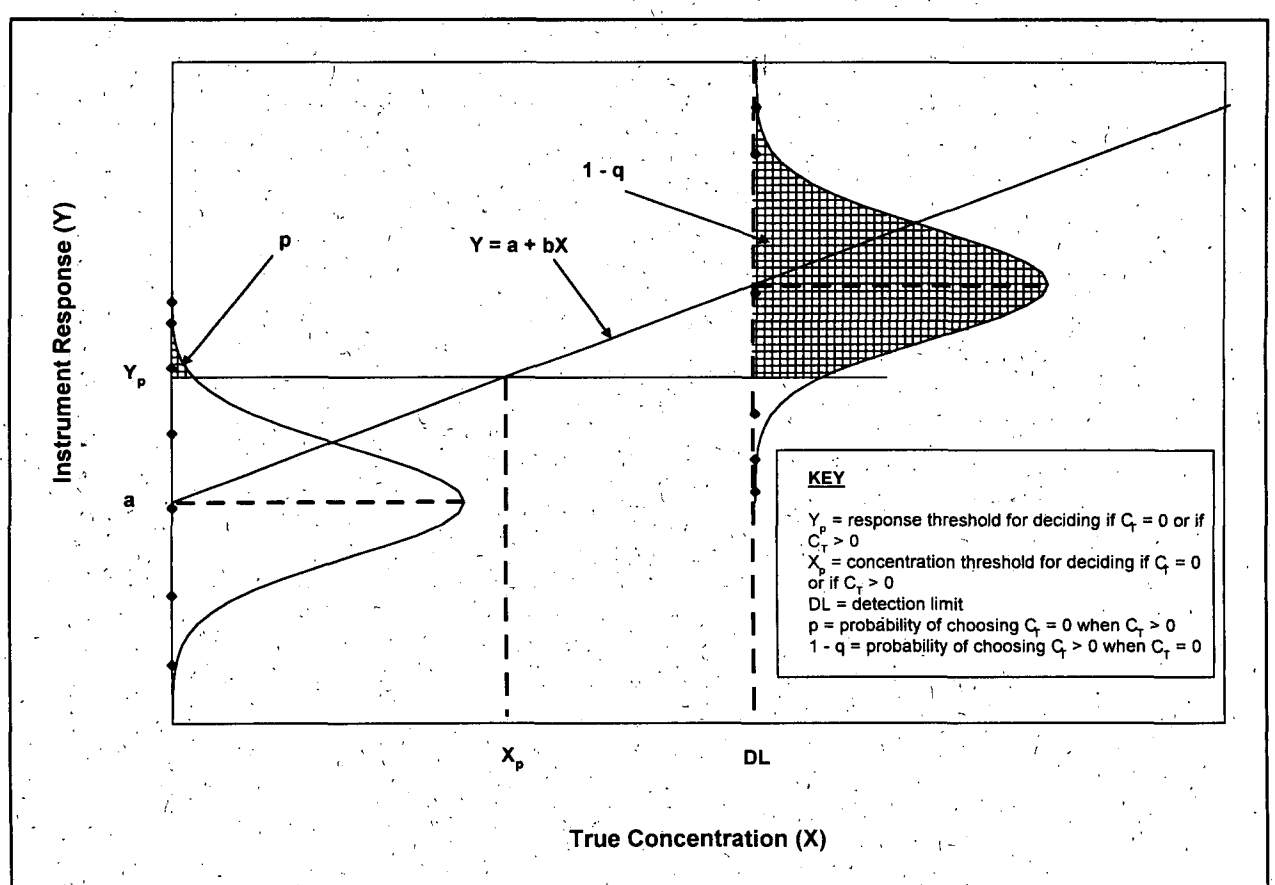


Figure E2. The Detection Limit as Defined by Choosing Between Hypotheses, $C_T = 0$ and $C_T > 0$.

Figure E2 shows how this decision problem leads to a DL definition. Suppose the value, Y_p, in instrument response units is selected as the threshold for deciding between $C_T = 0$ and $C_T > 0$. If the instrument response is greater than Y_p, then $C_T > 0$ is accepted. The probability that an instrument response would exceed Y_p when C_T is, in fact, zero is a small value, p due to the inherent analytical variability of the process. (Equivalently, p is the probability of erroneously concluding that $C_T > 0$.)

Now suppose that X_p, defined through the calibration line, is the corresponding threshold for the decision in concentration units. The DL is then defined as the true positive concentration where the probability of correctly deciding in favor of $C_T > 0$ is large (e.g., 0.95 or 0.99, conventionally denoted 1-q).

The various definitions of DL found in EPA literature are, for the most part, based on the approaches described in three articles; Hubaux and Vos, 1970; Glazer et al., 1981; and Clayton et al., 1987. These approaches are described in Section E2.2 below. The descriptions in Section 3.0 of other DL approaches found in EPA literature is facilitated by identifying them with these three approaches.

E2.2 Summary of Three Basic Articles

Before considering in detail the three referenced approaches to establishing a DL, a few general comparisons are noted. The approaches taken in Hubaux and Vos, 1970 and Clayton et al., 1987 are similar in that they explicitly involve a calibration line and account for variability in calibration data. These two approaches also reflect the statistical decision problem of choosing between two assertions, $C_T = 0$ and $C_T > 0$, for a sample with unknown concentration. Glazer et al. (1981), on the other hand, treat the calibration line as if it were known with certainty and address only a portion of the statistical decision problem.

Hubaux and Vos (1970) utilize the concept of confidence limits for predicted values from a regression line (the estimated calibration line) to define DLs. In Clayton et al., 1987 the approach is similar except the non-central t-distribution is employed where Hubaux and Vos use an approximation. The impact of this difference on the ultimate numerical value determined for the DL appears to be minimal.

2.2.1 Hubaux and Vos, 1970

This approach utilizes confidence limits for predicted values from a least squares fitted calibration line to establish numerical values for a "decision limit" and a "detection limit." The decision limit, denoted as Y_p in Figure 3, is the upper $(1-p)^{th}$ confidence limit for a predicted instrument response when the true concentration of the target analyte in the sample is zero. From the calibration line, the corresponding decision limit in concentration units is X_p. The detection limit (DL in the figure) is the concentration at which the decision limit, Y_p, is the lower $(1-q)^{th}$ confidence limit for an instrument response predicated for the fitted calibration line. The prediction formulae for these quantities are found using ordinary least squares linear regression to be:

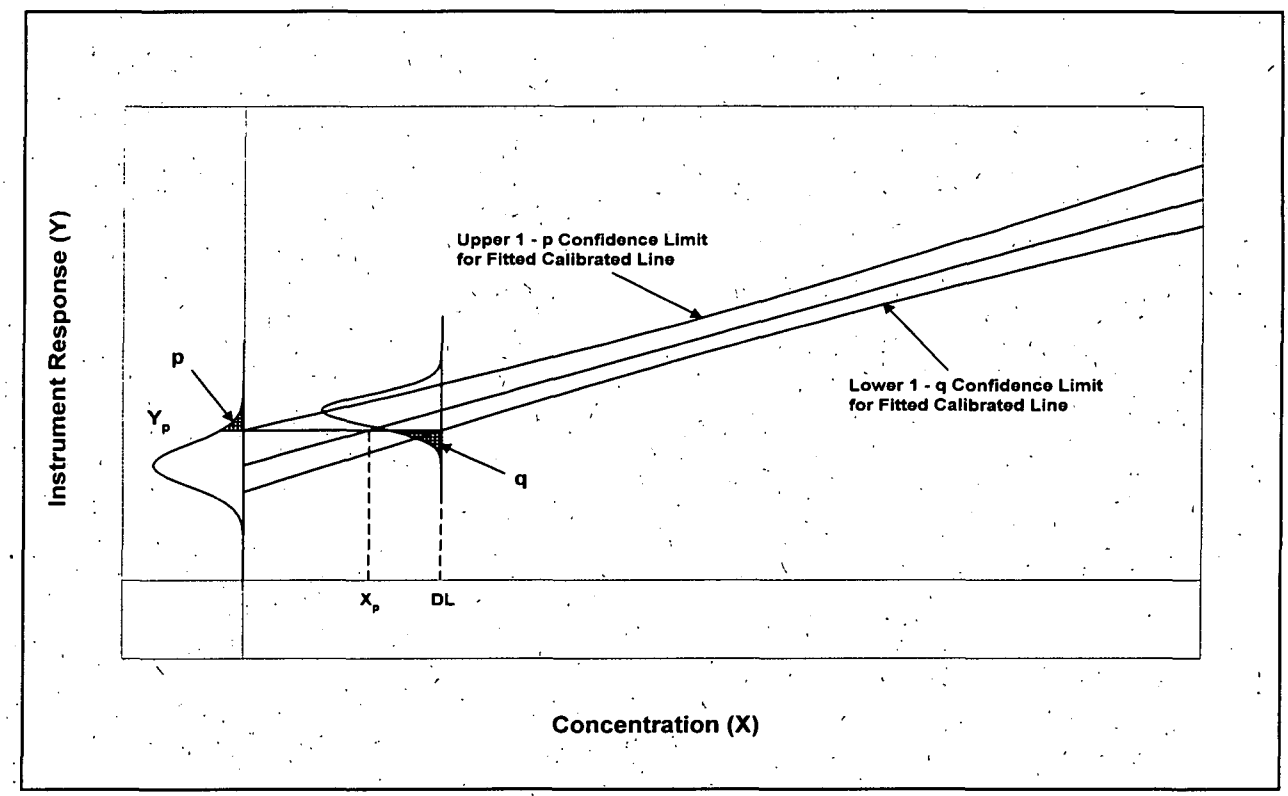


Figure E3. Detection Limit (DL) Definition by Hubaux and Vos (1970).

Decision Limit

$$Y_p = a + t_p s \sqrt{1 + \frac{1}{n} + \frac{\bar{X}^2}{Q^2}}$$

and from the regression line $Y_p = a + bXp$ it follows that

$$X_p = \frac{(Y_p - a)}{b}$$

The Detection Limit is then found by matching the Decision Limit (Y_p) to the lower qth percentile of the prediction formula for DL. That is to say

$$a + t_{1-p} s \sqrt{1 + \frac{1}{n} + \frac{\bar{X}^2}{Q^2}} = a + b(DL) + t_q s \sqrt{1 + \frac{1}{n} + \frac{(DL - \bar{X})^2}{Q^2}}$$

This reduces to a quadratic equation in DL which can then be solved for DL where

a - estimated intercept on the Y axis of the calibration line

b - estimated slope of the calibration line

n - number of calibration samples used to estimate the calibration line

t_{1-p} - (1-p)th percentile of the t distribution with n-2 degrees of freedom

t_q - qth percentile of the t distribution with n-2 degrees of freedom

$\bar{X}$ - Average of concentrations for calibrations samples

Q^2 - $\sum (X - \bar{X})^2$ for calibration samples

s - root mean square error for fitted calibration line

The article notes that the numerical value of the DL is influenced by a number of factors in addition to the inherent precision of the analytical method. Among these factors are the number of samples used for calibration, the calibration concentrations, and the replication rate for measurement of samples with unknown concentrations. If replication and averaging are employed, the formulas for computing the detection limit would be altered by replacing (1 + 1/n) wherever it appears by (1/r + 1/n), where r is the number of replicate measurements.

E2.2.2 Clayton, et al., 1987

In this approach, the DL is derived by directly solving the decision problem of choosing between $C_T = 0$ and $C_T > 0$. Y_p and X_p, referred to as threshold values in the Clayton discussion, are computed by the same formulas used in Habeux and Vos. For a sample with unknown concentration, the decision would be $C_T > 0$ if the instrument response where greater than Y_p, or equivalently the corresponding concentration estimate were greater than X_p. The detection limit is defined as the concentration at which the probability of choosing $C_T > 0$ over $C_T = 0$ is 1 - q. The computation involved the non-central t-distribution and the non-centrality parameter of the distribution is a key factor in computing the DL. Clayton et al. provide tables for determining Δ, the non-centrality parameter of the distribution corresponding to p, g, and the degrees of freedom of the t-distribution (n-2, where n is the number of samples used to estimate the calibration line). The detection limit is computed as:

$$DL = \Delta s \frac{\sqrt{1 + \frac{1}{n} + \frac{\lambda^2}{Q^2}}}{b}$$

where the value of Δ is obtained from tables provided by Clayton, et al. The other variables in the formula are the same as defined in Section 2.2.1. If replicate measurements of sample with an unknown concentration where averaged to estimate that concentration, the DL would be computed by replacing the factor 1 + 1/n wherever it appears by 1/r + 1/n, where r is the number of replicate measurements. As in the Hubaux and Vos approach, Clayton et al. note that the value determined for

the detection limit is affected by the number of samples used to estimate the calibration line, the concentrations selected for estimating the calibration line, and the number of replicate measurements used to estimate an unknown concentration.

2.2.3 Glaser et al., 1981 and 40 CFR 136 Appendix B

The detection limit is defined as:

> ... the minimum concentration of a substance that can be measured and reported with 99% confidence that the analyze is greater than zero and is determined from analysis of a sample in a given matrix containing the analyze.

This definition is deficient as an operational definition and, therefore, does not lead directly to a computational formula. A formula, however, is provided and the result is referred to as the MDL:

$$MDL = t_{0.99} * s$$

where

$t_{0.99}$ - 99th percentile of the t distribution with n-1 degrees of freedom.

s - Estimated standard deviation.

In this approach, the calibration line is treated as if it were known with certainty and the standard deviation, s, is computed directly from estimated concentrations, where Hubaux and Vos and Clayton et al. compute s from instrument responses. The procedure specifies that s should be computed from n (at least seven) aliquots, properly spiked with analyte and processed through the full analytical procedure. The resulting value of s is multiplied by the 99th percentile point of the t-distribution with n-1 degrees of freedom, $t_{0.99}$. For example, when $n = 7$, $t_{0.99} = 3.14$.

This DL is the threshold value in concentration units for deciding between $C_T = 0$ or $C_T > 0$. Since this approach does not explicitly incorporate the calibration line, it does not account for variability associated with the estimate of the calibration line. In addition, this approach does not acknowledge the effects on the DL of replication and averaging.

E3. DESCRIPTION OF DL DEFINITIONS FROM EPA LITERATURE

E3.1 American Chemical Society, subcommittee on Environmental Analytical Chemistry, 1980

Limit of detection (LOD) is defined as the instrument response for a field blank, denoted as S_b, plus a multiple of the instrument response standard deviation (σ) of the field blank measurements. The recommended multiple is 3:

$$LOD = S_b + 3\sigma$$

A limit of quantitation (LOQ) is defined as:

$$LOQ = S_b + 10\sigma$$

The discussion states that increasing the multiple of σ reduces false positive and false negative decisions and no formal justification for using multiples of 3 and 10 are provided.

E3.2 Gibbons et al., 1989

The DL defined here is based on the decision problem, choosing between $C_T = 0$ and $C_T > 0$, formulated and solved in Clayton et al., 1987. Gibbons et al. raise the question of whether the decision problem involves one future sample or many future samples. If only one future sample will be tested, the solution in Clayton et al. is correct. If many future samples will be tested, the Clayton et al. DL is too small to assure that detection will be accomplished with the specified probability of 1-q for all samples.

The solution proposed in this report is to derive a multiplier for the DL by formulating the decision problem as a tolerance limit problem. Simply stated, the requirement for testing $C_T = 0$ versus $C_T > 0$ is that the probability should be 0.99 that 99 percent of all future decisions be correct. Tables for values of the multiplier are provided.

E3.3 Bauer, 1990

This report, prepared as statistical support to the EPA Office of Solid Waste, recommends the Clayton et al. (1987) approach.

E3.4 Keith, 1991

This article includes the following definitions:

> Limit of Detection (LOD) - the lowest concentration level that can be determined to be statistically different from a blank at a specified level of confidence.
>
> Method Detection Limit (MDL) - the minimum concentration of a substance that can be measured and reported with 99% confidence that the analyze concentration is greater than zero. It is determined from analysis of a sample in a given matrix containing the analyte.
>
> Reliable Detection Limit (RDL) -the concentration level at which a detection decision is extremely likely. It is generally set higher than the MDL or LOD.
>
> Limit of Quantitation (LOQ) -the level above which quantitated results may be obtained with a specified degree of confidence.

The formula for the LOD is $S_b + 3\sigma$ where S_b is the instrument response for blanks and σ is the instrument response standard deviation. The definition of MDL provided here is the 40 CFR 136 Appendix B definition; the formula given for the MDL is 3σ. No guidance is provided for estimating a value for σ; however the approach in 40 CFR 136 Appendix B seems to be implied. The recommended RDL value is 6σ and the recommended LOQ is 10σ.

The new concept in this article is the LOQ. An LOQ equal to 10σ "... is recommended, corresponding to an uncertainty of ±30% in the measured value ... at the 99% confidence level." Stated differently, if <u>one</u> measurement were used to estimate the unknown concentration in a ample, the

estimation error would be less than ±30% with a confidence level greater than 99% if the true concentration being estimated is equal to 10σ, the LOQ. In mathematical terms, the estimation error is

$$\frac{X - C_T}{C_T}$$

and the corresponding confidence statement is slightly larger than 99%, in fact;

$$P[-0.30 < \frac{X - C_T}{C_T} < 0.30 \mid C_T = 10\sigma] = 0.997 .$$

where

C_T - true concentration
X - measurement used to estimate C_T
σ - analytical variability.

The following table shows other values that may be assigned to the LOQ corresponding to different estimation error and confidence level requirements.

Confidence Level (Probability)	Estimation Error (Percent)	LOQ (Concentration)
0.95	5	39.5σ
0.95	10	20.0σ
0.95	20	10.0σ
0.95	30	6.5σ
0.99	5	51.5σ
0.99	10	26.0σ
0.99	20	13.0σ
0.99	30	9.0σ

It should be noted that replication and averaging of measurements are usually employed to achieve estimation error goals. As an example, if n replicate measurements were specified, each LOQ in the table above would be divided by the square root of n.

E3.5 Meredith/Boli and Associates, Inc., 1992

This report argues for the use of a Practical Compliance Reporting Limit (PCRL) in the National Pollutant Discharge Elimination System permits process. The PCRL is defined to be the upper 95 percent confidence limit (UCL) of the MDL defined in 40 CFR 136 Appendix B, multiplied by 10. The formula for computing the upper 95 percent confidence limit is

$$UCL = \sqrt{\frac{n-1}{\chi^2_{0.025}}} * MDL$$

where

n - Number of aliquot used to compute the MDL

$\chi^2_{0.025}$ - 2.5th percentile of the Chi-Square distribution with n-1 degrees of freedom

MDL - $t_{0.99}$*s as defined in 40 CFR 136 Appendix B.

When n = 7, as recommended in 40 CFR 136 Appendix B, $\chi^2_{0.025} = 1.24$ and UCL = 2.2*MDL.

E3.6 Miller, 1992

This report adopts the definition of MDL in 40 CFR 136, Appendix B. One suggestion offered for an alternative to the interpretation of the MDL is to identify a "not enough information" region to complement the "detected region." If a measurement were in the "not enough information" region, additional (replicate) measurements then would be required in order to reach a decision (i.e., "detected" or "zero"). The implementation of this approach, described in Appendix I of the report, does not explicitly incorporate a calibration function. The concepts, however, controlling Type I and Type II statistical error rates for deciding between $C_T \doteq 0$ and $C_T > 0$, are similar to the approach in Hubaux and Vos, 1970, and Clayton et al., 1987.

E3.7 Telliard, 1992

This report describes, among other things, the Minimum Level (ML). As published in 40 CFR 136, October 26, 1984, the ML is defined as "the level at which the entire analytical system shall give recognizable signal and calibration points." This definition has subsequently been refined and related to the concentration of the lowest of the calibration standards analyzed. The refined definition has been expressed as "the concentration of the analyte in a sample that is equivalent to the concentration of the lowest of the initial calibration standards, assuming that all the method-specified sample weights and volumes have been employed."

The advantages to the use of the ML include the fact that the laboratory must have demonstrated this level of sensitivity during the routine analyses of calibration standards. The numerical value of the ML can be derived for any analytical method where the concentrations of the calibration standards are specified. If calculated for a range of similar methods, the ML offers a simple means of comparison between the alleged sensitivities of the methods.

The Office of Science and Technology within EPA's Office of Water uses the ML as a "standardizing reporting level" in its analytical contracts, thereby eliminating the single laboratory nature of the MDL.

E3.8 EPA Office of Solid Waste Quantitation/Detection Limits

This report summarizes a number of OS DL definitions. The MDL definition for SW-846 methods appears to be identical to the definition in 40 CFR 136 Appendix B.

SW-846's Estimated Quantitation Limit (EQL) is defined as the lowest concentration that can be reliably achieved within specified limits of precision and accuracy during routine laboratory operating conditions. The ECL is generally 5 to 10 times the MDL. However, it may be nominally chosen within these guidelines to simplify data reporting. For many analyses the ECL analyze concentration is selected for the lowest non-zero standard in the calibration curve.

E3.9 Grant, Hewitt, and Jenkins, 1991

MDL is defined as in 40 CFR 136 Appendix B. The decision problem of deciding between $C_T = 0$ and $C_T > 0$ is noted and a certified reporting limit (CRL) is defined in accordance with requirements of the U.S. Army Toxic and Hazardous Materials Agency to address the false positive decision error rate. The C.L. is established using the approach described in Hubaux and Vos, 1970.

E3.10 Diebold, 1991

This report, a "fact sheet" prepared in EPA Region IX, provides various definitions of DLs.

The Instrument Detection Limit (IDL) is the lowest amount of a substance that can be detected by the analytical instrument, such as gas chromatography (GC), above the background noise level of the instrument. The IDL is defined as "...three times the standard deviation of 7 replicate analyses of the substance (analyte) at the lowest concentration level that is statistically different from a blank." This definition reflects the approach in 40 CFR 136 Appendix B. The IDL is usually determined by analyzing solutions of the analyze in pure water. Since the IDL is dependent only on the instrument portion of detection, it does not reflect a measurement of the effects of sample preparation, concentration or dilution, or the sample matrix.

MDL is defined as the lowest amount of an analyze that can be detected using a specific analytical method. The implied computational procedure is that of 40 CFR 136 Appendix B.

Sample Quantitation Limit (SQL) is defined to be "...a sample specific quantitation limit that takes into account actual sample characteristics in addition to the detection ability of the analytical method. The SQL is obtained by adjusting the MDL to reflect sample-specific actions taken during analysis. An individual sample may require adjustments in preparation or analysis, such as dilution/concentration due to matrix effects or the high/low concentration of some analyses. The reported SQLs take into account sample characteristics, sample preparation, and analytical adjustments."

The Practical Quantitation limit (PQL) is intended to be a measurement concentration that is routinely achievable independent of time and laboratory. The PQL is defined as the lowest concentration that can be reliably quantified within specified limits of precision and accuracy during routine laboratory operating conditions. In the Agency's Safe Drinking Water Program, PQLs are determined by the following procedures:

> From multi-laboratory performance evaluation data, find a concentration that most good laboratories (e.g., 80% to 100%) could measure with error no greater than ±40%. If multi-laboratory data are not available, use 5 to 10 times the MDL defined in 49 CFR 136 Appendix B.

In the RCRA manual of test methods (SW-846), PQLs are provided for guidance and are determined by multiplying the MDL by a method-dependent matrix factor. For example in Method 8020, the matrix factor for groundwater and low-level soil is 10, while the factor for high-level soil and sludge is 1250. The PQLs listed for groundwater monitoring under RCRA are generally estimated at 10 times the MDL.

The terms Contract Required Detection and Quantitation Limit (CRDL and CRQL) are used by the Agency's Superfund Contract Laboratory Program (CLP). The CLP uses a CRDL for inorganics (e.g., metals) and a CRQL for organics (e.g., volatiles and pesticides). They refer to the minimum level of detection or quantitation, respectively, acceptable under the Contract Statement of Work (CSW). The CRDLs are the instrument detection limits obtained for the analyses in pure water (standards) and the CRQLs are typically 2 to 5 times the reported MDLs.

E4. COMPARISON OF DL APPROACHES

E4.1 Qualitative Comparisons

Among the variety of definitions for DLs recorded in Section E3, only a few represent distinctly different concepts. These are summarized in the following sub-sections.

E4.1.1 MDL defined in 40 CFR 136 Appendix B

This DL is the threshold value for deciding between $C_T = 0$ and $C_T > 0$ based on one measurement of a sample with unknown concentration. Concluding that $C_T > 0$ is correct is equivalent to "detection." The threshold value is intended to limit the false positive error rate for the decision (i.e., the Type I statistical error rate) to 0.01. The numerical value of this DL is the estimated analytical standard deviation for samples with low concentrations multiplied by a factor approximately equal to three.

E4.1.2 DL defined by Hubaux and Vos (1970) and Clayton et al. (1987)

This DL is the minimum true concentration associated with a large probability (e.g., 0.95) or 0.99) that $C_T > 0$ will be chosen over $C_T = 0$ when the decision is made by comparing a measured value to the threshold value defined in 4.1.1. This DL may be reduced if replicate measurements and averaging are employed, and adopted as an integral part of the analytical method. The numerical value of this DL is the standard deviation for the fitted calibration line multiplied by a factor that will be equal to, at least, six.

E4.1.3 LOQ at 10σ

This DL represents the smallest true concentration that can be estimated by a single measurement to within an error of ±30% with 99% confidence.

E4.1.4 PQL defined in Safe Drinking Water Program

This DL has been defined in various ways. The unique definition for purposes of this report is the one that accounts for laboratory differences by using the distribution of inter-laboratory test data. The data for defining this DL would consist of estimation errors for samples with low concentrations at a large number of laboratories. The lowest concentration with an estimation error of, say, less than ±30% in a large percentage of laboratories (e.g., 80 or 90 percent) would be the DL. It should be noted that replication and averaging may be used to reduce estimation error and, therefore, reduce the DL obtained by this approach if replication were adopted as an integral part of the analytical method.

E4.1.5 Tolerance Limit DL

The DLs described in Sections 4.1.1 to 4.1.4 above are derived from a structure where one future sample with unknown concentration is to be measured and tested to determine if the true concentration, C_T, is zero or if $C_T > 0$. If two future samples were to be tested, the probability of at least one incorrect decision would be larger than the probability of an incorrect decision when one sample is tested. The DL value, therefore, would have to be larger if decisions concerning $C_T = 0$ and $C_T > 0$ are anticipated for more than one future sample to assure that desired probabilities of correct decisions are achieved.

If the number of future samples under consideration were less than 20, the increase in the DL would be negligible (Gibbons et al., 1989). If the number of future samples expected were larger than 20, the DL would have to be increased to assure, for example, with 99% confidence that 99% of all future decisions concerning $C_T = 0$ versus $C_T > 0$ would be correct. The specified increase in the DL depends, in part, on the number of samples used to estimate σ for the analytical method. As an example, if the DL were based on an estimate of σ with six degrees of freedom, the DL for all future samples would be approximately 1.6 times the DL for one future sample determined by the Hubaux and Vos. (1970) or Clayton et al. (1987) methods.

E4.2 Matrices and Laboratories

Much of the controversy surrounding DLs derives from the concern that a DL based on samples of one type of matrix analyzed in one particular laboratory will not reflect the use of the analytical method for other matrices or in other laboratories. Every definition of DL involves, as a primary component, analytical variability, which is quantified as an estimated standard deviation. Since the analytical standard deviation for a particular analyze can be expected to vary by matrix, concentration, laboratory, and other factors, the concern that DLs will vary from one circumstance to another has merit. The DL, in almost all cases, is computed as the analytical standard deviation multiplied by a factor greater than one. The result may be viewed as an enlargement of standard deviation, which is intended to incorporate all the sources of measurement variability that may not have been operating in the experiment that produced the estimated standard deviation. It is doubtful if any one multiplicative factor can be justified for this purpose in all cases. DL values, therefore, will vary by matrix and laboratory, and separate DL determinations are likely to be required for different matrices and laboratories.

E5. BIBLIOGRAPHY

American Chemical Society, Subcommittee on Environmental Analytical Chemistry. 1980. Guidelines for data acquisition and data quality evaluation in environmental chemistry. *Anal. Chem.* 52:2242-2249.

Bauer, K.M. 1990. *Statistical Support to OSW.* Prepared by Midwest Research Institute for the Office of Solid Waste, U.S. EPA. June 29.

Clayton, C.A., Hines, J.W. and Elkins, P.D. 1987. Detection limits with specified assurance probabilities. *Anal. Chem.* 59:2506-2514.

Diebold, T. 1991. *Detection and Quantitation Limits.* U.S. EPA, Region IX, Quality Assurance Management Section.

Gibbons, R.D., Jarke, F.H., and Stoub, K.P. 1989. *Detection Limits: For Linear Calibration Curves with Increasing Variance and Multiple Detection Decisions.*

Glaser, J.A., Foerst, D.L., McKee, G.D., Quave, S.A. and Budde, W.L. 1987. Trace analyses for wastewaters. *Env. Sci. Tech.* 15:1426-1435.

Grant C.L., Hewitt, A.D. and Jenkins, T.F. 1991. Experimental comparison of EPA and USATHAMA detection and quantitation capability estimators. *Am. Lab.* (February): 15-33.

Hubaux, A. and Vos, G. 1970. Decision and detection limits for linear calibration curves. *Anal. Chem.* 42:849-855.

Keith, L.H. 1991. Report results right! *Chemtech* (June): 352-356; (August): 486-489.

Meredith/Boli and Associates, Inc. 1992. Discussion Paper: Quantification of Measurement Variability for Use in Setting Surface Water Discharge (NPDES) Permit Compliance Limits. Los Angeles, California.

Miller, D.P. 1992. *Method Detection Limit.* U.S. EPA, Region VII Laboratory.

Office of Solid Waste. *OSW Quantitation/Detection Limits.* Undated Information Sheet. Office of Solid Waste. U.S. EPA.

APPENDIX F

VERIFICATION AND VALIDATION

Data verification and validation are important parts of the Agency's QA Program because they impact directly on the assessment of data quality with respect to the planned use of the data. There is, however, no universal agreement on the precise definitions of the terms verification and validation. This appendix discusses different definitions and perspectives on data verification and validation, presents an overview of Data Validation Plans, and provides a brief example of data verification and validation principles applied to radiochemical data. Appendix F closes with a list of issues posed as questions for consideration within the environmental QA community.

F1. DEFINITIONS OF VERIFICATION AND VALIDATION

This section presents a sampling of definitions of the terms *verification* and *validation* taken from the literature. An analysis of these definitions leads to a synthesis of a general model for how to view the relationships among verification, validation, and Data Quality Assessment.

Definitions from the Literature

1. Webster's Dictionary

   Verification—The authentication of truth or accuracy by such means as facts, statements, citations, measurements, or attendant circumstances.

   Validation—An act, process, or instance of validating, where *validate* means:
   (1) to grant official sanction to, by, or as if by stamping or marking;
   (2) to corroborate or support on a sound basis or authority.

2. *EPA Requirements for Quality Management Plans*, EPA QA/R-2, Draft Interim Final, August 1994

   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, validation [*sic*] concerns the process of examining a result of a given activity to determine conformance to the stated requirements for that activity.

3. American National Standard ANSI/ASQC E4-1994

   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, validation [*sic*] concerns the process of examining a result of a given activity to determine conformance to the stated requirements for that activity.

4. Radiochemical Data Verification and Validation, 1995

Analytical Data Validation—a systematic process, performed external from the data generator, which applies a defined set of performance-based criteria to a body of data that may result in physical qualification of the data. Data validation occurs prior to drawing a conclusion from the body of data.

Analytical Data Verification—a process of evaluating the completeness, correctness, consistency, and compliance of a set of facts against a standard or contract. Data verification is defined as a systematic process, performed by either the data generator or by an entity external to the data generator.

5. *Environmental Sampling and Analysis: A Practical Guide*, Keith, L.H., 1991

Validation—an experimental process involving external corroboration by other laboratories (internal or external), methods, or reference materials to evaluate the suitability of methodology.

Verification—the general process used to decide whether a method is capable of producing accurate and reliable data.

6. EPA Internal QA Workgroup on QAPP Guidance

Validation—a systematic process that provides documented evidence with a high degree of assurance that a method, an instrument, or a system performs consistently, reliably, and accurately the function it is intended or designed to do, as defined in the project's Data Quality Objectives (DQOs).

Verification—a systematic process for evaluating compliance of a set of data to a set of standards to ascertain its completeness, correctness, and consistency. Verification is a process for determining that a given procedure produces the intended results within predefined limits, so that it will produce reliable data.

A General Model for Verification and Validation

Despite the diversity of wording, there seems to be general agreement in the literature on the meaning of the terms, at least with respect to key underlying concepts:

Validation—evaluation of the technical usability of the generated data.

Verification—determination of adherence to SOPs or contractual requirements.

It follows that verification is performed first, and involves a relatively objective or "mechanical" evaluation of whether or not the data collection plans and protocols were followed, and

that basic operations and calculations were performed correctly. Verification is followed by validation, which involves a higher level of scientific evaluation to determination if the protocols and procedures that were performed were appropriate for the actual situation encountered, and whether the results make sense in the context of the study objectives. The results of data validation determine whether the reported data values can be trusted in the final assessment phase, Data Quality Assessment (DQA). DQA is where the data set as a whole is evaluated to determine if the Data Quality Objectives (DQOs) were satisfied (i.e., whether scientific conclusions can be drawn or environmental management decisions can be made with acceptable confidence).

Figure F-1 shows how data verification, validation, and DQA can be viewed as an assessment hierarchy with overlapping boundaries. Verification is the lowest level, supporting subsequent validation and DQA activities, and relying on information provided in QAPP specifications for measurement protocols and performance. Validation is the middle level, supporting subsequent DQA activities, and relying on information from both the QAPP specification and from the DQOs for contextual meaning. DQA supports the decision making process at the top level, relying on valid data from the previous verification and validation activities, as well as information on context and assumptions to be evaluated from the DQOs. The distinction of where verification ends and validation begins is often blurred, as shown by the overlapping ovals in Figure F-1. Likewise, the early steps of DQA involve preliminary data analysis and evaluation of assumptions, which overlap with higher-level data validation activities. To the extent that these distinctions may be important for a project, one may appeal to the expert opinion of the data user as the deciding voice.

This general model is consistent with the perspective of R. Cohen in "Issues Regarding

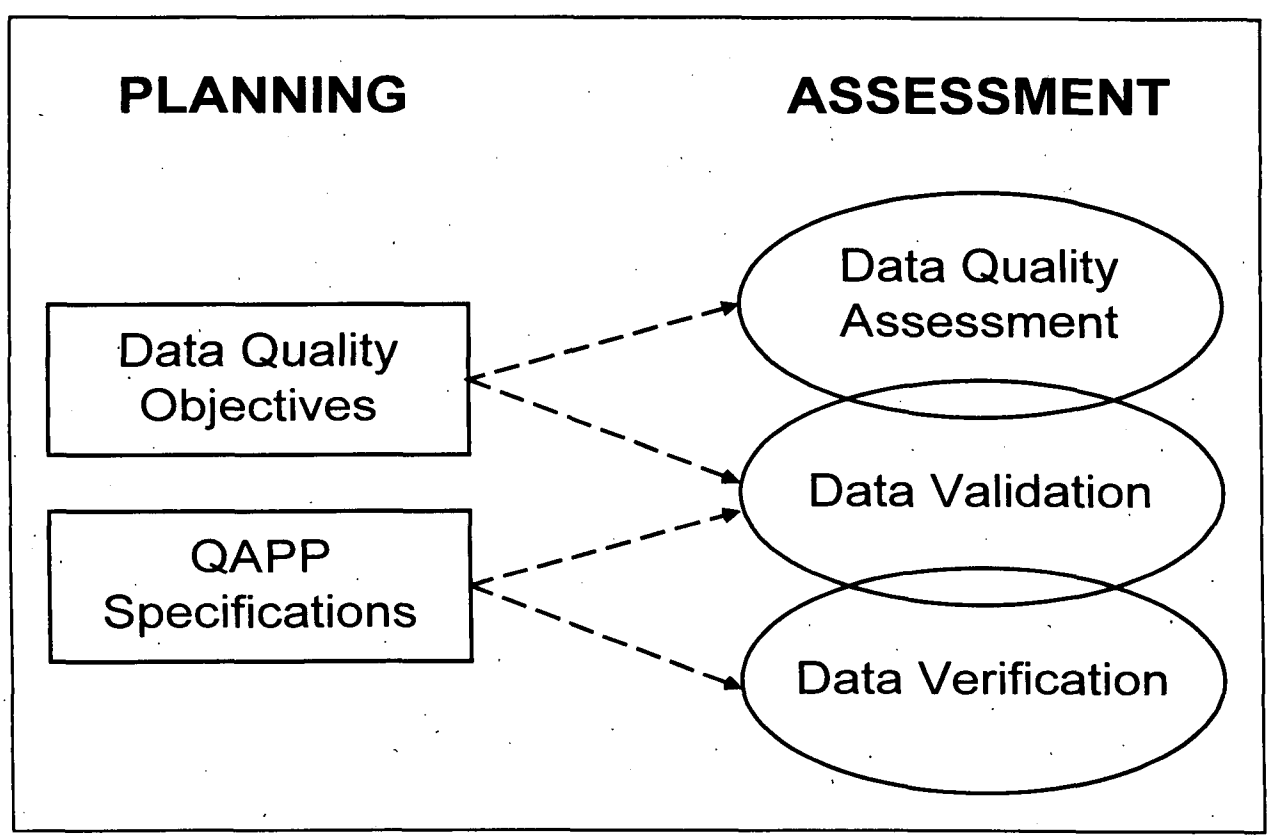


Figure F-1. Relationships Among Verification, Validation, and Data Quality Assessment

Validation of Environmental Data," presented at the Waste Testing and Quality Assurance Symposium, Washington, DC, 1995. Cohen characterizes verification as the quartet of:

- compliance with contractual specifications;
- completeness with respect to information for validation analysis;
- consistency of information from multiple data collection sites; and
- correctness in calculation of numerical results.

After data verification, the data may be assessed by a data validator with respect to quality control information, sampling collection techniques, analytical performance, and other related information. The data validator then assigns data qualifiers to each of the data points based on the impact of deviations from performance standards.

Many elements of the validation process are to be found in Appendix D, "Data Quality Indicators," and Appendix G, "Quality Control." Several organizations within EPA have combined the activities of verification and validation into a separate document, the Data Validation Plan, as a tool to assist data validator in their work, as described in the next section.

F2. DATA VALIDATION PLANS

Overview

The purpose of a Data Validation Plan (DVP) is to define and document the validation process prior to performing environmental measurements. The DVP serves as a mechanism for ensuring that all sampling and analysis requirements for the data's intended use are met.

The DVP provides a framework for implementing the validation scheme developed in the QAPP and a mechanism for consistent implementation of the data quality requirements specified in the DQO process. The DVP framework can be adapted for validating computer models and other hardware/software systems that support environmental data operations (this topic will be addressed in more detail in a future Appendix L on data management).

Scope and Inputs

The DVP should include a detailed implementation scheme that addresses the following areas:

- how the DQOs will be integrated with the validation plan;
- management and maintenance of the validation plan;
- resource requirements;
- agency policies;
- training (requirements, personnel needs, sources);
- validation process and methods (statistical procedures and methods for analyzing and evaluating data, such as calibration, evaluation of systematic and random error, technology standards, SOPs, sampling);
- QA/QC protocols;
- audits;
- data quality assessment; and
- level of compliance.

Typical basic inputs for the validation of data for air, drinking water, water and waste water, solid waste and hazardous waste are:

Initial Calibration/Frequency

- Calibration (continuing calibration)
- Calibration standards used include analytes of interest and concentration?

Standards/Required/Used/Frequency

- Internal standards (standard used, acceptance criteria met?)
- Standards level/Concentration
- Analytes (target, concentration, acceptance limits)
- Laboratory control

Qualitative or Quantitative test performed

- Method Blanks/blank matrix
- Frequency
- QA/QC criteria for acceptance

Samples analyzed/type (organic, inorganic, radionuclides, etc.)

- Sample blanks

Type of instruments checked

- Criteria used for accepting instrument performance (QA/QC, response factors, precision and accuracy, etc.)
- For software and hardware, criteria should be established to demonstrate suitability to meet the tests and challenges for the tasks expected for the system.

Actual Sample Analysis

- Sampling and analysis plan (sampling design, sample analysis, sampling execution)
- Holding time
- Volume/weight required/used
- Internal standard/blank (requirement met?)
- Surrogate (present, required QA/QC met? acceptable limits, recoveries met?)
- Analytes of interest (how analyzed/identified/quantification and/or qualitative criteria)
- Analytes of non-interest (how identified, analyzed, quantified or qualified, criteria used)
- Duplicates
- Method precision and accuracy

The DVP also should include budget projections, and a cost assessment to determine if projected cost match actual cost, determine any budget overrun, cost ceilings and whether or not these costs are justified.

The definitions of verification and validation are those used by the Internal QA Workshop on QAPP. Several organizations within the radiochemical environmental community use a more complex system, the Radiochemical Data Verification and Validation Procedure.

F3. RADIOCHEMICAL DATA VERIFICATION AND VALIDATION PROCEDURES

This section describes an example of how verification and validation procedures are applied to radiochemical data. Verification and validation of radiochemical data often is performed differently from that for traditional chemical analysis data due to the special nature of radiochemical measurement systems (i.e., the radioactive particle counting process employed in the measurement process is amenable to the calculation quantitative uncertainties for each measurement).

Overview of Activities

In this example, verification and validation are integrated into a sequence of clearly defined activities at each stage of the review process. The process is broken into 9 steps that address the following areas:

1. Custody of Samples and Sample Documentation
2. Holding Time and Turn-around Time
3. Sample Preservation
4. Instrument Calibration (12 sub-steps)
5. Quality-indicator Samples (6 sub-steps)
6. Chemical Yield Tracers and Carriers
7. Required Detection Limits
8. Nuclide Identification and Quantification (2 sub-steps)
9. Instrument Specific Sample Considerations (2 sub-steps)

Partial Example

The following is an example of the content for step 6:

F. Chemical Yield Tracers and Carriers

1. Verification

   Verify that for applicable analyses, one carrier or tracer recovery is reported for each sample. If a carrier or tracer percent recovery is not reported for each sample, contact the laboratory for submittal of this data. If the data can not be provided, state this as a non-correctable problem in the verification report.

   As yield decreases, the MDC may elevate to a point at which the RDL is exceeded, and analytical results are contractually noncompliant. If the laboratory has not initiated corrective action, for samples in which the MDC exceeds the RDL, the project may choose to contact the laboratory for sample rework. If rework is not feasible, indicate the noncompliant data in the verification report.

2. Validation

   Yield is validated based on percent recovery of the spiked nuclide. Low yield may be indicative of increased uncertainty in the sample result. Criteria for

qualification should be based on what magnitude of correction has been applied to the sample result (e.g., 20% recovery implies a sample result correction factor of 5), although a point of debate exists concerning useability of radionuclide data with yields near 0%. Yield criteria may also be established from existing sample yield data from previous sampling at the site, if these data are available.

Sample results should not be qualified based on yield alone. Sample yield should be evaluated in reference to chemical yield of quality-indicator samples. If yield is generally low throughout the preparation batch, but recoveries of target radionuclides in the LCS are acceptable, data may be accepted without qualification; however, if quality control sample yield is generally low, sample results with low yield may need qualification.

F4. UNRESOLVED ISSUES

This section raises some unresolved issues regarding data verification and validation, which are posed as questions for consideration within the QA community.

- Can the effects of differing matrices be quantified, and how can this be used to improve the comparability of different data sets?
- Should verification and validation issues be combined into a separate Data Validation Plan?
- How can the effects of deviations from verification standards (for example, exceeding a contractual holding time) and validation requirements (for example, recovery rates just outside the window of acceptability) be quantified and combined to make overall estimates of data quality?
- Is it necessary to break all aspects of a procedure into verification and validation instructions and guidance?
- Are verification and validation issues sufficiently well understood and documented such that no extra guidance is necessary?

APPENDIX G

QUALITY CONTROL FOR ENVIRONMENTAL STUDIES

G1. QUALITY CONTROL OPERATIONS

Quality Control (QC) plays an increasingly important role in environmental studies, especially when studies are conducted for the purpose of deciding what action to take to address an environmental problem. To minimize the chance of making an incorrect decision, data of adequate quality must be collected. QC programs can be designed and utilized to both lower the chances of making an incorrect decision, and to understand the level of uncertainty that surrounds the decision. QC operations provide the decision maker and data collectors with insight into where error is occurring, what the magnitude of that error is, and how it might impact the decision making process. This appendix provides a brief overview of this complex topic. It surveys the different types of QC samples that can be applied to environmental studies and evaluates how they are currently deployed as specified by EPA methods and regulations.

General Objectives

The two most important questions a manager should consider are:

What is the range of QC requirements for existing methods, and

What types of problems in environmental measurement systems do these requirements enable the Agency to detect?

Addressing these questions should provide the manager with the background needed for addressing the concept of a uniform, minimum, set of QC requirements for all environmental data collections. Understanding existing QC requirements for environmental data collection activities provides a framework for considering what set of QC requirements should be considered "core" irrespective of the end use of the data.

While it is difficult to define a standard of data quality irrespective of its use, core QC requirements can be established that will enable one to provide data of known quality in accordance with the Agency's QA Program. This program has the requirement that all environmental data collection efforts need information on bias, variability, and sample contamination. These error types are incurred throughout the data generation process including all sampling and analytical activities (i.e., sample collection, handling, transport and preparation; sample analysis; and subsampling). The principal issue centers on what level of detail in the error structure should QC operations be capable of revealing, given that it will be impractical to explore every known potential source of error.

Background

Many of the essential elements of a QAPP apply directly to sampling and analytical activities and include: QA objectives for measurement data specified in terms of precision, accuracy, bias, representativeness and comparability; sampling procedures; sample custody; calibration procedures and frequency; analytical procedures; internal quality control checks and frequency; performance and system audits and frequency; and specific routine procedures that should be used to assess data precision, accuracy and completeness of the specific measurement parameters involved.

There are no global QC requirements for EPA program offices, laboratories, and methods and

various program objectives and priorities warrant different levels of data quality and associated levels of QC. The program's Quality Assurance Officer or representative should have details on specific QC requirements.

Definitions and Terminology

In order to ensure that managers have a uniform perspective of QC requirements, it is necessary to discuss some basic terminology and definitions. Quality control and quality assurance, total study error and its components, types of QC operations, and Good Laboratory Practices will be discussed. Specific definitions of these terms and others are provided in Appendix I, "Additional Terms and Definitions." Table 1 summarizes the results of a study on how these terms are defined and used in EPA and non-EPA literature. Five commonly available sources are discussed in Table 1: Appendix I in EPA QA/G-5; *Definitions of Environmental Quality Assurance Terms* (1996) published by ASQC; *A Rationale for the Assessment of Errors in Sampling of Soils* by van Ee, Blume and Starks (1989); *Quality Assurance of Chemical Measurements* by Taylor (1987); and *Principles of Environmental Sampling* by Keith (1988).

Quality Control vs. Quality Assurance

EPA QA/G-5, van Ee, Blume and Starks, and Taylor provide somewhat similar definitions for both quality assurance and quality control. Quality control activities are designed to control the quality of a product so that it meets the user's needs. Quality assurance includes quality control as one of the activities needed to ensure that the product meets defined standards of quality.

These two terms have been defined in slightly different ways by other authors, but all are in agreement that quality control is a component of quality assurance. Many authors define quality control as "those laboratory operations whose objective is to ensure that the data generated by the laboratory are of known accuracy to some stated, quantitative degree of probability." (pp. 5-7, Dux 1986) The objective of quality control is not to eliminate or minimize errors, but to measure or estimate what they are in the system as it exists. The same authors then define quality assurance as the ability to prove that the quality of the data is as reported. Quality assurance relies heavily on documentation, including documentation of implemented quality control procedures, accountability, traceability, and precautions to protect raw data.

QC Samples

Table 1 offers a broad survey of commonly used QC terms, including the definitions of QC sample types that span the measurement process. The authors cited in Table 1 define different sample types in varied ways, however, the definitions are not in contradiction.

Good Laboratory Practices

The Food and Drug Administration (FDA) promulgated the first version of the Good Laboratory Practices (GLPs) in 1978. The Environmental Protection Agency (EPA) promulgated similar guidance requirements in 1983 for Resource Conservation Recovery Act (RCRA) and Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) compliance. The FIFRA GLPs were revised in 1988. Though much of the content relates to laboratory animal science, many requirements are relevant to the analytical chemist. The Good Laboratory Practice Standards for FIFRA (40 Code of Federal Regulations Part 160) and Toxic Substances Control Act (TSCA) (40 CFR 792) are similar. (pp. 176-177, Dux 1986) Selected topics of FIFRA subparts A through K appear below.

Subpart A	General Provisions.
Subpart B	Organization and Personnel. Includes: quality assurance unit.
Subpart C	Facilities. Includes: facilities for handling test, control, and reference substances; laboratory operations areas; and specimen and data storage facilities.
Subpart D	Equipment. Includes: maintenance and calibration of equipment.
Subpart E	Testing Facilities Operation. Includes: standard operation procedures; and reagents and solutions
Subpart F	Test, Control, and Reference Substances. Includes: characterization and handling; and mixtures of substances with carriers.
Subpart G	Protocol for and Conduct of a Study.
Subpart H	Reserved.
Subpart I	Reserved.
Subpart J	Records and Reports. Includes: reporting of study results; storage and retrieval of records and data; and retention of records.

Good laboratory practices are defined similarly by the Agency and by Taylor (1987) as an acceptable way to perform some basic laboratory operation or activity that is known or believed to influence the quality of its outputs.

G2. QC REQUIREMENTS IN EXISTING PROGRAMS

To identify QC requirements for this section, standard EPA method references, such as SW-846, and the Code of Federal Regulations (CFR) were consulted together with information on non-EPA methods identified through a computerized literature search. Within the EPA literature, some of the major programs were reviewed, including Drinking Water, Air, and the Contract Laboratory Program (CLP). Different types of methods, such as gas chromatography (GC), atomic absorption (AA), and inductively coupled plasma (ICP), and different media were included in this process but it was not intended to be exhaustive.

Summary of QC Requirements by Program and Method

Table 2 presents the frequency of QC requirements for different selected programs and Table 3 presents information for methods. In cases where different programs use dissimilar terms for similar QC samples, the table uses the term from the program or method.

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
Blank Sample	A clean sample or a sample of matrix processed so as to measure artifacts in the measurement (sampling and analysis) process.	Blanks provide a measure of various cross-contamination sources, background levels in reagents, decontamination efficiency, and other potential error that can be introduced from sources other than the sample. A rinsate blank (decontamination sample) measures any chemical that may have been on the sampling and sample preparation tools after the decontamination process is completed.	The measured value obtained when a specified component of a sample is not present during measurement. Measured value/signal for the component is believed to be due to artifacts; it should be deducted from a measured value to give a net value due to the component contained in a sample. The blank measurement must be made to make the correction process valid.	Samples expected to have negligible or unmeasurable amounts of the substance of interest. They are necessary for determining some of the uncertainty due to random errors. Three kinds required for proper quality assurance: equipment blanks, field blanks. and sampling blanks.
Blind Sample	A subsample submitted for analysis with a composition and identity known to the submitter but unknown to the analyst. Used to test analyst or laboratory proficiency in execution of the measurement process.	Single-Blind Samples: Field Rinsate Blanks, Preparation Rinsate Blank, Trip Blank	A sample submitted for analysis whose composition is known to the submitter but unknown to the analyst. One way to test the proficiency of a measurement process.	

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
Calibration Standard	A substance or reference material used to calibrate an instrument. (calibration check standard, reference standard, quality control check sample)		In physical calibration, an artifact measured periodically, the results of which typically are plotted on a control chart to evaluate the measurement process.	Or quality control calibration standard (CCS). In most laboratory procedures, a solution containing the analyte of interest at a low but measurable concentration. Standard deviation of the CCSs is a measure of instrument precision unless the CCS is analyzed as a sample, in which case it is a measure of method precision.
Check sample		Example: ICP Interference Check Sample - Part A contains potential interfering analytes. Part B contains both the analytes of interest and the target analytes. Part A and B are analyzed separately to determine the potential for interferences.		
Check Standard	A substance or reference material obtained from a source independent from the source of the calibration standard; used to prepare check samples. (control standard)			Laboratory control standards are certified standards, generally supplied by an outside source. They are used to ensure that the accuracy of the analysis is in control.

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
Double Blind Samples		Samples that can not be distinguished from routine samples by analytical laboratory. Examples: Field Evaluation Samples, Low Level Field Evaluation Samples, External Laboratory Evaluation Samples, Low Level External Laboratory Evaluation Samples, Field Matrix Spike, Field Duplicate, Field Split	A sample known by the submitter but submitted to an analyst so that neither its composition nor its identification as a check sample are known to the analyst.	
Duplicate Measure-ment			A second measurement made on the same (or identical) sample of material to assist in the evaluation of measurement variance.	
Duplicate Sample	Two samples taken from and representative of the same population and carried through all steps of the sampling and analytical procedures in an identical manner. Used to assess variance of the total method including sampling and analysis.	Field duplicate - an additional sample taken near the routine field sample to determine total within-batch measurement variability. Analytical laboratory duplicate - a subsample of a routine sample analyzed by the same method. Used to determine method precision. It is non-blind so can only be used by the analyst in internal control, not an unbiased estimate of analytical precision.	A second sample randomly selected from a population of interest to assist in the evaluation of sample variance.	

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
Error	The difference between an observed or corrected value of a variable and a specified, theoretically correct, or true value.		Difference between the true or expected value and the measured value of a quantity or parameter.	
Field Blank				Used to estimate incidental or accidental contamination of a sample during the collection procedure. One should be allowed per sampling team per day per collection apparatus. Examples include matched-matrix blank, sampling media or trip blank, equipment blank.
Good Laboratory Practices (GLPs)	Either general guidelines or formal regulations for performing basic laboratory operations or activities that are known or believed to influence the quality and integrity of the results.		An acceptable way to perform some basic operation or activity in a laboratory that is known or believed to influence the quality of its outputs. GLPs ordinarily are essentially independent of the measurement techniques used.	
Instrument Blank				Also called system blank. Used to establish baseline response of an analytical system in the absence of a sample. Not a simulated sample but a measure of instrument or system background response.

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
Method Blank				One of the most important in any process. DDI water processed through analytical procedure as a normal sample. After use to determine the lower limit of detection, a reagent blank is analyzed for each 20 samples and whenever a new batch of reagents is used.
Non-Blind Sample		QC samples with a concentration and origin known to the analytical laboratory. Examples: Laboratory Control Sample, Pre-digest Spike, Post-digest Spike, Analytical Laboratory Duplicate, Initial Calibration Verification and Continuing Calibration Verification Solutions, Initial Calibration Blank and Continuing Calibration Blank Solution, CRDL Standard for ICP and AA, Linear Range Verification Check Standard, ICP Interference Check Sample.		

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
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. (Defined in EPA QA/G-5, App. 1)			
Quality assessment	Assessment is the evaluation of environmental data to determine if they meet the quality criteria required for a specific application.	The overall system of activities that provides an objective measure of the quality of data produced.	The overall system of activities whose purpose is to provide assurance that the quality control activities are done effectively. It involves a continuing evaluation of performance of the production system and the quality of the products produced.	
Quality Assessment Sample (QAS)		Those samples that allow statements to be made concerning the quality of the measurement system. Allow assessment and control of data quality to assure that it meets original objectives. Three categories: double blind, single-blind, and non-blind.		

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
Quality assurance (QA)	An integrated system of activities involving planning, quality control, quality assessment, reporting and quality improvement to ensure that a product or service meets defined standards of quality with a stated level of confidence.	A system of activities whose purpose is to provide to the producer or user of a product or service the assurance that it meets defined standards of quality. It consists of two separate, but related activities, quality control and quality assessment.	Same as Van Ee.	
Quality control (QC)	The overall system of technical activities whose purpose is to measure and control the quality of a product or service so that it meets the needs of users. The aim is to provide quality that is satisfactory, adequate, dependable, and economical.	The overall system of activities whose purpose is to control the quality of the measurement data so that they meet the needs of the user.	The overall system of activities whose purpose is to control the quality of a product or service so that it meets the needs of users. The aim is to provide quality that is satisfactory, adequate dependable, and economic.	
Quality Control Sample	An uncontaminated sample matrix spiked with known amounts of analytes from a source independent from the calibration standards. Generally used to establish intralaboratory or analyst specific precision and bias or to assess performance of all or part of the measurement system. (Laboratory control sample) (Defined in EPA QA/G-5, App. 1)	A sample of well-characterized soil, whose analyte concentrations are known to the laboratory. Used for internal laboratory control. Also called QC audit sample.	A material of known composition that is analyzed concurrently with test samples to evaluate a measurement process.	Used in quality control procedures to determine whether or not the analytical procedures is in control.

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
Reagent Blank	A sample consisting of reagent(s), without the target analyte or sample matrix, introduced into analytical procedure at the appropriate point and carried through all subsequent steps to determine contribution of the reagents and the involved analytical steps to error in the observed value. (analytical blank, laboratory blank) (Defined in EPA QA/G-5, App. 1)			Also called Method blank. Used to detect and quantitate contamination introduced during sample preparation and analysis. Contains all reagents used in sample preparation and analysis and is carried through the complete analytical procedure.
Sample Preparation Blank				Required when methods like stirring, mixing, blending, or subsampling are used to prepare sample prior to analysis. One should be prepared per 20 samples processed.
Sampling Equipment Blank				Used to determine types of contaminants introduced through contact with sampling equipment; also to verify the effectiveness of cleaning procedures. Prepared by collecting water or solvents used to rinse sampling equipment.

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
Solvent Blank				Used to detect and quantitate solvent impurities; the calibration standard corresponds to zero analyte concentration. Consists only of solvent used to dilute the sample.
Spiked Sample	A sample prepared by adding a known mass of target analyte to a specified amount of matrix sample for which an independent estimate of target analyte concentration is available. Spiked samples are used, for example, to determine the effect of the matrix on a method's recovery efficiency. (matrix spike)	A sample prepared by adding a known amount of reference chemical to one of a pair of split samples. Comparing the results of the analysis of a spiked member to that of the non-spiked member of the split measures spike recovery and provides a measure of the analytical bias. Field matrix spike - a routine sample spiked with the contaminant of interest in the field.		Matrix control or field spike -for sample matrices where a complex mixture (e.g. sediments, sludges) may interfere with analysis, a field spike may be required to estimate the magnitude of those interferences. Losses from transport, storage treatment, and analysis can be assessed by adding a known amount of the analyte of interest to the sample in the field.

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
Split Sample	Two or more representative portions taken from a sample or subsample and analyzed by different analysts or laboratories. Split samples are used to replicate the measurement of the variable(s) of interest.	Samples can provide: a measure of within-sample variability; spiking materials to test recovery; and a measure of analytical and extraction errors. Where the sample is split determines the components of variance that are measured. Field split - a sample is homogenized and spilt into two samples of theoretically equal concentration at the sampling site. Indicate within batch measurement error. Also called replicates.	A replicate portion or subsample of a total sample obtained in such a manner that is not believed to differ significantly from other portions of the same sample.	
Total Measurement Error	The sum of all the errors that occur from the taking of the sample through the reporting of results; the difference between the reported result and the true value of the population that was to have been sampled.			
Transport Blank				Used to estimate sample contamination from the container and preservative during transport and storage of the sample. One should be allowed per day per type of sample.

Table 1
Comparison of QC Terms

Terms	ASQC, Definitions of Environmental Quality Assurance Terms or EPA QA/G-5 App. I	van Ee, Blume and Starks *A Rationale for the Assessment of Errors in the Sampling of Soils*	John Keenan Taylor *Quality Assurance of Chemical Measurements*	Lawrence H. Keith, ed. *Principles of Environmental Sampling*
Trip Blank	A clean sample of matrix that is carried to the sampling site and transported to the laboratory for analysis without having been exposed to sampling procedures. (Defined in EPA QA/G-5, App. 1)	Used when volatile organics are sampled. Consist of actual sample containers filled with ASTM Type II water, kept with routine samples throughout sampling event, packaged for shipment with routine samples and sent with each shipping container to the laboratory. Used to determine the presence or absence of contamination during shipment.		A type of field blank as called, sampling media blank. To detect contamination associated with the sampling media such as filters, traps, and sample bottles. Consists of sampling media used for sample collection.

Table 2
QA Requirements for Programs

Potential Problems: / QC Samples to Identify Potential Problems:	Contamination		Calibration Drift	Bias		Imprecision		
	Blanks		Calibration Check Samples	Spike	Standard	Replicate	Collocated	Other
CLP Organics: 1991 Statement of Work, Exhibit E	Volatiles	A method blank once every 12 hours.	Continuing calibration standard every 12 hours. BFB analysis once every 12 hours.	Matrix spike with every case, batch, 20 samples, or 14 days.	3 system monitoring compounds added to every sample.	Matrix spike duplicate with every case, batch, 20 samples, or 14 days.		
	Semi-volatiles	A method blank with every batch.	DFTPP analysis once every 12 hours. Continuing calibration standard every 12 hours.	Matrix spike with every case, batch, 20 samples, or 14 days.	8 surrogates spiked into each sample.	Matrix spike duplicate with every case, batch, 20 samples, or 14 days.		
	Pesti-cides/ Aroclor	Instrument blank at start of analyses and every 12 hours. Method blank with each case, 14 days, or batch. Sulfur blanks are sometimes required.	Performance evaluation mixture to bracket 12 hour periods.	Matrix spike with every 20 samples.	2 surrogates added to each sample.	Matrix spike duplicate with every 20 samples.		

Table 2
QA Requirements for Programs

Potential Problems: / QC Samples to Identify Potential Problems:	Contamination	Calibration Drift	Bias		Imprecision		
	Blanks	Calibration Check Samples	Spike	Standard	Replicate	Collocated	Other
CLP Inorganics: 1991 Statement of Work, Exhibit E	Initial calibration blank; then continuing calibration blank 10% or every 2 hours. Preparation Blank with every batch.	Initial calibration verification standard; then continuing calibration verification 10% or every 2 hours.	1 spike for every batch. Method of standard additions for AA if spikes indicate problem.	Interference check sample for ICP 2 x /8 hours. Laboratory control sample with each batch.	1 duplicate/ batch. For AA, duplicate injections.		
PSD 40 CFR Part 58 Appendix B				For SO_2, NO_2, O_3, and CO, response check 1/ sampling quarter. For TSP and lead, sample flow check 1/sampling quarter. For lead, check with audit strips 1/quarter.		For TSP and lead, collocated sample 1/week or every 3rd day for continuous sampling.	For SO_2, NO_2, O_3, and CO, precision check once every 2 weeks.
SLAMS 40 CFR Part 58 Appendix A				For automated SO_2, NO_2, O_3, and CO response check for at least 1 analyzer (25% of all) each quarter. For manual SO_2 and NO_2, analyze audit standard solution each day samples are analyzed (at least 2x/quarter). For TSP, PM_{10}, and lead, sample flow rate check at least 1 analyzer/quarter (25% of all analyzers). For lead, check with audit strips 1/quarter.		For manual methods, including lead, collocated sample 1/week.	For auto-mated SO_2, NO_2, O_3, and CO precision check once every 2 weeks.

Table 2
QA Requirements for Programs

Potential Problems:	Contamination	Calibration Drift	Bias		Imprecision		
QC Samples to Identify Potential Problems:	Blanks	Calibration Check Samples	Spike	Standard	Replicate	Collocated	Other
A Rationale for the Assessment of Errors in the Sampling of Soils, by van Ee, Blume, and Starks	Preparation rinsate blanks and field rinsate blanks discussed, but no frequency given.			At least 21 pairs of field evaluation samples. At least 20 pairs of external laboratory evaluation samples if estimating components of variance is important.	At least 20 pairs or 10 triples of field duplicates. At least 20 pairs of preparation splits if estimating variance is important.		

Table 3
QC Requirements for Methods

Potential Problems:	Contamination	Calibration Drift	Bias		Imprecision		
QC Samples to Identify Potential Problems:	Blanks	Calibration Check Samples	Spike	Standard	Replicate	Collocated	Other
SW-846 Method 7000 (Proposed Update I) Atomic Absorption	Reagent blank as part of daily calibration.	Midrange standard analyzed every 10 samples.	One spiked matrix sample analyzed every 20 samples or analytical batch. Method of standard additions required for difficult matrices.		One replicate sample every 20 samples or analytical batch; one spiked replicate sample for each matrix type.		

Table 3
QC Requirements for Methods

Potential Problems:	Contamination	Calibration Drift	Bias		Imprecision		
QC Samples to Identify Potential Problems:	Blanks	Calibration Check Samples	Spike	Standard	Replicate	Collocated	Other
SW-846 Method 8000 (Proposed Update I) Gas Chromatography	Reagent blank before sample analysis and for each batch of up to 20 samples.	A daily calibration sample analyzed.	One matrix spike for each batch of up to 20 samples.	QC check sample required, but frequency not specified.	One replicate or matrix spike replicate for each analytical batch of up to 20 samples.		
503.1 Volatile Aromatic and Unsaturated Organic Compounds in Water by Purge and Trap GC (from PB89-220461)	Laboratory reagent blank with each batch. Field reagent blank with each set of field samples.	Calibration verified daily with 1 or more calibration standards.	Laboratory fortified blank with each batch or 20 samples.	Quality control sample analyzed at least quarterly.	Samples collected in duplicate. Laboratory fortified blanks analyzed in duplicate at least quarterly.		
200 Atomic Absorption Methods (from EPA-600-4-79-020)	Reagent blank at least daily.	Daily checks at least with reagent blank and 1 standard. Verification with an additional standard every 20 samples.		Analysis of an unknown performance sample at least once per year.			
624-Purgeables 40 CFR Part 136, Appendix A	Reagent water blank daily.	Analyze BFB every day analyses are performed.	Spike a minimum of 5% of samples.	Surrogate standards used with all samples. Analyze quality control check samples as 5% of analyses.			
1624-Volatile Organic Compounds by Isotope Dilution GC/MS 40 CFR Part 136, Appendix A	Blanks analyzed initially and with each sample lot.	Aqueous standard with BFB, internal standards, and pollutants is analyzed daily. A standard used to compare syringe injection with purge and trap.	All samples spiked with labeled compounds.		8 aliquots of the aqueous performance standard analyzed initially.		

Table 3
QC Requirements for Methods

Potential Problems:	Contamination	Calibration Drift	Bias		Imprecision		
QC Samples to Identify Potential Problems:	Blanks	Calibration Check Samples	Spike	Standard	Replicate	Collocated	Other
TCLP-Fed. Reg., Vol 55, No. 126 Friday, June 29, 1990	One blank for every 20 extractions.		One matrix spike for each waste type and for each batch.				
SW-846 Method 6010 (Proposed Update I) Inductively Coupled Plasma Atomic Emission Spectroscopy	At least one reagent blank with every sample batch.	Verify calibration every 10 samples and at the end of the analytical run with a blank and standard.	Spiked replicate samples analyzed at a frequency of 20%.	An interference check sample analyzed at the beginning and end of each run or 8-hour shift.	One replicate with every batch or 20 samples. Also spiked replicates analyzed, as discussed under "Spikes".		

Comparing Various QC Requirements

QC requirements for Program Offices

Table 2 shows that QC requirements vary considerably and are established by the Program Office responsible for the data collection activity. Ambient air monitoring methods (Office of Air Quality Planning and Standards) require periodic analysis of standards for assessment of accuracy (combination of imprecision and bias) and for manual methods, collocated samples for the assessment of imprecision. *Prevention of Significant Deterioration* (PSD) and *State and Local Air Monitoring Stations* (SLAMS) make a unique distinction in defining two terms: precision checks and accuracy checks. They entail essentially the same QC requirements, but are checked by different parties; the accuracy check is essentially an external audit while the precision check is an internal QC operation.) It should be noted that some water methods require additional QC operations for GC/MS than for other methods (e.g., tuning, Isotopic dilution).

In general, the wet chemistry analytical methods (TCLP being a preparation method) require periodic analysis of blanks and calibration standards. Most require analysis of matrix spikes and replicate samples, the exceptions being the 200 Series (no spikes or replicates) and the 600 series (GC/MS require no replicates).

While the QC operations for the PSD and SLAMS methods appear minimal, these monitoring programs require active QA programs which include procedures for zero/span checks. (The zero check may be considered a blank sample, while the span check may be considered a calibration check sample.)

The Program Office Quality Assurance Officer or representative should have details on specific QC requirements.

Organized by type of potential problem

Table 3 lists the QC requirements of various EPA measurement methods and presents the required frequencies for different kinds of QC operations. The table is divided into four sections, one for each general type of QC problem:

- *Contamination*: This occurs when the analyte of interest or an interferant is introduced through any of a number of sources, including contaminated sample equipment, containers, and reagents. The contaminant can be the analyte of interest or another chemical which interferes with the measurement of the analyte or causes loss or generation of the analyte.

- *Calibration Drift*: This is caused by changes in the measurement system over time, such as a (systematic) change in instrument response when challenged by a known standard.

- *Bias*: Can be regarded as a systematic error caused by contamination and calibration drift, and also by numerous other causes such as extraction efficiency by the solvent, matrix effect and losses during shipping/handling.

- *Imprecision*: This is a random error, observed as different results from repeated measures of the same or identical samples.

For internal consistency, the names of QC operations used in Table 3 are those given in the specific reference methods.

Using QC Data

The relationships between monitoring design specifications and the final use of the data described above incorporate two significant assumptions: (I) laboratory measurements, through use of internal standards or other adjustments that are integral to the analytical protocol, are unbiased; and (ii) the variance structure of these measurements does not change over time. Bias enters as a consequence of under-recovery of the contaminant of interest during the sample preparation stage of the analytical protocol, and as undetected drift in calibration parameters. The variance of measurements also may change over time due to unintentional changes in the way samples are prepared and degradation of electro-mechanical instrumentation used to analyze the samples. QC samples are intended to detect bias and variability changes and should be specified in the sampling plan.

QC samples that address bias are calibration check standards (CCSs) and spiked samples (performance check samples or PCSs). CCSs typically consist of reagent water samples spiked with the concentrations used to develop the calibration curve. Measurements obtained by analyzing these samples, which reflect the existing calibration relationship, are compared to the actual concentrations that were added to the samples. If the difference exceeds a pre-specified calibration test limit, the measurement system would be considered to be "out of control" and the calibration function would be re-estimated.

Detecting a change in calibration parameters is a statistical decision problem in detecting a material change in the calibration function. In many QC programs, CCSs typically are analyzed at the beginning and end of each shift, and after any other QC sample has detected a failure. By definition, significant change in the calibration parameters would lead to biased measurements of field samples and this can be detected through use of statistical tests.

The spiked sample is another type of QC sample used to detect bias. It typically has the same matrix characteristics found in field samples, but has been spiked (as soon after the sample is taken as is practical) with a known concentration of the target contaminant. Because spiked samples are intended to detect recovery changes, they are processed through the same preparation steps as field samples and the spiked sample measurement is used to form an estimate of recovery. Significant changes lead to the conclusion that measurements of field samples would be biased.

The second of the two monitoring program assumptions identified at the beginning of this section is a constant variance structure for monitoring data over time. Measurements from split (or duplicate) field samples provide a check on this variance assumption. Changes in measurement variability; for example a uniform increase in the standard deviation or changes in the way variability depends on concentration, would have a direct impact on subsequent investigations.

Classification of QC Samples: Control versus Assessment

QC programs are designed foremost to detect a measurement process entering an "out of control" state so corrective measures can be initiated. QC samples used in this way are performing a "control" function. Each of the three types of QC samples previously discussed, CCSs, spiked samples, and split (or duplicate) samples, may be used for control. In addition, spiked samples and split samples also may be used to estimate measurement bias and variability. QC samples that also can be used to estimate measurement parameters are sometimes referred to as quality assessment samples. This should not be confused with the much larger Data Quality Assessment Program; see also EPA QA/G-9, *Guidance for Data Quality Assessment.*

QC samples that are used for control must be analyzed and reported soon after they are obtained if their intervention potential is to be realized. Among the three types of QC samples discussed above, CCSs are the most likely to be effective for control purposes. Spiked samples and split samples generally would not be effective for control purposes, in part, because they are analyzed "blind" and therefore the results could not be reviewed immediately. Spiked samples and split samples, however, may be used for control if consecutive batches of similar field samples were to be analyzed.

Spiked samples and split samples can be effective quality assessment samples. For example, spiked samples may be used to estimate bias. The estimate would be applied as a bias correcting adjustment to individual measurements or to batches of measurements before the measurements are used in compliance tests. The adjustment would improve the test by eliminating bias. However, the variance of the adjusted estimate used in the test would be greater than the variance of the unadjusted estimate.

Split (or duplicate) samples also can be used as quality assessment samples, but their application in the monitoring program is not as constructive as the application of spiked samples. Split samples lead to an estimate of the measurement replication component of variability. (The variance of a measurement has, at a minimum, a sampling component and a measurement replication component, which is sometimes referred to as measurement error. If the sampling design involves stratification, the variance will include additional components.) If the estimate based on split samples suggests a measurement replication standard deviation larger than the value assumed in establishing the original sampling design, a loss in efficiency will result.

QC data collection and analysis does add cost to a monitoring program but is often not fully used for improving data collection activities.

References

American Society for Quality Control. Environmental Restoration Committee. Terms and Definitions Task Group. 1996. Definitions of Environmental Quality Assurance Terms. Milwaukee, WI: Quality Press.

Dux, James P. 1986. *Handbook of Quality Assurance for the Analytical Chemistry Laboratory.*

Federal Insecticide, Fungicide and Rodenticide Act (FIFRA). 1989. Good Laboratory Practices Standards. Final Rule. Federal Register, vol. 54, no. 158, August. [6]

Good Laboratory Practices: An Agrochemical Perspective. 1987. Division of Agrochemicals, 194th Meeting of the American Chemical Society.

Keith, Lawrence H., ed. 1988. *Principles of Environmental Sampling.* American Chemical Society.

Taylor, John Keenan. 1987. *Quality Assurance of Chemical Measurements.* Chelsea, MI: Lewis Publishers, Inc

van Ee, J. Jeffrey, Louis J. Blume and Thomas H. Starks. 1989. *A Rationale for the Assessment of Errors in Sampling of Soils.* EPA/600/X-89/203.

APPENDIX H

REPRESENTATIVENESS OF ENVIRONMENTAL DATA

H1. INTRODUCTION

This appendix discusses the concept of representativeness and is intended to help environmental scientists and engineers understand how representativeness relates to the development of Data Quality Objectives (DQOs) and Quality Assurance Project Plans (QAPPs). After introducing some basic terms and concepts, this appendix presents an overview of how representativeness is addressed in EPA regulations. Next, a review of a variety of scientific perspectives on the meaning of representativeness taken from the literature and ongoing work of consensus standards-setting bodies is provided. Finally, a conceptual model for defining and evaluating representativeness is presented. The conceptual model, called the Cycle of Representativeness, is general enough to apply to a broad variety of environmental studies.

What Is Representativeness?

Representativeness is one of the Data Quality Indicators (DQIs) (see also Appendix D), which are quantitative and qualitative descriptors used to determine whether or not data satisfy performance criteria specified in the QAPP. Representativeness is defined in *American National Standard: Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs* (ANSI/ASQC E4-1994) as follows:

> The measure of the degree to which data accurately and precisely represent a characteristic of a population, parameter variations at a sampling point, a process condition, or an environmental condition.

To determine whether or not data are representative, one must clarify the context and objectives of the study and consider many qualitative and quantitative factors throughout the planning, implementation, and assessment of data collection activities. One may conceive of representativeness as being applied at two scales—macroscopic and microscopic. In general, macroscopic issues deal with the following questions:

- How well does a sampled population represent the target population of interest?
- How well do the sampling units actually selected for measurement represent the sampled population?

Microscopic issues address these following questions:

- How well does a physical sample or specimen represent a sampling unit?
- How well does a data value represent a physical sample or specimen?

Why Does One Need to Consider Representativeness?

Representativeness arises as an issue because the population of interest is virtually always heterogeneous. Sampling and analysis of a heterogeneous population involves unavoidable errors that introduce bias and imprecision, which distort the picture of how the true environmental conditions fluctuate over space and time. If investigators fail to collect samples and obtain measurements that faithfully represent the target population, they may make dubious decisions based on an incorrect picture of the true state of nature.

From a very practical standpoint, it is important to consider representativeness because some EPA regulations require "representative samples" to be taken to support compliance monitoring. Regulatory perspectives are addressed in a later section.

H2. REGULATORY PERSPECTIVES

Many environmental regulations address the collection of "representative samples." However, representativeness is not used the same way in the different regulations. In fact, there is no universal definition of representativeness in the environmental regulations as presented in the Code of Federal Regulations (CFRs). Because there are often specific complex legal and procedural implications associated with collecting representative samples for the different regulatory statutes, this guidance recommends that investigators consult with relevant program officials and QA managers or coordinators to determine the applicable programmatic requirements for collecting representative samples. This section of the appendix discusses only some of the general issues to consider when collecting representative samples to support regulatory enforcement decisions.

Resource Conservation and Recovery Act (RCRA)

In 40 CFR 260.10, a representative sample is defined as "a sample of a universe or whole (e.g., waste pile, lagoon, ground water) which can be expected to exhibit the average properties of the universe or whole." In other words, representative samples are used to establish the hazardous characteristics of the waste. In some circumstances, specific methods that should be used for sample collection are detailed. For example, methods used to collect representative samples of certain types of waste are specified in 40 CFR 261 Appendix I. Other methods used to collect representative samples are prescribed in Chapter 9 of EPA Manual SW-846, *Test Methods for Evaluating Solid Waste Physical/Chemical Methods*. Investigators should note, however, that this chapter provides sampling guidance that has been deemed only advisory or not applicable in many enforcement cases.

Superfund Amendments and Reauthorization Act of 1986 (SARA)

The regulations for the Superfund program do not discuss the collection of representative samples, per se. However, there are numerous Superfund guidance documents that address technical and procedural sampling issues that affect representativeness. Investigators should follow general Agency and specific Office of Emergency and Remedial Response QA/QC requirements throughout the planning, implementation, and assessment of data.

Toxic Substances Control Act (TSCA)

There is no overall definition of representativeness in TSCA. However, the statute discusses specific instances in which investigators must consider representativeness when collecting samples. For example, in 40 CFR 763, "Asbestos," when investigators are collecting air samples, they should collect "a minimum of 5 samples per ambient air positioned at locations representative of the air entering the abatement site." Note that the regulation does not specifically state where representative locations might be.

In another example, when pumping and collecting ground water for anaerobic microbiological transformation rate data, "the pumping mechanism should be flushed with enough ground water to insure that a representative ground water sample is obtained" (40 CFR 766.16). Note that "representative" is not defined specifically, but the intent of using the term is to ensure that the ground water sample is not biased from having been in the pump mechanism for a long time.

Air Programs

No definition of what is a representative sample is provided in the CFRs for Air Programs. However, there are a few cases where representative sampling is required. For example, representativeness is used to describe the type of data for modeling and determining where to site monitoring stations (e.g., meteorological data "used as input to a dispersion model should be selected on the basis of spatial and climatological (temporal) representativeness. . ." [40 CFR 51 Appendix W]).

Water Programs

In general, the regulations for water programs use representativeness in the context of permitting. For example, when monitoring an outfall in certain cases, "samples should be representative of daily operations" (40 CFR 403.12 (b)). To demonstrate continued compliance, investigators will collect samples that are "representative of conditions" (40 CFR 403.12 (g)). Note that the regulation does not define what is "representative," leaving investigators to determine what sampling design will allow them to collect samples that are representative.

In some cases, the regulations detail specific instances when a certain method should be used to ensure representative samples are collected. For example, when sampling effluent under the National Pollutant Discharge Elimination System (NPDES), if the use of an autosampler is infeasible, then four grab samples are defined to be a "representative sample of the effluent being discharged" (40 CFR 122.21).

Summary and Recommendations Regarding Regulatory Issues

To be defensible, it is a necessary condition that sampling always be correct from a scientific and statistical standpoint. However, technical correctness may not be a sufficient condition where procedural requirements for a particular program must be followed to ensure legal defensibility. The investigator should, at a minimum, consult with the QA manager or coordinator for the applicable program to ensure that sampling and measurement protocols are being selected and addressed in the QAPP in a way that is consistent with relevant regulations, policies, and guidelines, including QA/QC requirements.

H3. SCIENTIFIC PERSPECTIVES

Because representativeness does not have a single unambiguous definition in the scientific and statistical literature, it is useful to consider a variety of perspectives on what representativeness means from a technical standpoint.

H3.a Kruskal and Mosteller's Papers on Representative Sampling

Kruskal and Mosteller (1979), two eminent statisticians, presented a series of three papers in which they examined how "representativeness" was misused in scientific, statistical, and everyday writing, with the intention of clarifying the technical meaning of the term. The following discussion is summarized from this series of papers.

They note that some of the confusion in how the term is used arises because "representative sampling" does not have a standard definition and is often used differently in various contexts. They present the nine ways in which the term is commonly applied:

1. as a "seal of approval,"

2. to denote the "absence of selective forces,"
3. as a "miniature or small replica of the population,"
4. as being a "typical or ideal case,"
5. to denote "coverage of a population,"
6. as a "vague term to be made more precise,"
7. as a "specific sampling method,"
8. as "permitting good estimation," and
9. as "good enough for a particular purpose."

Seal of approval. In the first case, writers use the term "representative sample" to give credence or an undeserved "seal of approval" to their work, for example:

> ". . .private and municipal museums are, if my sampling has been representative, a little better than all but the most prestigious state museums." [Douglas J. Stewart, "Two cheers for the tombaroli," *The New Republic*, 28 April 1973, p.21]

> "Fifteen samples of consumer spackling and patching compounds were purchased at hardware stores in the New York City area . . . (O)ur analysis of [the] fifteen representative samples . . . has shown that five contained appreciable amounts of . . . asbestos minerals." [A.N. Rohl, A.M. Langer, I.J. Selikoff, and W.J. Nicholson, "Exposure to asbestos in the use of consumer spackling, patching, and taping compounds," *Science*, 15 August 1975, pp. 551, 553.]

In both examples, the authors have not explained what processes took them from the target population to the actual sampled population. Rather, the term "representative" was used to convince the reader to have faith in the methods that were used and, by doing so, convince the reader of the truthfulness of the author's conclusions, which were based on the study's results.

Absence of selective forces. In the second case, representativeness is used to mean that the sampling method excluded selective forces that might over-represent some portion of the population. However, the principal flaw of this concept of representativeness is that unless a probability-based survey design is being used, investigators cannot be sure that they have eliminated selective forces. Kruskal and Mosteller present the example of the *Literary Digest* election poll that predicted incorrectly the winner of the 1936 presidential race because the magazine's inference was based on a "representative sample" that actually was "a sample that over-represented Republican voters, who were at that time far more likely to respond to the *Digest*'s poll than Democratic voters."

Miniature replica of the population. In the third case, a representative sample is used to refer to a miniature of the population. However, this concept is flawed for several reasons. First of all, the notion that a representative set of samples forms a miniature version of the population implies that individual units within a class are identical, and that the various classes are perfectly mixed throughout the population so that the samples exhibit the same relative frequency distribution as the population. As a practical matter, one rarely knows what the true population frequency distribution is like; therefore, one is unable to evaluate how close the set of samples are to achieving the goal of a miniature replica.

Typical or ideal case. In the fourth case, a representative sample is intended to describe either a typical or ideal sample. A problem with this definition is that it implies that one specimen was collected, which does not indicate that the sample was collected using a probability-based design. Furthermore, the term "ideal" often implies that a superlative (e.g., "best," "worst,"or "perfect") specimen was selected from the population. Kruskal and Mosteller illustrate this point with the example of Emerson's book, *Representative Men*, which contains essays on men such as Plato, Goethe, and Napoleon. In this case,

"representative" is used to connote some ideal type, which Emerson judged that each man "represents," such as the philosopher, the writer, and the "man of the world."

Coverage of the population. In the fifth case, representativeness is used to mean that a sample has wide coverage. In this sense, representative samples are supposed to include a sample of at least one member from each class of a relevant partition of the population, but such sampling would not give investigators an indication of the proportions or relative frequencies in each class and lead to biased estimates.

Vague term to be made more precise. In the sixth case, representativeness is used vaguely to describe a sampling scheme, and it is not readily apparent whether or not investigators are using a sampling design that will produce statistically representative samples. For example, Kruskal and Mosteller cite one study where the investigators wrote of planning "nationally representative, internationally comparable, scientifically designed and conducted sample surveys." From the outset of this article, the reader would not be certain whether the author was misusing the term in one of the many ways presented previously or whether the author truly had collected representative samples.

Specific sampling method. This involves the use of "representative sampling" in place of "probabalistic sampling," "random sampling," "stratified sampling," "quota sampling," or "purposive sampling." However, because of the variety of meanings attributed to "representative sampling," Kruskal and Mosteller suggest that writers clarify the exact statistical sampling plan to be used, as the properties of the derived estimates vary substantially.

Permitting good estimation. In this case, representative sampling is used to imply a satisfactory estimation of population characteristics without defining "good." Kruskal and Mosteller suggest that the "virtue" of sampling is better described in "terms of little or no bias, in terms of low sampling error, or in yet other terms."

Good enough for a particular purpose. Representativeness is sometimes misused by authors in the literature to mean "good enough for our purposes." In this case, data are representative if they help prove or disprove an investigator's assertion. Kruskal and Mosteller illustrate this situation with the following example: "if the physicians thought that all patients with a particular kind of burn developed a particular symptom, but a sample showed that a number did not, that [the sample] would be good enough to settle the particular issue."

In conclusion, Kruskal and Mosteller recommended that one should "avoid the term in statistical and other scientific writing, just as one tries to avoid praising the accuracy of results of unknown precision." Clearly, the term "representativeness" can be misused unless properly defined.

H3.b Gy's Theory of Sampling

One of the more important scientific perspectives on representativeness is provided by Pierre Gy, a French mining engineer who developed a comprehensive theory of sampling of particulate materials. Although developed to aid in the proper sampling and estimation of ore content for the mining industry, the concepts and techniques are applicable to a wide range of environmental problems. This section provides an overview of Gy's theory of sampling based on the work of Francis Pitard, a colleague of Gy who has written a comprehensive text on the subject in English (Pitard 1992).

Representativeness is defined unambiguously as the quality of an estimate that has acceptable bias and precision, expressed as the mean squared error of an estimate in relation to the true parameter value. As such, this use of the term "representative" falls squarely into Kruskal and Mosteller's classification scheme as "permitting good estimation." One of the most important contributions of Gy's theory is that it provides a theoretically sound yet practical basis for determining the amount of material to be taken when sampling, and the methods by which the sampling should be done, to provide reliable estimates of average conditions within the population (or subpopulation) of interest.

In the development that follows, the key concepts underlying Gy's theory of sampling are explained in the context of environmental sampling. This section starts with some observations regarding how Gy's theory and practice fits within the field of environmental sampling. Second, Gy's classification of types of sampling problems is explained, which sets the stage for a discussion of the central concept of heterogeneity and its different types. Finally, Gy's classification of the types of errors that arise when sampling heterogeneous populations is addressed and followed by a discussion of the notion of correct sampling.

Importance of Gy's theory for environmental sampling. Gy's theory of sampling is important for the field of environmental sampling for several reasons. First, Gy's theory picks up where traditional statistical sampling theory leaves off: obtaining measurements of sampling units. Although statistical sampling theory provides many approaches for how to select sampling units from a population to support valid inferences, Gy's theory provides a comprehensive and systematic approach for how to properly obtain measurements of the sampling units while respecting the very same principle of equiprobable selection that underlies most traditional statistical sampling designs. Consequently, Gy's theory provides a basis for linking microscopic sampling protocol design issues (such as the quantity of sample support) with macroscopic sampling design issues (such as how many samples to take) so that the overall sampling design is more fully integrated throughout all stages. This approach also provides one of the key links for ensuring that the protocols and methods specified in the QAPP are consistent with and based upon the study's Data Quality Objectives.

Another key feature of Gy's theory of sampling is that it provides a systematic approach for minimizing bias and variation due to field sampling activities. Field sampling traditionally has been the greatest challenge for QA/QC, due to the difficulty of controlling the sampling process under the great variety of conditions encountered in the field. The theoretical and practical aspects of Gy's theory inherently reduce variation caused by inadequate sampling practices and minimize the chance that the sampling process will over- or under-select parts of the population, which may lead to undetected bias in the results.

Classification of sampling lots. At its core, Gy's theory of sampling is consistent with classical statistical principles that rely on a random sampling process whereby each member (or unit) of a population has an equal probability of being selected. It is useful, then, to begin by considering the types of random sampling processes one may encounter, which are described as "sampling lots." Gy's theory classifies the types of sampling problems by considering the number of dimensions presented by the problem from a statistical estimation standpoint.

Zero dimensional lot: a set of population units where the order in which units are selected is unimportant; this is a population in the sense of randomly selecting different entities and counting the total of each characteristic observed without regard to the order in which the entities were selected.

One dimensional lot: a set of population units where the order in which units are selected is very important and which yields an ordered set of samples identified with time or position such as a time series.

Two dimensional lot: a set of population units where each unit is selected from a two-dimensional domain, such as a mapped surface having latitude and longitude coordinates.

Three dimensional lot: a set of population units where each unit is selected from a three-dimensional domain, such as a mapped volume having latitude, longitude, and elevation coordinates (or height, width, and depth).

Real-life environmental problems usually are four-dimensional, in the sense that pollutants are distributed in three dimensions, and the distribution changes over the fourth dimension of time. However, these problems often can be reduced to fewer dimensions through simplifying assumptions or decomposition. For example, a three-dimensional problem can be transformed into a series of two-dimensional problems if it is possible to consider "slicing" the three-dimensional space into two or more "slabs" of appropriate thickness, then investigating each slab as a two-dimensional problem. Likewise, a two-dimensional problem can be decomposed into a number of one-dimensional problems by considering one-dimensional transect lines that run through the two-dimensional space, either orthogonally, in parallel, or radially about a point, as appropriate to the problem. Even a one-dimensional problem requires a simplifying assumption when applying the principle in the real world, in that a sample taken at a given "slice" in time or location must be taken in a manner such that the "slice" is considered as a zero-dimensional lot.

These transformations to lower dimensions are important tools of analysis because of the practical difficulties and theoretical complexities of sampling two- and three-dimensional lots. To obtain samples one must identify a module of observation that has a shape, size, and orientation that is appropriate for that type of sampling problem. The appropriate module of observation for a three-dimensional lot is a sphere or cube. However, it is almost always practically impossible to obtain a spherical sample from a volume of real material. Fortunately, one can decompose the three-dimensional lot into one or more two-dimensional slabs. The correct module of observation for a two-dimensional lot is a cylinder with a circular cross section, extending through the entire thickness of the slab. This can be achieved in practice using coring devices.

Understanding heterogeneity. If all the units that make up a population are exactly alike in their characteristics, then the population is said to be homogeneous. However, this is an ideal condition that is almost never encountered in the real world. In virtually all environmental problems, the units of a population differ in ways that are relevant to sampling, analysis, and estimation; therefore, the population is said to be heterogeneous. This quality of heterogeneity is observed as variability in measurement values from one location to the next in time and/or space. Intuitively, then, the notion of heterogeneity is central to the concept of representativeness because the greater the heterogeneity of the population, the more difficult it is to define and achieve a "representative sample."

In Gy's theory of sampling, two general types of heterogeneity are defined for all populations, and three categories of heterogeneity are defined particularly for one-dimensional lots. The two general types of heterogeneity are constitution heterogeneity and distribution heterogeneity:

constitution heterogeneity: this is the heterogeneity that is inherent to the composition of the population, in the sense that it is a measure of how characteristics vary from one unit of the population to the next. If we were interested in contamination at several sites, then a

site with chromium, arsenic, and lead would have more Constitution Heterogeneity (CH) than a site with only chromium and arsenic; another site that contained pure contaminants and a number of their combinations would have the greatest CH of the three sites.

distribution heterogeneity: this is the heterogeneity that is due to the manner in which the units of the population are distributed over space or time, in the sense that different types of units may be evenly mixed in space or time, versus the condition where similar or identical units are clustered together. Distribution heterogeneity is usually caused by physical forces or chemical reactions acting on the geometry, density, size, composition, and other qualities of the population units. In the site example, if all the chromium was at the 2-feet depth and all of the arsenic was at the surface, then that site would have greater Distribution Heterogeneity (DH) than a site where both contaminants were randomly mixed throughout.

Specific to sampling from a one-dimensional lot, such as a waste stream flowing in a pipe, Gy's theory classifies heterogeneity into three types:

short-range: this type of heterogeneity represents random fluctuations over small distances or time intervals. These fluctuations are described in terms of constitution heterogeneity and distribution heterogeneity, as discussed above.

long-range: this type of heterogeneity represents non-random fluctuations over larger distances that can be attributed to trends caused by human activities or natural processes.

periodic: this type of heterogeneity represents cyclic fluctuations that may be caused by human activities, daily or seasonal variations, or other natural processes.

The heterogeneity of a one-dimensional lot is usually studied by constructing a variogram, which is a tool from the field of geostatistics that measures the amount of heterogeneity or variation as a function of distance between population units in time or space. A variogram is related to a correlogram which measures serial correlation; see 2.3.8.2 of EPA QA/G-9 *Data Quality Assessment* for further discussion. Usually, the closer in time or space two units are, the more alike they will be; therefore, the measure of heterogeneity or variance will be smaller. As the separation distance becomes zero (which might represent co-located samples), the amount of heterogeneity observed is due to the short-range CH and DH (geostatisticians sometimes call this the "nugget effect," which represents the amount of heterogeneity at the y-intercept of the variogram). As separation distance increases, the heterogeneity (variance) increases until leveling off at some maximum value. The distance at which the maximum heterogeneity is reached is called the *range* of the variogram. Samples that are separated by a distance at least as large as the variogram's range usually are considered to be statistically independent.

The heterogeneity of two- and three-dimensional lots has been studied extensively within the field of geostatistics. However, by transforming the three-dimensional problem into sampling of one or more two-dimensional lots, practical geostatistical sampling programs can be developed.

Classification of errors. Gy's theory of sampling presents a classification of seven types of errors that account for the different types of heterogeneity encountered when sampling from zero- and one-dimensional lots. The total error is the sum of the seven types of errors. The last four types of errors are due to the selection processes involved in choosing which population units will be characterized; the first three types of errors are due to practical imperfections in the implementation of the selection scheme:

1. *Preparation error:* this error is caused by contamination, loss, or transformation of material, or by human blunders, so that the material that is analyzed or measured no longer reflects the true characteristics or constituents that were originally obtained in the sample. Many quality control protocols are intended to minimize this type of error.

2. *Increment extraction error:* this is a materialization error that results from imperfect collection of material defined by the sample increment delimitation. This error also can be caused by incorrect choice or use of sampling equipment which by its very application alters the physical characteristics of the sample.

3. *Delimitation error:* this is an implementation or "materialization" error that results from failure to use the correct type of sampling device to obtain material that will make up the sample. This error occurs when the module of observation is incorrectly defined ("delimited") in terms of its shape and orientation; hence, the material obtained in the sample does not respect the condition of equiprobable selection for that type of sampling lot or problem.

4. *Periodic fluctuation error:* a non-random selection error due to cyclic variations over intervals of space or time. Often the investigator is interested in adjusting for or "canceling out" the periodic fluctuations so that other long-term trends can be detected more clearly or confidently.

5. *Long-range fluctuation error:* a non-random selection error due to trends or other systematic variations over larger distances or time intervals. Often this long-range fluctuation is what an investigator is trying to understand through modeling the processes that describe pollution transport and fate in the environment.

6. *Grouping and segregation error (GE):* a short-range selection error that is due to the distribution heterogeneity of the population. The grouping and segregation error cannot be larger than the fundamental error and will depend on the size and configuration of potential groupings of particles or units of the population. The grouping and segregation error is reduced as the heterogenous population units become more well-mixed or as the number of increments[1] making up a sample is increased.

7. *Fundamental error (FE):* another short-range random selection error that is due to the constitution heterogeneity of the population. The fundamental error represents the theoretical lower bound on the total error and is a function of the quantity of material used to make up a sample (sometimes called *sample support*), as well as the maximum particle size for soils and other particulate materials. In general, the fundamental error is reduced as the sample support increases or the maximum particle size decreases.

In Gy's system, the error attributable to the bias and imprecision of the analytical instrument or measurement device is not considered as part of the sampling theory but is acknowledged as part of the overall estimation error (overall estimation error corresponds to decision error from Data Quality Objectives, which applies when estimating characteristics of a population). Figure H-1 shows Gy's categorization of errors in a tree diagram. Note that all of the types of errors described in Gy's theory are relevant to the preparation steps of many laboratory analytical method protocols (even though they are carried out at a smaller scale than field sampling); hence, Gy's classification differs from other schemes that group errors introduced through analytical method preparation steps (Gy's preparation errors) with

[1] In Gy's theory, a sample is made up of one or more *increments* that are combined to form a physical sample. This is analogous to sample compositing, but at a smaller scale.

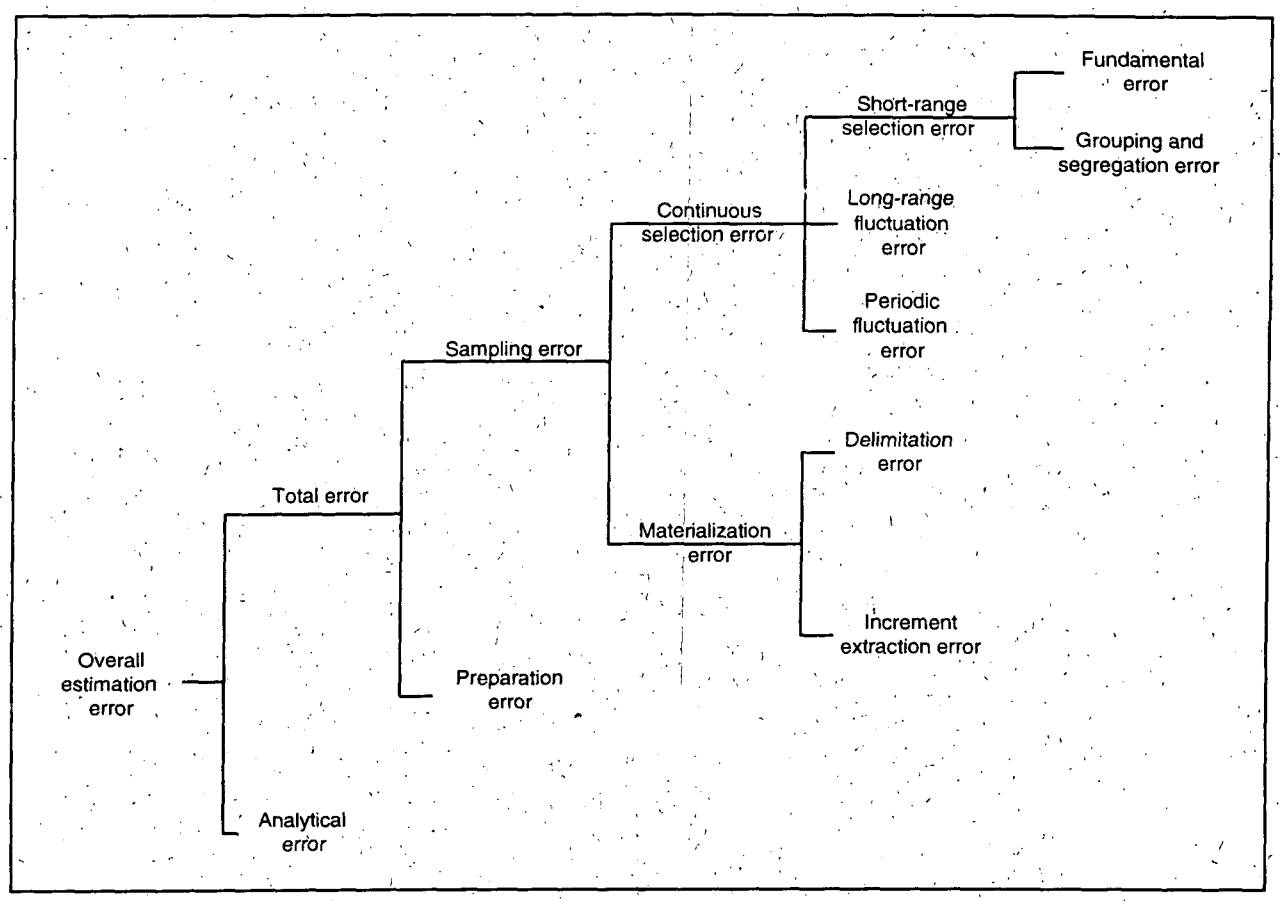


Figure H-1. Gy's classification of errors

analytical instrument errors. This point indicates that the full life cycle of obtaining samples in the field, through implementation of analytical methods in the lab, involves multiple stages or iterations of sample selection, delimitation, extraction, and preparation, usually at increasingly smaller scales.

Correct sampling. The fundamental notion in Gy's theory of sampling is that there is a "correct" approach to sampling that respects the principles of equiprobable selection in the context of obtaining samples from heterogeneous populations. When a sampling protocol adheres to the principles of correct sampling, then the results should fall within the pre-specified goals of precision and bias in a repeatable manner. Bias, in particular, is difficult to detect even with a fully operational quality system in place and can have devastating effects on the accuracy of conclusions drawn from analysis of erroneous data. Only correct sampling significantly reduces the chance of biased results.

The principles of correct sampling follow directly from the concepts of heterogeneity discussed previously, and the notion of minimizing the components of total error. The essence of correct sampling practice is in planning and implementing protocols that respect the principle of equiprobable selection given the nature of the heterogeneity encountered in the target population. To accomplish this, one must address the following issues:

- Define the sampling problem correctly in terms of a zero-, one-, or two-dimensional lot. Consider the assumptions or practical issues that must be addressed in simplifying the problem from higher dimensions to tractable zero-, one-, or two-dimensional problems.

- Conduct a preliminary study of the target population to determine the nature of the

heterogeneity relevant to the sampling problem. In all cases, an understanding of the short-range constitution heterogeneity and distribution heterogeneity will provide information on the quantity of material needed to optimize the amount of sample support (see Pitard 1992a, Chapter 11). Depending on the nature of the sampling problem, information about variography and periodic fluctuations may be important for optimizing the sampling protocol.

- Ensure that the sample increment is correctly delimited by selecting the correct module of observation, given how the distribution heterogeneity occurs throughout the dimensions of the sampling lot. For one-dimensional lots, this involves designing the sampling protocol to account for the geometry and dynamics of the material flow (see Pitard 1992b, Chapter 14).

- Ensure that the sample increment material is correctly extracted by selecting and using sampling tools and equipment so that the correctly delimited increments are actually available for analysis. This involves designing the sampling protocol to account for the geometry and physics of the increment extraction process (see Pitard 1992b, Chapter 15).

- Ensure that the sample maintains its integrity by specifying and implementing sample preparation protocols that minimize the chance for material loss, contamination, or transformations.

Conclusion

Gy's theory is not without tradeoffs and areas requiring further research. Because the theory was developed for the mining industry, the detailed procedures work for the sampling of particulate materials but are not as well defined for other media. Another drawback of the theory is that it requires an up-front investment in a pilot investigation of the heterogeneity of the population. However, this investigation is usually a sound investment that pays significant dividends not only in a more efficient sampling design but also in a more thorough understanding of the nature of the environmental problem. Nonetheless, the requirements of Gy's theory are well suited for iterative investigation strategies that have gained favor in recent years, which often incorporate early pilot studies.

In general, Gy's theory of sampling provides an important scientific perspective on representativeness by linking the statistical concept of equiprobable selection of population units to the practical issues of sample collection and measurement. The practice of environmental sampling can be improved dramatically by respecting the notion of correct sample increment delimitation and extraction, and minimizing fundamental error by taking multiple sample increments. Future research is targeted toward extending Gy's theory to non-particulate media and the design of better sampling tools and equipment.

H4. THE CYCLE OF REPRESENTATIVENESS

This section describes a conceptual model for defining and evaluating representativeness within the context of an environmental study. This model is based on a framework for the planning, implementation, and assessment of data collection activities that is often referred to as the data life cycle.

H4.a. Probabilistically Based Sampling

When data are to be collected using a probability-based sampling design, the different components of representativeness also can be illustrated by a cycle (see Figure H-2). There are five stages in the cycle of representativeness:

- Defining the objectives of the study, which help define the target population,
- Identifying practical constraints and key assumptions, which help specify the sampled population (population of inference),
- Developing and optimizing a sampling and measurement design,
- Implementing the design and obtaining measurements (observations), and
- Conducting statistical analysis of the data.

In Figure H-2, the hexagonal boxes express a hierarchy of entities that represent something of interest to the data user, starting at the highest level of the target population and working down to the level of data. The rectangular boxes are processes (such as developing a statistical sampling design or actually collecting samples) that help investigators get from one entity to the next. Solid arrows direct investigators from one stage to the next in the cycle. The return arrows (dashed lines) are part of the Data Quality Assessment (DQA) Process and help investigators determine the degree to which their data are representative of sampling units, the sampled population, and the target population.

H4.b. Judgementally Based Sampling

Sometimes investigators may find that the study objectives call for a judgmental sampling approach, which does not employ a probabilistic scheme for selecting sampling units. Whenever judgmental sampling is used, investigators will be limited in their ability to describe and defend the representativeness of the data. Usually, investigators will only be able to draw defensible inferences about the sampling units, assuming that a valid measurement protocol was implemented correctly. Any extrapolation from the data to characteristics of the target population will be based on professional judgment rather than reproducible statistical inference; the extrapolation becomes vulnerable to challenge, depending on the credibility of the investigator. Figure H-3 illustrates this point by showing how judgmental sampling "short-circuits" the cycle of representativeness.

H4.c. A Discussion of the Cycle of Representativeness

This section discusses each stage of the cycle of representativeness as represented in Figure H-2.

Defining the Study Objectives and Target Population

In this step, investigators develop a clear statement of the study objectives in terms of what the data user really would like to know to inform the decision at hand or the study question under consideration.

As part of this definition of the *target population*, the investigator must define an elementary module of observation that will serve as the basic conceptual units making up the population. That is, the target population can be viewed as the sum or union of all elementary *population units* that one can (at least in theory) select for observation.

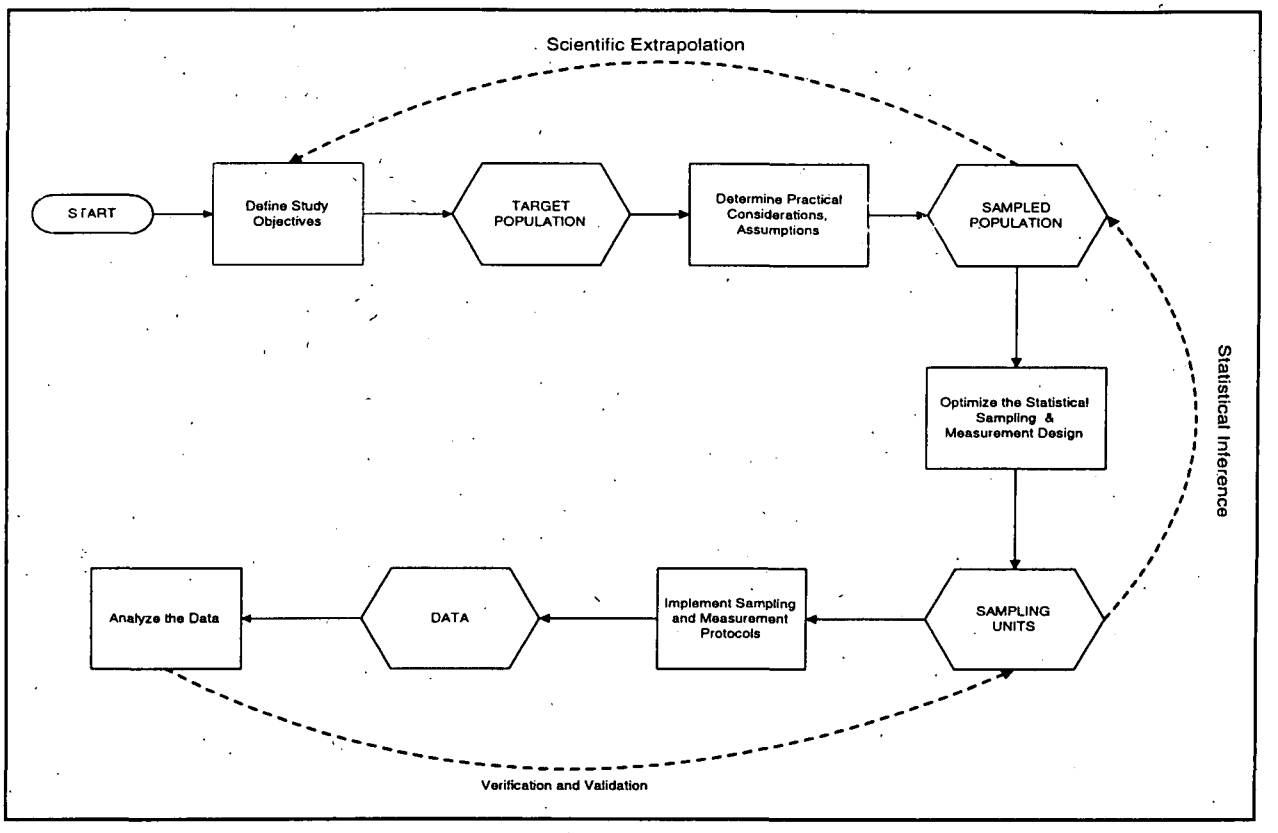


Figure H-2. Cycle of Representativeness

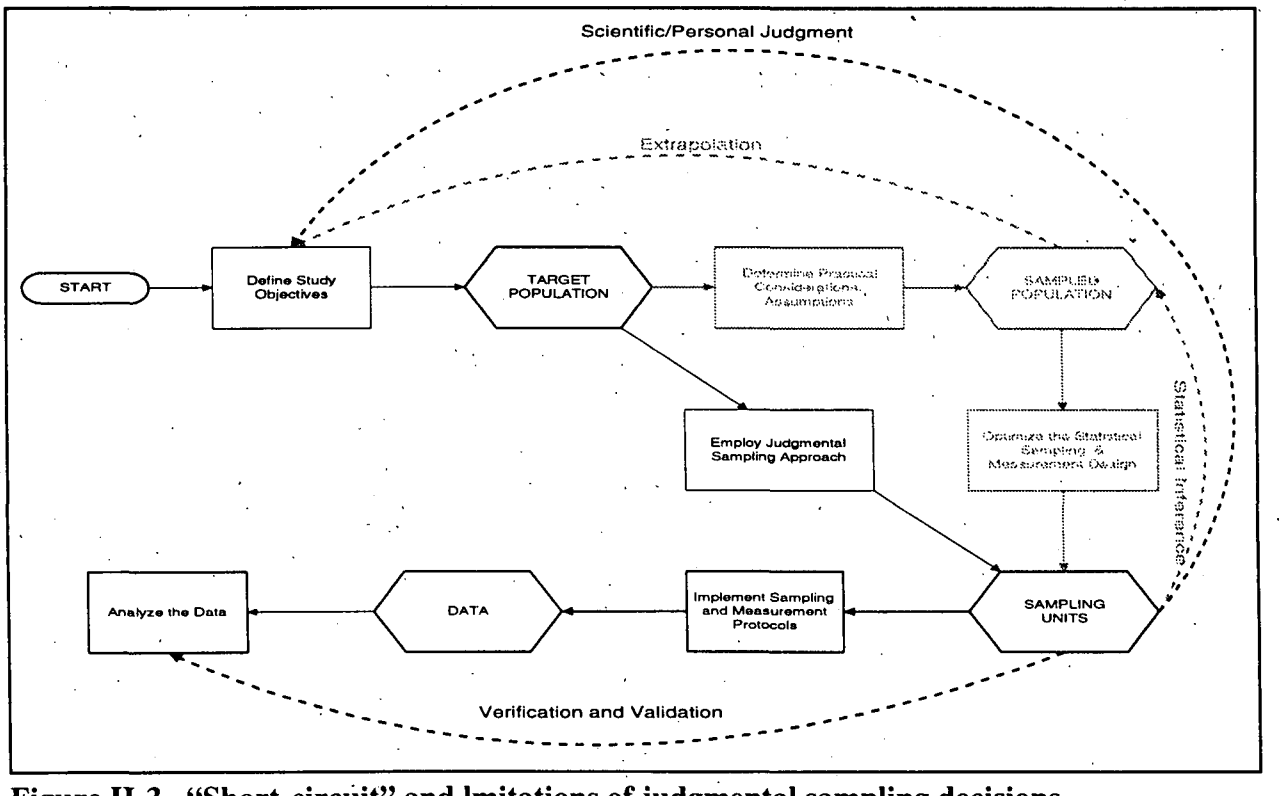


Figure H-3. "Short-circuit" and limitations of judgmental sampling decisions

EPA recommends that investigators use a systematic planning method to develop study objectives and define the target population. EPA developed the Data Quality Objectives (DQO) Process to help investigators clearly define qualitative and quantitative performance criteria for environmental data, which facilitates the development of sampling designs that satisfy study objectives (EPA 1994; see also QAPP element A7 for a discussion of DQOs). The DQO Process addresses the cycle of representativeness by helping to clarify study objectives and define the target population.

Determining Practical Constraints, Assumptions, and Specifying the Sampled Population

Once the target population has been defined in theory, the investigator must consider the practical constraints and requirements of sampling, and determine whether the entire set of population units making up the target population will be available to select and measure or observe.

After considering these assumptions and practical constraints, the investigator is in a position to define the *sampled population*, which is simply that subset (or in some cases a surrogate) of the target population that actually will be available to the sampling and measurement process. The sampled population also may be called the population of inference.

Optimizing the Sampling and Measurement Design

In this activity, investigators develop a statistical sampling/measurement design that will (a) define the process for selecting sampling units from the sampled population, and (b) define a measurement protocol for obtaining measurement values or observations from the sampling units. The process for selecting sampling units uses probability-based sampling designs (see also EPA QA/G-5S, *Sampling Designs to Support QAPPs*).

Probability-based sampling designs help ensure that the selection of sampling units will lead to valid and defensible inferences about the sampled population. As shown in Figure H-3, when a non-probability-based sampling approach (i.e., judgmental approach) is used, then the cycle of representativeness is "short circuited," and a much greater leap of faith in scientific judgment is required to link the sampling units to conclusions about the target population.

Implementing the Sampling and Measurement Design

This stage involves the implementation of the sample collection and measurement protocols to produce data. Quality control protocols are critical here to ensure that samples are obtained correctly and their integrity maintained. Quality assurance is important throughout for establishing and documenting the procedures used, anomalies encountered, and corrective actions taken, so as to establish a chain of defensibility that will allow evaluation of data validity.

Analyzing the Data

In this activity, investigators determine how well their data represent sampling units, the sampled population, and (through extrapolation) the target population. This stage is where the cycle of representativeness moves back upward in the hierarchy of entities, allowing the investigator to use the data to draw conclusions about the target population. The three key stages are described in the following subsections.

Determine how well the data represent the sampling units. This stage concerns the issues of data verification and validation (see also Appendix F), and the determination of whether the measurement protocols were implemented properly. Additionally, the underlying assumptions of the measurement protocol are evaluated using the routine QC data and other information.

Determine how well the sampling units represent the sampled population. Using the statistical techniques of Data Quality Assessment (EPA QA/G-9), the investigator determines if the underlying assumptions of the DQOs and the sampling design were satisfied, and uses the tools of statistical inference to draw conclusions about the sampled population. The strength of these conclusions can be quantified in terms of confidence or probability intervals for estimates, and probabilities of decision errors for hypothesis tests. If a judgmental sampling approach was used, this stage is short-circuited, and quantitative statements about the strength of conclusions are extremely difficult to defend.

Determine how well the sampling data represent the target population. This stage is the domain of scientific extrapolation, where the investigator determines the extent to which the study results for the sampled population can be extrapolated to the target population. These extrapolations are based on an evaluation of how strongly the study results support or verify the assumptions linking the sampled population to the target population.

H5. CONCLUSIONS

Representativeness is a quality that must be addressed primarily through scientific and statistical evaluation. Ultimately it is the scientific/statistical perspective on representativeness that bears most directly on the quality of risk management decision making, which lies at the heart of virtually all environmental laws and regulations.

At the time of the publication of this appendix, an American Society for Testing and Materials technical subcommittee (D34.02) is developing a standard guide for "Representative Sampling for Management of Waste and Contaminated Media," which will apply primarily to investigations of waste at RCRA and CERCLA facilities.

In closing, it is instructive to cite a definition of a "representative sample" from the International Statistical Institute's *A Dictionary of Statistical Terms* (Marriott 1990):

> **representative sample** In the widest sense, a sample which is representative of a population. Some confusion arises according to whether 'representative' is regarded as meaning 'selected by some process which gives all samples an equal chance of appearing to represent the population'; or, alternatively, whether it means 'typical in respect of certain characteristics, however chosen'.
>
> On the whole, it seems best to confine the word 'representative' to samples which turn out to be so, however chosen, rather than apply it to those chosen with the object of being representative.

On the whole, then, it seems best to admit that the word 'representative' is not easily defined with the clarity and brevity usually sought in scientific or statistical subject matter.

H6. REFERENCES

American Society for Quality Control. 1994. *American National Standard: Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs* (ANSI/ASQC E4-1994). Milwaukee, WI: American Society for Quality Control.

Englund, E. J. 1987. Spatial autocorrelation: Implications for sampling and estimation. In *ASA/EPA Conferences on Interpretation of Environmental Data: III Sampling and Site Selection in Environmental Studies* (14-15 May), 31-39. EPA/230-08-88-035.

U.S. Environmental Protection Agency. 1994. *Guidance for the Data Quality Objectives Process.* EPA QA/G-4.

Gagner, S. D., and A. B. Crockett. 1996. Compositing and subsampling of media related to waste management activities. In *The Twelfth Annual Waste Testing and Quality Assurance Symposium*, 22-29. American Chemical Society and U. S. Environmental Protection Agency, Washington, DC.

Gilbert, R. O. 1987. *Statistical Methods for Environmental Pollution Monitoring.* New York: Van Nostrand Reinholt Co.

Jenkins, T. F., C. L. Grant, G. S. Brar, P. G. Thorne, P. W. Schumacher, and T. A. Ranney. 1996. Sample representativeness: a necessary element in explosives site characterization. In *The Twelfth Annual Waste Testing and Quality Assurance Symposium*, 30-36. American Chemical Society and U. S. Environmental Protection Agency, Washington, DC.

Kruskal, W., and F. Mosteller. 1979. Representative sampling, I: Non-scientific literature. *International Statistical Review* 47, 13-24.

Kruskal, W., and F. Mosteller. 1979. Representative sampling, II: Scientific literature, excluding statistics. *International Statistical Review* 47, 111-127.

Kruskal, W., and F. Mosteller. 1979. Representative sampling, III: The current statistical literature. *International Statistical Review* 47, 245-265.

Maney, J. P. 1996. Sampling strategies: impact of heterogeneity on waste characterization. In *The Twelfth Annual Waste Testing and Quality Assurance Symposium*, 1-14. American Chemical Society and U. S. Environmental Protection Agency, Washington, DC.

Marriott, F.H.C. 1990. *A Dictionary of Statistical Terms.* 5th ed. New York: John Wiley.

Pitard, Francis F. 1992a. *Pierre Gy's Sampling Theory and Sampling Practice: Volume I, Heterogeneity and Sampling.* Boca Raton, FL: CRC Press.

Pitard, Francis F. 1992b. *Pierre Gy's Sampling Theory and Sampling Practice: Volume II, Sampling Correctness and Sampling Practice.* Boca Raton, FL: CRC Press.

Splitstone, D. E. 1987. Sampling design: some very practical considerations. In *ASA/EPA Conferences on Interpretation of Environmental Data: III Sampling and Site Selection in Environmental Studies* (14-15 May), 15-22. EPA/230-08-88-035.

APPENDIX I

ADDITIONAL TERMS AND DEFINITIONS

Acceptance criteria - specified limits placed on characteristics of an item, Process, or service defined in requirements documents. (ASQC Definitions)

Accuracy - Accuracy is a measure of the closeness of an individual measurement or the average of a number of measurements to the true value. Accuracy includes a combination of random error (precision) and systematic error (bias) components which are due to sampling and analytical operations; EPA recommends that his term not be used and that precision and bias be used to convey the information usually associated with accuracy. Refer to Appendix D Data Quality Indicators for a more detailed definition.

Activity - an all-inclusive term describing a specific set of operations of related tasks to be performed, either serially or in parallel (e.g., research and development, field sampling, analytical operations, equipment fabrication), that in total, result in a product or service.

Analysis matrix spike - the subjection of a prepared sample, extract or digestate that has been fortified (spiked) with a known amount of the analyte of interest, to the determinative step of an analytical method to estimate the bias imparted by the instrumental or determinative procedure.

Assessment - the evaluation process used to measure the performance or effectiveness of a system and its elements. As used here, assessment is an all-inclusive term used to denote any of the following: audit, performance evaluation, management systems review, peer review, inspection, or surveillance.

Audit (quality) - a systematic and independent examination to determine whether quality activities and related results comply with planned arrangements and whether these arrangements are implemented effectively and are suitable to achieve objectives.

Audit of data quality (ADQ) - a qualitative and quantitative evaluation of the documentation and procedures associated with environmental measurements to verify that the resulting data are of acceptable quality.

Authenticate - the act of establishing an item as genuine, valid, or authoritative.

Bias - the systematic or persistent distortion of a measurement process which causes errors in one direction (i.e., the expected sample measurement is different from the sample's true value). Refer to Appendix D Data Quality Indicators for a more detailed definition.

Blank - a sample that has not been exposed to the analyzed sample stream in order to monitor contamination during sampling, transport, storage, or analysis. The blank is subjected to the usual analytical or measurement process to establish a zero baseline or background value and is sometimes used to adjust or correct routine analytical results.

Blunder - mistakes that occur on occasion and produce erroneous results. Refer to Appendix D Data Quality Indicators for a more detailed definition.

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.

Calibration drift - the deviation in instrument response from a reference value over a period of time before recalibration.

Certification - the process of testing and evaluation against specifications designed to document, verify, and recognize the competence of a person, organization, or other entity to perform a function or service usually for a specified time.

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

Characteristic - any property of attribute of a datum, item, process, or service that is distinct, describable, and/or measurable.

Collocated samples - two or more portions collected at the same point in time and space so as to be considered identical.

Comparability - a measure of the confidence with which one data set or method can be compared to another.

Completeness - a measure of the amount of valid data obtained from a measurement system compared to the amount that was expected to be obtained under correct, normal conditions. Refer to Appendix D Data Quality Indicators for a more detailed definition.

Computer program - a sequence of instructions suitable for processing by a computer. Processing may include the use of an assembler, compiler, an interpreter, or a translator to prepare the program for execution. A computer program may be stored on magnetic media, and be referred to as "software" or may be stored permanently on computer chips, and be referred to as "firmware." Computer programs covered by this Standard are those used for design analysis, data acquisition, data reduction, data storage (data bases), operation or control, and data base or document control registers when used as the controlled source of quality information.

Confidence interval - the numerical interval constructed around a point estimate of a population parameter, combined with a probability statement (the confidence coefficient) linking it to the population's true parameter value. If the same confidence interval construction technique and assumptions are used to calculate future intervals, they will include the unknown population parameter with the same specified probability.

Confidentiality procedure - a procedure used to protect confidential business information (including proprietary data and personnel records) from unauthorized access.

Configuration - the functional, physical, and procedural characteristics of an item, experiment, or document.

Conformance - an affirmative indication or judgement that a product or service has met the requirements of the relevant specification, contract, or regulation; also the state of meeting the requirements.

Consensus standard - a standard established by a group representing a cross section of a particular industry or trade, or a part thereof.

Contractor - any organization or individual that contracts to furnish services or items or perform work.

Corrective action - measures taken to rectify conditions adverse to quality and, where possible, to preclude their recurrence.

Correlation coefficient - a number between -1 and 1 that indicates the degree of linearity between two variables or sets of numbers. The closer to -1 or +1, the stronger the linear relationship between the two (i.e., the better the correlation.) Values close to zero suggest no correlation between the two variables. The most common correlation coefficient is the product-moment, a measure of the degree of linear relationship between two variables.

Data of known quality - data that have the qualitative and quantitative components associated with their derivation documented appropriately for their intended use, and when such documentation is verifiable and defensible.

Data Quality Assessment (DQA) - 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 statistics and qualitative descriptors that are used to interpret the degree of acceptability or utility of data to the user. The principal data quality indicators are bias, precision, accuracy (precision and bias are preferred), comparability, completeness, representativeness, and statistical confidence.

Data Quality Objectives (DQOs) - qualitative and quantitative statements derived from the DQO Process that clarify study technical and quality 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 Quality Objectives Process - a systematic strategic planning tool based on the scientific method that identifies and defines the type, quality, and quantity of data needed to satisfy a specified use. The key elements of the process include:

- concisely defining the problem,
- identifying the decision to be made,
- identifying the key inputs to that decision,
- defining the boundaries of the study,
- developing the decision rule,
- specifying tolerable limits on potential decision errors, and
- selecting the most resource efficient data collection design.

Data Quality Objectives are the qualitative and quantitative outputs from the DQO Process. (See also Graded Approach)

Data reduction - the process of transforming the number of data items by arithmetic or statistical calculations, standard curves, concentration factors, etc., and collation into a more useful form. Data reduction is irreversible and generally results in a reduced data set and an associated loss of detail.

Data usability - the process of ensuring or determining whether the quality of the data produced meets the intended use of the data.

Deficiency - an unauthorized deviation from acceptable procedures or practices, or a defect in an item.

Demonstrated capability - the capability to meet procurement technical and quality specifications through evidence presented by the supplier to substantiate its claims and in a manner defined by the customer.

Design - specifications, drawings, design criteria, and performance requirements. Also the result of deliberate planning, analysis, mathematical manipulations, and design processes.

Design change - any revision of alteration of the technical requirements defined by approved and issued design output documents and approved and issued changes thereto.

Design review - a documented evaluation by a team, including personnel such as the responsible designers, the client for the work or product being designed, and a QA representative, but other than the

original designers, to determine if a proposed design will meet the established design criteria and perform as expected when implemented.

Detection Limit (DL)- the lowest concentration or amount of the target analyte that can be determined to be different from zero by a single measurement at a stated level of probability. (See also Appendix F)

Document - any written or pictorial information describing, defining, specifying, reporting, or certifying activities, requirements, procedures; or results.

Document control - the policies and procedures used by an organization to ensure that its documents and their revisions are proposed, reviewed, approved for release, inventoried, distributed, archived, stored, and retrieved in accordance with the organization's requirements.

Duplicate samples - two samples taken from and representative of the same population and carried through all steps of the sampling and analytical procedures in an identical manner. Duplicate samples are used to assess variance of the total method including sampling and analysis. See Collocated sample.

Environmental conditions - the description of a physical medium (e.g., air, water, soil, sediment) or biological system expressed in terms of its physical, chemical, radiological, or biological characteristics.

Environmental data - any parameters or pieces of information collected or produced from measurements, analyses, or models of environmental processes, conditions, and effects of pollutants on human health and the ecology, including results from laboratory analyses or from experimental systems representing such processes and conditions.

Environmental data operations - work performed to obtain, use, or report information pertaining to environmental processes and conditions.

Environmental monitoring - the process of measuring or collecting environmental data.

Environmental processes - manufactured or natural processes that produce discharges to or that impact the ambient environment.

Environmental programs - an all-inclusive term pertaining to any work or activities involving the environment, including but not limited to: characterization of environmental processes and conditions; environmental monitoring; environmental research and development; the design, construction, and operation of environmental technologies; and laboratory operations on environmental samples.

Environmental technology - an all-inclusive term used to describe pollution control devices and systems, waste treatment processes and storage facilities, and site remediation technologies and their components that my be utilized to remove pollutants or contaminants from or prevent them from entering the environment. Examples include wet scrubbers (air), soil washing (soil), granulated activated carbon unit (water), and filtration (air, water). Usually, this term will apply to hardware-based systems; however, it will also apply to methods or techniques used for pollution prevention, pollutant reduction, or containment of contamination to prevent further movement of the contaminants, such as capping, solidification or vitrification, and biological treatment.

Estimate - a characteristic from the sample from which inferences on parameters are made.

Evidentiary records - records identified as part of litigation and subject to restricted access, custody, use, and disposal.

Expedited change - an abbreviated method of revising a document at the work location where the document is used when the normal change process would cause unnecessary or intolerable delay in the work.

Field blank - a blank used to provide information about contaminants that may be introduced during sample collection, storage, and transport. A clean sample, carried to the sampling site, exposed to sampling conditions and returned to the laboratory and treated as an environmental sample.

Field (matrix) spike - a sample prepared at the sampling point (i.e., in the field) by adding a known mass of target analyte to a specified amount of sample. Field matrix spikes are used, for example, to determine the effect of the sample preservation, shipment, storage and sample preparation on analyte recovery efficiency (analytical bias).

Field split samples - two or more representative portions taken from the same sample and submitted for analysis to different laboratories to estimate interlaboratory precision.

Financial assistance - the process by which funds are provided by one organization (usually government) to another organization for the purpose of performing work or furnishing services or items. Financial assistance mechanisms include grants, cooperative agreements, and government interagency agreements.

Finding - an assessment conclusion that identifies a condition having a significant effect on an item or activity. An assessment finding may be positive or negative, and is normally accompanied by specific examples of the observed condition.

Goodness-of-fit test - the application of the chi-square distribution in comparing the frequency distribution of a statistic observed in a sample with the expected frequency distribution based on some theoretical model

Grade - the category or rank given to entities having the same functional use but different requirements for quality.

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. (See Data Quality Objectives Process)

Guidance - suggested practice that is not mandatory, intended as an aid or example in complying with a standard or requirement.

Guideline - a suggested practice that is non-mandatory in programs intended to comply with a standard.

Hazardous waste - any waste material that satisfies the definition of "hazardous waste" as given in 40 CFR part 261, "Identification and Listing of Hazardous Waste."

Holding time - the period a sample may be stored prior to its required analysis. While exceeding the holding time does not necessarily negate the veracity of analytical results, it causes the qualifying or "flagging" of the data for not meeting all of the specified acceptance criteria.

Identification error - misidentification of an analyte. Results in the contaminant of concern not being identified and the measured concentration being incorrectly assigned to another contaminant.

Independent assessment - an assessment performed by a qualified individual, group, or organization that is not a part of the organization directly performing and accountable for the work being assessed.

Inspection - examination or measurement of an item or activity to verify conformance to specific requirements.

Internal standard- a standard added to a test portion of a sample in a known amount and carried through the entire determination procedure as a reference for calibration and controlling the precision and bias of the applied analytical method.

Item - an all-inclusive term used in place of the following: appurtenance, facility, sample, assembly, component, equipment, material, module, part, product, structure, subassembly, subsystem, system, unit, documented concepts, or data.

Laboratory split samples - two or more representative portions taken from the same sample and analyzed by different laboratories to estimate the interlaboratory precision or variability and data comparability.

Limit of quantitation - the minimum concentration of an analyte or category of analytes in a specific matrix that can be identified and quantified above the method detection limit and within specified limits of precision and bias during routine analytical operating conditions.

Management - those individuals directly responsible and accountable for planning, implementing, and assessing work.

Management system - a structured non-technical system describing the policies, objectives, principles, organizational authority, responsibilities, accountability, and implementation plan of an organization for conducting work and producing items and services.

Management Systems Review (MSR) - the qualitative assessment of a data collection operation and/or organization(s) to establish whether the prevailing quality management structure, policies, practices, and procedures, are adequate for ensuring that the type and quality of data needed are obtained.

Matrix spike - a sample prepared by adding a known mass of target analyte to a specified amount of matrix sample for which an independent estimate of target analyte concentration is available. Spiked samples are used, for example, to determine the effect of the matrix on a method's recovery efficiency.

May - when used in a sentence denotes permission but not a requirement.

Mean (arithmetic) - the sum of all the values of a set of measurements divided by the number of values in the set; a measure of central tendency.

Mean squared error - a statistical term for variance added to the square of the bias.

Measurement and testing equipment (M&TE) - tools, gauges, instruments, sampling devices or systems used to calibrate, measure, test, or inspect in order to control or acquire data to verify conformance to specified requirements.

Memory effects error - the effect that a relatively high concentration sample has on the measurement of a lower concentration sample of the same analyte when the higher concentration sample precedes the lower concentration sample in the same analytical instrument.

Method - a body of procedures and techniques for performing an activity (e.g., sampling, chemical analysis, quantification) systematically presented in the order in which they are to be executed.

Mid-range check - a standard used to establish whether the middle of a measurement method's calibrated range is still within specifications.

Mixed waste - hazardous waster material as defined by 40 CFR 261 (RCRA) and mixed with radioactive waste subject to the requirements of the Atomic Energy Act.

Must - when used in a sentence denotes a requirement that has to be met.

Nonconformance - a deficiency in characteristic, documentation, or procedure that renders the quality of an item or activity unacceptable or indeterminate; nonfulfillment of a specified requirement.

Objective evidence - any documented statement of fact, other information, or record, either quantitative or qualitative, pertaining to the quality of an item or activity, based on observations, measurements, or tests which can be verified.

Observation - an assessment conclusion that identifies a condition (either positive or negative) which does not represent a significant impact on an item or activity. An observation may identify a condition which does not yet cause a degradation of quality.

Organization - a company, corporation, firm, enterprise, or institution, or part thereof, whether incorporated or not, public or private, that has its own functions and administration.

Organization structure - the responsibilities, authorities, and relationships, arranged in a pattern, through which an organization performs its functions.

Outlier - an observed value that appears to be discordant from the other observations in a sample. One of a set of observations that appears to be discordant from the others.

Parameter - a quantity, usually unknown, such as a mean or a standard deviation characterizing a population. Commonly misused for "variable", "characteristic" or "property."

Peer review - a documented critical review of work generally beyond the state of the art or characterized by the existence of potential uncertainty. The peer review is conducted by qualified individuals (or organization) who are independent of those who performed the work, but are collectively equivalent in technical expertise (i.e., peers) to those who performed the original work. The peer review is conducted to ensure that activities are technically adequate, competently performed, properly documented, and satisfy established technical and quality requirements. The peer review is an in-depth assessment of the assumptions, calculations, extrapolations, alternate interpretations, methodology, acceptance criteria, and conclusions pertaining to specific work and of the documentation that supports them. Peer reviews provide an evaluation of a subject where quantitative methods of analysis or measures of success are unavailable or undefined, such as in research and development.

Performance evaluation (PE) - 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.

Pollution prevention - an organized, comprehensive effort to systematically reduce or eliminate pollutants or contaminant prior to their generation or their release or discharge to the environment.

Population - the totality of items or units of material under consideration.

Precision - a measure of mutual agreement among individual measurements of the dame property, usually under prescribed similar conditions expressed generally in terms of variance. Refer to Appendix D Data Quality Indicators for a more detailed definition.

Procedure - a specified way to perform an activity.

Process - a set of interrelated resources and activities which transforms inputs into outputs. Examples of processes include analysis, design, data collection, operation, fabrication, and calculation.

Project - an organized set of activities within a program.

Qualified data - any data that have been modified or adjusted as part of statistical or mathematical evaluation, data validation, or data verification operations.

Qualified services - an indication that suppliers providing services have been evaluated and determined to meet the technical and quality requirements of the client as provided and approved procurement documents and demonstrated by the supplier to the client's satisfaction.

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 (QA) - an integrated system of management activities involving planning, implementation, assessment, reporting, and quality improvement to ensure that a process, item, or service is of the type and quality needed and expected by the client.

Quality assurance program description/plan - see quality management plan.

Quality Assurance Project Plan (QAPP) - a formal document describing in comprehensive detail the necessary QA, QC, and other technical activities that must be implemented to ensure that the results of the work performed will satisfy the stated performance criteria.

Quality control (QC) - 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 requirements established by the customer; operational techniques and activities that are used to fulfill requirements for quality.

Quality control sample - an uncontaminated sample matrix spiked with known amounts of analytes from a source independent from the calibration standards. It is generally used to establish intra laboratory or analyst specific precision and bias or to assess the performance of all or a portion of the measurement system.

Quality improvement - a management program for improving the quality of operations. Such management programs generally entail a formal mechanism for encouraging worker recommendations with timely management evaluation and feedback or implementation.

Quality management - that aspect of the overall management system of the organization that determines and implements the quality policy. Quality management includes strategic planning, allocation or resources, and other systematic activities (e.g., planning, implementation, and assessment) pertaining to the quality system.

Quality management plan (QMP) - a formal document that describes the quality system in terms of the organizational structure, functional responsibilities of management and staff, lines of authority, and required interfaces for those planning, implementing, and assessing all activities conducted.

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, and assessing work performed by the organization and for carrying out required QA and QC.

Radioactive waste - waste material containing radionuclides, or contaminated by radionuclides, subject to the requirements of the Atomic Energy Act.

Readiness review - a systematic, documented review of the readiness for the start-up or continued use of facility, process, or activity. Readiness reviews are typically conducted before proceeding beyond project milestones and prior to initiation of a major phase of work.

Reagent blank - a blank that contains any reagents used in the sample preparation and analysis procedure.

Record (quality) - a document that furnishes objective evidence of the quality of items or activities and that has been verified and authenticated as technically complete and correct. Records may include photographs, drawings, magnetic tape, and other data recording media.

Recovery - whether or not the methodology measures all of the analyte that is contained in the sample. Refer to Appendix D Data Quality Indicators for a more detailed definition.

Remediation - the process of reducing the concentration of a contaminant (or contaminants) in air, water, or soil media to a level that poses an acceptable risk to human health.

Repeatability - the degree of agreement between independent test results produced by the same analyst, using the same test method and equipment on random aliquots of the same sample within a short time period.

Reporting limit - the lowest concentration or amount of the target analyte required to be reported from a data collection project. Reporting limits are generally greater than detection limits and are usually not associated with a probability level.

Representativeness - a measure of the degree to which data accurately and precisely represent a characteristic of a population, parameter variation at a sampling point, a process condition, or an environmental condition. Refer also to Appendix D and Appendix H.

Reproducibility - the precision, usually expressed as variance, that measures the variability among the results of measurements of the same sample at different laboratories.

Requirement - a formal statement of a need and the expected manner in which it is to be met.

Research (applied) - a process, the objective of which is to gain knowledge or understanding necessary for determining the means by which a recognized and specific need may be met.

Research (basic) - a process, the objective of which is to gain fuller knowledge or understanding of the fundamental aspects of phenomena and of observable facts without specific applications toward processes or products in mind.

Research development/demonstration - systematic use of the knowledge and understanding gained from research and directed toward the production of useful materials, devices, systems, or methods, including prototypes and processes.

Round-robin study - a method validation study involving an undefined number of laboratories or analysts, all analyzing the same sample(s) by the same method. In a round-robin study all results are compared and used to develop summary statistics such as interlaboratory precision and method bias or recovery efficiency.

Ruggedness study - the carefully ordered testing of an analytical method while making slight variations in test conditions (as might be expected in routine use) to determine how such variations affect test results. If a variation affects the results significantly, the method restrictions are tightened to minimize this variability.

Scientific method - the principles and processes regarded as necessary for scientific investigation, including rules for concept or hypothesis formulation, conduct of experiments, and validation of hypotheses by analysis of observations.

Self-assessment - assessments of work conducted by individuals, groups, or organizations directly responsible for overseeing and/or performing the work.

Sensitivity - the capability of a method or instrument to discriminate between measurement responses representing different levels of a variable of interest. Refer to Appendix D Data Quality Indicators for a more detailed definition.

Service - the result generated by activities at the interface between the supplier and the customer, and the supplier internal activities to meet customer needs. Such activities in environmental programs include design, inspection, laboratory and/or field analysis, repair, and installation.

Shall - denotes a requirement that is mandatory whenever the criterion for conformance with the specification requires that there by no deviation. This does not prohibit the use of alternative approaches or methods for implementing the specification so long as the requirement is fulfilled.

Should - denotes a guideline or recommendation whenever noncompliance with the specification is permissible.

Significant condition - any state, status, incident, or situation of an environmental process or condition, or environmental technology in which the work being performed will be adversely affected sufficiently to require corrective action to satisfy quality objectives or specifications and safety requirements.

Software life cycle - the period of time that starts when a software product is conceived and ends when the software product is no longer available for routine use. The software life cycle typically includes a requirement phase, a design phase, an implementation phase, a test phase, an installation and check-out phase, and operation and maintenance phase, and sometimes a retirement phase.

Source reduction - any practice that reduces the quantity of hazardous substances, contaminants, or pollutants.

Span check - a standard used to establish that a measurement method is not deviating from its calibrated range.

Specification - a document stating requirements and which refers to or includes drawings or other relevant documents. Specifications should indicated the means and criteria for determining conformance.

Spike - a known quantity of a chemical that is added to a sample for the purpose of determining (1) the concentration of an analyte by the method of standard additions, or (2) analytical recovery efficiency, based on sample matrix effects and analytical methodology.

Split samples - two or more representative portions taken from one sample in the field or in the laboratory and analyzed by different analysts or laboratories. Split samples are quality control samples that are used to assess analytical variability and comparability.

Standard deviation - the most common measure of the dispersion or imprecision of observed values expressed as the positive square root of the variance. See Variance.

Standard operating procedure (SOP) - a written document that details the method for an operation, analysis, or action with thoroughly prescribed techniques and steps, and that is officially approved as the method for performing certain routine or repetitive tasks.

Supplier - any individual or organization furnishing items or services or performing work according to a procurement document or financial assistance agreement. This is an all-inclusive term used in place of any of the following: vendor, seller, contractor, subcontractor, fabricator, or consultant.

Surrogate spike or analyte - a pure substance with properties that mimic the analyte of interest. It is unlikely to be found in environmental samples and is added to them to establish that the analytical method has been performed properly.

Surveillance (quality) - continual or frequent monitoring and verification of the status of an entity and the analysis of records to ensure that specified requirements are being fulfilled.

Technical review - a documented critical review of work that has been performed within the state of the art. The review is accomplished by one or more qualified reviewers who are independent of those who performed the work, but are collectively equivalent in technical expertise to those who performed the original work. The review is an in-depth analysis and evaluation of documents, activities, material, data, or items that require technical verification or validation for applicability, correctness, adequacy, completeness, and assurance that established requirements are satisfied.

Technical systems audit (TSA) - a thorough, systematic, on-site, qualitative audit of facilities, equipment, personnel, training, procedures, recordkeeping, data validation, data management, and reporting aspects of a system.

Traceability - the ability to trace the history, application, or location of an entity by means of recorded identifications. In a calibration sense, traceability relates measuring equipment to national or international standards, primary standards, basic physical constants or properties, or reference materials. In a data collection sense, it relates calculations and data generated throughout the project back to the requirements for he quality of the project.

Trip blank - a clean sample of matrix that is carried to the sampling site and transported to the laboratory for analysis without having been exposed to sampling procedures.

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. See also Appendix G.

Variance (statistical) - a measure of the dispersion of a set of values.

Verification - confirmation by examination and provision of objective evidence that specified requirements have been fulfilled. In design and development, validation concerns the process of examining a result of a given activity to determine conformance to the stated requirements for that activity. See also Appendix G.

APPENDIX J

QAPP SOFTWARE AVAILABILITY

This Appendix has three sections:

1. Overview of Potential Need for Software in QAPP Preparation,
2. Existing Software, and
3. Software Availability and Source.

The information presented in this Appendix is only a subset of what is available to the QA Manager. Mention of certain products or software does not constitute endorsement, but only that some potentially useful material can be obtained from those products.

J1. OVERVIEW OF POTENTIAL NEED FOR SOFTWARE IN QAPP PREPARATION

The software needs are categorized under the four classes of QAPP elements. Within each category is an explanation of the general functions of a software tool that could prove useful in preparing, reviewing, or implementing a QAPP. In addition, the QAPP elements to which the software would apply are listed.

Class A: Project Management

This category of software would be used to produce planning documentation, such as assisting in the preparation of the QAPP document. In addition, this type of software could be used to produce other project documentation such as Standard Operating Procedures (SOPs), Quality Assurance Management Plans (QAMPs), and Data Quality Objectives (DQOs) reports.

GENERAL SOFTWARE FUNCTIONS	QAPP ELEMENTS
Provide the user guidance on what to address in each QAPP element and serve as a template for the production of the QAPP document.	All elements
Generate flowcharts to assist in preparing project organization charts and in illustrating processes that occur in the project, such as sample collection and analysis or data management.	A4, B10
Identify training or certification required for personnel in given program areas.	A9
Provide applicable regulatory standards (e.g., action or clean-up levels) for the various program areas (e.g., air, water, and solid waste).	A6
Provide guidance on implementing the DQO Process.	A5, A6, A7

Class B: Measurement and Data Acquisition

This type of software could be used to assist in the design of a sampling plan. In addition, this software could provide information on analytical methods and sample collection and handling.

GENERAL SOFTWARE FUNCTIONS	QAPP ELEMENTS
Assist in the development of sampling designs that will meet specified DQOs. The software should handle a variety of general design types with and without compositing, such as simple random sampling, grid sampling, and stratified sampling.	B1
Provide information on analytical procedures and sampling methods for various contaminants and media. This software could provide QC data for the analytical method (method detection limit (MDL), precision, and bias), references to standard methods, SOPs where calibration and maintenance information could be found.	B2, B4, B5, B6, B7
Assist in tracking samples and assisting with documenting sample handling and custody.	B3
Integrating QC design and sampling design to meet DQOs and facilitate DQA.	B1, B5, B10

Class C: Assessment and Oversight

This software would assist in assessment and oversight activities.

GENERAL SOFTWARE FUNCTIONS	QAPP ELEMENTS
Produce checklists, checklist templates, or logic diagrams (such as problem diagnostics) for technical systems audits, management systems reviews, and audits of data quality.	C1
Perform data quality assessment and facilitate corrective actions during the implementation phase as preliminary or field screening data become available.	C1, C2

Class D: Data Validation and Usability

This software would assist in validating data and assessing its usability.

GENERAL SOFTWARE FUNCTIONS	QAPP ELEMENTS
Assist in performing data validation and usability.	D2
Assist in performing data quality assessment.	D3

J2. EXISTING SOFTWARE

This information is summarized as a list of identified software; a more detailed description of each item is found in Section J3. A variety of commercial software packages are available to assist in statistical analysis, laboratory QC, and related activities, but this Appendix focuses on software used specifically by those preparing, implementing, and reviewing QAPPs.

Template Software

Several applications have been implemented in word processing software that provide guidance on how to complete each QAPP element and have provided a template for the discussion portion. Four examples of these applications are:

- Quality Systems and Implementation Plan, Section J3, No. 2;
- Quality Integrated Work Plan Template, Section J3, No. 3;
- QAPP Template, Section J3, No. 4; and
- Region 5 QAPP Template, Section J3, No. 5.

A more sophisticated application, Quality Assurance Sampling Plan for Environmental Response (QASPER), was identified that combines a template with links to a variety of lists that provide the user response options. Section J3, No. 1.

Flowcharting Software

Various flowcharting software is commercially available. One example found in QA/QC literature is allCLEAR III, Section J3, No. 6. Other more sophisticated packages link the flowchart diagrams to active databases or simulation modeling capabilities.

Regulatory Standards Software

This software provides regulatory limits under the various statutes for a wide variety of contaminants:

- Environmental Monitoring Methods Index (EMMI), Section J3, No. 7; and
- Clean-Up Criteria for Contaminated Soil and Groundwater (an example of a commercially available product), Section J3, No. 10.

Sampling Design Software

A variety of software has been developed to assist in the creation of sampling designs:

- DEFT, Section J3, No. 11;
- GeoEASE, Section J3, No. 12;
- ElipGrid, Section J3, No. 13;
- DRUMs, Section J3, No. 14; and
- DQOPro, Section J3, No. 15.

In addition, there are many statistical packages that support sampling design.

Analytical Methods Software

This software provides information on method detection limits and method summaries for a wide variety of analytical methods:

- EMMI, Section J3, No. 7; and
- EPA's Sampling and Analysis Methods Database, Section J3, No. 9.

DQO Guidance Software

DQOES provides guidance and generates documentation for performing the DQO Process, Section J3, No. 20.

Data Validation Software

Research Data Management and Quality Control System (RDMQ), Section J3, No. 16, is a data management system that allows for the verification, flagging and interpretation of data.

Data Quality Assessment Software

Several software packages have been developed to perform data quality assessment tasks. Examples of this software include:

- DataQUEST, Section J3, No. 17;
- ASSESS, Section J3, No. 18; and
- RRELSTAT, Section J3, No. 19.

Note that most commercially available statistical packages (not listed above) perform a variety of DQA tasks.

QAPP Review

QATRACK, Section J3, No. 20, is used to track QAPPs undergoing the review process.

SOFTWARE NEED	QAPP ELEMENTS	EXISTING SOFTWARE
PROJECT MANAGEMENT		
Template Guidance	All elements	QASPER, QSIP, QWIP, QAPP Template
Flowcharting	A4, B10	allCLEAR III
Training/Certification Requirements	A9	None identified
Regulatory Standards	A6	EMMI, Clean-Up Criteria for Contaminated Soil and Ground Water
DQO Guidance	A5, A6, A7	DQOES
MEASUREMENT AND DATA ACQUISITION		
Sample Design	B1	DEFT, GeoEASE, ElipGrid, DRUMs, DQOPRO, miscellaneous statistical packages
Analytical and Sampling Procedures	B2, B4, B5, B6, B7	EMMI, EPA's Sampling and Analysis Database
Sample Tracking, Documenting Sample Handling and Custody	B3	None identified
Integrating QC Design and Sampling Design to Meet DQOs and Facilitate DQA.	B1, B5, B10	DQOPRO
ASSESSMENT AND OVERSIGHT		
Checklists	C1	None identified
Data Quality Assessment	C1, C2	DataQUEST, ASSESS, RRELSTAT
DATA VALIDATION AND USABILITY		
Data Validation	D2	RDMQ
Data Quality Assessment	D3	DataQUEST, ASSESS, RRELSTAT, miscellaneous statistical packages

J3. SOFTWARE AVAILABILITY AND SOURCES

The wide variety of existing software has potential to meet the needs identified for preparing QAPPs. As illustrated in the table, at least one example of a software tool was identified that could potentially be applied to aspects of QAPP preparation or implementation for all but three of the need areas. The capabilities of the existing software should match the QAPP needs as most of the software was developed for use with a QAPP or for environmental data collection or analysis. Software not designed for these uses could be modified or used to form the basis of an application that is more tailored to QAPP preparation or implementation.

1. Quality Assurance Sampling Plan for Environmental Response (QASPER) Version 4.0

Sponsoring Organization:	EPA
Implementing Software:	Clipper 5.2
Information Source:	Randall Romig, EPA, Region 6, (214) 665-8346 and *Quality Assurance Sampling Plan for Environmental Response (QASPER Version 4.0 User's Guide), latest version is QASPER Version 4.1,* January 1995. William Coakley, EPA, (908) 906-6921.

QASPER allows the creation and editing of a Quality Assurance Sampling Plan for Environmental Response. The plan template consists of 11 sections: (1) title page, (2) site background, (3) data use objectives, (4) sampling design, (5) sampling and analysis, (6) standard operating procedures, (7) quality assurance requirements, (8) data validation, (9) deliverables, (10) project organization and responsibilities, and (11) attachments. While preparing the plan, the user may enter the required information or select from the options provided in a variety of "picklists". The picklists cover topics such as holding times, methods, preservatives, and sampling approaches. The user may add or delete options from the picklists. QASPER also provides various utility functions such as backup, restore, export, and import a plan. Output may be directed to a file or a printer.

2. Quality Systems and Implementation Plan (QSIP)

Sponsoring Organization:	EPA
Implementing Software:	WordPerfect 5.1/5.2 (A more recent version and implementation may be available)
Information Source:	Gene Tatsch, RTI, CEMQA, QAD, (919) 541-6930 and QSIP template, Ron Patterson, EPA, ORD, NERL-RTP, APRD, (919) 541-3779

QSIP is a Work Plan with all the applicable QA elements identified and integrated at the point(s) where they apply. QSIP is intended as a combined Work Plan and QAPP. QSIP utilizes the comment feature of Word Perfect. The commented text provides guidance on what information to supply in the various sections of QSIP. An asterisk indicates where in the template the users should enter their discussion. The comments are not printed, leaving the preparer's discussion only in the final document.

Sections 1, 2, and parts of 3 relate to management functions and address the Quality Assurance aspects of the overall project. Sections 3-7 relate to the technical functions specific to each work effort crucial to the accomplishment of the overall project. These sections address the Quality Control aspects of the critical work activities being performed under the project. The seven sections of the template are: (1) Project Planning and Organization, (2) Management Assessment and Communications Plan, (3) Project Implementation Plan, (4) Data Acquisition and Management, (5) Records Usage and Management, (6) Routine Controls and Procedures, and (7) Technical Assessment and Response.

3. Quality Integrated Work Plan Template for R&D and Monitoring Projects

Sponsoring Organization:	North American Research Strategy for Tropospheric Ozone (NARSTO)
Implementing Software:	Word Perfect 6.1
Information Source:	Ron Patterson, EPA, ORD, NERL-RTP, APRD, (919) 541-3779 and NARSTO homepage

The Quality Integrated Work Plan (QIWP) template is a tool designed to assist with the planning, managing, and implementing a specific monitoring or R&D project. The QIWP template is formatted with comment boxes that provide guidance on the information to provide in each section. When activated, the text in the comment boxes will appear on screen; however, they will not appear in a printout. An asterisk indicates where the user should begin entering the discussion for each section. The QIWP document control format is already setup in the template header. When a particular element is considered not applicable, the rationale for that decision must be stated in response to that element. Once satisfied with the information entered under all elements of the template, the resulting printout is the combined project work plan and quality assurance plan. In addition, a printout of the QIWP template, prior to entering project related information, can be used as a checklist for planning and review purposes.

Other software packages available are Quality Integrated Work Plan Template for Model Development Projects and Quality Integrated Work Plan Template for Model Application Projects.

4. QAPP Template

Sponsoring Organization:	EPA
Implementing Software:	Word Perfect 6.1
Information Source:	Joe Livolsi, EPA, NHEERL, AED, Narragansett (401) 782-3163 and QAPP template

This package contains an annotated template containing instructions for completing each section of the QAPP. The users are also instructed where to insert their discussions within the template. After completing the QAPP, the italicized instructions are not printed, leaving only the preparer's discussion. In addition, a table of contents is automatically generated. The template describes the information that should be provided under the main topics of project management, measurement/data acquisition, data, assessment/oversight, and references. The project management section covers the introduction, goals of the project, organization of the project participants and QA, and DQOs. The measurement/data acquisition section discusses the topics to address to describe the statistical research design and sampling. This section also covers the elements related to sample analysis: description of the instrument, calibration, quality control, consumables, and preventative maintenance. The data section provides for a discussion of the data management procedures. The assessment/oversight section covers audits and QA reports. The next section is a list of references. Finally, six tables are provided as examples for displaying information on the following topics: (1) measurement quality criteria; (2) sample collection, handling, and preservation; (3) instrument data and interferences; (4) instrument calibration, (5) quality control checks; and (6) preventive maintenance.

5. Region 5 QAPP Template

Sponsoring Organization:	EPA
Implementing Software:	Word Perfect 5.1/5.2
Information Source:	George Schupp, EPA Region 5, (312) 886-6221 and QAPP template

This software consists of two model documents (one for Superfund sites and one for RCRA sites) that describe the preparation of a QAPP in a series of elements. Each element contains two types of information: (1) content requirements that are presented as smaller text and (2) structural guidance that is presented as larger text and headed by appropriate section number. This information is intended to show to the QAPP preparer the requirements that must be described in each element and the level of detail that is typically needed to gain Region 5 approval. Example text is provided that should be deleted and replaced

with the specific site information. Alternative text specific to RCRA/Superfund sites, and general notes, are indicated in bold print. Some of the example language is applicable to a broad range of sites and may be considered "boiler-plate." Text with a dark background indicates boiler-plate language.

A TSCA Model Plan template is also available which attempts to be a comprehensive guide of all the data gathering activities for FY 94 Title IV grantees. In this template, headers are provided in "background" format, and text that may apply to specific situations is in italic font. Open spaces indicate where the preparer's input is required.

6. allCLEAR

Sponsoring Organization:	Commercial
Implementing Software:	Proprietary
Information Source:	American Society for Quality Control Quality Press, Publications Catalogue, (800) 248-1946

This software enables the creation of simple process diagrams, organizational charts, or decision trees. It also creates diagrams from text outlines, spreadsheets, and database information.

7. Environmental Monitoring Methods Index (EMMI)

Sponsoring Organization:	EPA
Implementing Software:	Proprietary
Information Source:	DynCorp Environmental Technical Support, (703) 519-1222

This software consists of an analytical methods database containing over 4200 analytes, 3400 analytical and biological methods, and 47 regulatory and non-regulatory lists. EMMI cross-references analytes, methods, and lists and has information about related laws, organizations, and other chemical databases. The information does not include measurement method performance such as precision and bias.

8. EPA's Sampling and Analysis Methods Database, 2nd Edition

Sponsoring Organization:	EPA
Implementing Software:	Proprietary
Information Source:	Larry Keith, Radian Corporation, (512) 454-4797 and documentation.

This software has a menu driven program allowing the user to search a database of 178 EPA-approved analytical methods with more than 1300 method and analyte summaries. The database covers industrial chemicals, pesticides, herbicides, dioxins, and PCBs and focusses on water, soil matrices, and quality parameters. The software generates reports that are stand-alone documents that can be browsed, printed, or copied to files. Each report contains information for initial method selection such as applicable matrices, analytical interferences and elimination recommendations, sampling and preservation requirements, method detection limits, and precision, accuracy, and applicable concentration ranges.

9. CleanUp Criteria for Contaminated Soil and Groundwater

Sponsoring Organization:	Commercial
Implementing Software:	Proprietary
Information Source:	American Society for Quality Control Quality Press, Publications Catalogue, (800) 248-1946

This software consists of a one volume document and diskette summarizing cleanup criteria developed by EPA, all 50 state regulatory agencies, and select countries outside the United States.

10. Decision Error Feasibility Trials (DEFT)

Sponsoring Organization:	EPA, QAD
Implementing Software:	Microsoft C
Information Source:	QAD (202) 260-5763 (Guidance Document G4-D)

This package allows quick generation of cost information about several simple sampling designs based on the DQO constraints. The DQO constraints can be evaluated to determine their appropriateness and feasibility before the sampling and analysis design is finalized.

This software supports the *Guidance for the Data Quality Objectives Process*, EPA QA/G-4 that provides general guidance to organizations on developing data quality criteria and performance specifications for decision making. The *Data Quality Objectives Decision Error Feasibility Trials (DEFT) User's Guide*, contains detailed instructions on how to use DEFT software and provides background information on the sampling designs that the software uses.

11. GeoEAS

Sponsoring Organization:	EPA
Implementing Software:	Fortran
Information Source:	*GEO-EAS 1.2.1 User's Guide*, EPA/600/8-91/008, April, 1991, Evan Englund, (702) 798-2248

Geostatistical Environmental Assessment Software (Geo-EAS) is a collection of interactive software tools for performing two-dimensional geostatistical analyses of spatially distributed data. Programs are provided for data file management, data transformations, univariate statistics, variogram analysis, cross validation, kriging, contour mapping, post plots, and line/scatter plots. Users may alter parameters and re-calculate results or reproduce graphs, providing a "what if" analysis capability.

Software and the user's guide can be downloaded through the ORD World Wide Web site at http://www.epa.gov/ORD/ or http://www.epa.gov/ORD/nerl.htm.

12. ELIPGRID-PC

Sponsoring Organization:	DOE
Implementing Software:	CA-Clipper
Information Source:	*ELIPGRID-PC: UPGRADED VERSION*, ORNL/TM-13103, Jim Davidson, ORNL/GJ, (970) 248-6259

ELIPGRID-PC calculates the probabilities related to hitting a single hot spot. The user has the following options: (1) calculating the probability of detecting a hot spot of given size and shape when using a specified grid, (2) calculating the grid size required to find a hot spot of given size and shape with specified confidence, (3) calculating the size of the smallest hot spot likely to be hit with a specified sampling grid, (4) calculating a grid size based on fixed sampling cost, and (5) displaying a graph of the probability of hitting a hot spot versus sampling costs.

13. DQOPRO

Sponsoring Organization:	Radian International
Implementing Software:	Visual Basic
Information Source:	Larry Keith, Radian International, (512) 454-4797 and documentation

This software consists of a series of three computer programs that calculate the number of samples needed to meet specific DQOs. DQOPRO provides answers for three objectives: (1) determining the rate at which an event occurs, (2) determining an estimate of an average within a tolerable error, and (3) determining the sampling grid necessary to detect "hot-spots." The three programs that make up DQOPRO are described below.

(1) Success-Calc is used to determine the number of samples needed to detect a specified characteristic in a population of samples. For example, the software may be used to calculate the number of QC samples (such as method blanks or matrix spikes) needed in order to assure that no more than a specified rate (e.g., 5%) of false positive or false negative detections will occur in the environmental samples associated with the QC samples. Or, the software may be used to calculate the number of samples needed to ensure detection of any other characteristic of interest that occurs in more than a specified portion of the population. In addition, Success-Calc also calculates the maximum and minimum proportions corresponding to the observed (sample) proportion. Inputs include the maximum percentage of the selected characteristic that is allowed to go undetected, the desired probability of detecting that characteristic if it occurs in more than the maximum percentage specified, and how many samples, if any, that will be allowed to fail the specified criteria.

(2) Enviro-Calc is used to calculate how many environmental samples will need to be collected and analyzed in order to meet a specified tolerable error (e.g., for the average concentration calculated from environmental samples to be within plus or minus 10% of the true average with a confidence level of 95%). Inputs include the maximum tolerable error, the desired confidence level, and the expected relative standard deviation (RSD) or standard deviation (SD) of the sampling and analysis measurement results.

(3) HotSpot-Calc is used to determine the grid size needed to detect the presence of a single localized spot of pollutants ("hot spot") of a specified size and shape with a specified probability of missing its detection if it is present. Once the grid size is calculated, then the number of samples needed are automatically calculated by dividing the sampling area by the square of the grid size. Inputs include the shape of the grid that will be used (e.g., triangle, square or rectangle), the size and shape of the spot (e.g.,

circle, ellipse, or long ellipse), the acceptable probability of missing it (e.g., 10%, 20%, etc.), and the size of the area to be sampled.

14. Research Data Management and Quality Control System (RDMQ)

Sponsoring Organization:	Environment Canada and EPA
Implementing Software:	SAS
Information Source:	Mike Papp, EPA, OAQPS, (919) 541-2408 and documentation

This software is a data management system that allows for the verification, flagging, and interpretation of data. RDMQ is a menu-driven application with facilities for loading data, applying quality control checks, viewing and changing data, producing tabular and graphical reports, and exporting data in ASCII files. RDMQ provides a shell environment that allows the end-user to perform these tasks in a structured manner.

The user creates the databases and quality control checks through a user friendly interface. During the quality control process, every datum is assigned one or more validity flags based on the results of the quality control checks. These flags are stored in the same dataset as the sample values. The user defines a flag to indicate a "warning" or "corrective action required." This flagging method allows the end-user to zero in on the anomalies in the data, which streamlines the QC process and ensures that quality control is applied in a consistent and thorough manner.

RDMQ provides a number of tools for viewing the measurement values and their corresponding flags. The role of the user is to decide whether a data value flagged with a "warning" or "corrective action required" flag should be corrected (changes to data are recorded in an audit log) or whether the flag should remain in the database permanently and be passed on to the users of the data. The user also has the capability of adding manual flags to a value. When initially defining a flag, the user may record the usual cause and suggested corrective action. This information is easily available to the user during the QC process, which can be very helpful if the person doing the QC is different from the person who has defined the flags.

Once the data have been quality controlled, they can be exported in comma-delimited ASCII files. Features included in RDMQ are:

(a) input of measurement data from instruments and samples (including QC information such as field blanks and diagnostics);

(b) data quality control including flag assignments to every value of every variable;

(c) corrections to measurements, e.g., blank corrections and calibrations (this requires a customized SAS program);

(d) archiving of data files with the ability to extract subsets for research and interchange with other agencies;

(e) open-ended design to accommodate additions to QA/QC checks, new variables, and sampling intervals;

(f) data visualization (as an integral part of the quality control process);

(g) data import and export in ASCII files;

(h) QA/QC checks separated into modules that can be maintained by the user;

(i) user configurable outlier checking;

(j) audit trail of data changes, with reporting facility; and

(k) system-generated reports documenting the flag and variable definitions.

15. DataQUEST

Sponsoring Organization:	EPA
Implementing Software:	MicroSoft C
Information Source:	QAD, (202) 260-5763 (Guidance G-9D)

This tool is designed to provide a quick and easy way for managers and analysts to perform baseline Data Quality Assessment. The goal of the system is to allow those not familiar with standard statistical packages to review data and verify assumptions that are important in implementing the DQA Process. This software supports the *Guidance for Data Quality Assessment,* EPA QA/G-9 that demonstrates the use of the DQA Process in evaluating environmental data sets.

16. ASSESS 1.01a

Sponsoring Organization:	EPA
Implementing Software:	Fortran 77
Information Source:	Software and documentation, Jeff van Ee, (702) 798-2367

This software tool was designed to calculate variances for quality assessment samples in a measurement process. The software performs the following functions: (1) transforming the entire data set, (2) producing scatter plots of the data, (3) displaying error bar graphs that demonstrate the variance, and (4) generating reports of the results and header information.

17. RRELSTAT

Sponsoring Organization:	EPA
Implementing Software:	C or FORTRAN
Information Source:	Philip C.L. Lin, (513) 569-7324

This set of computer programs provides 22 statistical tests for solving sampling and related statistical problems. The programs are designed so that persons without an in-depth understanding of statistics can easily use them. Specific, detailed written instructions for application of these programs are also provided in each of the programs on the disc. The introduction screen helps guide the user to the appropriate program through a series of questions and answers.

18. QATRACK

Sponsoring Organization:	EPA
Implementing Software:	MicroSoft Access
Information Source:	Mike Papp, EPA, GLNPO, (919) 541-2408 and documentation

This software provides a database that tracks QAPPs requiring approval. Data are entered into QATRACK during the assistance agreement start-up stage, as soon as the QA manager reviews and signs the agreement. Users can edit the data, query the database to perform data reviews, and archive files once the QAPP is approved.

APPENDIX K

CALCULATION OF STATISTICAL QUANTITIES

This appendix is taken directly from Sections 2.2 and 2.3 of EPA QA/G-9 *Guidance for Data Quality Assessment.*

2.2.1 Measures of Relative Standing

Sometimes the analyst is interested in knowing the relative position of one of several observations in relation to all of the observations. Percentiles are one such measure of relative standing that may also be useful for summarizing data. A percentile is the data value that is greater than or equal to a given percentage of the data values. Stated in mathematical terms, the p^{th} percentile is the data value that is greater than or equal to p% of the data values and is less than or equal to (1-p)% of the data values. Therefore, if 'x' is the p^{th} percentile, then p% of the values in the data set are less than or equal to x, and (100-p)% of the values are greater than or equal to x. A sample percentile may fall between a pair of observations. For example, the 75^{th} percentile of a data set of 10 observations is not uniquely defined. Therefore, there are several methods for computing sample percentiles, the most common of which is described in Box 2.2-1.

Important percentiles usually reviewed are the quartiles of the data, the 25^{th}, 50^{th}, and 75^{th} percentiles. The 50^{th} percentile is also called the sample median (section 2.2.2), and the 25^{th} and 75^{th} percentile are used to estimate the dispersion of a data set (section 2.2.3). Also important for environmental data are the 90^{th}, 95^{th}, and 99^{th} percentile where a decision maker would like to be sure that 90%, 95%, or 99% of the contamination levels are below a fixed risk level.

Box 2.2-1: Directions for Calculating the Measure of Relative Standing (Percentiles) with an Example

Let $X_1, X_2, ..., X_n$ represent the n data points. To compute the p^{th} percentile, y(p), first list the data from smallest to largest and label these points $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ (so that $X_{(1)}$ is the smallest, $X_{(2)}$ is the second smallest, and $X_{(n)}$ is the largest). Let t = p/100, and multiply the sample size n by t. Divide the result into the integer part and the fractional part, i.e., let nt = j + g where j is the integer part and g is the fraction part. Then the p^{th} percentile, y(p), is calculated by:

If g = 0, $y(p) = (X_{(j)} + X_{(j+1)})/2$; otherwise, $y(p) = X_{(j+1)}$

Example: The 90^{th} and 95^{th} percentile will be computed for the following 10 data points (ordered from smallest to largest) : 4, 4, 4, 5, 5, 6, 7, 7, 8, and 10 ppb.

For the 95th percentile, t = p/100 = 95/100= .95 and nt = (10)(.95) = 9.5 = 9 + .5. Therefore, j = 9 and g = .5. Because g = .5 ≠ 0, $y(95) = X_{(j+1)} = X_{(9+1)} = X_{(10)}$ = 10 ppm. Therefore, 10 ppm is the 95^{th} percentile of the above data. For the 90^{th} percentile, t = p/100 = 90/100 = .9 and nt = (10)(.9) = 9. Therefore j = 9 and g = 0. Since g = 0, $y(90) = (X_{(9)} + X_{(10)}) / 2 = (8 + 10) / 2 = 9$ ppm.

A quantile is similar in concept to a percentile; however, a percentile represents a percentage whereas a quantile represents a fraction. If 'x' is the p^{th} percentile, then at least p% of the values in the data set lie at or below x, and at least (100-p)% of the values lie at or above x, whereas if x is the p/100 quantile of the data, then the fraction p/100 of the data values lie at or below x and the fraction (1-p)/100 of the data values lie at or above x. For example, the .95 quantile has the property that .95 of the observations lie at or below x and .05 of the data lie at or above x. For the example in Box 2.2-1, 9 ppm would be the .95 quantile and 10 ppm would be the .99 quantile of the data.

2.2.2 Measures of Central Tendency

Measures of central tendency characterize the center of a sample of data points. The three most common estimates are the mean, median, and the mode. Directions for calculating these quantities are contained in Box 2.2-2; examples are provided in Box 2.2-3.

The most commonly used measure of the center of a sample is the sample mean, denoted by $\bar{x}$. This estimate of the center of a sample can be thought of as the "center of gravity" of the sample. The sample mean is an arithmetic average for simple sampling designs; however, for complex sampling designs, such as stratification, the sample mean is a weighted arithmetic average. The sample mean is influenced by extreme values (large or small) and nondetects (see section 4.7).

The sample median ($\tilde{x}$) is the second most popular measure of the center of the data. This value falls directly in the middle of the data when the measurements are ranked in order from smallest to largest. This means that ½ of the data are smaller than the sample median and ½ of the data are larger than the sample median. The median is another name for the 50th percentile (section 2.2.1). The median is not influenced by extreme values and can easily be used in the case of censored data (nondetects).

The third method of measuring the center of the data is the mode. The sample mode is the value of the sample that occurs with the greatest frequency. Since this value may not always exist, or if it does it may not be unique, this value is the least commonly used. However, the mode is useful for qualitative data.

2.2.3 Measures of Dispersion

Measures of central tendency are more meaningful if accompanied by information on how the data spread out from the center. Measures of dispersion in a data set include the range, variance, sample standard deviation, coefficient of variation, and the interquartile range. Directions for computing these measures are given in Box 2.2-4; examples are given in Box 2.2-5.

The easiest measure of dispersion to compute is the sample range. For small samples, the range is easy to interpret and may adequately represent the dispersion of the data. For large samples, the range is not very informative because it only considers (and therefore is greatly influenced) by extreme values.

The sample variance measures the dispersion from the mean of a data set. A large sample variance implies that there is a large spread among the data so that the data are not clustered around the mean. A small sample variance implies that there is little spread among the data so that most of the data are near the mean. The sample variance is affected by extreme values and by a large number of nondetects. The sample standard deviation is the square root of the sample variance and has the same unit of measure as the data.

The coefficient of variation (CV) is a unitless measure that allows the comparison of dispersion across several sets of data. The CV is often used in environmental applications because variability (expressed as a standard deviation) is often proportional to the mean.

When extreme values are present, the interquartile range may be more representative of the dispersion of the data than the standard deviation. This statistical quantity does not depend on extreme values and is therefore useful when the data include a large number of nondetects.

Box 2.2-2: Directions for Calculating the Measures of Central Tendency

Let $X_1, X_2, ..., X_n$ represent the n data points.

Sample Mean: The sample mean $\bar{x}$ is the sum of all the data points divided by the total number of data points (n):

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Sample Median: The sample median ($\tilde{X}$) is the center of the data when the measurements are ranked in order from smallest to largest. To compute the sample median, list the data from smallest to largest and label these points $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ (so that $X_{(1)}$ is the smallest, $X_{(2)}$ is the second smallest, and $X_{(n)}$ is the largest).

If the number of data points is odd, then $\tilde{X} = X_{([n+1]/2)}$

If the number of data points is even, then $\tilde{X} = \frac{X_{(n/2)} + X_{([n/2]+1)}}{2}$

Sample Mode: The mode is the value of the sample that occurs with the greatest frequency. The mode may not exist, or if it does, it may not be unique. To find the mode, count the number of times each value occurs. The sample mode is the value that occurs most frequently.

Box 2.2-3: Example Calculations of the Measures of Central Tendency

Using the directions in Box 2.2-2 and the following 10 data points (in ppm): 4, 5, 6, 7, 4, 10, 4, 5, 7, and 8, the following is an example of computing the sample mean, median, and mode.

Sample mean:

$$\bar{X} = \frac{4 + 5 + 6 + 7 + 4 + 10 + 4 + 5 + 7 + 8}{10} = \frac{60}{10} = 6\ ppm$$

Therefore, the sample mean is 6 ppm.

Sample median: The ordered data are: 4, 4, 4, 5, 5, 6, 7, 7, 8, and 10. Since n=10 is even, the sample median is

$$\tilde{X} = \frac{X_{(10/2)} + X_{([10/2]+1)}}{2} = \frac{X_{(5)} + X_{(6)}}{2} = \frac{5 + 6}{2} = 5.5\ ppm$$

Thus, the sample median is 5.5 ppm.

Sample mode: Computing the number of times each value occurs yields:

4 appears 3 times; 5 appears 2 times; 6 appears 1 time; 7 appears 2 times; 8 appears 1 time; and 10 appears 1 time.

Because the value of 4 ppm appears the most times, it is the mode of this data set.

Box 2.2-4: Directions for Calculating the Measures of Dispersion

Let $X_1, X_2, ..., X_n$ represent the n data points.

Sample Range: The sample range (R) is the difference between the largest value and the smallest value of the sample, i.e., R = maximum - minimum.

Sample Variance: To compute the sample variance (s^2), compute:

$$s^2 = \frac{\sum_{i=1}^{n} X_i^2 - \frac{1}{n}(\sum_{i=1}^{n} X_i)^2}{n-1}$$

Sample Standard Deviation: The sample standard deviation (s) is the square root of the sample variance, i.e.,

$$s = \sqrt{s^2}$$

Coefficient of Variation: The coefficient of variation (CV) is the standard deviation divided by the sample mean (section 2.2.2), i.e., CV = s/ $\bar{x}$. The CV is often expressed as a percentage.

Interquartile Range: Use the directions in section 2.2.1 to compute the 25th and 75th percentiles of the data (y(25) and y(75) respectively). The interquartile range (IQR) is the difference between these values, i.e.,
IQR = y(75) - y(25).

Box 2.2-5: Example Calculations of the Measures of Dispersion

In this box, the directions in Box 2.2-4 and the following 10 data points (in ppm): 4, 5, 6, 7, 4, 10, 4, 5, 7, and 8, are used to calculate the measures of dispersion. From Box 2.2-2, $\bar{x}$ = 6 ppm.

Sample Range: R = maximum - minimum = 10 - 4 = 6 ppm

Sample Variance:

$$s^2 = \frac{[4^2+5^2+...+7^2+8^2] - \frac{(4+5+...+7+8)^2}{10}}{10-1} = \frac{396 - \frac{(60)^2}{10}}{9} = 4\ ppm^2$$

Sample Standard Deviation: $s = \sqrt{s^2} = \sqrt{4} = 2\ ppm$

Coefficient of Variation: $CV = s / \bar{X} = 2ppm / 6ppm = \frac{1}{3} = 33\%$

Interquartile Range: Using the directions in section 2.2.1 to compute the 25th and 75th percentiles of the data (y(25) and y(75) respectively): y(25) = $X_{(2+1)} = X_{(3)}$ = 4 ppm and y(75) = $X_{(7+1)} = X_{(8)}$ = 7 ppm. The interquartile range (IQR) is the difference between these values: IQR = y(75) - y(25) = 7 - 4 = 3 ppm

2.2.4 Measures of Association

Data often include measurements of several characteristics (variables) for each sample point and there may be interest in knowing the relationship or level of association between two or more of these variables. One of the most common measures of association is the correlation coefficient. Directions and an example for calculating a correlation coefficient are contained in Box 2.2-6.

The correlation coefficient measures the linear relationship between two variables. A linear association implies that as one variable increases so does the other linearly, or as one variable decreases the other increases linearly. Values of the correlation coefficient close to +1 (positive correlation) imply that as one variable increases so does the other, the reverse holds for values close to -1. A value of +1 implies a perfect positive linear correlation, i.e., all the data pairs lie on a straight line with a positive slope. A value of -1 implies perfect negative linear correlation. Values close to 0 imply little correlation between the variables.

The correlation coefficient does not imply cause and effect. The analyst may say that the correlation between two variables is high and the relationship is strong, but may not say that one variable causes the other variable to increase or decrease without further evidence and strong statistical controls. The correlation coefficient does not detect nonlinear relationships so it should be used only in conjunction with a scatter plot (section 2.3.7.2). A scatter plot can be used to determine if the correlation coefficient is meaningful or if some measure of nonlinear relationships should be used. The correlation coefficient can be significantly changed by extreme values so a scatter plot should be used first to identify such values.

Box 2.2-6: Directions for Calculating the Correlation Coefficient with an Example

Let $X_1, X_2, \ldots, X_n$ represent one variable of the n data points and let $Y_1, Y_2, \ldots, Y_n$ represent a second variable of the n data points. The Pearson correlation coefficient, r, between X and Y is computed by:

$$r = \frac{\sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\left[\left[\sum_{i=1}^{n} X_i^2 - \frac{(\sum_{i=1}^{n} X_i)^2}{n}\right]\left[\sum_{i=1}^{n} Y_i^2 - \frac{(\sum_{i=1}^{n} Y_i)^2}{n}\right]\right]^{1/2}}$$

Example: Consider the following data set (in ppb): Sample 1 — arsenic (X) = 4.0, lead (Y) = 8.0; Sample 2 - arsenic = 3.0, lead = 7.0; Sample 3 - arsenic = 2.0, lead = 7.0; and Sample 4 - arsenic = 1.0, lead = 6.0.

$$\sum_{i=1}^{n} X_i = 10,\ \sum_{i=1}^{n} Y_i = 28,\ \sum_{i=1}^{n} X_i^2 = 30,\ \sum_{i=1}^{n} Y_i^2 = 198,\ \sum_{i=1}^{n} X_i Y_i = (4x8) + \ldots + (1x6) = 73.$$

and $$r = \frac{73 - \frac{(10)(28)}{4}}{\left[\left[30 - \frac{(10)(10)}{4}\right]\left[198 - \frac{(28)(28)}{4}\right]\right]^{1/2}} = 0.949$$

Since r is close to 1, there is a strong linear relationship between these two contaminants.

2.3 GRAPHICAL REPRESENTATIONS

2.3.1 Histogram/Frequency Plots

Two of the oldest methods for summarizing data distributions are the frequency plot (Figure 2.3-1) and the histogram (Figure 2.3-2). Both the histogram and the frequency plot use the same basic principles to display the data: dividing the data range into units, counting the number of points within the units, and displaying the data as the height or area within a bar graph. There are slight differences between the histogram and the frequency plot. In the frequency plot, the relative height of the bars represents the relative density of the data. In a histogram, the area within the bar represents the relative density of the data. The difference between the two plots becomes more distinct when unequal box sizes are used.

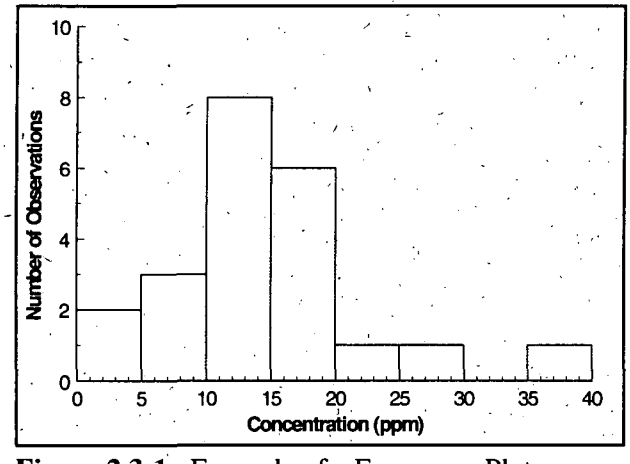


Figure 2.3-1. Example of a Frequency Plot

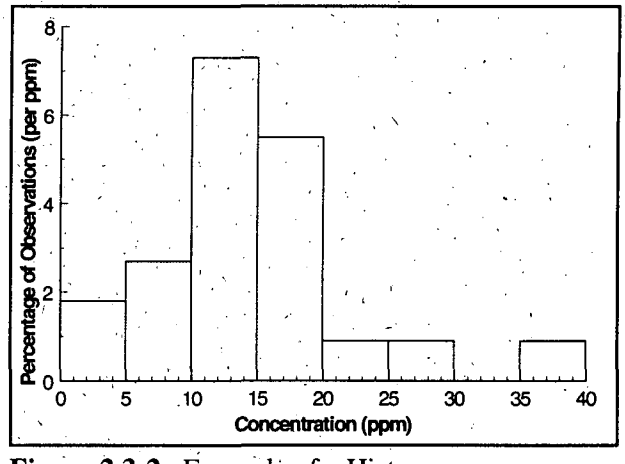


Figure 2.3-2. Example of a Histogram

The histogram and frequency plot provide a means of assessing the symmetry and variability of the data. If the data are symmetric, then the structure of these plots will be symmetric around a central point such as a mean. The histogram and frequency plots will generally indicate if the data are skewed and the direction of the skewness.

Directions for generating a histogram and a frequency plot are contained in Box 2.3-1 and an example is contained in Box 2.3-2. When plotting a histogram for a continuous variable (e.g., concentration), it is necessary to decide on an endpoint convention; that is, what to do with cases that fall on the boundary of a box. With discrete variables, (e.g., family size) the intervals can be centered in between the variables. For the family size data, the intervals can span between 1.5 and 2.5, 2.5 and 3.5, and so on, so that the whole numbers that relate to the family size can be centered within the box. The visual impression conveyed by a histogram or a frequency plot can be quite sensitive to the choice of interval width. The choice of the number of intervals determines whether the histogram shows more detail for small sections of the data or whether the data will be displayed more simply as a smooth overview of the distribution.

Box 2.3-1: Directions for Generating a Histogram and a Frequency Plot

Let $X_1, X_2, ..., X_n$ represent the n data points. To develop a histogram or a frequency plot:

STEP 1: Select intervals that cover the range of observations. If possible, these intervals should have equal widths. A rule of thumb is to have between 7 to 11 intervals. If necessary, specify an endpoint convention, i.e., what to do with cases that fall on interval endpoints.

STEP 2: Compute the number of observations within each interval. For a frequency plot with equal interval sizes, the number of observations represents the height of the boxes on the frequency plot.

STEP 3: Determine the horizontal axis based on the range of the data. The vertical axis for a frequency plot is the number of observations. The vertical axis of the histogram is based on percentages.

STEP 4: For a histogram, compute the percentage of observations within each interval by dividing the number of observations within each interval (Step 3) by the total number of observations.

STEP 5: For a histogram, select a common unit that corresponds to the x-axis. Compute the number of common units in each interval and divide the percentage of observations within each interval (Step 4) by this number. This step is only necessary when the intervals (Step 1) are not of equal widths.

STEP 6: Using boxes, plot the intervals against the results of Step 5 for a histogram or the intervals against the number of observations in an interval (Step 2) for a frequency plot.

Box 2.3-2: Example of Generating a Histogram and a Frequency Plot

Consider the following 22 samples of a contaminant concentration (in ppm): 17.7, 17.4, 22.8, 35.5, 28.6, 17.2 19.1, <4, 7.2, <4, 15.2, 14.7, 14.9, 10.9, 12.4, 12.4, 11.6, 14.7, 10.2, 5.2, 16.5, and 8.9.

STEP 1: This data spans 0 - 40 ppm. Equally sized intervals of 5 ppm will be used: 0 - 5 ppm; 5 - 10 ppm; etc. The endpoint convention will be that values are placed in the highest interval containing the value. For example, a value of 5 ppm will be placed in the interval 5 - 10 ppm instead of 0 - 5 ppm.

STEP 2: The table below shows the number of observations within each interval defined in Step 1.

STEP 3: The horizontal axis for the data is from 0 to 40 ppm. The vertical axis for the frequency plot is from 0 - 10 and the vertical axis for the histogram is from 0% - 10%.

STEP 4: There are 22 observations total, so the number observations shown in the table below will be divided by 22. The results are shown in column 3 of the table below.

STEP 5: A common unit for this data is 1 ppm. In each interval there are 5 common units so the percentage of observations (column 3 of the table below) should be divided by 5 (column 4).

STEP 6: The frequency plot is shown in Figure 2.3-1 and the histogram is shown in Figure 2.3-2.

Interval	# of Obs in Interval	% of Obs in Interval	% of Obs per ppm
0 - 5 ppm	2	9.10	1.8
5 - 10 ppm	3	13.60	2.7
10 - 15 ppm	8	36.36	7.3
15 - 20 ppm	6	27.27	5.5
20 - 25 ppm	1	4.55	0.9
25 - 30 ppm	1	4.55	0.9
30 - 35 ppm	0	0.00	0.0
35 - 40 ppm	1	4.55	0.9

2.3.2 Stem-and-Leaf Plot

The stem-and-leaf plot is used to show both the numerical values themselves and information about the distribution of the data. It is a useful method for storing data in a compact form while, at the same time, sorting the data from smallest to largest. A stem-and-leaf plot can be more useful in analyzing data than a histogram because it not only allows a visualization of the data distribution, but enables the data to be reconstructed and lists the observations in the order of magnitude. However, the stem-and-leaf plot is one of the more subjective visualization techniques because it requires the analyst to make some arbitrary choices regarding a partitioning of the data. Therefore, this technique may require some practice or trial and error before a useful plot can be created. As a result, the stem-and-leaf plot should only be used to develop a picture of the data and its characteristics. Directions for constructing a stem-and-leaf plot are given in Box 2.3-3 and an example is contained in Box 2.3-4.

Each observation in the stem-and-leaf plot consist of two parts: the stem of the observation and the leaf. The stem is generally made up of the leading digit of the numerical values while the leaf is made up of trailing digits in the order that corresponds to the order of magnitude from left to right. The stem is displayed on the vertical axis and the data points make up the leaves. Changing the stem can be accomplished by increasing or decreasing the digits that are used, dividing the groupings of one stem (i.e., all numbers which start with the numeral 6 can be divided into smaller groupings), or multiplying the data by a constant factor (i.e., multiply the data by 10 or 100). Nondetects can be placed in a single stem.

A stem-and-leaf plot roughly displays the distribution of the data. For example, the stem-and-leaf plot of normally distributed data is approximately bell shaped. Since the stem-and-leaf roughly displays the distribution of the data, the plot may be used to evaluate whether the data are skewed or symmetric. The top half of the stem-and-leaf plot will be a mirror image of the bottom half of the stem-and-leaf plot for symmetric data. Data that are skewed to the left will have the bulk of data in the top of the plot and less data spread out over the bottom of the plot.

2.3.3 Box and Whisker Plot

A box and whisker plot or box plot (Figure 2.3-3) is a schematic diagram useful for visualizing important statistical quantities of the data. Box plots are useful in situations where it is not necessary or feasible to portray all the details of a distribution. Directions for generating a box and whiskers plot are contained in Box 2.3-5, and an example is contained in Box 2.3-6.

A box and whiskers plot is composed of a central box divided by a line and two lines extending out from the box called whiskers. The length of the central box indicates the spread of the bulk of the data (the central 50%) while the length of the whiskers show how stretched the tails of the distribution are. The width of the box has no particular meaning; the plot can be made quite narrow without affecting its visual impact. The sample median is displayed as a line through the box and the sample mean is displayed using a '+' sign. Any unusually small or large data points are displayed by a '*' on the plot. A box and whiskers plot can be used to assess the symmetry of the data. If the distribution is symmetrical, then the box is divided in two equal halves by the median, the whiskers will be the same length and the number of extreme data points will be distributed equally on either end of the plot.

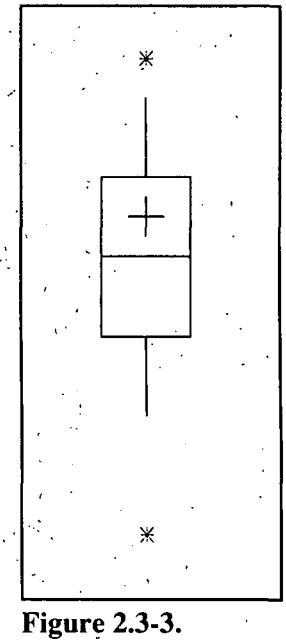


Figure 2.3-3. Example of a Box and Whisker Plot

Box 2.3-3: Directions for Generating a Stem and Leaf Plot

Let $X_1, X_2, ..., X_n$ represent the n data points. To develop a stem-and-leaf plot, complete the following steps:

STEP 1: Arrange the observations in ascending order. The ordered data is usually labeled (from smallest to largest) $X_{(1)}, X_{(2)}, ..., X_{(n)}$.

STEP 2: Choose either one or more of the leading digits to be the stem values. As an example, for the value 16, 1 could be used as the stem as it is the leading digit.

STEP 3: List the stem values from smallest to largest at the left (along a vertical axis). Enter the leaf (the remaining digits) values in order from lowest to highest to the right of the stem. Using the value 16 as an example, if the 1 is the stem then the 6 will be the leaf.

Box 2.3-4: Example of Generating a Stem and Leaf Plot

Consider the following 22 samples of trifluorine (in ppm): 17.7, 17.4, 22.8, 35.5, 28.6, 17.2 19.1, <4, 7.2, <4, 15.2, 14.7, 14.9, 10.9, 12.4, 12.4, 11.6, 14.7, 10.2, 5.2, 16.5, and 8.9.

STEP 1: Arrange the observations in ascending order: <4, <4, 5.2, 7.7, 8.9, 10.2, 10.9, 11.6, 12.4, 12.4, 14.7, 14.7, 14.9, 15.2, 16.5, 17.4, 17.7, 19.1, 22.8, 28.6, 35.5.

STEP 2: Choose either one or more of the leading digits to be the stem values. For the above data, using the first digit as the stem does not provide enough detail for analysis. Therefore, the first digit will be used as a stem; however, each stem will have two rows, one for the leaves 0 - 4, the other for the leaves 5 - 9.

STEP 3: List the stem values at the left (along a vertical axis) from smallest to largest. Enter the leaf (the remaining digits) values in order from lowest to highest to the right of the stem. The first digit of the data was used as the stem values; however, each stem value has two leaf rows.

```
0 (0, 1, 2, 3, 4)   | <4   <4
0 (5, 6, 7, 8, 9)   | 5.2  7.7  8.9
1 (0, 1, 2, 3, 4)   | 0.2  0.9  1.6  2.4  2.4  4.7  4.7  4.9
1 (5, 6, 7, 8, 9)   | 5.2  6.5  7.4  7.7  9.1
2 (0, 1, 2, 3, 4)   | 2.8
2 (5, 6, 7, 8, 9)   | 8.6
3 (0, 1, 2, 3, 4)   |
3 (5, 6, 7, 8, 9)   | 5.5
```

Note: If nondetects are present, place them first in the ordered list, using a symbol such as <L. If multiple detection limits were used, place the nondetects in increasing order of detection limits, using symbols such as <L1, <L2, etc. If the first stem extends from zero to a value above the detection limit, then nondetects can be placed in this interval, as shown in the example above. Otherwise, special intervals dedicated to nondetects can be used.

Box 2.3-5: Directions for Generating a Box and Whiskers Plot

STEP 1: Set the vertical scale of the plot based on the maximum and minimum values of the data set. Select a width for the box plot keeping in mind that the width is only a visualization tool. Label the width w; the horizontal scale then ranges from -½W to ½W.

STEP 2: Compute the upper quartile (Q(.75), the 75th percentile) and the lower quartile (Q(.25), the 25th percentile) using Box 2.2-1. Compute the sample mean and median using Box 2.2-2. Then, compute the interquartile range (IQR) where IQR = Q(.75) - Q(.25).

STEP 3: Draw a box through points (-½W, Q(.75)), (-½W, Q (.25)), (½W, Q(.25)) and (½W, Q(.75)). Draw a line from (½W, Q(.5)) to (-½W, Q(.5)) and mark point (0, X) with (+).

STEP 4: Compute the upper end of the top whisker by finding the largest data value X less than Q(.75) + 1.5(Q(.75) - Q(.25)). Draw a line from (0, Q(.75)) to (0, X).

Compute the lower end of the bottom whisker by finding the smallest data value Y greater than Q(.25) - 1.5(Q(.75) - Q(.25)). Draw a line from (0, Q(.25)) to (0, Y).

STEP 5: For all points X* > X, place an asterisk (*) at the point (0, X*).

For all points Y* < Y, place an asterisk (*) at the point (0, Y*).

Box 2.3-6. Example of a Box and Whiskers Plot

Consider the following 22 samples of trifluorine (in ppm) listed in order from smallest to largest: 4.0, 6.1, 9.8, 10.7, 10.8, 11.5, 11.6, 12.4, 12.4, 14.6, 14.7, 14.7, 16.5, 17, 17.5, 20.6, 20.8, 25.7, 25.9, 26.5, 32.0, and 35.5.

STEP 1: The data ranges from 4.0 to 35.5 ppm. This is the range of the vertical axis. Arbitrarily, a width of 4 will be used for the horizontal axis.

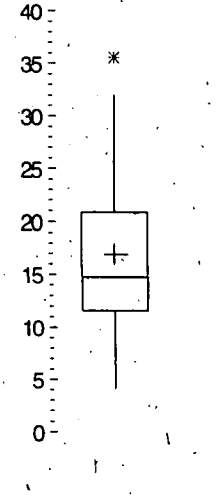
STEP 2: Using the formulas in Box 2.2-2, the sample mean = 16.87 and the median = 14.70. Using Box 2.2-1, Q(.75) = 20.8 and Q(.25) = 11.5. Therefore, IQR = 20.8 - 11.5 = 9.3.

STEP 3: In the figure, a box has been drawn through points (-2, 20.8), (-2, 11.5), (2, 11.5), (2, 20.8). A line has been drawn from (-2 , 14.7) to (2, 14.7), and the point (0, 16.87) has been marked with a '+' sign.

STEP 4: Q(.75) + 1.5(9.3) = 34.75. The closest data value to this number, but less than it, is 32.0. Therefore, a line has been drawn in the figure from (0, 20.8) to (0, 32.0).

Q(.25) - 1.5(9.3) = -2.45. The closest data value to this number, but greater than it, is 4.0. Therefore, a line has been drawn in the figure from (0, 4) to (0, 11.5).

STEP 5: There is only 1 data value greater than 32.0 which is 35.5. Therefore, the point (0, 35.5) has been marked with an asterisk. There are no data values less than 4.0.



2.3.4 Ranked Data Plot

A ranked data plot is a useful graphical representation that is easy to construct, easy to interpret, and makes no assumptions about a model for the data. The analyst does not have to make any arbitrary choices regarding the data to construct a ranked data plot (such as cell sizes for a histogram). In addition, a ranked data plot displays every data point; therefore, it is a graphical representation of the data instead of a summary of the data. Directions for developing a ranked data plot are given in Box 2.3-7 and an example is given in Box 2.3-8.

A ranked data plot is a plot of the data from smallest to largest at evenly spaced intervals (Figure 2.3-4). This graphical representation is very similar to the quantile plot described in section 2.3.5. A ranked data plot is marginally easier to generate than a quantile plot; however, a ranked data plot does not contain as much information as a quantile plot. Both plots can be used to determine the density of the data points and the skewness of the data; however, a quantile plot contains information on the quartiles of the data whereas a ranked data plot does not.

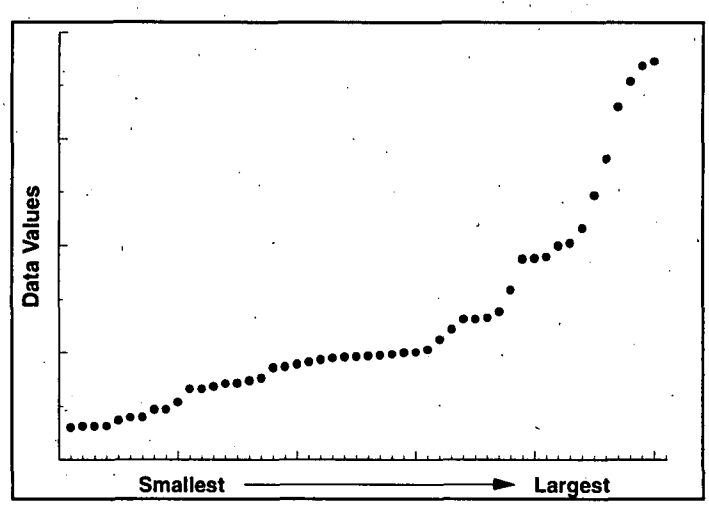


Figure 2.3-4. Example of a Ranked Data Plot

A ranked data plot can be used to determine the density of the data values, i.e., if all the data values are close to the center of the data with relatively few values in the tails or if there is a large amount of values in one tail with the rest evenly distributed. The density of the data is displayed through the slope of the graph. A large amount of data values has a flat slope, i.e., the graph rises slowly. A small amount of data values has a large slope, i.e., the graph rises quickly. Thus the analyst can determine where the data lie, either evenly distributed or in large clusters of points. In Figure 2.3-4, the data rises slowly up to a point where the slope increases and the graph rises relatively quickly. This means that there is a large amount of small data values and relatively few large data values.

A ranked data plot can be used to determine if the data are skewed or if they are symmetric. A ranked data plot of data that are skewed to the right extends more sharply at the top giving the graph a convex shape. A ranked data plot of data that are skewed to the left increases sharply near the bottom giving the graph a concave shape. If the data are symmetric, then the top portion of the graph will stretch to upper right corner in the same way the bottom portion of the graph stretches to lower left, creating a s-shape. Figure 2.3-4 shows a ranked data plot of data that are skewed to the right.

Box 2.3-7: Directions for Generating a Ranked Data Plot

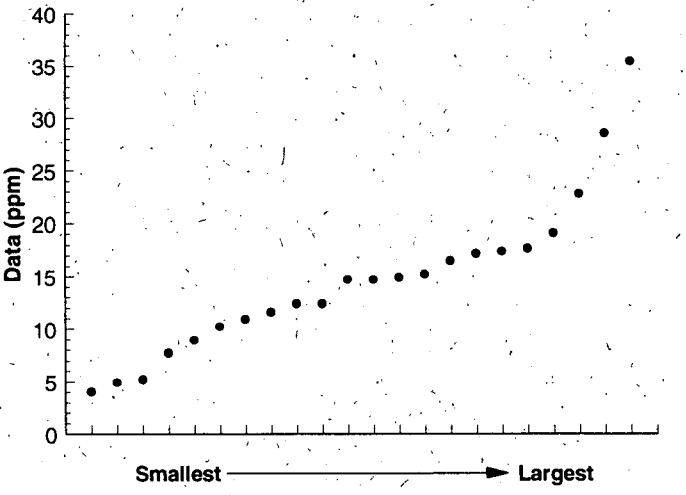
Let $X_1, X_2, ..., X_n$ represent the n data points. Let $X_{(l)}$, for l=1 to n, be the data listed in order from smallest to largest so that $X_{(1)}$ (l = 1) is the smallest, $X_{(2)}$ (l = 2) is the second smallest, and $X_{(n)}$ (l = n) is the largest. To generate a ranked data plot, plot the ordered X values at equally spaced intervals along the horizontal axis.

Box 2.3-8: Example of Generating a Ranked Data Plot

Consider the following 22 samples of triflourine (in ppm): 17.7, 17.4, 22.8, 35.5, 28.6, 17.2 19.1, 4.9, 7.2, 4.0, 15.2, 14.7, 14.9, 10.9, 12.4, 12.4, 11.6, 14.7, 10.2, 5.2, 16.5, and 8.9. The data listed in order from smallest to largest $X_{(l)}$ along with the ordered number of the observation (l) are:

l	$X_{(l)}$	l	$X_{(l)}$
1	4.0	12	14.7
2	4.9	13	14.9
3	5.2	14	15.2
4	7.7	15	16.5
5	8.9	16	17.2
6	10.2	17	17.4
7	10.9	18	17.7
8	11.6	19	19.1
9	12.4	20	22.8
10	12.4	21	28.6
11	14.7	22	35.5

A ranked data plot of this data is a plot of the pairs (l, $X_{(l)}$). This plot is shown below:



2.3.5 Quantile Plot

A quantile plot (Figure 2.3-5) is a graphical representation of the data that is easy to construct, easy to interpret, and makes no assumptions about a model for the data. The analyst does not have to make any arbitrary choices regarding the data to construct a quantile plot (such as cell sizes for a histogram). In addition, a quantile plot displays every data point; therefore, it is a graphical representation of the data instead of a summary of the data.

A quantile plot is a graph of the quantiles (section 2.2.1) of the data. The basic quantile plot is visually identical to a ranked data plot except its horizontal axis varies from 0.0 to 1.0, with each point plotted according to the fraction of the points it exceeds. This allows the addition of vertical lines indicating the quartiles or, any other quantiles of interest. Directions for developing a quantile plot are given in Box 2.3-9 and an example is given in Box 2.3-10.

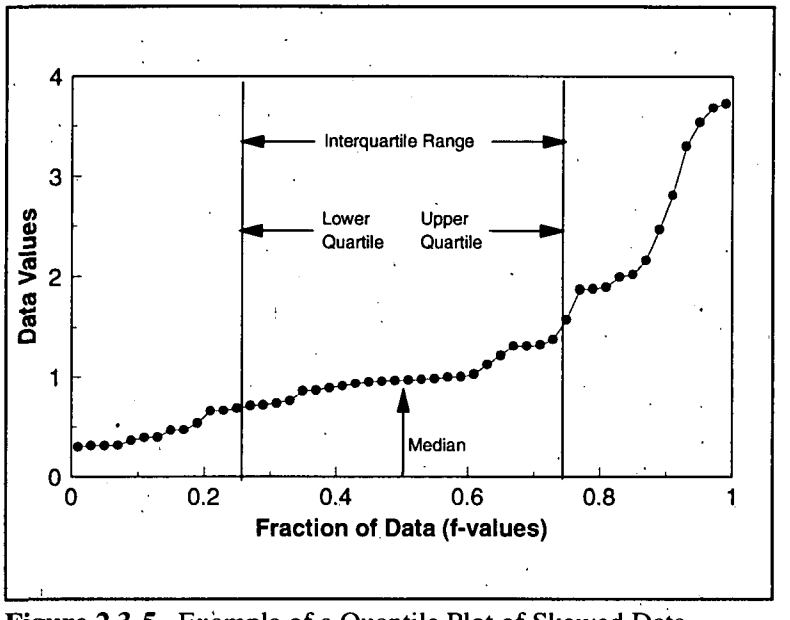


Figure 2.3-5. Example of a Quantile Plot of Skewed Data

A quantile plot can be used to read the quantile information such as the median, quartiles, and the interquartile range. In addition, the plot can be used to determine the density of the data points, e.g., are all the data values close to the center with relatively few values in the tails or are there a large amount of values in one tail with the rest evenly distributed? The density of the data is displayed through the slope of the graph. A large amount of data values has a flat slope, i.e., the graph rises slowly. A small amount of data values has a large slope, i.e., the graph rises quickly. A quantile plot can be used to determine if the data are skewed or if they are symmetric. A quantile plot of data that are skewed to the right is steeper at the top right than the bottom left, as in Figure 2.3-5. A quantile plot of data that are skewed to the left increases sharply near the bottom left of the graph. If the data are symmetric then the top portion of the graph will stretch to the upper right corner in the same way the bottom portion of the graph stretches to the lower left, creating an s-shape.

Box 2.3-9: Directions for Generating a Quantile Plot

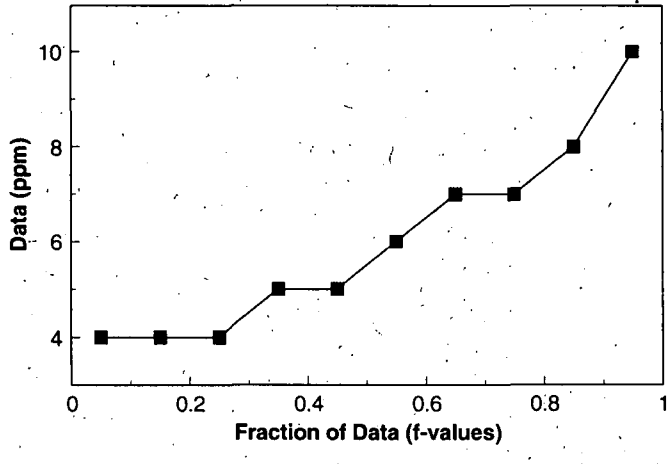
Let $X_1, X_2, ..., X_n$ represent the n data points. To obtain a quantile plot, let $X_{(i)}$, for
i = 1 to n, be the data listed in order from smallest to largest so that $X_{(1)}$ (i = 1) is the smallest, $X_{(2)}$ (i = 2) is the second smallest, and $X_{(n)}$ (i = n) is the largest. For each i, compute the fraction $f_i = (i - 0.5)/n$. The quantile plot is a plot of the pairs
$(f_i, X_{(i)})$, with straight lines connecting consecutive points.

Box 2.3-10: Example of Generating a Quantile Plot

Consider the following 10 data points: 4 ppm, 5 ppm, 6 ppm, 7 ppm, 4 ppm, 10 ppm, 4 ppm, 5 ppm, 7 ppm, and 8 ppm. The data ordered from smallest to largest, $X_{(i)}$, are shown in the first column of the table below and the ordered number for each observation, i, is shown in the second column. The third column displays the values f_i for each i where $f_i = (i - 0.5)/n$.

$X_{(i)}$	i	f_i	$X_{(i)}$	i	f_i
4	1	0.05	6	6	0.55
4	2	0.15	7	7	0.65
4	3	0.25	7	8	0.75
5	4	0.35	8	9	0.85
5	5	0.45	10	10	0.95

The pairs $(f_i, X_{(i)})$ are then plotted to yield the following quantile plot:



Note that the graph curves upward; therefore, the data appear to be skewed to the right.

2.3.6 Normal Probability Plot (Quantile-Quantile Plot)

There are two types of quantile-quantile plots or q-q plots. The first type, an empirical quantile-quantile plot (section 2.3.7.4), involves plotting the quantiles of two data variables against each other. The second type of a quantile-quantile plot, a theoretical quantile-quantile plot, involves graphing the quantiles of a set of data against the quantiles of a specific distribution. The following discussion will focus on the most common of these plots for environmental data, the normal probability plot (the normal q-q plot); however, the discussion holds for other q-q plots. The normal probability plot is used to roughly determine how well the data set is modeled by a normal distribution. Formal tests are contained in Chapter 4, section 2. Directions for developing a normal probability plot are given in Box 2.3-11 and an example is given in Box 2.3-12.

A normal probability plot is the graph of the quantiles of a data set against the quantiles of the normal distribution using normal probability graph paper (Figure 2.3-6). If the graph is linear, the data may be normally distributed. If the graph is not linear, the departures from linearity give important information about how the data distribution deviates from a normal distribution.

If the graph of the normal probability plot is not linear, the graph may be used to determine the degree of symmetry (or asymmetry) displayed by the data. If the data are skewed to the right, the graph is convex. If the data are skewed to the left, the graph is concave. If the data in the upper tail fall above and the data in the lower tail fall below the quartile line, the data are too slender to be well modeled by a normal distribution, i.e., there are fewer values in the tails of the data set than what is expected from a normal distribution. If the data in the upper tail fall below and the data in the lower tail fall above the quartile line, then the tails of the data are too heavy to be well modeled using a normal distribution, i.e., there are more values in the tails of the data than what is expected from a normal distribution. A normal probability plot can be used to identify potential outliers. A data value (or a few data values) much larger or much smaller than the rest will cause the other data values to be compressed into the middle of the graph, ruining the resolution.

Box 2.3-11: Directions for Constructing a Normal Probability Plot

Let $X_1, X_2, ..., X_n$ represent the n data points.

STEP 1: For each data value, compute the absolute frequency, AF_i. The absolute frequency is the number of times each value occurs. For distinct values, the absolute frequency is 1. For non-distinct observations, count the number of times an observation occurs. For example, consider the data 1, 2, 3, 3. The absolute frequency of value 1 is 1 and the absolute frequency of value 2 is 1. The absolute frequency of value 3 is 2 since 3 appears 2 times in the data set.

STEP 2: Compute the cumulative frequencies, CF_i. The cumulative frequency is the number of data points that are less than or equal to X_i, i.e., $CF_i = \sum_{j=1}^{i} AF_j$. Using the data given in step 2, the cumulative frequency for value 1 is 1, the cumulative frequency for value 2 is 2 (1+1), and the cumulative frequency for value 3 is 4 (1+1+2).

STEP 3: Compute $Y_i = 100\ x\ \frac{CF_i}{(n+1)}$ and plot the pairs (Y_i, X_i) using normal probability paper (Figure 2.3-6). If the graph of these pairs approximately forms a straight line, then the data are probably normally distributed. Otherwise, the data may not be normally distributed.

Box 2.3-12: Example of Normal Probability Plot

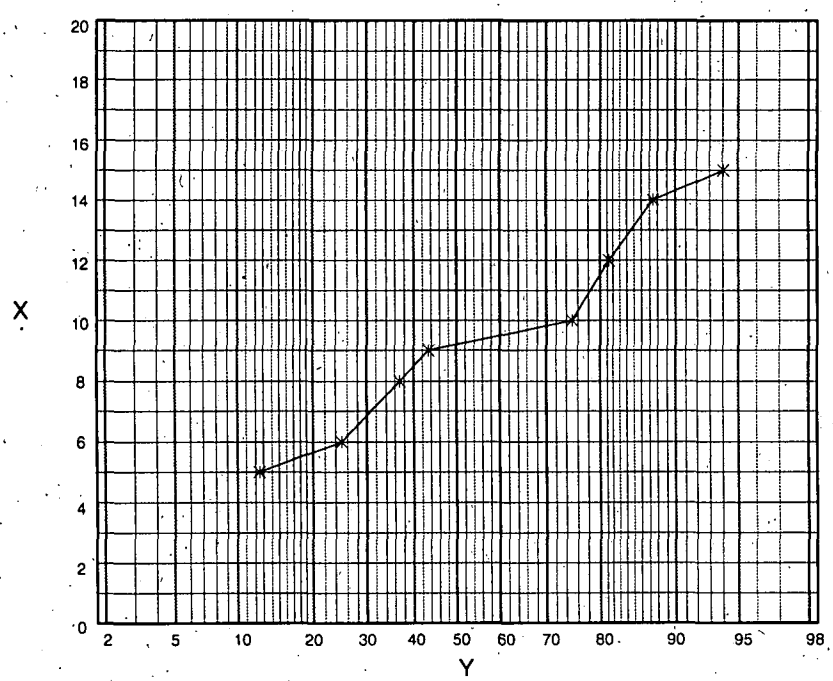
Consider the following 15 data points: 5, 5, 6, 6, 8, 8, 9, 10, 10, 10, 10, 10, 12, 14, and 15.

STEP 1: Because the value 5 appears 2 times, its absolute frequency is 2. Similarly, the absolute frequency of 6 is 2, of 8 is 2, of 9 is 1, of 10 is 5, etc. These values are shown in the second column of the table below.

STEP 2: The cumulative frequency of the data value 8 is 6 because there are 2 values of 5, 2 values of 6, and 2 values of 8. The cumulative frequencies are shown in the 3rd column of the table.

STEP 3: The values $Y_i = 100\ x\ (\frac{CF_i}{n+1})$ for each data point are shown in column 4 of the table below. A plot of these pairs (Y_i, X_i) using normal probability paper is also shown below.

I	Individual X_i	Absolute Frequency AF_i	Cumulative Frequency CF_i	Y_i
1	5	2	2	12.50
2	6	2	4	25.00
3	8	2	6	37.50
4	9	1	7	43.75
5	10	5	12	75.00
6	12	1	13	81.25
7	14	1	14	87.50
8	15	1	15	93.75



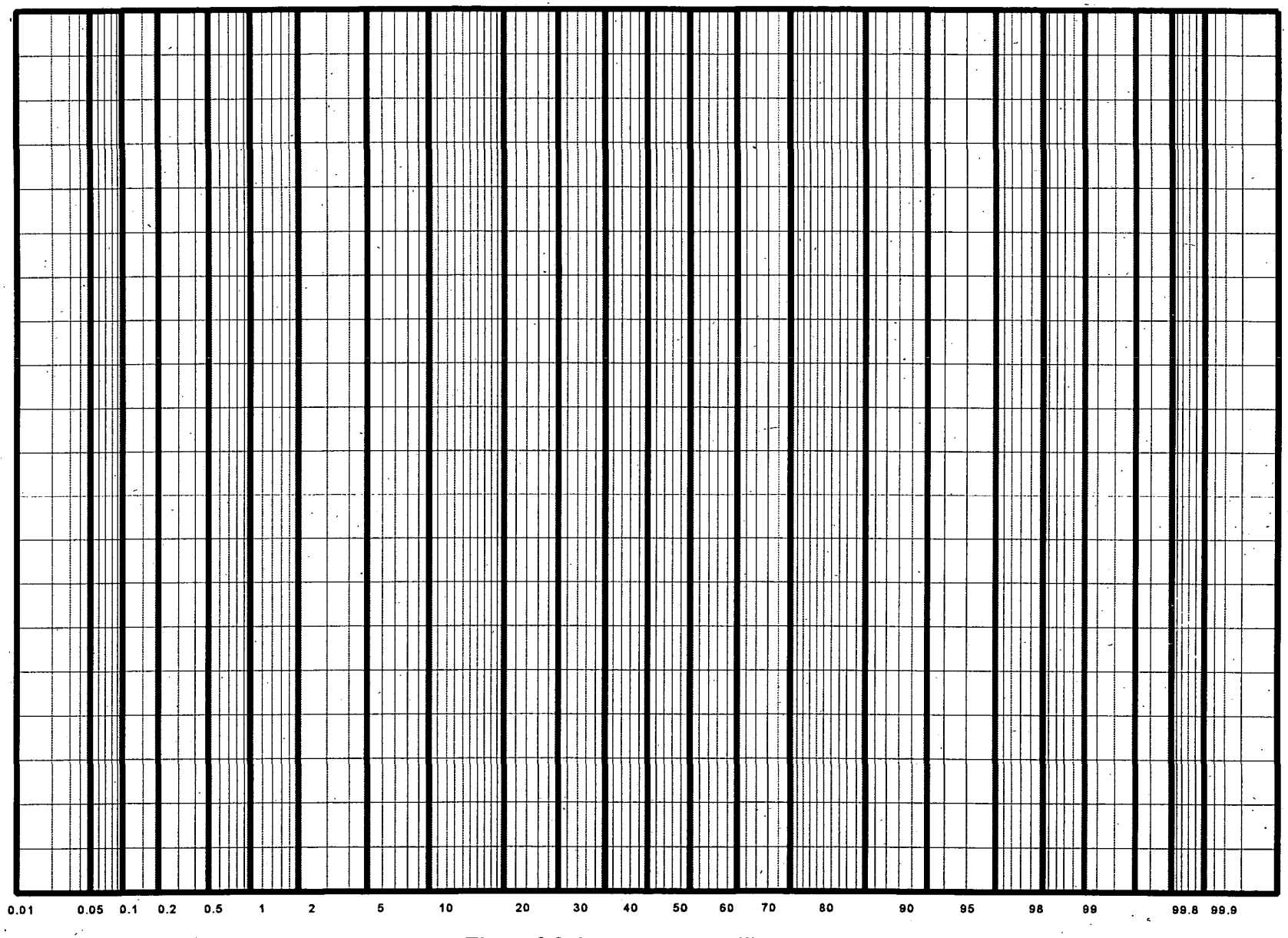


Figure 2.3-6. Normal Probability Paper

APPENDIX L

DATA MANAGEMENT (RESERVED)

This appendix will be completed at a later date.