Designation: D 5792 – 95

# Standard Practice for Generation of Environmental Data Related to Waste Management Activities: Development of Data Quality Objectives[1]

This standard is issued under the fixed designation D 5792; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ϵ) indicates an editorial change since the last revision or reapproval.

## 1. Scope

1.1 This practice covers the process of development of data quality objectives (DQOs) for the acquisition of environmental data. Optimization of sampling and analysis design is a part of the DQO process. This practice describes the DQO process in detail. The various strategies for design optimization are too numerous to include in this practice. Many other documents outline alternatives for optimizing sampling and analysis design. Therefore, only an overview of design optimization is included. Some design aspects are included in the practice's examples for illustration purposes.

1.2 DQO development is the first of three parts of data generation activities. The other two aspects are (*1*) implementation of the sampling and analysis strategies and (*2*) data quality assessment. This guide should be used in concert with Practice D 5283, which outlines the quality assurance (QA) processes specified during planning and used during implementation.

1.3 Environmental data related to waste management activities include, but are not limited to, the results from the sampling and analyses of air, soil, water, biota, or waste samples, or any combinations thereof.

1.4 The DQO process should be developed and initiated prior to the application of planning, implementation, and assessment of sampling and analysis activities.

1.5 This practice presents extensive requirements of management, designed to ensure high-quality environmental data. The words "must" and "shall" (requirements), "should" (recommendation), and "may" (optional), have been selected carefully to reflect the importance placed on many of the statements in this practice. The extent to which all requirements will be met remains a matter of technical judgement.

1.6 The values stated in SI units are to be regarded as the standard. The values given in parentheses are for information only.

1.7 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

## 2. Referenced Documents

2.1 *ASTM Standards:*

C 970 Guide for Sampling Special Nuclear Materials in Multi-Container Lots[2]

C 1215 Guide for Preparing and Interpreting Precision and Bias Statements in Test Method Standards Used in the Nuclear Industry[2]

D 5283 Practice for Generation of Environmental Data Related to Waste Management Activities: Quality Assurance and Quality Control Planning and Implementation[3]

## 3. Terminology

3.1 *Definitions*—Where applicable, the originating reference is associated with the definition and follows the definition in boldface type.

3.1.1 *accuracy (see bias)*—(*1*) bias; (*2*) the closeness of a measured value to the true value; (*3*) the closeness of a measured value to an accepted reference or standard value.

3.1.1.1 *Discussion*—For many investigators, *accuracy* is attained only if a procedure is both precise and unbiased (see *bias*). Because this blending of *precision* and *accuracy* can lead to confusion, ASTM requires a statement on *bias* instead of *accuracy*. **D 5283**

3.1.2 *action level*—the numerical value that causes the decision maker to choose one of the alternative actions (for example, compliance or noncompliance). It may be a regulatory threshold standard, such as maximum contaminant level for drinking water, a risk-based concentration level, a technological limitation, or reference-based standard. **EPA QA/G-4 (1)**[4]

3.1.3 *bias (see accuracy)*—the difference between the population mean of the test results and an accepted reference value.

3.1.3.1 *Discussion*—*Bias* represents a constant error as opposed to a *random error*. A method *bias* can be estimated by the difference (or relative difference) between a measured

[1] This practice is under the jurisdiction of ASTM Committee D34 on Waste Management and is the direct responsibility of Subcommittee D34.01.01 on Planning for Sampling.

Current edition approved Nov. 10, 1995. Published January 1996.

[2] *Annual Book of ASTM Standards*, Vol 12.01.

[3] *Annual Book of ASTM Standards*, Vol 11.04.

[4] The boldface numbers in parentheses refer to the list of references at the end of this practice.

average and an accepted standard or reference value. The data from which the estimate is obtained should be statistically analyzed to establish *bias* in the presence of *random error*. A thorough *bias* investigation of a measurement procedure requires a statistically designed experiment to repeatedly measure, under essentially the same conditions, a set of standards or reference materials of known value that cover the range of application. *Bias* often varies with the range of application and should be reported accordingly. **C 1215, D 5283**

3.1.4 *confidence interval*—an interval used to bound the value of a population parameter with a specified degree of confidence (this is an interval that has different values for different samples).

3.1.4.1 *Discussion*—When providing a *confidence interval*, analysts should give the number of observations on which the interval is based. The specified degree of confidence is usually 90, 95, or 99 %. The form of a *confidence interval* depends on underlying assumptions and intentions. *Confidence intervals* are usually taken to be symmetric, but that is not necessarily so, as in the case of *confidence intervals* for *variances*. **C 1215**

3.1.5 *confidence level*—the probability, usually expressed as a percent, that a *confidence interval* will contain the parameter of interest (see discussion of *confidence interval*).

3.1.6 *data quality objectives (DQOs)*—qualitative and quantitative statements derived from the DQO process describing the decision rules and the uncertainties of the decision(s) within the context of the problem(s).

3.1.6.1 *Discussion*—DQOs clarify the study objectives, define the most appropriate type of data to collect, determine the most appropriate conditions from which to collect the data, and establish 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 used to develop a sampling and analysis design.

3.1.7 *data quality objectives process*—a quality management tool based on the scientific method and developed by the U.S. Environmental Protection Agency (EPA) 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), decision criteria (action level), and decision maker's acceptable decision error rates. The products of the DQO process are the DQOs.

3.1.7.1 *Discussion*—DQOs result from an iterative process between the decision makers and the technical team to develop qualitative and quantitative statements that describe the problem and the certainty and uncertainty that decision makers are willing to accept in the results derived from the environmental data. This acceptable level of uncertainty should then be used as the basis for the design specifications for project data collection and data assessment. All of the information from the first six steps of the DQO process are used in designing the study and assessing the data adequacy. **EPA QA/G-4**

3.1.8 *decision error*

3.1.8.1 *false negative error*—this occurs when environmental data mislead decision maker(s) into not taking action specified by a decision rule when action should be taken.

3.1.8.2 *false positive error*—this occurs when environmental data mislead decision maker(s) into taking action specified by a decision rule when action should not be taken.

3.1.9 *decision rule*—a set of directions in the form of a conditional statement that specify the following: (*1*) how the sample data will be compared to the action level, (*2*) which decision will be made as a result of that comparison, and (*3*) what subsequent action will be taken based on the decisions.

3.1.10 *precision*—a generic concept used to describe the dispersion of a set of measured values.

3.1.10.1 *Discussion*—It is important that some quantitative measure be used to specify *precision*. A statement such as "the *precision* is 1.54 g" is useless. Measures frequently used to express *precision* are standard deviation, relative standard deviation, variance, repeatability, reproducibility, confidence interval, and range. In addition to specifying the measure and the *precision*, it is important that the number of repeated measurements upon which the estimated *precision* is based also be given. **D 5283**

3.1.11 *quality assurance (QA)*—an integrated system of management activities involving planning, quality control, quality assessment, reporting, and quality improvement to ensure that a process or service (for example, environmental data) meets defined standards of quality with a stated level of confidence. **EPA QA/G-4**

3.1.12 *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. **EPA QA/G-4**

3.1.13 *random error*—(*1*) the chance variation encountered in all measurement work, characterized by the random occurrence of deviations from the mean value; (*2*) an error that affects each member of a set of data (measurements) in a different manner. **D 5283**

3.1.14 *risk*—the probability or expectation that an adverse effect will occur.

3.1.14.1 *Discussion*—*Risk* is frequently used to describe the adverse effect on health or on economics. Health-based *risk* is the probability of induced diseases in persons exposed to physical, chemical, biological, or radiological insults over time. This *risk* probability depends on the concentration or level of the insult, which is expressed by a mathematical model describing the dose and *risk* relationship. *Risk* is also associated with economics when decision makers have to select one action from a set of available actions. Each action has a corresponding cost. The *risk* or expected loss is the cost multiplied by the probability of the outcome of a particular action. Decision makers should adopt a strategy to select actions that minimize the expected loss.

3.1.15 *standard deviation*—the square root of the sum of the squares of the individual deviations from the sample average divided by one less than the number of results involved.

$$S = \sqrt{\frac{\sum_{j=1}^{n} (X_j - \bar{X})^2}{n - 1}}$$

where:
$S$ = sample standard deviation,
$n$ = number of results obtained,

$X_j$ = $j$th individual result, and
$\bar{X}$ = sample average.

3.1.15.1 *Discussion*—The use of the *standard deviation* to describe *precision* implies that the uncertainty is independent of the measurement value. The practice of associating the ± symbol with *standard deviation* (or RSD) is not recommended. The ± symbol denotes an interval. The *standard deviation* is not an interval, and it should not be treated as such. **D 5283**

## 4. Summary of Practice

4.1 This practice describes the process of developing and documenting the DQO process and the resulting DQOs. This practice also outlines the overall environmental study process as shown in Fig. 1. It must be emphasized that any specific study scheme must be conducted in conformity with applicable agency and company guidance and procedures.

4.2 For example, the investigation of a Superfund site would include feasibility studies and community relations plans, which are not a part of this practice.

## 5. Significance and Use

5.1 Environmental data are often required for making regulatory and programmatic decisions. Decision makers must determine whether the levels of assurance associated with the data are sufficient in quality for their intended use.

5.2 Data generation efforts involve three parts: development of DQOs and subsequent project plan(s) to meet the DQOs, implementation and oversight of the project plan(s), and assessment of the data quality to determine whether the DQOs were met.

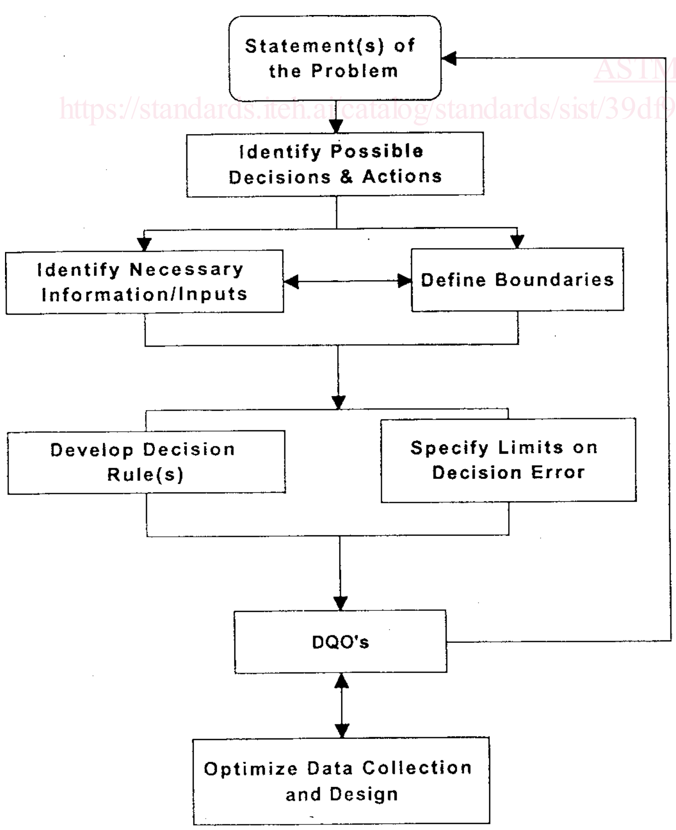


FIG. 1 DQO Process

5.3 To determine the level of assurance necessary to support the decision, an iterative process must be used by decision makers, data collectors, and users. This practice emphasizes the iterative nature of the process of DQO development. Objectives may need to be reevaluated and modified as information related to the level of data quality is gained. This means that DQOs are the product of the DQO process and are subject to change as data are gathered and assessed.

5.4 This practice defines the process of developing DQOs. Each step of the planning process is described.

5.5 This practice emphasizes the importance of communication among those involved in developing DQOs, those planning and implementing the sampling and analysis aspects of environmental data generation activities, and those assessing data quality.

5.6 The impacts of a successful DQO process on the project are as follows: (*1*) a consensus on the nature of the problem and the desired decision shared by all the decision makers, (*2*) data quality consistent with its intended use, (*3*) a more resource-efficient sampling and analysis design, (*4*) a planned approach to data collection and evaluation, (*5*) quantitative criteria for knowing when to stop sampling, and (*6*) known measure of risk for making an incorrect decision.

## 6. Data Quality Objective Process

6.1 The DQO process is a logical sequence of seven steps that leads to decisions with a known level of uncertainty (Fig. 1). It is a planning tool used to determine the type, quantity, and adequacy of data needed to support a decision. It allows the users to collect proper, sufficient, and appropriate information for the intended decision. The output from each step of the process is stated in clear and simple terms and agreed upon by all affected parties. The seven steps are as follows:

*(1)* Stating the problem,
*(2)* Identifying possible decisions,
*(3)* Identifying inputs to decisions,
*(4)* Defining boundaries,
*(5)* Developing decision rules,
*(6)* Specifying limits on decision errors, and
*(7)* Optimizing data collection design.

All outputs from steps one through six are assembled into an integrated package that describes the project objectives (the problem and desired decision rules). These objectives summarize the outputs from the first five steps and end with a statement of a decision rule with specified levels of the decision errors (from the sixth step). In the last step of the process, various approaches to a sampling and analysis plan for the project are developed that allow the decision makers to select a plan that balances resource allocation considerations (personnel, time, and capital) with the project's technical objectives. Taken together, the outputs from these seven steps comprise the DQO process. The relationship of the DQO process to the overall project process is shown in Fig. 2. At any stage of the project or during the field implementation phase, it may be appropriate to reiterate the DQO process, beginning with the first step based on new information. See Refs **(2, 3)** for examples of the DQO process.

6.2 *Step 1—Stating the Problem*:

6.2.1 *Purpose*—The purpose of this step is to state the

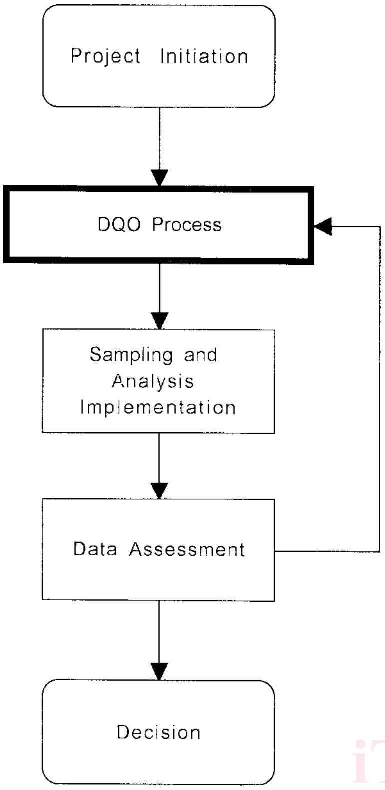


**FIG. 2 DQOs Process and Overall Decision Process**

problem clearly and concisely. The first indication that a problem (or issue) exists is often articulated poorly from a technical perspective. A single event or observation is usually cited to substantiate that a problem exists. The identity and roles of key decision makers and technical qualifications of the problem-solving team may not be provided with the first notice. Only after the appropriate information and problem-solving team are assembled can a clear statement of the problem be made.

6.2.2 *Activities*:

6.2.2.1 *Assembling of all Pertinent Information*—The necessary first action to describe a problem is to verify the conditions that indicate a problem exists. The pertinent information should be assembled during this phase of problem definition. A key source is any historical record of events at the site where the problem is believed to exist. This enables the decision makers to understand the context of the problem. A series of questions need to be developed concerning the problem.

*(1)* What happened (or could happen) that suggests a problem?

*(2)* When did it (could it) happen?

*(3)* How did it (could it) happen?

*(4)* Where did it (could it) happen?

*(5)* Why did it (could it) happen?

*(6)* How bad is (might be) the result or situation?

*(7)* How fast is (might be) the situation changing?

*(8)* What is (could be) the impact on human health and the environment?

*(9)* Who was (could be) involved?

*(10)* Who knows (should know) about the situation?

*(11)* Has anything been (might anything be) done to mitigate the problem?

*(12)* What contaminants are (could be) involved?

*(13)* How reliable is the information?

*(14)* What regulations could or should apply?

*(15)* Is there any information that suggests there is not a problem?

This list of potential information is not exhaustive, and there may be other data applicable to the definition of the problem.

6.2.2.2 *Identification of the DQO Team*—Even as information is being gathered, it is necessary to begin assembling a team of decision makers and technical support personnel to organize and evaluate the information. These individuals become the core of the DQO team and may be augmented by others as information and events dictate. The identities and roles of the DQO team members are usually determined by the decision makers who have either jurisdiction over the site and personnel or financial resources that will be used in resolving the problem. The DQO team is usually made up of the following key individuals:

*(1) Site Owners or Potentially Responsible Parties*—These individuals have authority to commit personnel and financial resources to resolve the problem and have a vital interest in the definition of the problem and possible decisions.

*(2) Representatives of Regulatory Agencies*—These individuals are usually responsible for enforcing the standards that have been exceeded, leading to classifying the observations or events as a problem. Additionally, they have an active role in characterizing the extent of the problem, approving any proposed remedial action, and concurring that the action mitigated the problem.

*(3) Project Manager*—This individual generally has the responsibility for overseeing resolution of the problem. This person may represent either the regulatory agency or the potentially responsible parties.

*(4) Technical Specialists*—These individuals have the expertise to assess the information and data to determine the nature and extent of the potential problem and may become key players in the design and implementation of proposed decisions.

It is important that these individuals be assembled early in the process and remain actively involved to foster good communications and to achieve consensus among the DQO team on important decision-related issues.

6.2.3 *Outputs*:

6.2.3.1 *Statement of Problem and Context*—Once the initial information and data have been collected, organized, and evaluated, the conclusions of the DQO team should be documented. If it is determined that no problem exists, the conclusion must be supported by a summary of the existing conditions and the standards or regulatory conditions that apply to the problem.

*(1)* If a problem is found to exist, the reasons must be stated clearly and concisely. Any standards or regulatory conditions that apply to the situation must be cited. If the initial investigation concludes that the existing conditions are the result of a series of problems, the DQO team should attempt to define as many discrete problems (or issues) as possible.

*(2)* The following are examples of problem statements:
(*a*) A former pesticide formulation facility is for sale, but it is unknown whether it meets local environmental standards for property transfer.
(*b*) An industrial site is known to be contaminated with low levels of lead, but it is unknown whether levels are below risk-based standards.
(*c*) Most of a vacant lot is believed to be uncontaminated with PCBs (<2 ppm), but it is unknown whether abandoned, leaky transformers in the vacant lot make it necessary to remove any of the top layer of soil.
(*d*) The former industrial site has contaminated soil areas that may be contaminating ground water, and it is necessary to decide which type of monitoring program will satisfy local health requirements.
(*e*) The city would like to use local ground water on an athletic field near a Superfund site, but must know how this water will impact the health of the athletes and spectators.

*(3)* Complex problems should be broken down into manageable smaller problems that are linked together to form the final decision. As an example, the sale of a piece of property may involve solving the following problems:
(*a*) Is the site contaminated? If yes, then,
(*b*) Is off-site disposal required? If no, then
(*c*) Which of two allowable on-site treatment options should be used?

6.2.3.2 *Identification of Resources*—As the nature and magnitude of the problem is being documented, the decision makers should be conferring to determine the type and amount of resources that can be committed. Preliminary budget, personnel assignments, and schedule should be established. Preliminary milestones, timelines, and approvals should be documented and concurred upon by affected decision makers. The DQO team leader and technical specialists should be included in these discussions where possible. At a minimum, they should be kept informed of these issues so their impact can be anticipated in the definition of the problem.

*(1)* Fig. 3 shows the primary components of the problem statement step. After this step is completed, the DQO team moves on to the next step, where the process to resolve the problem continues.

*(2)* It is important to remember that the DQO process is an iterative one. New information is collected as projects proceed. The DQO team members associated with the problem-statement step should remain involved with the DQO process. If new data, unavailable to the DQO team during the development of the problem statement, demonstrates that the statement is incomplete or otherwise inadequate, the problem statement should be reconsidered.

6.3 *Step 2—Identifying Possible Decisions*:

6.3.1 *Purpose*—The purpose of this step is to identify the possible decision(s) that will address the problem once it has been stated clearly. Multiple decisions are required when the problem is complex. Information required to make decisions and to define the domain or boundaries of the decision will be determined in later steps (6.4 and 6.5, respectively). Each potential decision is tested to ensure that it is worth pursuing further in the process. A series of one or more decisions will result in actions that resolve the problem. The activities that lead to identifying the decision(s) are shown in Fig. 3 and discussed in 6.3.2.

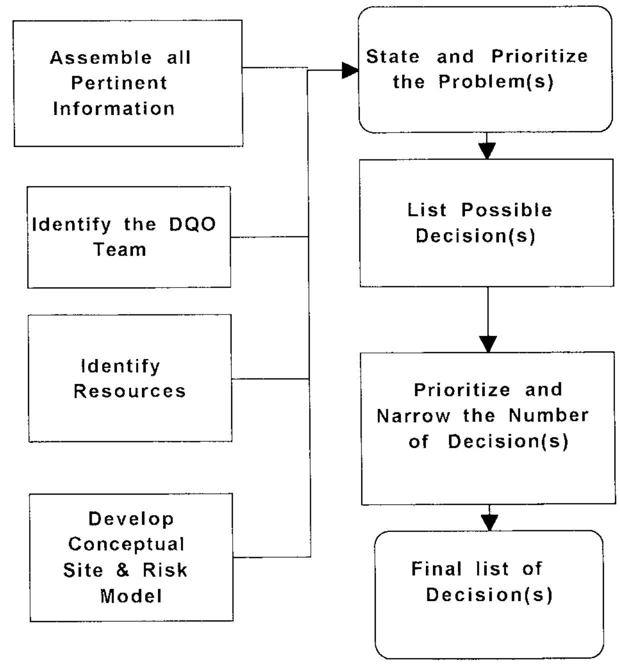


**FIG. 3 Stating the Problem and Identifying the Decisions**

6.3.2 *Activities*:

6.3.2.1 *Listing of Possible Decisions*—All possible decisions concerning the problem should be listed. Choices should not be eliminated at this time. Possible decision statements are presented in the form of a series of questions that, when answered, result in actions that will resolve the problem. Examples of questions related to problems given in 6.2.3 (Step 1) are as follows:

*(1)* Are possible contaminants on the site below regulatory thresholds?

*(2)* Must all of the surface soil be remediated to less than 5 ppm lead?

*(3)* Can only locations with PCB levels above 2 ppm be remediated?

*(4)* Will a ground water monitoring program at the site capable of detecting contaminants at the 5-ppm level satisfy regulatory requirements?

*(5)* Will a single monitoring point on or near the athletic field be sufficient?

6.3.3 *Output*—After all possible decisions that might be made have been documented, those determined to be most appropriate to resolve the problem should be prioritized by the DQO team in decreasing order of level of effort (available resources and technical challenge). Justification for the rankings should be provided. The recommended sequence in which the decisions are made should also be listed. In cases in which a complex decision statement has been broken down into a series of simpler decisions, the DQO team should identify whether the individual decisions should be addressed sequentially or in parallel. After the possible decisions have been identified, the DQO team focuses on gathering the information necessary to formulate the decision statements in Step 3 (6.4).

6.4 *Step 3—Identifying Inputs to Decisions*:

6.4.1 *Purpose*—The answers to each of the questions identified by the previous step in the DQO process must be resolved with data. Fig. 4 shows the key activities that lead to development of the data requirements. This sequence of activities must be performed for each question. Note that the limits of the study (or boundary conditions) are determined in a parallel step identified as "define boundaries" in Fig. 1. This is another type of data requirement and is discussed in 6.4.

6.4.2 *Activities*:

6.4.2.1 *Determination of Data Requirements*—At this stage of the process, it is important to carefully examine the complete set of data requirements needed to support each of the decisions. Each possible decision to be made should be considered independently of others to ensure that no omissions have occurred. After all possible questions concerning the decisions have been considered, group the data requirements together to determine overall data needs for the project. It may be possible to plan efficiencies in collecting and processing data to meet multiple needs and thereby lower overall project costs or reduce the time necessary to meet important milestones, or both.

(*1*) When considering whether specific information is needed for making a decision, test the data to ensure that it is appropriate for the decision statement. If no use of the data can be identified, it may be extraneous to the needs.

(*2*) The following list is indicative of some of the information needs that may be considered for each decision. It is not inclusive of all important data, but it provides examples common to many environmental problems.

*(a)* (*a*) What regulatory limits may be associated with the problem or regulatory issue?

*(b)* (*b*) Does contamination exceed regulatory limits?

*(c)* (*c*) What tests must be performed for the type of waste in question?

*(d)* (*d*) What are the hydrogeological considerations?

*(e)* (*e*) What populations are at risk?

*(f)* (*f*) What are the ecological considerations?

*(g)* (*g*) What process knowledge is available?

*(h)* (*h*) What historical/background data (past uses or spills) are available?

*(i)* (*i*) What are the budget constraints?

*(j)* (*j*) What is the time schedule?

*(k)* (*k*) What potential health, political, and social factors must be considered?

*(l)* (*l*) What is the potential for legal action?

*(m)* (*m*) Who is the end-user of the data?

*(n)* (*n*) What data validation criteria will be used?

*(o)* (*o*) What, if any, limitations exist on the data collection process (detection limits, matrix interferences, or no known measurement technology)?

6.4.3 *Outputs*:

6.4.3.1 The DQO team must specify data needs for each problem/decision that has been identified in the first two steps.

6.4.3.2 List the types of data required. Some example data types include, but are not limited to, the following:

*(1)* Chemical,

*(2)* Physical (including site hydrogeology and meteorology),

*(3)* Biological,

*(4)* Toxicological,

*(5)* Historical,

*(6)* Economic (time, budget, and manpower),

*(7)* Demographic,

*(8)* Toxicity characteristics, and

*(9)* Fate and transport model output.

6.4.3.3 *Listing of Data Generation Activities*—Determine which data can be acquired from historical records and which new data must be obtained in the field or laboratory, or both. If the DQO team determines that no new data are necessary to make a decision, they should document their reasoning. If new information is necessary, activities that will be required to generate inputs (data) affecting the decision should be listed. Examples of these include, but are not limited to, the following:

*(1)* Assembly of historical data,

*(2)* Sampling and chemical analysis,

*(3)* Physical testing, and

*(4)* Modeling.

6.4.3.4 *Definition of Data Use(s)*—Each set of data will be used for some purpose. This purpose must be defined. For example, will action levels for contaminants be determined by a risk-based calculation, by reference dose, or by pre-defined threshold values established by regulators? If so, ensure that data requirements are consistent with the criteria against which

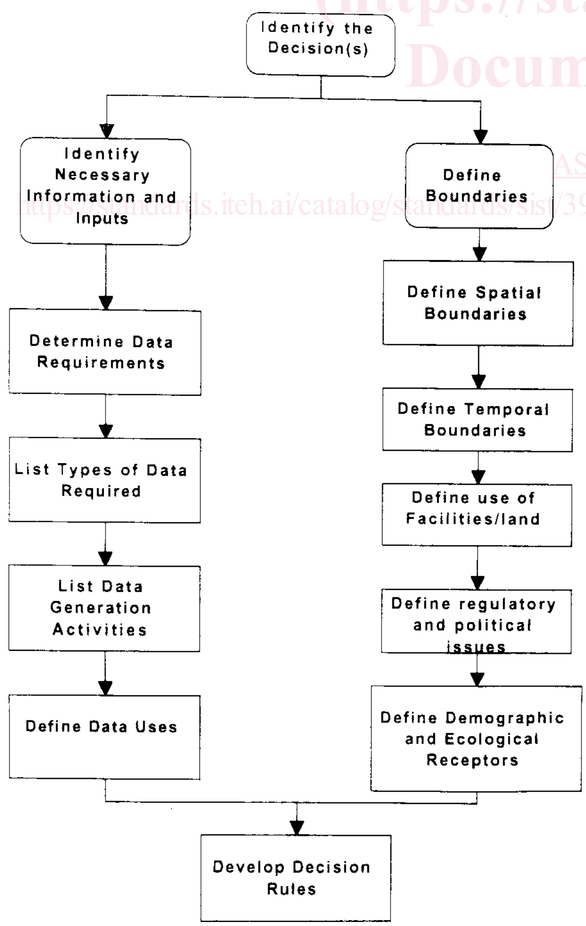


**FIG. 4 Determination of Information Inputs and Study Boundaries**