

Guidance for  
Planning for Data Collection  
in Support of  
Environmental Decision Making  
Using the Data Quality Objectives Process

EPA QA/G-4

Quality Assurance Management Staff  
U S Environmental Protection Agency  
Washington, DC 20460

\*\*\* INTERIM FINAL \*\*\*

**ADMIN RECORD**

DOCUMENT CLASSIFICATION  
REVIEW WAIVER PER  
CLASSIFICATION OFFICE

FILED

FOREWORD

Chapter

- Forward
- List of Figures
- Introduction
- 1 Step 1
- 2 Step 2
- 3 Step 3
- 4 Step 4
- 5 Step 5
- 6 Step 6
- 7 Step 7
- References

The U.S. Environmental Protection Agency (EPA) has developed the Data Objectives (DQO) Process as an important tool for project managers and planners to determine the type, quantity, and quality of data needed to make decisions. Data sampling or monitoring activities are used extensively in problem definition, rule enforcement decisions. These activities are supported through implementation of mandatory Agency-wide Quality System which requires all organizations to develop management processes and structures for assuring that the data collected meet needed and expected quality for their desired use. The DQO Process has emerged as an effective means by which managers and technical staff may plan and design more and more timely sampling and analysis programs. The DQO Process relies heavily on customer and supplier communication and understanding to determine the types of data needed, the quantity of data needed, and the quality of the data necessary, and to assure that the customer (whether internal or external) is satisfied with the results.

The purpose of this document is to provide general guidance to organizations developing data quality criteria and performance specifications for data operations. This guidance assumes that an appropriate Quality System has been established and is being used.

This document is one of a series of quality management requirements and guidance documents that have been prepared by the Quality Assurance Management Staff (QAMS) to assist users in implementing the Agency mandatory Quality System.

Appendices

- A Beyond the Data Quality Objectives
- B DQO Checklist
- C Glossary

## List of Figures

- 1: The Data Quality Objectives Process . . . . .
- 2 Repeated Application of the DQO Process . . . . .
- 6-1 An Example of a Design Performance Goal Diagram -  
Baseline Condition Parameter Exceeds Action Level . . . . .
- 6-2 An Example of a Design Performance Goal Diagram -  
Baseline Condition Parameter is Less Than Action Level . . . . .
- 7-1 An Example of a Power Curve . . . . .
- B-1. Design Performance Goal Diagram for Cadmium Compliance Testing . . . . .

## List of Tables

- 6-1 Decision Error Limits Table Corresponding to Figure 6-1
- 6-2 Decision Error Limits Table Corresponding to Figure 6-2 . . . . .

## INTRODUCTION

Each year the U.S. Environmental Protection Agency (EPA) and the regulated community spend approximately \$5 billion collecting environmental data for scientific research, regulatory decision making, and regulatory compliance. While these activities are necessary for effective environmental protection, it is the goal of EPA and the regulated community to minimize expenditures related to data collection by eliminating unnecessary, duplicative, or overly precise data. At the same time, they would like to collect data of sufficient quantity and quality to support defensible decision making. The most efficient way to accomplish both of these goals is to begin by ascertaining the type, quality, and quantity of data necessary to address the problem before the study begins. To facilitate this determination, the Quality Assurance Management Staff (QAMS) of EPA has developed the Data Quality Objectives (DQO) Process, a systematic planning tool based on the Scientific Method for determining the type, quantity, and quality of data that will be sufficient and appropriate for the data's intended use. By using the DQO Process to plan environmental data collection efforts, EPA can improve the effectiveness, efficiency, and defensibility of decisions in a resource-effective manner.

What are DQOs? DQOs are qualitative and quantitative statements derived from the outputs of each step of the DQO Process that:

- 1) Clarify the study objective,
- 2) Define the most appropriate type of data to collect,
- 3) Determine the most appropriate conditions from which to collect the data, and
- 4) Specify acceptable levels of decision errors that will be used as the basis for establishing the quantity and quality of data needed to support the decision.

The DQOs are then used to develop a scientific and resource-effective sampling design.

What is the DQO Process? The DQO Process is a series of planning steps based on the Scientific Method that is designed to ensure that the type, quantity, and quality of environmental data used in decision making are appropriate for the intended application.

The DQO Process was developed by EPA to help Agency personnel avoid collecting data that are inconsequential to decision making. The Process allows the decision makers to define their data requirements and acceptable levels of decision errors during planning before they collect data. Application of the DQO Process should result in data collection designs that will yield results of appropriate quality for defensible decision making.

The DQO Process consists of seven distinct steps, as shown in Figure 1. The DQO Process is also iterative, the outputs from one step may influence prior steps and cause them to be redefined. This will ultimately lead to a more efficient data collection design. Each of the seven steps are described briefly below. A more detailed description can be found in the subsequent chapters of this guidance.

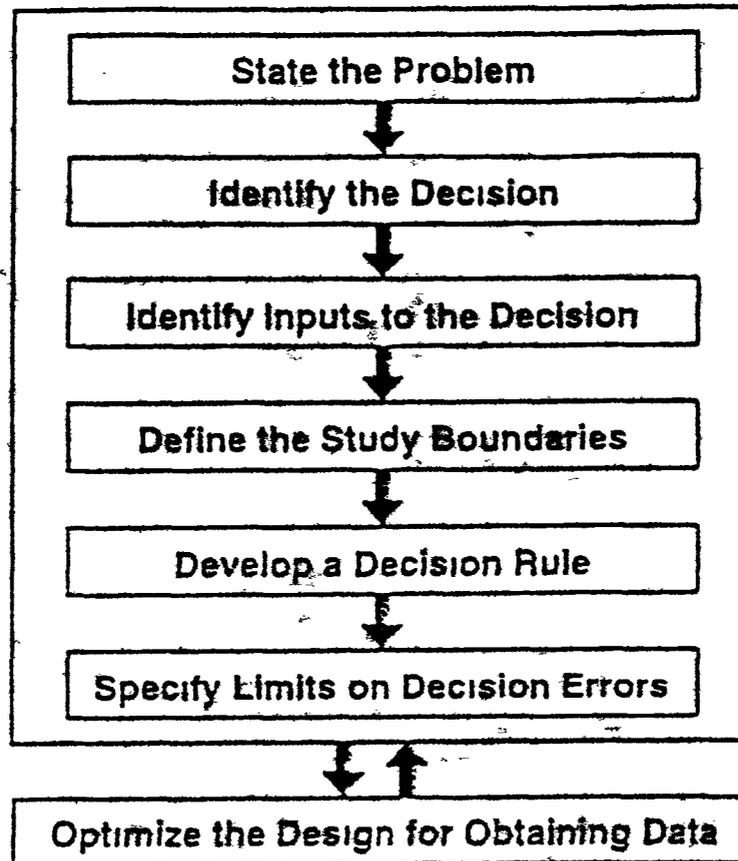


Figure 1 The Data Quality Objectives Process

- Step 1. State the Problem Concisely describe the problem to be studied. Review prior studies and existing information to gain an acceptable understanding of the problem.
- Step 2. Identify the Decision Identify the decision that will solve the problem using new data.
- Step 3. Identify the Inputs to the Decision Identify the information that needs to be learned and the measurements that need to be taken in order to resolve the decision.

- Step 4 Define the Study Boundaries Specify the conditions (time periods and situations) to which decisions will apply and within which the data should be collected
- Step 5 Develop a Decision Rule Integrate the outputs from previous steps into an if then statement that defines the conditions that would cause the decision maker to choose among alternative actions
- Step 6 Specify Acceptable Limits on Decision Errors Define the decision maker's acceptable decision error rates<sup>1</sup> based on a consideration of the consequences of making an incorrect decision
- Step 7 Optimize the Design Evaluate information from the previous steps and generate alternative sampling designs. Choose the most resource-efficient design that meets all DQOs

#### What is the value of using the DQO Process?

- The DQO Process is a planning tool that can save resources by making data collection operations more efficient and cost-effective. Good planning will streamline the study process and increase the likelihood of collecting appropriate and useful data.
- The structure of the DQO Process provides a convenient way to document activities and decisions and to communicate the study design to others.
- The DQO Process enables data users and relevant technical experts to participate in planning and to specify their particular needs prior to data collection. The DQO Process also fosters communication among all participants, one of the central tenets of quality management practices.
- The DQO Process provides a method for defining the quality of data in relation to the intended use of the data. This is done by considering the consequences of decision errors and then placing limits on the likelihood of those decision errors. A statistical sampling design then can be generated to provide the most efficient method for controlling decision errors and satisfying the DQOs.
- The DQO Process helps to focus studies by encouraging data users to clarify vague objectives and to limit the number of decisions that will be made.

---

<sup>1</sup> A decision error rate is the probability of making an incorrect decision based on data that inaccurately estimate the state of nature.

**When should the DQO Process be used?** The DQO Process should be used during the planning stage of any study that requires data collection, before the data are collected. In general, EPA's policy is to use the DQO Process to plan all data collection efforts that will require or result in a substantial commitment of resources. The Quality Management Plans (QMPs) of the Agency's National Program Offices, Regional Offices, and Research and Development organizations will specify which studies require DQOs

Even when the entire DQO Process is not utilized, it can and should be used as a guide for planning a study. This will help ensure that the data will be useful for supporting a decision

**Who participates in the DQO Process?** A DQO planning team generally consists of senior program staff, technical experts, senior managers, someone with statistical expertise, and a Quality Assurance (QA)/Quality Control (QC) advisor, such as a QA Manager. It is important that all of these people, including managers, participate (or stay informed) from the beginning of the DQO Process so that it can proceed efficiently

**What is a study design?** A study design specifies the final configuration of the environmental monitoring effort to satisfy the DQOs. It designates the types of samples or monitoring information to be collected; where, when, and under what conditions they should be collected, what variables are to be measured, and the QA/QC components that ensure acceptable sampling error and measurement error to meet the decision error rates specified in the DQOs

**Where does the DQO Process fit into EPA's Quality System?** DQOs are used to arrive at the QC requirements for data collection, sampling, and analysis. These QC requirements become part of a Quality Assurance Project Plan (QAPP)

**What projects are covered by this guidance?** This guidance document applies to all efforts to collect environmental data for Agency decisions and monitoring studies where the results will be used to make decisions. This guidance may not apply to data collection activities that will not lead to immediate actions or that are exploratory in nature.

**How is this guidance structured?** This guidance contains eight chapters, three appendices and a list of references. Each of the remaining chapters describes one of the seven steps of the DQO Process. Each chapter is divided into three sections as follows

- (1) **Outputs from this step** - This section identifies and describes the products from the DQO Process step described
- (2) **Background** - This section provides background information on the DQO Process step including the rationale for the activities in the step
- (3) **Activities** - This section describes the activities comprising the DQO Process step, including how inputs to the step are used.

Appendix A provides a brief overview of the Data Quality Assessment (DQA) Process that is used to evaluate data once it has been collected. Appendix B is a case study application of the DQO planning process to an environmental problem. Appendix C provides a glossary of terms used in this guidance.

How should this guidance be used? This guidance should be used as a tool to structure the planning activities for collecting environmental data. It should be used to organize meetings, focus the collection of background information, and facilitate communication among experts from many different fields and between the technical experts, the program managers, and the decision makers.

The DQO Process consists of seven steps. In most cases, each successive step derives information from the previous ones, thus, each step should be completed in the order that it appears in the guidance. The DQO Process is iterative, however, so it may be useful to refine the outputs from previous steps. Iteration is encouraged since it leads to a more focused study with a greater chance of meeting its objectives. Above all, every step should be completed before data collection begins.

The DQO Process can be used repeatedly throughout the life cycle of a project to help make separate decisions. Often, the decisions that are made early in the study will be less serious in nature and will require a limited evaluation effort. In some cases, these decisions will be purely qualitative rather than quantitative. As the study nears conclusion and the possibility of making a decision error becomes more crucial, however, the level of effort needed to resolve a decision generally will become greater. Figure 2 illustrates this point graphically.

# REPEATED APPLICATION OF THE DQO PROCESS FOR SEPARATE STUDY DECISIONS

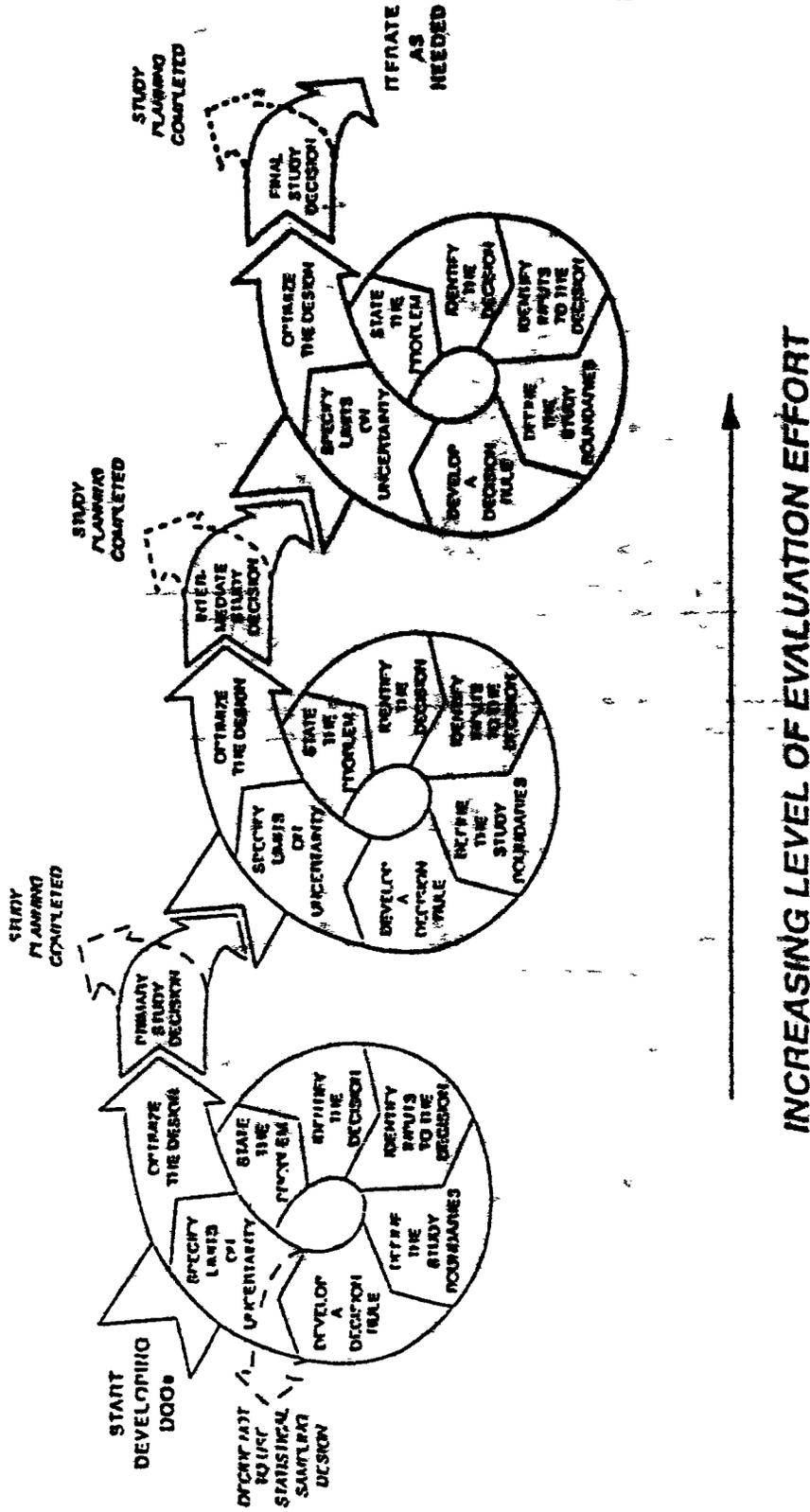
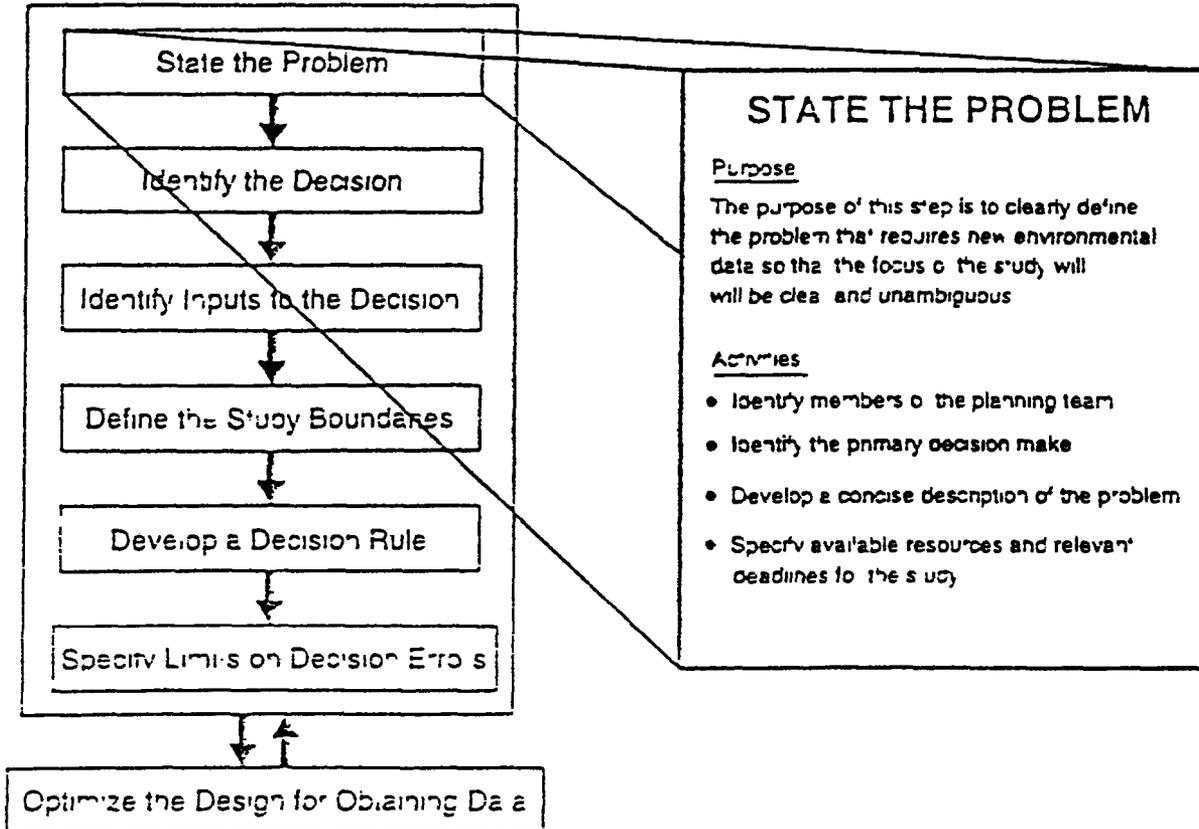


Figure 2. Repeated Application of the DQO Process

## CHAPTER 1

### STEP 1 STATE THE PROBLEM

#### THE DATA QUALITY OBJECTIVES PROCESS



#### Outputs from this step

- A concise description of the problem
- A list of the planning team members and identification of the decision maker
- A summary of available resources and relevant deadlines for the study

## Background

The purpose of this step is to assemble a planning team of study experts, data users and managers who will help plan the study design using the DQO Process, then clearly define the problem that is being evaluated so that the study will be focused and unambiguous.

## Activities

**Identify members of the planning team** The planning team is the group that will develop DQOs for the study. The team should include representatives from all groups who will be involved in the project, including but not limited to samplers, chemists and other scientists and engineers, modelers, technical project managers, administrative and executive managers, QA/QC experts (such as a QA Manager), data users and decision makers. A reasonable effort should be made to include any decision makers who may use the study findings later. A statistician (or someone knowledgeable and experienced with environmental statistical design) should also be included on this team.

**Identify the primary decision maker of the planning team and define each member's role and responsibility during the DQO Process** The planning team generally has a leader, referred to as the 'decision maker.' The decision maker has the ultimate authority for making final decisions based on the recommendations of the planning team. The decision maker is often the person with the most authority over the study, and is responsible for assigning the roles and responsibilities to the planning team members.

## **Develop a concise description of the problem**

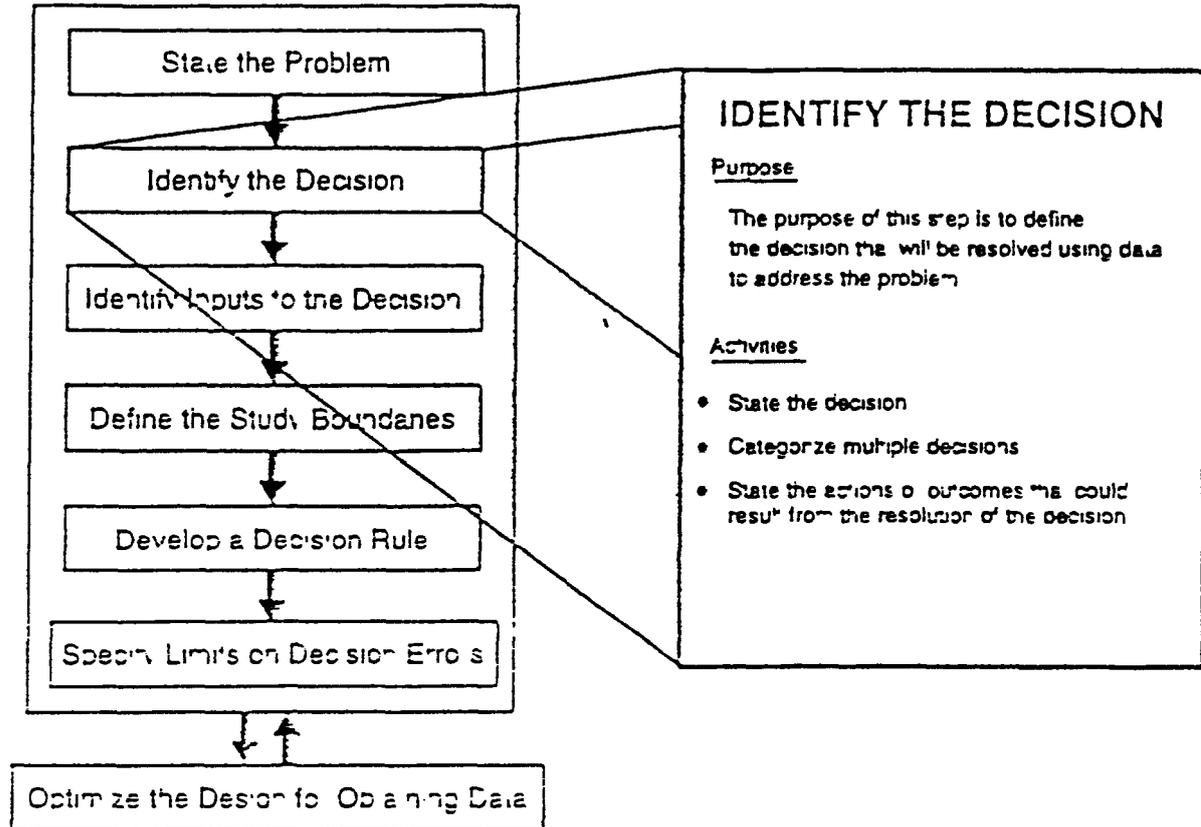
- Describe the problem as it is currently understood by summarizing existing information and explaining the conflicts or uncertainties that the study seeks to resolve.
- Conduct literature searches and explore ongoing studies to ensure that the problem is correctly defined and has not previously been solved. Organize and review relevant information, including preliminary studies, and indicate the source and reliability of the information.
- If the problem is complex, consider breaking it into more manageable pieces. Identify those pieces that could be addressed by separate studies. Assign priorities to each piece of the problem.
- Describe initial ideas or approaches for resolving the problem. For example, one approach might involve modeling, another might be empirical (involving actual measurements) and a third might be a combination of these two approaches.

**Specify the available resources and relevant deadlines for the study** Include the anticipated budget, available personnel and contractual vehicles (if applicable). Also specify any deadlines for completion of the study.

## CHAPTER 2

## STEP 2 IDENTIFY THE DECISION

## THE DATA QUALITY OBJECTIVES PROCESS

Outputs from this step

- A statement of the decision that must be resolved using data in order to address or solve the problem
- A list of possible actions or outcomes that would result from each resolution of the decision statement

## Background

The goal of this step is to define a decision statement that will be resolved using data in order to address the problem identified in Step 1 **State the Problem**. The decision statement (the decision) is the pronouncement of what question must be resolved or answered using the data. An example of a decision is "Is the concentration of contaminants above or below the regulatory standard?" Stating the decision will help focus the efforts of the planning team onto the common objective of the study.

In addition to the decision, the actions or outcomes that would result from the resolution of the decision will also be defined in this step. There are at least two alternative actions or outcomes that are associated with each decision (e.g., taking action or not taking action).

If the decision is not obvious, then it may be helpful to think about the alternative actions in order to identify the decision. If specific actions cannot be identified, then the study may fall in the category of exploratory research.

## Activities

**State the decision**. Identify the appropriate decision that must be resolved in order to address or solve the problem. The decision should be phrased as a question to allow the decision maker to choose between alternative actions or outcomes based on the data collected. An example of the decision statement is "Is the permittee out of compliance?" There are two alternative outcomes to this statement: the permittee is out of compliance or the permittee is in compliance. It is imperative to specifically phrase the decision and avoid general statements of goals or objectives. If necessary, consult with the decision maker to translate general goals into a specific decision.

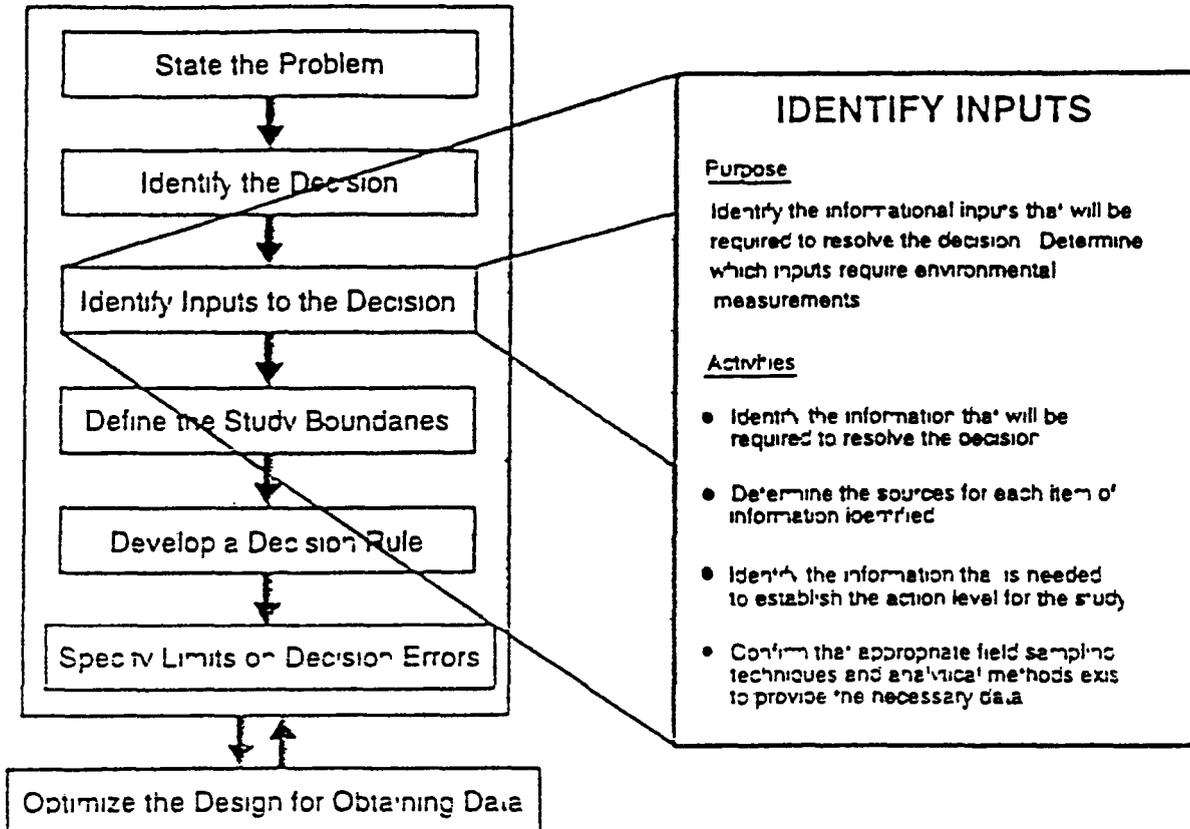
**Categorize multiple decisions**. If several separate decisions must be defined in order to address the problem, then specify all of these decisions and identify the sequence in which they should be resolved.

**State the actions or outcomes that could result from the resolution of the decision**. State the actions or outcomes that could result from each resolution of the decision statement. An example of this is "If the facility is found to be out of compliance, then enforcement action will occur, if the facility is found to be in compliance, then no action will be taken."

## CHAPTER 3

## STEP 3 IDENTIFY THE INPUTS TO THE DECISION

## THE DATA QUALITY OBJECTIVES PROCESS

Outputs from this step

- A list of informational inputs needed to resolve the decision
- The list of environmental variables or characteristics that will be measured

## Background

For this step the planning team will examine the decision specified in Step 2. Identify the Decision and identify the informational inputs that will be required to resolve the decision. The goal is to determine which inputs will require new environmental measurements.

## Activities

Identify the information that will be required to resolve the decision. Define the data that will be necessary to support the decision; for example, ambient ozone monitoring data of concentrations are necessary to determine if the ozone levels are above the regulatory standard.

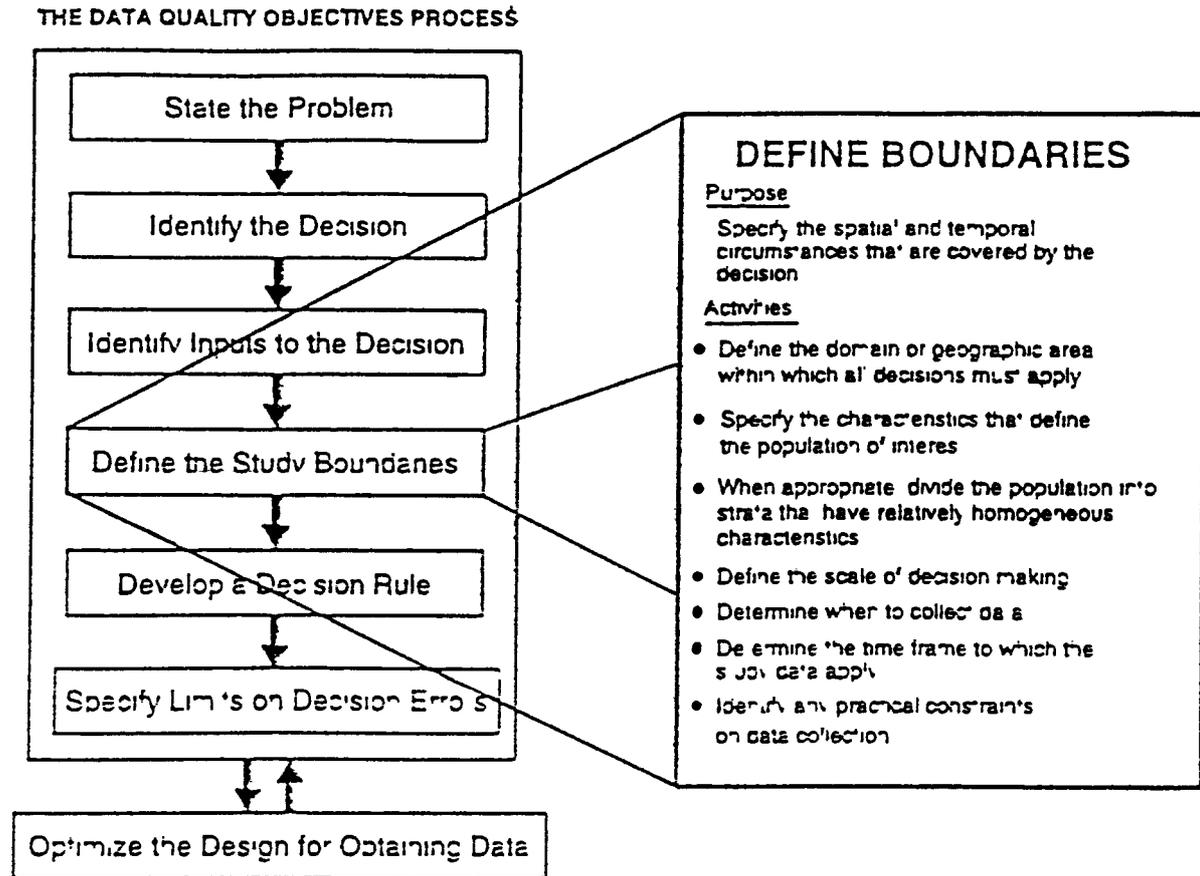
Determine the sources for each item of information identified above. List the sources of the variables that need to be measured to provide sufficient information to resolve the decision. Determine whether relevant data already exist. Also, investigate alternative data sources or the use of surrogates.

Identify the information that is needed to establish the action level for the study. The action level is the measurement threshold which provides the criterion for choosing between alternative actions. The simplest example of an action level is a standard or a regulatory threshold. An action level can also be based on risk or exposure assessment, technological limits or reference-based standards. Subsequent steps in the DQO Process require that the action level be set precisely; here, the basis for setting the action level is identified.

Confirm that appropriate field sampling techniques and analytical methods exist to provide the necessary data. Determine whether the types of data necessary to resolve the decision can be collected. If it is not possible to collect the necessary data, determine if it is reasonable to make assumptions or draw conclusions from other sources of information. If no practical source of information exists that is applicable for resolving the decision, consider shifting the effort to develop the research tools needed to address the problem or consider not conducting the study at this time.

## CHAPTER 4

## STEP 4 DEFINE THE BOUNDARIES OF THE STUDY

Outputs from this step

- Characteristics that define the domain of the study
- A detailed description of the spatial and temporal boundaries of the decision
- Any practical constraints that may interfere with the study

## Background

Data are difficult to interpret unless they are drawn from a well-defined sampling area. The purpose of this step is to clearly define the set of circumstances (boundaries) that will be covered by the decision. These will include:

- Spatial boundaries that define what should be studied and from where the samples should be taken, and
- Temporal boundaries that describe when samples should be taken and what time frame the study data should represent.

These boundaries will be used to ensure that the study design incorporates the time periods in which the study should be implemented, areas that should be sampled, and the time period to which the study results should apply. This will help ensure that the study data are representative of the objects or people being studied. Defining boundaries before the data are collected can also prevent inappropriate pooling of data in a way that masks useful information.

Practical constraints that could interfere with sampling are also identified in this step. A practical constraint is any hindrance or obstacle that may interfere with the full implementation of the study design.

## Activities

Define the spatial boundary of the decision.

- (1) Define the domain or geographic area within which all decisions must apply. The domain or geographic area is a region distinctively marked by some physical features (i.e. volume, length, width, boundary) to which the decision will apply. Some examples are the metropolitan city limits, the property boundaries, and the natural habitat range of a particular species.
- (2) Specify the characteristics that define the population of interest. The "population" is a statistical term that refers to the total collection of objects or people to be studied and from which the sample is to be drawn. For instance, a population may be PCB concentrations in soil at a Superfund site, blood lead levels in children under the age of seven, or the hourly ozone concentrations within a city. Clearly define the attributes that make up the population by stating them in a way that makes the focus of the study unambiguous. For example, the top 12 inches of soil is less ambiguous than merely "surface soil."

- (3) When appropriate, divide the population into strata that have relatively homogeneous characteristics. Using existing information stratify each medium or set of objects into subsets or categories that exhibit relatively homogeneous properties such as contaminant concentrations. Stratification is desirable for studying sub-populations or reducing the complexity of the problem by breaking it into more manageable pieces. The decision maker can choose to make separate decisions about each stratum or the entire population.
- (4) Define the scale of decision making. Define the smallest, most appropriate subset of the population about which decisions will be made to meet the goals of the study. For example, in a study of contaminated soil where the goal is to protect future residents from exposure and where the future land use is residential, the planning team may set the scale of decision making to a 14 by 14 area if the children derive most of their exposure from an outdoor play area of this size. Consequently, decisions that will be made at the site would be protective of children, a sensitive population in exposure assessment.

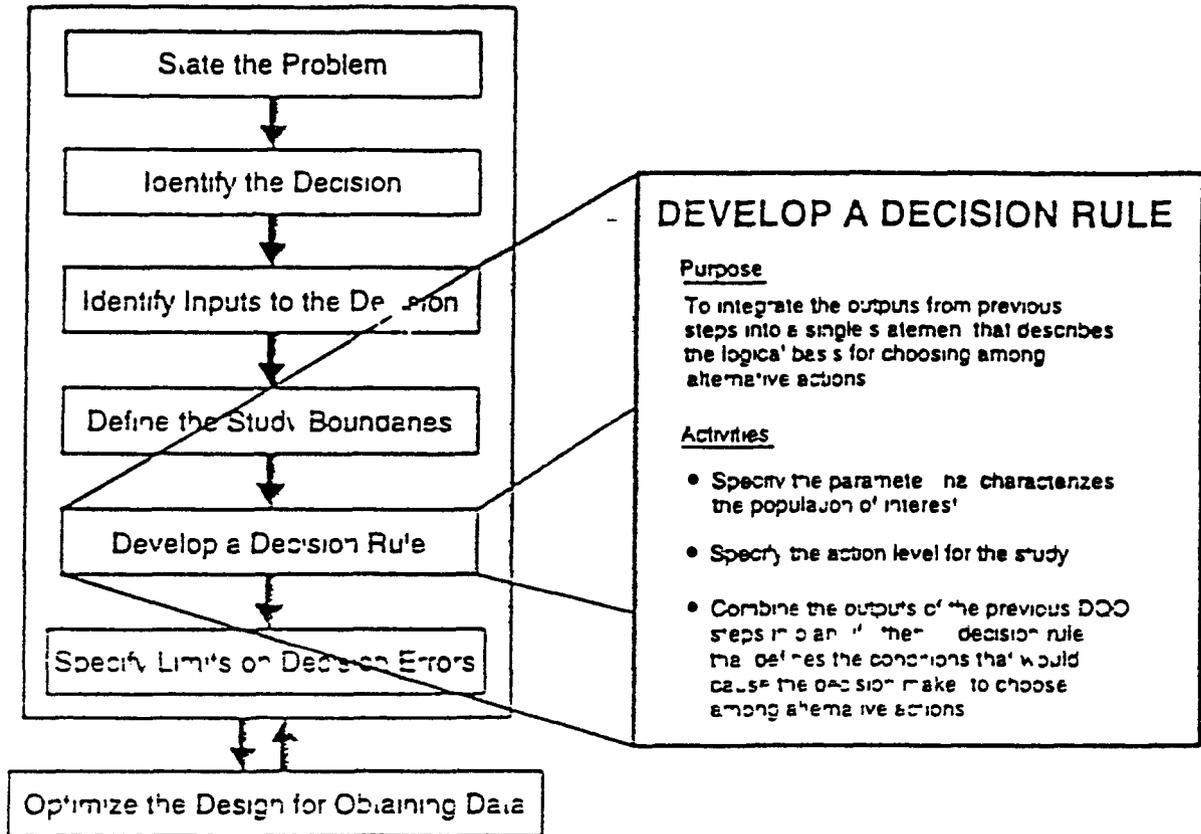
#### Define the temporal boundary of the decision

- (1) Determine when to collect data. Conditions may vary over the course of a study due to weather or other factors. Moreover, the study decision may be influenced by the seasons. For example, a study to measure ambient airborne particulate matter may give misleading information if the sampling is conducted in the wetter winter months rather than the drier summer months. Therefore the planning team must determine the most appropriate time period to collect data that will reflect the conditions that are of interest.
- (2) Determine the time frame to which the study data apply. It may not be possible to collect data over the full time period to which the decision will apply. Therefore the planning team must determine the most appropriate time frame that the data should reflect (e.g. the study data will reflect the condition of contaminant leaching into ground water over a period of a hundred years).

Identify any practical constraints on data collection. These constraints include seasonal or meteorological conditions when sampling is not possible and the unavailability of personnel, time or equipment. For example, it could occur that surface soil samples could not be taken beyond the east boundaries of a site under investigation because access to that area had not been granted by the owner of the adjacent property.

STEP 5 DEVELOP A DECISION RULE

THE DATA QUALITY OBJECTIVES PROCESS



Output from this step

- An "if then" statement that defines the conditions that would cause the decision maker to choose among alternative courses of action

## Background

The primary purpose of this step is to integrate the outputs from previous steps into a single statement that describes the logical basis for choosing among alternative actions. There are three main elements to a decision rule:

- (1) the *parameter of interest*, a descriptive measure (such as a mean, median, or proportion) that specifies the characteristic or attribute that the decision maker would like to know about the statistical population. The chosen parameter of interest for a population should be thought of as the true characteristic of a statistical population, however the population parameter can only be estimated using environmental data.
- (2) the *action level*, a measurement threshold that provides the criterion for choosing among alternative actions. The action level can be based on regulatory standards, a risk or exposure assessment, technology-based limits, or reference-based standards.
- (3) the *alternative actions*, which describe the courses of action that the decision maker would take, depending on the true value of the parameter of interest.

These three elements are combined in the following way. If the parameter of interest (e.g., true mean concentration of lead in the surface soil) is greater than the action level (e.g., 1 mg/kg) then the correct choice is alternative action A (e.g., remove the soil from the site), otherwise the correct choice is alternative action B (e.g., leave the soil in place).

## Activities

Specify the parameter that characterizes the population of interest. The planning team should specify the parameter of interest (such as the mean, median, or percentile) that would characterize the population and would be used to make the decision if its true value were known. For example, to determine if the contamination level at a given site exceeds an action level, the planning team must specify the parameter that will be evaluated with respect to the action level (e.g., the mean concentration). Some regulations specify the parameter, but if that is not the case it may be necessary to consult with a statistician to help select a parameter that is consistent with the intended application. Be aware that the parameter that is chosen in this step may be changed later in the DQO Process as more information becomes available or if estimating the parameter becomes unwieldy.

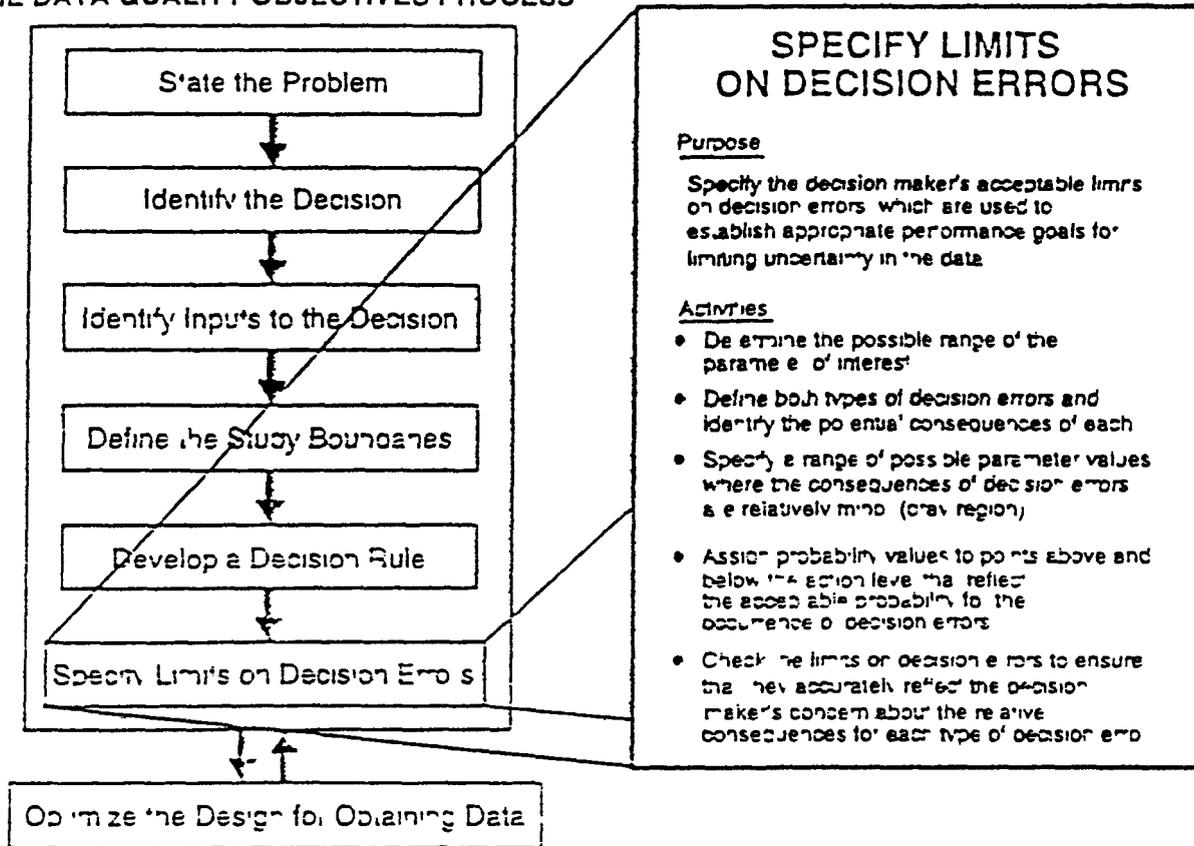
Specify the action level for the study. The decision maker should specify the numerical value that would cause him/her to choose between two or more actions. For example, the decision maker would choose one action if the parameter that characterizes the population of interest is above 1 mg/L and a different action otherwise.

Develop a decision rule. Develop a decision rule as an if-then statement that incorporates the parameter of interest, the action level, and the action(s) that would result from the decision. For example, If the mean concentration of cadmium in the waste ash exceeds 10 mg/Kg, then the waste ash will be considered to pose an unacceptable risk and will be disposed of in a RCRA hazardous waste landfill; otherwise, the waste ash will be considered not to pose a risk and will be disposed of in a municipal landfill.

CHAPTER 6

STEP 6 SPECIFY ACCEPTABLE LIMITS ON DECISION ERRORS

THE DATA QUALITY OBJECTIVES PROCESS



Output from this step

- The decision maker's acceptable decision error rates based on a consideration of the consequences of making an incorrect decision

## Background

Decision makers are interested in knowing the true state of some feature of the environment. Since measurement data can only estimate this state, however, decisions that are based on measurement data could be in error (decision error). Therefore, the goal of the planning team is to design a sampling plan that reduces the chance of making a decision error to an acceptable level. This step of the DQO Process will help the decision maker define what constitutes acceptable limits on the probability of making a decision error.

There are two reasons why the decision maker cannot know the true value of a population parameter:

- (1) The population of interest almost always varies over time and space. Limited sampling will miss some features of this natural variation because it is usually impossible or impractical to measure every point of a population. Sampling error occurs when sampling is unable to capture the complete scope of natural variability that exists in the true state of the environment.
- (2) Because of measurement error, analytical measurements can only estimate the true population parameter. Measurement error is a combination of random and systematic errors that inevitably arise during the various steps of the measurement process, such as sample collection, sample handling, sample preparation, sample analysis, data reduction, and data handling.

The combination of sampling error and measurement error is called total study error, which is directly related to decision error. Basing decisions on unavoidably imperfect measurement data will lead to the possibility of making a decision error.

The probability of decision errors can be controlled by adopting a scientific approach. The scientific method employs a system of decision making that controls decision errors through the use of statistical hypothesis testing. In hypothesis testing, the data are used to select between one condition of the environment (the baseline condition or null hypothesis,  $H_0$ ) and an alternative condition (the alternative hypothesis,  $H_1$ ). The baseline condition is presumed to be true in the absence of strong evidence to the contrary. This feature provides a way to guard against making a decision error that the decision maker considers to have the greatest undesirable consequences.

A decision error occurs when the data lead the decision maker to believe that the null hypothesis is true when it is not, or that the alternative hypothesis is true when it is not. These two types of decision errors are classified as false positive errors and false negative errors. They are described below.

**False Positive Error** — The false positive error occurs when the data mislead the decision maker into believing that the burden of proof has been satisfied and that the null hypothesis ( $H_0$ ) should be rejected. Consider an example where the decision maker presumes that a certain waste is hazardous (i.e., the baseline condition or null hypothesis is the waste is hazardous). If the data lead the decision maker to wrongly conclude that the waste is non-hazardous when it actually is, then the decision maker would be making a false positive error. A statistician usually refers to the false positive error as alpha ( $\alpha$ ) the level of significance the size of the critical region or a Type I error depending upon its precise use.

**False Negative Error** — The false negative error occurs when the data mislead the decision maker into wrongly concluding that the burden of proof has not been satisfied so that the null hypothesis ( $H_0$ ) is accepted. In the waste example, the data lead the decision maker to wrongly conclude that the waste is hazardous when it truly is not. A statistician usually refers to a false negative error as beta ( $\beta$ ) or a Type II error depending on its precise use. It is also known as the complement of the power of a test.

The definition of false positive and false negative errors depends on the viewpoint of the decision maker and the actions that are taken. Consider the viewpoint where a person has been presumed to be innocent until proven guilty (i.e.,  $H_0$  is innocent,  $H_1$  is guilty). A false positive error would be convicting an innocent person (i.e., taking action when action is not warranted); a false negative error would be not convicting the guilty person (i.e., not taking action when action is warranted). From a decision maker's viewpoint, where a criminal is guilty until proven innocent (i.e.,  $H_0$  is guilty,  $H_1$  is innocent) the errors are reversed. Here, the false positive error would be releasing the guilty person (i.e., taking action when it is not warranted) and the false negative error is to not release the innocent person (i.e., not take action when it is warranted).

While the possibility of decision errors can never be totally eliminated, it can be reduced. To reduce decision errors, the planning team must develop an adequate estimate of the population parameter (i.e., reduce total study error). This can be accomplished by collecting a large number of samples (to reduce sampling error) and by analyzing individual samples several times using more precise laboratory methods (to reduce measurement error). Better sampling designs can also be developed to collect data that more accurately and efficiently represent the population of interest. Reducing decision errors, however, generally increases costs. In many cases, reducing decision errors is unnecessary for making a reasonable decision. For instance, if the consequences of decision errors are minor, a reasonable decision could be made based on relatively crude data. Similarly, if the consequences of decision errors are severe, the decision maker will want to develop a sampling design that eliminates as much sampling and measurement error as possible (within budget constraints) in order to provide more exact estimates.

To minimize the effort that is spent unnecessarily reducing decision errors, the planning team must determine whether reducing sampling and measurement errors is necessary to meet the limits on decision errors. Therefore, they must first define what



decision error occurs the decision maker is guarding against making the more severe decision error

- (4) *Assign the terms 'false positive' and 'false negative' to the proper decision errors* The false positive decision error occurs when the data mislead the decision maker into believing that the burden of proof has been satisfied so that  $H_0$  is erroneously rejected (i.e. the decision maker wrongly concludes that the waste is nonhazardous when it actually is). The false negative decision error occurs when the data mislead the decision maker into wrongly concluding that the burden of proof has not been satisfied, so that  $H_0$  is erroneously accepted (i.e. the decision maker concludes that the waste is hazardous when it truly is not).

Define the potential consequences of decision errors at several points within the false positive and false negative domains. In the waste example, the consequences of a false positive decision error when the true parameter value is merely 10% above the action level may be minimal because it would cause only a moderate increase in the risk to human health. On the other hand, the consequences of a false positive error when the true parameter is ten times the action level may be severe because it could greatly increase the exposure risk to humans, as well as cause severe damage to a local ecosystem.

Specify a range of possible parameter values where the consequences of decision errors are relatively minor (gray region). The gray region is a range of points (bounded on one side by the action level) where the consequences of a false negative decision error are relatively minor. Establish the general location of the gray region by evaluating the consequences of wrongly concluding that the baseline condition (the null hypothesis) is true.

The gray region or area of uncertainty establishes the minimum distance from the action level to which the decision maker would like to control decision errors. In statistics this value is called delta ( $\Delta$ ) and is an essential part of the calculations needed to determine the number of samples that need to be collected. The width of the gray region reflects the decision maker's concern for decision errors. A more narrow gray region implies a desire to conclusively detect the condition when the true parameter value is close to the action level. When the sample estimate of the parameter falls within the gray region, the decision maker may have a high probability of making a decision error (i.e. the data may be 'too close to call') and may wrongly conclude that the baseline condition is true.

The gray region is an area where it will not be feasible or reasonable to control the false negative decision error rate to low levels because the resources that would be required would exceed the expected costs of the consequences of making the decision error. For example, when testing whether a parameter (such as the mean concentration) exceeds the action level, if the true parameter is near the action level, then the imperfect data will tend to be clustered around the action level with some values above the action level and some below. In order to determine with confidence whether the true value of the parameter is above or

below the action level, the decision maker would need to collect a large amount of data, increase the precision of the measurements, or both. If taken to an extreme, the cost of collecting data can exceed the cost of making a decision error, especially where the consequences of the decision error may be relatively minor. Therefore, the decision maker should establish the gray region or the region where it is not crucial to control the false negative decision errors, by balancing the resources needed to "make a close call" versus the consequences of making that decision error.

Assign probability values to points above and below the action level that reflect the acceptable probability for the occurrence of decision errors. Assign probability values to points above and below the action level that reflect the decision maker's acceptable limits for making an incorrect decision. The most stringent limits on decision errors that are typically encountered for environmental data are 01 (1%) for both the false positive and false negative decision errors ( $\alpha$  and  $\beta$ ). This guidance recommends using .01 as the starting point for setting decision error rates. The most frequent reasons for setting limits greater than 01 are that the consequences of the decision errors may not be severe enough to warrant setting decision error rates that are thus stringent. The value of 01 should not be considered a prescriptive value for setting decision error rates, nor should it be considered as the policy of EPA to encourage the use of any particular decision error rate. Rather, it should be viewed as a starting point from which to develop limits on decision errors that are applicable for each situation. If the decision maker chooses to relax the decision error rates from 01 for false positive and false negative decision errors, the planning team should document the reasoning behind setting the decision error rate and what the potential impacts may be on cost, resource expenditure, human health, and ecological concerns.

Repeat this activity for both sides of the gray region. Generally, the acceptable limits for making a decision error should decrease further away from the action level, as the consequences of decision error become more severe.

Check the limits on decision errors to ensure that they accurately reflect the decision maker's concerns about the relative consequences for each type of decision error. The acceptable limits on decision errors should be smaller (i.e. have the lowest probability of error) for cases where the decision maker has greatest concern for decision errors. This means that if one type of error is more serious than another, then its acceptable limits should be smaller (more restrictive). In addition, the limits on decision errors are usually larger (higher probability of error can be tolerated) near the action level, since the consequences of decision errors are generally less severe as the action level is approached. Verify that the decision maker's acceptable limits on decision errors are consistent with these principles.

The combined information from the activities section of this chapter can be graphed onto a Design Performance Goal Diagram or charted in a "Decision Error Limits Table" (see Figures 6-1 and 6-2 and Tables 6-1 and 6-2 below). Both are useful tools for visualizing and evaluating all of the outputs from this step. Figure 6-1 and Table 6-1 illustrate the case where the baseline conclusion (null hypothesis) is that the parameter of

interest exceeds the action level (e.g. the waste is hazardous) Figure 6-2 and Table 6.2 illustrate the case where the baseline condition is that the parameter is less than the action level (e.g. the waste is non-hazardous)

The Design Performance Goal Diagram (which is sometimes called a "Decision Performance Curve") can be refined by breaking the steps of decision errors into smaller units. This would have the effect of adding rows of information to its corresponding Decision Error Limits Table. The diagram will be used in the final step of the DQO Process (Step 7 Optimize the Design) in order to construct a statistically based evaluation of how well the sampling design will meet the DQOs. This evaluation involves the construction of a Power Curve which is a graphical descriptor of a sampling design's expected performance. If the Power Curve lies within the acceptable regions of the Design Performance Goal Diagram then the corresponding sampling design satisfies the decision maker's acceptable limits on decision errors.

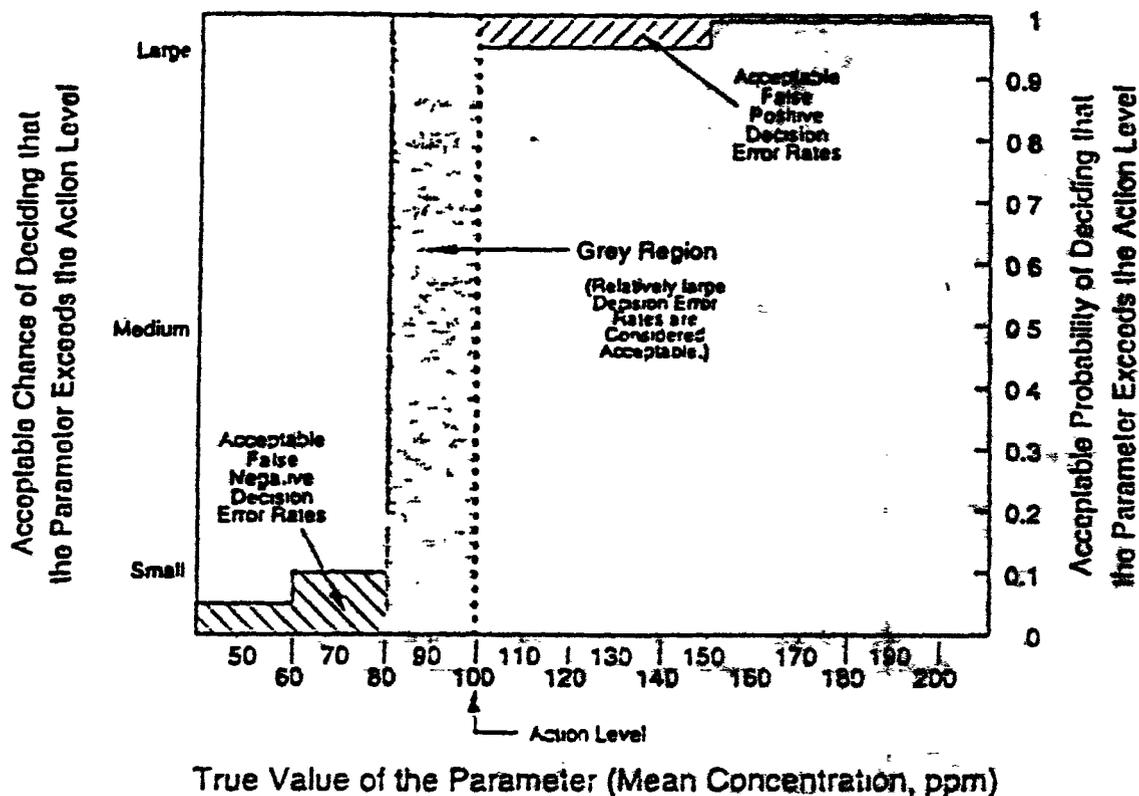


Figure 6-1 An Example of a Design Performance Goal Diagram (Baseline condition parameter exceeds action level).

<u>True concentration</u>	<u>Correct decision</u>	<u>Acceptable probability of making an incorrect decision (a decision error)</u>
50 to 60 ppm	does not exceed action level	5%
60 to 80	"	10%
80 to 100	"	grey region—no probability specified
100 to 150	exceeds action level	5%
150 to 200		1%

Table 6-1 Decision Error Limits Table Corresponding to Figure 6-1

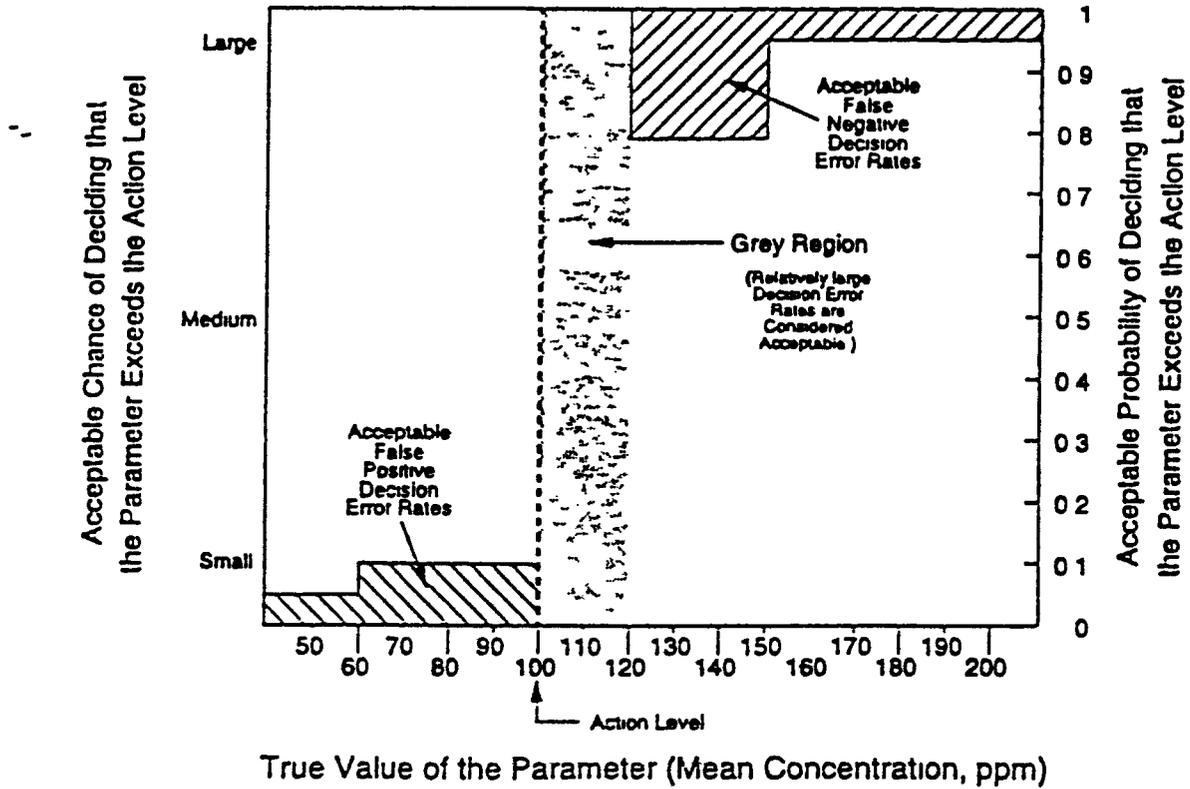


Figure 6-2 An Example of a Design Performance Goal Diagram (Baseline condition parameter less than action level)

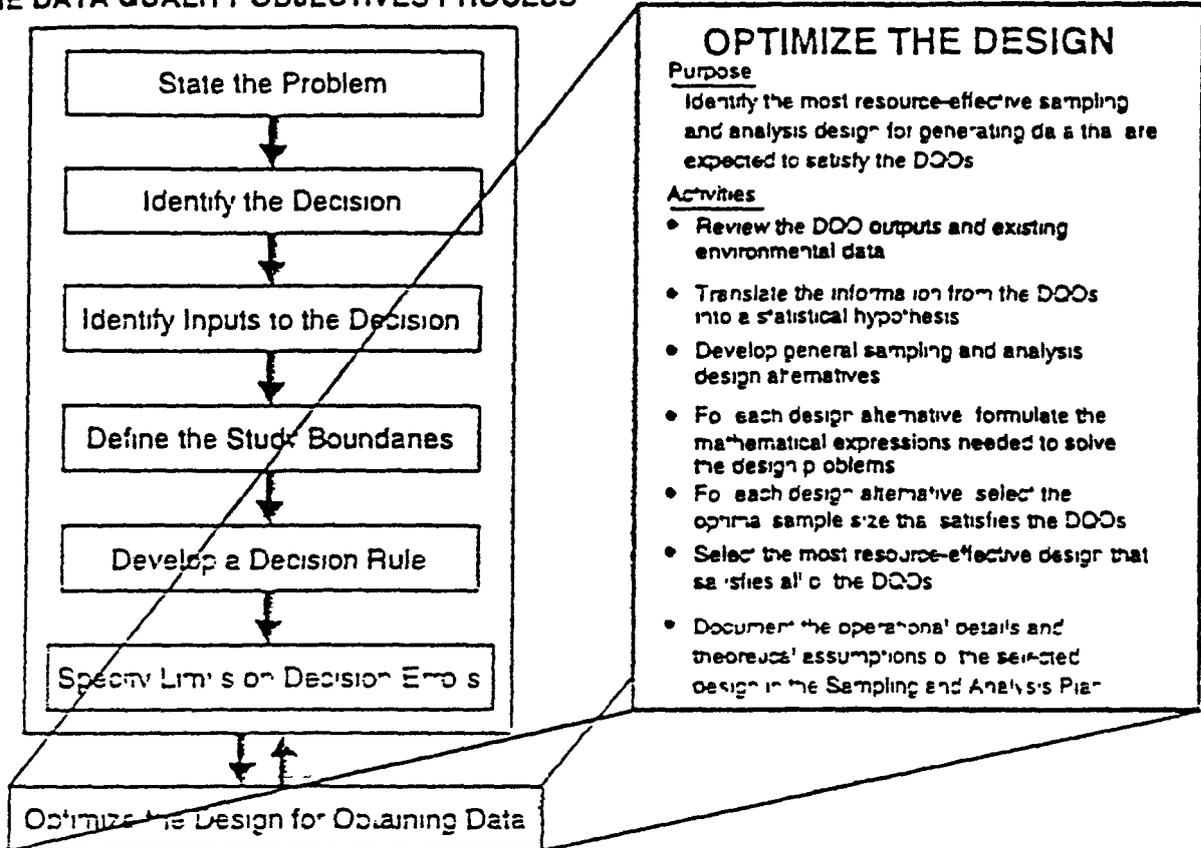
<u>True concentration</u>	<u>Correct decision</u>	<u>Acceptable probability of making an incorrect decision (a decision error)</u>
50 to 60 ppm	does not exceed action level	5%
60 to 100		10%
100 to 120		gray region—no probability specified
120 to 150	exceeds action level	20%
150 to 200		5%

Table 6 2 Decision Error Limits Table Corresponding to Figure 6 2

CHAPTER 7

STEP 7 OPTIMIZE THE DESIGN

THE DATA QUALITY OBJECTIVES PROCESS



Output from this step

- The most resource-effective design for the study that is expected to achieve the DQOs, selected from a group of alternative designs generated during this step

## Background

In this step statistical techniques are used to develop alternative environmental sampling designs and evaluate their efficiency at meeting the DQOs. To develop the optimal design for this study, it may be necessary to work through this step more than once after revisiting previous steps of the DQO Process.

The objective of this step is to identify the most resource-effective sampling design expected to generate data that satisfy the DQOs specified in the preceding steps. While this step does not explain the exact procedures for developing a sampling design, it does provide a broad overview of the steps that need to be accomplished in order to reach this goal. Because the role of statistics is very important in developing a study design that should achieve the acceptable decision error rates, a statistician or someone with statistical expertise should be consulted if they are not already on the planning team.

## Activities

Review the DQO outputs and existing environmental data. Review the background information on the project or program along with the DQO outputs, this provides a succinct collection of information on the context, requirements, and constraints of the sampling design.

Develop general sampling and analysis design alternatives. Develop alternative sampling and analysis designs based on the DQO outputs and a general understanding of the statistical parameters of the contaminants and media. Examples of general design alternatives include

- simple random sampling
- sequential random sampling
- systematic sampling
- composite sampling (in conjunction with another sampling design)

For each design alternative, formulate the mathematical expressions needed to solve the design problem. Develop the three types of mathematical expressions needed to optimize the sampling design as follows:

- (1) Determine a suggested method for testing the statistical hypothesis (e.g. a t-test).
- (2) Develop a statistical model that describes the relationship of the measured value to the "true" value. Often the model will describe the components of error or bias that are believed to exist in the measured value.
- (3) Develop a cost function that relates the numbers of samples to the total cost of sampling and analysis.

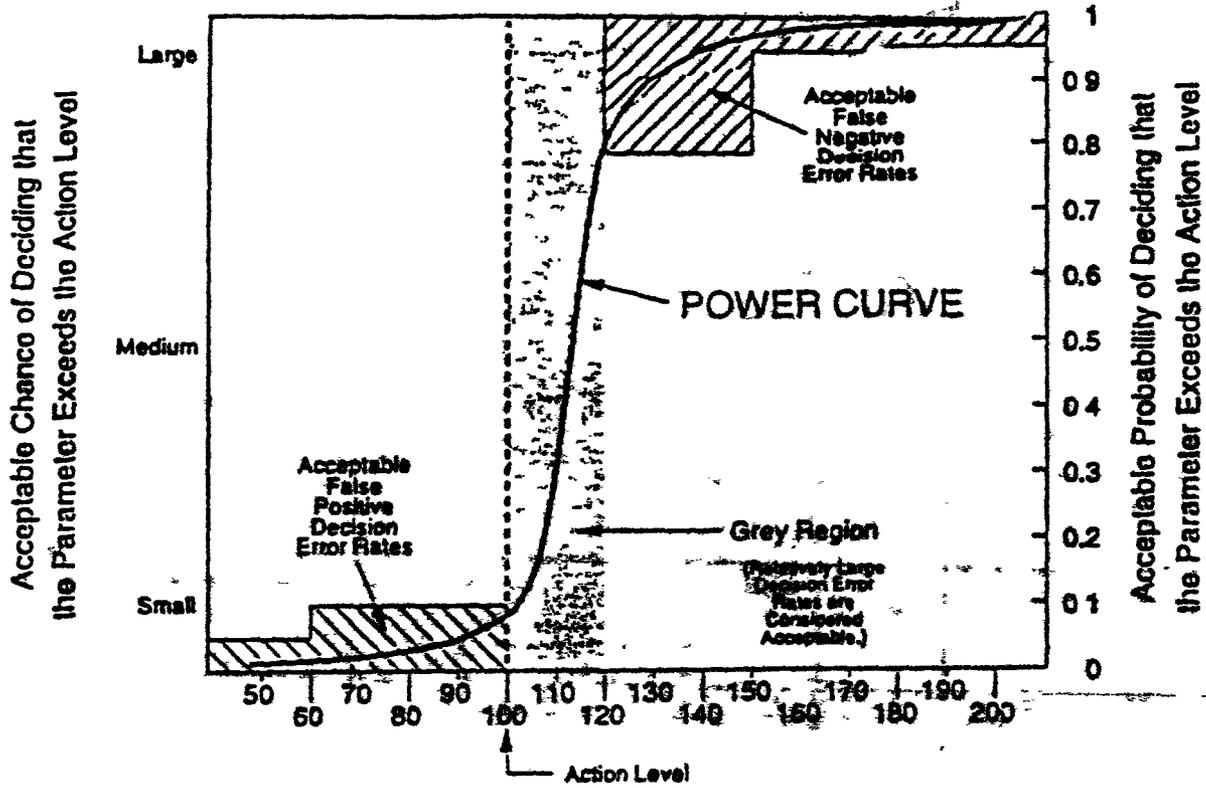
For each design alternative, select the optimal sample size that satisfies the DQOs. Using the mathematical expressions from the previous activity, solve for the optimal sample size that satisfies the DQOs, including the decision maker's limits on decision errors.

If no design will meet the limits on decision errors within the budget or other constraints, then the planning team will need to relax one or more constraints. For example,

- increase the acceptable decision error rates,
- increase the width of the gray region,
- relax other project constraints such as the number of available personnel,
- increase funding for sampling and analysis, or
- change the boundaries, it may be possible to reduce sampling and analysis costs by changing or eliminating subgroups that will require separate decisions.

Select the most resource-effective design that satisfies all of the DQOs. Evaluate the design options based on cost and ability to meet the DQO constraints. Choose the one that provides the best balance between cost (or expected cost) and ability to meet the DQOs. A Power Curve is a useful statistical tool used to evaluate whether the design has the ability to meet the DQOs. An example of a Power Curve is shown in Figure 7-1.

Document the operational details and theoretical assumptions of the selected design in the sampling and analysis plan. Document the selected design's key features that must be implemented properly to allow for efficient and valid statistical interpretation of the data. It is particularly important to document the statistical assumptions that could be violated through errors in or practical constraints on field sample collection procedures or laboratory methods.



True Value of the Parameter (Mean Concentration, ppm)

Figure 7-1 An example of a Power Curve

## REFERENCES

- Cochran, W 1977 Sampling Techniques New York John Wiley
- Desu, M M and D Raghavarao 1990 Sample Size Methodology San Diego, CA Academic Press
- Gilbert Richard O 1987 Statistical Methods for Environmental Pollution Monitoring New York Van Nostrand Reinhold
- U S Environmental Protection Agency 1993 Guidance for Environmental Data Quality Assessment. EPA QA/G-9
- U S Environmental Protection Agency 1992 Methods for Evaluating the Attainment of Cleanup Standards Volume 2 Ground Water EPA 230-R-92-014 Office of Policy, Planning and Evaluation
- U S Environmental Protection Agency 1989 Methods for Evaluating the Attainment of Cleanup Standards Volume 1 Soils and Solid Media EPA 230/02-89-042 Office of Policy, Planning and Evaluation
- U S Environmental Protection Agency April 1984 Order 5360 1, Policy and Program Requirements to Implement the Mandatory Quality Assurance Program U S Environmental Protection Agency

## APPENDIX A

BEYOND THE DQO PROCESS EVALUATION OF THE DESIGN  
USING THE DATA QUALITY ASSESSMENT PROCESS

Once the DQO Process has been completed the planning team will have the information needed to choose the sampling design that best meets the needs of their study. The needs of the planning team have not been fully met however until the sampling data are analyzed to ensure that any decision made from the data will meet the decision error rates specified in the DQO Process. This analysis is part of a related process called Data Quality Assessment (DQA).

The DQA Process is used to assess the scientific and statistical quality of data for a specified purpose. During the DQA Process the data will be analyzed scientifically to inspect for technical anomalies and to judge that the context of the data is correct. At the same time the data will be evaluated statistically to confirm that the statistical model was correct, select a statistical test, and validate that statistical test by verifying assumptions such as distribution and independence. The outcome of the DQA analysis will determine whether a decision can be made using the existing data or whether additional sampling data must be collected. The DQA Process is also useful for determining whether a sampling design is appropriate for similar studies.

DQA guidance from EPA's Quality Assurance Management System is titled "Guidance for Environmental Data Quality Assessment" EPA QA/G-9.

## APPENDIX B

## DQO CASE STUDY CADMIUM-CONTAMINATED FLY ASH WASTE

Introduction

Appendix B presents an example of the DQO outputs for a realistic case study. The case study has been chosen because it is simple and straightforward and because the outputs are uncomplicated. Although some of the outputs from this example may seem intuitive, this is not often the case in practice. For many studies the DQO Process is complicated and thought-provoking. Even so, some steps will require more effort than others. Keep in mind that all of the steps in the DQO Process are necessary to develop a statistical study design. Once all of the steps have been completed and thoroughly thought out, then development of the most resource-effective study design can proceed.

Background

A hazardous waste incineration facility located in the Midwest routinely removes fly ash from its flue gas scrubber system and disposes of it in a local sanitary landfill. Previously it was determined that the ash was nonhazardous under RCRA regulations. The incinerator however recently began treating a new waste stream. A local environmental public interest group has asked that the ash be retested against RCRA standards before it is disposed of. The group is primarily concerned that the ash could contain hazardous levels of cadmium from the new waste sources. The facility manager has agreed to test the ash and has decided to employ the DQO Process to help guide decision making throughout the project.

Cadmium is primarily used as corrosion protection on metal parts of cars and electrical appliances. It is also used in some batteries. Cadmium and cadmium salts have toxic effects for humans through both ingestion and inhalation exposures. Ingestion exposure usually causes mild to severe irritation of the gastrointestinal tract, which can be caused by concentrations as low as 0.1 mg/kg/day. Chronic (long term) inhalation exposure can cause increased incidence of emphysema and chronic bronchitis, but the primary target is the kidney. Severe and chronic inhalation exposure have been shown to cause increased incidence of kidney stones and kidney dysfunction.

Under the current Code of Federal Regulations (CFR) Part 261, the method for determining if a solid substance is a hazardous waste under RCRA is to sample a representative portion of the waste and perform a Toxicity Characteristic Leaching Procedure (TCLP). During this process the solid fly ash will be extracted or dissolved using an acid solution. The extraction liquid (the TCLP leachate) will then be subjected to test for specific metals and compounds. For this example, the only concern is with the concentration of cadmium in the leachate. The primary benefit of the DQO Process will be to

establish the sampling design needed to determine if the waste is hazardous under RCRA regulations within acceptable decision error rates

As a precursor to the DQO Process the incineration company has conducted a pilot study of the fly ash to determine the variability in the concentration of cadmium between loads of ash leaving the facility. They have determined that each load is fairly homogeneous. There is a high variability between loads however, due to the nature of the waste stream. Most of the fly ash produced is nonhazardous and may be disposed of in a sanitary landfill. Thus the company has decided that testing each individual waste load before it leaves the facility would be the most economical. Then they could send loads of ash that exceeded the regulated cadmium concentrations to the higher cost RCRA landfills and continue to send the others to the sanitary landfill.

### DQO Development

The following is a representative example of the output from each step of the DQO Process

**State the Problem** — a description of the problem(s) and specifications of available resources and relevant deadlines for the study

- (1) *Identify the members of the planning team* — The members of the planning team will include the incineration plant manager, a representative of the environmental public interest group, a representative of the community where the ash is currently being disposed, a statistician, a quality assurance officer, a representative of EPA that works within RCRA, and a chemist with sampling experience.
- (2) *Identify the primary decision maker* — There will not be a primary decision maker. Separate decisions will either be allocated to members of the planning team or made by consensus.
- (3) *Develop a concise description of the problem* — The problem is to determine which loads are to be sent to a RCRA landfill and which are to be sent to a sanitary landfill.
- (4) *Specify available resources and relevant deadlines for the study* — The waste generator (the incineration company) wishes to spend as little as possible to determine the concentration of cadmium. The project will not be constrained by cost, however. The environmental public interest group has threatened to file a lawsuit for violation of environmental regulations if testing does not proceed within a reasonable time frame.

Identify the Decision — a statement of the decision that will use environmental data and the actions that could result from this decision

(1) *State the decision(s)* — Does the concentration of cadmium in the waste fly ash exceed the RCRA regulatory standard?

(2) *State the actions that could result from the decision* —

(a) The waste fly ash could be disposed of in a RCRA landfill

(b) The waste fly ash could be disposed of in a sanitary landfill

Identify the Inputs to the Decision — a list of the environmental variables or characteristics that will be measured and other information needed to make the decision

(1) *Identify the information that will be required to make a decision* — To evaluate the problem the planning team must collect samples of fly ash waste and subject them to the TCLP extraction and analyze for cadmium

(2) *Determine the sources for each item of information identified* — The concentration of cadmium should be measured in samples of the fly ash using the test methods listed in 40 CFR Pt 261, App II. Some existing pilot study data provide information about variability, but it does not provide enough information to resolve the decision.

(3) *Identify the information that is needed to establish the action level for the study* — The action level is the RCRA standard for cadmium (10 mg/L in the TCLP leachate) using the test methods listed in 40 CFR, Pt 261, App II

(4) *Confirm that appropriate field sampling techniques and analytical methods exist to provide the necessary data* — Cadmium can be measured in the solid waste according to the methodologies specified in 40 CFR Pt 261 App II. The detection limit for this method is below the standard.

Define the Boundaries of the Study — a detailed description of the spatial and temporal boundaries of the decision characteristics that define the population of interest and any practical considerations for the study

(1) *Define the spatial boundary of the decision* —

a) *Define the domain or geographic area within which all decisions must apply* — Decisions will apply to each container load of waste ash

b) *Specify the characteristics that define the population that will be studied* Samples of waste fly ash from the hazardous waste incinerator will be analyzed. The fly ash should not be mixed with any other constituents except water that is used for dust control. Each load of ash should fill at least 70% of the waste trailer. In cases where the trailer is filled less than 70%, the trailer must wait on-site until more ash is produced and fills the trailer to the appropriate capacity.

c) *When appropriate, divide the population into strata that have relatively homogeneous characteristics* Stratification is not necessary since the waste ash is relatively homogeneous within each container.

d) *Define the scale of decision making* The scale of decision making will be each container of waste ash.

(2) *Define the temporal boundary of the decision* —

a) *Determine when to collect data* Contained in the trucks, the waste does not pose a threat to humans or the environment. Additionally, since the fly ash is not subject to change, disintegration or alteration, the chemical properties of the waste do not warrant any temporal constraints. To expedite decision making however, the planning team has placed deadlines on sampling and reporting. The fly ash waste will be tested within 48 hours of being loaded onto waste hauling trailers. The analytical results from each sampling round should be completed and reported within 5 working days of sampling. Until analysis is complete, the trailer cannot be used. This poses an economic impact on the owners of the incinerator.

b) *Determine the time frame to which the study data apply* It will be assumed that the sampling data represents both the current and future concentration of cadmium within the ash.

(3) *Identify practical constraints or data collection* — The most important practical consideration that could interfere with the study is the ability to take samples from the fly ash that is stored in waste hauling trailers. Although the trailers have open access special procedures and methods will have to be implemented in order for the samples to be representative of the entire depth of the ash. It has been suggested that core samples may be one practical solution to this problem. Additionally, to get the best sample from each truck and to minimize the cost of composing of core samples has been suggested.

**Develop a Decision Rule** — Combine the outputs of the previous DQO steps into an if-then decision rule that defines the conditions that would cause the decision maker to choose among alternative actions.

- (1) *Specify the parameter that characterizes the population of interest* — The mean concentration of cadmium for each container will be evaluated with respect to the RCRA regulatory standard
- (2) *Specify the action level for the study* — The action level for this problem will be the RCRA regulatory standard for cadmium of 10 mg/L in the TCLP leachate using the test methods listed in 40 CFR, Pt 261, App II
- (3) *Develop a decision rule (an if then statement)* — If the mean concentration of cadmium from the fly ash waste is greater than 10 mg/L (using the TCLP method as defined in 40 CFR 261) then the waste will be considered to be hazardous and will be disposed of at a RCRA landfill

If the mean concentration of cadmium from the fly ash waste is less than 10 mg/L (using the TCLP method as defined in 40 CFR 261) then the waste will be considered to be non-hazardous and will be disposed of in a sanitary landfill

*Specify Acceptable Limits on Decision Errors* — the decision maker's acceptable decision error rates based on a consideration of the consequences of making an incorrect decision

- (1) *Determine the possible range of the parameter of interest* — The range of cadmium concentration is expected to be from 0.2 mg/L
- (2) *Define both types of decision errors and identify the potential consequences of each* — The action level has already been defined in Step 5. Develop a Decision Rule of this DQO example. Therefore the purpose of this stage of the Process is to define the acceptable probability of making an incorrect decision on either side of the action level

a) *Define both types of decision error and establish which decision errors have the more severe consequences* The planning team has decided that the two decision errors are (i) deciding that the waste is hazardous when it is truly not and (ii) deciding that the waste is not hazardous when it truly is

The consequences of deciding that the waste is hazardous when it truly is not will be that the incinerator will have to pay more for the disposal of the fly ash at a RCRA facility than at a sanitary landfill

The consequences of deciding that the waste is not hazardous when it truly is will be that the incinerator company will dispose of the waste in a sanitary landfill which could possibly endanger human health and the environment. In this situation they may also be liable for future damages and environmental clean-up costs. Additionally the reputation of the incinerator company may be compromised jeopardizing its future

The planning team has concluded that the second decision error has the more severe consequences

b) *Establish the true state of nature for each decision error*

The true state of nature when the more severe decision error occurs (the waste is decided to be nonhazardous when it truly is) is that the waste is hazardous

The true state of nature when the less severe decision error occurs (the waste is decided to be hazardous when it truly is not) is that the waste is nonhazardous

c) *Define the true state of nature for the more severe decision error as the baseline condition or the null hypothesis ( $H_0$ ) and define the true state for the less severe decision error as the alternative hypothesis ( $H_1$ ).*

~~The baseline condition or null hypothesis ( $H_0$ ) is the waste is hazardous.~~

The alternative hypothesis ( $H_1$ ) is the waste is not hazardous

d) *Assign the terms "false positive" and "false negative" to the proper decision errors*

The false positive decision error occurs when the data mislead the decision maker into believing that the burden of proof has been satisfied, so the  $H_0$  is erroneously rejected. For this example the false positive decision error occurs when the decision maker decides the waste is not hazardous when it truly is.

The false negative decision error occurs when the data mislead the decision maker into wrongly concluding that the burden of proof has not been satisfied so that  $H_0$  is erroneously accepted. For this example the false negative decision error occurs when the decision maker decides that the waste is hazardous when it truly is not.

(3) *Specify a range of possible values of the parameter of interest where the consequences of decision errors are relatively minor (gray region) —* The gray region is the area adjacent to the action level where the planning team feels that the consequences of a false negative decision error are minimal. To decide how to set the width of the gray region, the planning team must decide where the consequences of a false negative decision error are minimal. Below the action level, even if the concentration of cadmium were very close to the action level, the monetary costs of disposing of the waste at a RCRA facility are the same as if the waste had a much lower concentration of cadmium. Clearly, a false negative decision error (to the left

of the action level) will cause the incinerator company and their customers to bear the cost of unnecessary expense (i.e. sending nonhazardous waste to a RCRA facility). The planning team has specified a width of 0.25 mg/L for this gray region based on their preferences to detect decision errors at a concentration of 0.75 mg/L (see Figure B-1).

(4) *Assign probability values to points above and below the action level that reflect the acceptable probability for the occurrence of decision errors.* For this example RCRA regulations allow a 5% decision error rate at the action level. The planning team has set the decision error rate to 5% from 1 mg/L to 1.5 mg/L, and 1% from 1.5 mg/L to 2 mg/L as the consequences of health effects from the waste disposed in the municipal landfill increased. On the other side of the action level the planning team has set the acceptable probability of making a false negative error at 20% when the true parameter is from 25 to 75 mg/L and 10% when it is below 25 mg/L based on both experience and an economic analysis that shows that these decision error rates are reasonable to balance the cost of sampling versus the consequence of sending clean ash to the RCRA facility (See Figure B-1).

**Optimize the Design** — The decision maker(s) will select the most resource-effective sampling and analysis design for generating data that are expected to satisfy the DQOs.

Optimizing the design is the one step of the DQO Process that will most likely be completed by a statistician or someone who has sampling design expertise. Using the case study as an example, the following section has been included to provide the reader with a background on the overall process that the statistician might follow in order to optimize the final sampling design.

### Background

Developing a sampling design requires an understanding of the sampled medium and the information that was generated in previous DQO steps. The statistician's job is to review the background information, determine the appropriate statistical application to adequately solve the problem, and develop one or more appropriate sampling designs. Once this is complete, the statistician will compare the cost and performance of the different sampling designs. This process can be broken down into five distinct steps:

- (1) Review the DQO outputs and existing environmental data.
- (2) Develop general sampling and analysis design alternatives and for each design alternative formulate the mathematical expressions needed to solve the design problems.
- (3) For each design alternative, select the optimal sample size that satisfies the DQOs.

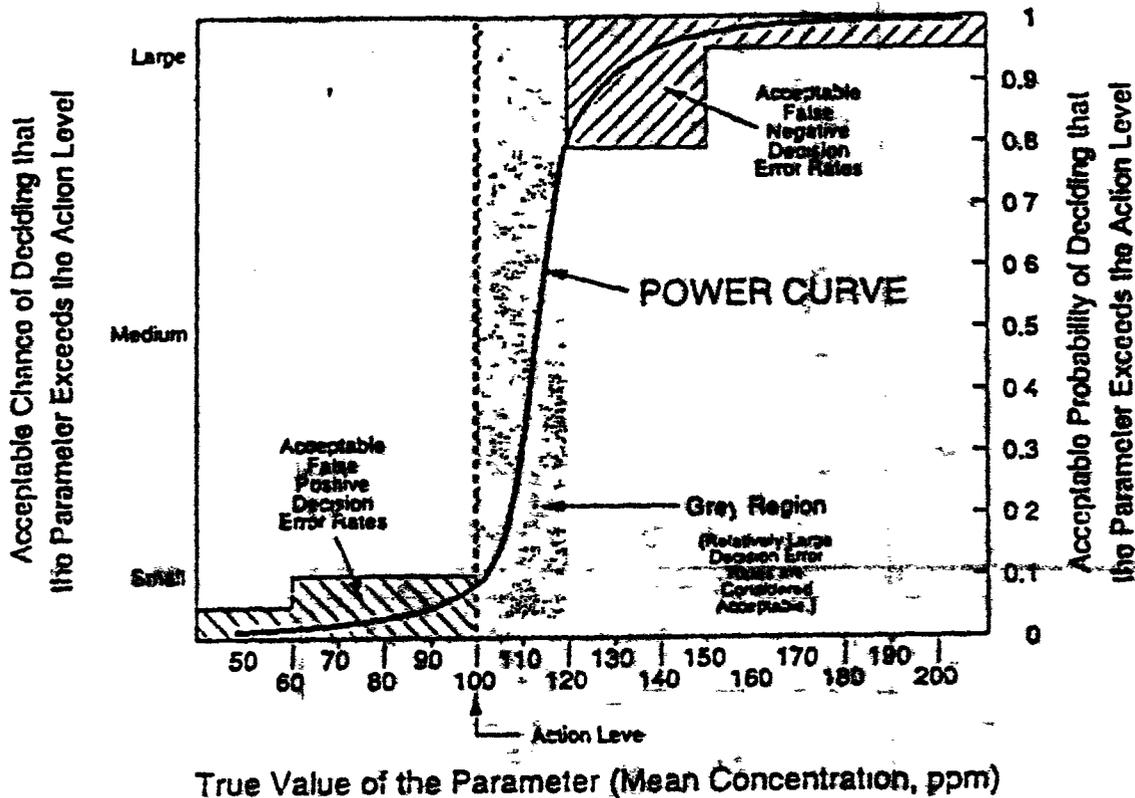


Figure B - 1. Design Performance Goal Diagram for Cadmium Compliance Testing

- (4) Select the most resources-effective design that satisfies all of the DQOs
- (5) Document the operational details and theoretical assumptions of the selected design in the sampling and analysis plan
- (1) Review the DQO outputs and existing environmental data.

The statistician has participated in the DQO Process, so there is no need to review the DQO outputs further. The only existing data relevant to this problem are the pilot study data. Based on the pilot study, the incineration company has determined that each load of ash is fairly homogenous, and has estimated the standard deviation in the concentration of cadmium within loads of ash to be 0.6 mg/L.

- (2) Develop general sample and analysis design alternatives and for each design alternative, formulate the mathematical expressions needed to solve the design problems.

Generally the design alternatives are based on a combination of design objectives developed in previous DQO steps and knowledge of statistical parameters

about the medium or contaminant. Below are four examples of possible designs that could apply to the case study.

- (a) **Simple Random Sampling** — The simplest type of probability sample is the simple random sample. With this type of sampling every possible point in the sampling medium has an equal chance of being selected. For this reason the distribution of the sample measurement is expected to be representative of the medium. Simple random samples are used primarily when the variance of the medium is relatively small and the cost of analysis is relatively inexpensive. Simple random sample locations are generally developed through the use of a random number table or through computer generation of pseudo-random numbers. The number of samples is fixed.

In the case of the cadmium-contaminated ash a fixed number of grab samples would be taken and analyzed separately. For this problem, standard lab splits and QC samples would be taken according to standard procedures for the RCRA program. The samples would be chosen randomly in three dimensions. A t-test is suggested as a possible method for testing the statistical hypothesis.

For each container, the following model is applicable for a simple random sample design:

$$y_i = \mu + b + \eta_i + \delta_i$$

where

- $y_i$  =  $i^{\text{th}}$  observation
- $\mu$  = true mean
- $b$  = bias in measurement
- $\eta_i$  = deviation from  $\mu$  for the  $i^{\text{th}}$  observation (natural variability)
- $\delta_i$  = deviation from  $\mu + \eta_i$  for the  $i^{\text{th}}$  observation (analytic/measurement error)

In this simplistic example the bias is assumed to be zero ( $b=0$ ) and that  $\eta_i$  and  $\delta_i$  cannot be distinguished during sampling and analysis. Therefore the model will reduce to

$$y_i = \mu + \epsilon_i$$

where

- $y_i$  =  $i^{\text{th}}$  observation
- $\mu$  = true mean
- $\epsilon_i$  = total deviation from  $\mu$  for the  $i^{\text{th}}$  observation

and each  $\epsilon_i$  is independently distributed with a mean of 0 and a variance of  $\sigma^2$ .

- (b) **Composite Sampling** — This type of sampling consists of taking multiple samples, physically combining (compositing) them, and drawing one or more subsamples for analysis. Composite samples are taken primarily when an average concentration is sought and there is no need to detect peak concentrations. By compositing the samples, researchers are able to sample a large number of locations than if compositing was not used, and also reduce the cost of analysis by combining the samples together.

In the case of the cadmium-contaminated ash, a fixed number of grab samples would be taken and composited. The number of grab samples contained in a composite sample ( $g$ ) is also fixed. To determine sampling locations within the composite, a container would be divided into " $g$ " equal-volume strata and samples would be chosen randomly within each strata. Strata would be based on location to ensure full coverage of each container. Standard lab splits and QC samples would be taken according to standard procedures for the RCRA program. A  $t$ -test is suggested as the possible method for testing the statistical hypothesis. For each container, the following model is applicable for a composite sample design.

$$y_i = \mu + \epsilon_i + \delta_i$$

where

- $y_i$  =  $i^{\text{th}}$  observation ( $i^{\text{th}}$  composite sample)  
 $\mu$  = true mean  
 $\epsilon_i$  = deviation from  $\mu$  for the  $i^{\text{th}}$  observation due to natural variability  
 $\delta_i$  = deviation from  $\mu$  for the  $i^{\text{th}}$  observation due to variability within each composite

Each  $\epsilon$  and  $\delta$  is assumed to be independently distributed with a mean of 0 and variances of  $\sigma_\epsilon^2$  and  $\sigma_\delta^2$  respectively.

- (c) **Sequential Sampling** — Sequential sampling involves making several rounds of sampling and analysis. A statistical test is performed after each analysis to arrive at one of three possible decisions: reject the null hypothesis, accept the null hypothesis, or collect more samples. This strategy is applicable when sampling and/or analysis costs are high, when information concerning sampling and/or measurement variability is lacking, when the waste and site characteristics of interest are stable over the time frame of the sampling effort, and when the objective of the sampling is to test a single hypothesis. By taking samples in sequence, the researcher can hold down the cost of sampling and analysis.

In the case of the cadmium-contaminated ash, a sequential procedure on sample can be performed. The samples in each sampling round would be chosen randomly in three dimensions. The decision to stop sampling is

not been made before the number of samples required for the simple random sample are taken sampling would stop at this point and the simple random sample test would be performed. Standard lab splits and QC samples would be taken according to standard procedures for the RCRA program. An approximate ratio test would be used after each round of sampling is complete to decide whether or not to conclude that the waste is hazardous or to continue sampling. For each container, the following model is applicable for a systematic sample design

$$y_{ij} = \mu + \epsilon_{ij}$$

where

- $y_{ij}$  =  $i^{\text{th}}$  observation from the  $j^{\text{th}}$  sampling round
- $\mu$  = true mean
- $\epsilon_{ij}$  = deviation from  $\mu$  for the  $i^{\text{th}}$  observation (natural variability) from the  $j^{\text{th}}$  sampling round

Each  $\epsilon$  is assumed to be independently distributed with a mean of 0 and a variance of  $\sigma^2$

- (d) Stratified Random Sampling — Stratified sampling involves dividing the study area into two or more non-overlapping subsets (strata) which cover the entire volume to be sampled. These strata should be defined so that physical samples within a stratum are more similar to each other than to samples from other strata. Sampling depth, concentration level, previous cleanup attempts, and confounding contaminants can be used as the basis for creating strata. Once the strata have been defined, each stratum is then sampled separately using one of the above designs. Stratification can be used to ensure that important areas of a site are represented in the sample. In addition, a stratified random sample may provide more precise estimates of contaminant levels than those obtained from a simple random sample. Even with imperfect information, a stratified sample can be more cost-effective.

Since the incineration company has already determined that each load of ash is fairly homogenous, stratification does not have any advantages over a simple random sample. In addition, since the company has decided to test each waste load individually before it leaves the facility, stratifying each waste load would be difficult and unnecessary. Therefore, this sampling design will not be considered.

(3) For each design alternative, select the optimal sample size that satisfies the DQOs

The formula for determining the sample size (number of samples to be collected) is chosen based on the hypothesis test and sample design. Standard formulas can be found in several references, including

Cochran, W 1977 Sampling Techniques. New York. John Wiley

Desu, M M and D Raghavarao 1990 Sample Size Methodology. San Diego, CA, Academic Press

Gilbert, Richard O 1987 Statistical Methods for Environmental Pollution Monitoring. New York. Van Nostrand Reinhold.

U S Environmental Protection Agency 1989 Methods for Evaluating the Attainment of Cleanup Standards Volume 1 Soils and Solid Media. EPA 230/02-89-042 Office of Policy, Planning and Evaluation.

U S Environmental Protection Agency 1992 Methods for Evaluating the Attainment of Cleanup Standards Volume 2. Ground Water. EPA 230-R-92-014 Office of Policy Planning and Evaluation

These formulas can also be found in many basic statistics textbooks. Different formulas are necessary for each sampling design, for each parameter and for each statistical test. These formulas are generally a function of  $\alpha$ ,  $\beta$  the detection difference  $\Delta$ , and the standard deviation  $\sigma$

For example, a formula for computing the sample size necessary to meet the DQO constraints for comparing a mean against a regulatory threshold, using Normality and sampling using simple random samples is

$$n = \frac{\sigma^2 (Z_{1-\alpha} - Z_{1-\alpha})^2}{(U - RT)^2} + 0.5 (Z_{1-\alpha})^2 = \frac{\sigma^2 (Z_{1-\alpha} + Z_{1-\alpha})^2}{\Delta^2} + 0.5 (Z_{1-\alpha})^2$$

where

$n$  = number of samples

$z$  = standard normal values (from standard statistical tables),

$U$  = minimum desired detection point (i.e. edge of gray region),

$RT$  = regulatory threshold for testing against a standard (i.e. action level)

Simple Random Sample — Using the formula above  $n$  was determined that 37 samples are necessary to achieve the specified limits on decision errors. In addition this sampling plan satisfies all the DQOs including budget, schedule, and practical constraints.

Composite Sampling — To determine sample sizes for a composite sample  $n$  is necessary to compute the number of composite samples  $r$ , the number of samples  $g$  within each composite, and the number of subsamples,  $m$ , to be measured for each composite. Usually  $m=1$ , however, since this design is to be used repeatedly it is suggested that two subsamples from each composite sample be measured to estimate  $\sigma_s^2$ , which can then be used to re-optimize the number of samples  $m$  and  $g$ .

For a composite sample, with random sample locations it has been determined that eight composite samples of eight samples each are sufficient in order to meet the limits on decision errors that have been specified. It is recommended that two subsamples be taken from each composite sample to give an estimate of the within-composite variability. Then the sample design can be re-estimated and made more efficient. This design is more than sufficient to achieve the specified limits on decision errors. In addition this sampling and analysis plan satisfies all the DQOs including budget, schedule, and practical constraints.

Sequential Sampling — For a sequential sample there is no fixed sample size, but the expected sample sizes can be computed. The average sample size for concluding the waste is hazardous is 16. The average sample size for concluding the waste is non-hazardous is 22. In addition this sampling plan satisfies all the DQOs including budget, schedule, and practical constraints.

(4) Select the most resources-effective design that satisfies all of the DQOs.

Compare the overall efficiency of each model and choose the one that will solve the problem most effectively.

#### Models Applied to the Case Study

Each of the following designs will be evaluated for the case study example.

Simple Random Sampling — A simple random sampling scheme can be implemented for each load of fly ash by first generating three-dimensional random sampling points. This can most easily be done by using a computer. Samples can then be taken using a special grab sampler which will be forced into the ash, opened to take the sample, then closed and removed. The difficulty with this type of sampling scheme is that it is difficult to measure sampling locations in three dimensions and it may be difficult to gain access to the correct sampling locations.

This design meets all of the required limits on decision errors. Since 37 samples need to be taken and analyzed, the cost of this design is

$$\begin{aligned} \text{Cost}_{\text{SRS}} &= 37 \times \$10 + 37 \times \$150 \\ &= \$370 + \$5550 = \$5920 \end{aligned}$$

Composite Sampling — Composite sampling will be performed similarly to simple random sampling except that after 8 samples are collected, they will be lumped together and homogenized. Sample aliquots for analysis will be drawn from the homogenized mixture. This process will be repeated eight times until 64 samples have been collected.

This design meets all of the required limits on decision errors. The cost of this design is based on the cost of selecting (\$10) and analyzing (\$150) a sample. Eight samples will be used to make each composite sample for a sampling cost of \$80. Two subsamples will be analyzed from this composite sample for a cost of \$300. Therefore, each composite sample will cost \$380. The total cost of this design is

$$\text{Cost}_{\text{CS}} = 8 \times \$380 = \$3040$$

Sequential Sampling — Sequential sampling will be performed similarly to random sampling. The primary difference is that the ultimate number of samples will be determined by the results of one or more sampling rounds.

This design has the potential to reduce the number of samples required in the simple random sampling design and still meet the decision error limits. The possible costs of the two decisions are used below.

$$\begin{aligned} \text{The ash is hazardous} & 16 \times (\$160) = \$2,560 \\ \text{The ash is non-hazardous} & 22 \times (\$160) = \$3,520 \end{aligned}$$

To determine the expected cost, estimate the number of loads of ash that should be sent to a RCRA facility versus the number of loads that can be sent to a municipal facility. Suppose 25% of the loads are hazardous and should be sent to a RCRA facility. Then the expected cost ( $EC_{\text{SS}}$ ) of this design should be

$$\begin{aligned} EC_{\text{SS}} &= 0.25 \times (\text{cost of RCRA disposal}) + (0.75 \times \text{cost of municipal disposal}) \\ &= 0.25 \times (\$2,560) + 0.75 \times (\$3,520) = \$3,280 \end{aligned}$$

Because the simple random sampling design requires that many samples be taken and analyzed, it is inefficient for the goals of this study. Sampling will cost almost as much to determine whether the waste is hazardous or nonhazardous as it would cost to send all the waste to a RCRA hazardous waste landfill.

The sequential sampling design is more cost-effective than the simple random sampling design. The potential savings over sending all waste to a RCRA hazardous waste facility is  $\$6,750 - \$3,280 = \$3,470$ . The site owner has expressed disapproval for this sampling plan because of the time it may take before a decision can be made. If the ash was not homogenous within a container, however, this sampling design may be the design of choice.

The composite sample design is the best option. It is the most cost-effective design and requires the least amount of time to implement. In addition, the use of strata ensures full coverage of each container. It is recommended that each of the eight composite samples have two subsamples analyzed. In the future, after sufficient data have been collected to estimate the proportion of hazardous containers to non-hazardous containers and to estimate the variability within each composite sample, it will be possible to re-optimize the sample size.

- (5) Document the operational details and theoretical assumptions of the selected design in the sampling and analysis plan.

A composite sample design should be used to determine whether each container of ash should be sent to a RCRA landfill or to a municipal landfill. Eight composite samples consisting of eight grab samples should be taken from each container and two subsamples from each composite will be analyzed at the laboratory. To form the composite samples, the containers will be divided into eight strata of equal size and one grab sample will be taken randomly within each stratum and composited. Sample locations will be generated randomly using computer-generated random numbers. The model assumes that the variability within a composite sample is negligible. Data from the subsamples can be used to test this assumption and make corrections to the model.

#### Beyond the DQO Process - Evaluation of the Design using the DQA Process

For this study, the data were collected using the composite sampling design. Once the samples were collected and analyzed, the data was evaluated statistically and scientifically using the DQA Process to inspect for scientific anomalies, confirm that the model assumptions were correct, select a statistical test, and verify that the test assumptions such as distribution and independence can be met. For this study, a t-test satisfied the DQO and inspection of the data indicated that there was no reason to believe that the data was not normally distributed or that there was correlation between data points. After three weeks of sampling, approximately 30% of the waste loads leaving the incinerator were found to have hazardous concentrations of cadmium in the ash. The sampling design was determined to be cost effective because the combined cost of sampling and disposal was less than sending all of the waste to a RCRA landfill.

## APPENDIX C

## GLOSSARY OF TERMS

**action level** the numerical value that causes the decision maker to choose one of the alternative actions (e.g. compliance or noncompliance) It may be a regulatory threshold standard such as a Maximum Contaminant Level for drinking water, a risk-based concentration level, a technological limitation, or reference-based standard

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

**boundaries** the area or volume (spatial boundary) and the time period (temporal boundary) to which the decision will apply Samples are collected within these boundaries to be representative of the population of interest for the decision

**Data Quality Assessment (DQA)** a process of statistical and scientific evaluation that is used to assess the validity and performance of the data collection design and statistical tests and to establish whether a data set is adequate for its intended use

**Data Quality Objectives (DQOs)** qualitative and quantitative statements derived from the outputs of each step of the DQO Process which specify the study objectives, domain limitations, the most appropriate type of data to collect, and specify the levels of decision error that will be acceptable for the decision

**Data Quality Objectives Process** A Quality Management tool based on the Scientific Method and developed by the U.S. Environmental Protection Agency to facilitate the planning of environmental data collection activities The DQO Process enables planners to focus their planning efforts by specifying the use of the data (the decision), the decision criteria (action level), and the decision maker's acceptable decision error rates The products of the DQO Process are the DQOs

**decision errors**

**false positive error** — The false positive error occurs when data mislead a decision maker into believing that the burden of proof in a hypothesis test has been satisfied so that the null hypothesis is erroneously rejected A statistician usually refers to the false positive error as alpha ( $\alpha$ ) the level of significance size of the critical region or a Type I error depending upon its precise use

**false negative error** — The false negative error occurs when data mislead the decision maker into wrongly concluding that the burden of proof in a hypothesis test has not been satisfied so that the null hypothesis is accepted. A statistician usually refers to this as beta ( $\beta$ ), or a Type II error, depending on its precise use. It is also known as the complement of Power.

**defensible** the ability to withstand any reasonable challenge related to the veracity or integrity of laboratory documents and derived data.

**gray region** an area that is adjacent to or contains the action level, and where the consequences of making a decision error are relatively small.

**limits on decision errors** the acceptable decision error rates established by the decision maker. Economic, health, ecological, political, and social consequences should be considered when setting limits on decision errors.

**mean** the arithmetic average of a set of values.

**measurement error** the difference between the true or actual state and that which is reported from measurements.

**median** the middle value for an ordered set of  $n$  values; represented by the central value when  $n$  is odd or by the average of the two most central values when  $n$  is even.

**medium** a substance (e.g., air, water, soil) which serves as a carrier of the analytes of interest.

**natural variability** the variability that is inherent or natural to the media, objects, or people being studied.

**parameter** a numerical descriptive measure of a population.

**percentile** a value on a scale of 100 that indicates the percentage of a distribution that is equal to or below it.

**planning team** the group of people that will carry out the DQO Process. Members include the decision maker (senior manager), representatives of other data users, senior program and technical staff, senior managers (decision makers), someone with statistical expertise, and a QA/QC advisor (such as a QA Manager).

**population** the total collection of objects or people to be studied and from which a sample is to be drawn.

power function the probability of rejecting the null hypothesis ( $H_0$ ) over the range of the population. The power function is used to assess the goodness of a test or to compare two competing tests.

quality assurance (QA) an integrated system of management activities involving planning, quality control, quality assessment, reporting, and quality improvement to ensure that a product or service (e.g., environmental data) meets defined standards of quality with a stated level of confidence.

Quality Assurance Project Plan (QAPP) a formal technical document containing the detailed procedures for assuring the quality of environmental data prepared for each EPA environmental data collection activity and approved prior to collecting the data.

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.

Quality Management Plan (QMP) a formal document describing the management policies, objectives, principles, organizational authority, responsibilities, accountability, and implementation protocols of an agency, organization, or laboratory for ensuring quality in its products and utility to its users. In EPA, QMPs are submitted to QAMS for approval.

range the numerical difference between the minimum and maximum of a set of values.

sample a single item or specimen from a larger whole or group, such as any single sample of any medium (air, water, soil, etc.).

sample a group of samples from a statistical population whose properties are studied to gain information about the whole.

sample variance a measure of the dispersion of a set of values.

sampling the process of obtaining a subset of measurements from a population.

sampling error the error due to observing only a limited number of the total possible values that make up the population being studied. It should be distinguished from errors due to imperfect selection bias in response and errors of observation, measurement, or recording, etc.

standard deviation the square root of the variance.

statistic a function of the sample measurements, e.g., the sample mean or standard deviation.

**study design:** A study design specifies the final configuration of the environmental monitoring effort to satisfy the DQOs. It includes the types of samples or monitoring information to be collected, where, when, and under what conditions they should be collected, what variables are to be measured, and the Quality Assurance and Quality Control (QA/QC) components that ensure acceptable sampling error and measurement error to meet the decision error rates specified in the DQOs. The study design is the principal part of the QAPP.

**total study error** the sum of all the errors that are incurred during the process of sample design through data reporting. Total study error is related to decision error.

**true** being in accord with the actual state of affairs.

**Type I error** an error that can occur during a statistical hypothesis test. A Type I error occurs when a decision maker rejects the null hypothesis (decides that the null hypothesis is false) when it is actually true.

**Type II error** an error that can occur during a statistical hypothesis test. A Type II error occurs when the decision maker accepts the null hypothesis (decides that the null hypothesis is true) when it is actually false.

**uncertainty** a measure of the total variability associated with sampling and measurement that includes the two major error components: systematic error (bias) and random error (imprecision).