

# APPENDIX B: THE DATA QUALITY OBJECTIVES PROCESS

## B1.0 Introduction

MARLAP's objective in this appendix is to provide information about the basic framework of the DQO process (ASTM 5792; EPA, 2000; NRC, 1998; MARSSIM, 1997). The DQO planning process empowers both data users and data suppliers to take control and resolve issues in a stepwise fashion. It brings together at the right time all key players from the data user and data supplier constituencies and enables each participant to play a constructive role in clearly defining:

- The problem that requires resolution;
- What type, quantity, and quality of data the decision maker needs to resolve that problem;
- Why the decision maker needs that type and quality of data;
- How much risk of making a wrong decision is acceptable; and
- How the decision maker will use the data to make a defensible decision.

The DQO Process provides a logic for setting well-defined, achievable objectives and developing a cost-effective, technically sound sampling and analysis design. It balances the data user's tolerance for uncertainty with the available resources for obtaining data. The number of visible and successful applications of the DQO process has proven its value to the environmental community. The DQO process is adaptable depending on the complexity of the project and the input from the decision makers. Some users have combined DQO planning with remedy selection for restoration projects (e.g., DOE's SAFER—see Appendix A.5). Other users have integrated the project scoping meetings with the DQO Process. Much of the information that is developed during the DQO process is useful for the development of the project plan documents (Chapter 4) and the implementation of the data validation process (Chapter 8) and the data quality assessment (DQA) process (Chapter 9).

Since its inception, the term "data quality objectives" has been adopted by many organizations, and the definition has been adapted and modified (see box on next page). Throughout this document, MARLAP uses EPA's (2000) definition of DQOs: "Qualitative and quantitative statements derived from the DQO process that clarify study objectives, define the appropriate type of data, and specify the tolerable levels of potential decision errors that will be used as the basis for establishing the quality and quantity of data needed to support decisions."

**Definitions of Data Quality Objectives**

- (1) Statements on the level of uncertainty that a decision maker is willing to accept in the results derived from environmental data (ASTM 5283; EPA, 1986).
- (2) Qualitative and quantitative statements derived from the DQO process that clarify study objectives, define the appropriate type of data, and specify the tolerable levels of potential decision errors that will be used as the basis for establishing the quality and quantity of data needed to support decisions (EPA, 2000).
- (3) 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) (ASTM D5792).
- (4) The qualitative and quantitative statements that specify the quality of the data required to support decisions for any process requiring radiochemical analysis (radioassay) (ANSI 42.23).

**B2.0 Overview of the DQO Process**

The DQO process (Figure B1) consists of seven steps (EPA, 2000). In general, the first four steps of the DQO Process require the project planning team to define the problem and qualitatively determine required data quality. Once these steps have been addressed adequately, the last three steps of the process establish quantitative performance measures for the decision and the data.

The last step of the process involves developing the data collection design based on the DQOs, which is dependent on a clear understanding of the first six steps.

Although the DQO process is described as a sequence of steps, it is inherently iterative. The output from each step influences the choices that will be made in subsequent steps. For instance, a decision rule cannot be created without first knowing the problem and desired decision. Similarly, optimization of the sampling and analysis design generally cannot occur unless it is clear what is being optimized—the results of the preceding steps. Often the outputs of one step will trigger the need to

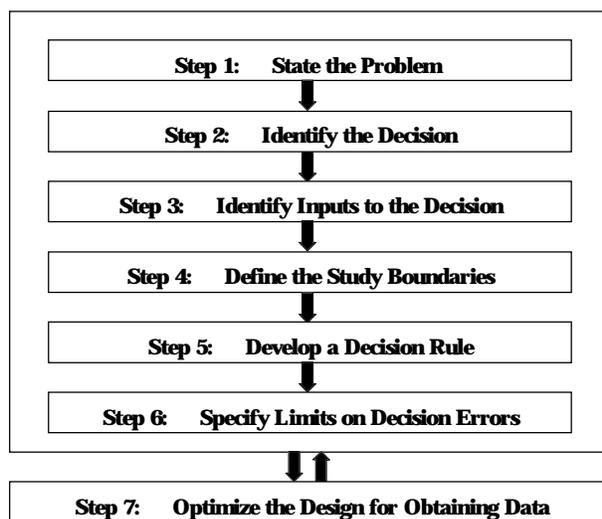


Figure B1—Seven steps of the DQO process.

65 rethink or address issues that were not evaluated thoroughly in prior steps. These iterations lead  
66 to a more focused sampling and analysis design for resolving the defined problem. The first six  
67 steps should be completed before the sampling and analysis design is developed, and every step  
68 should be completed before data collection begins. The DQO process is considered complete  
69 with the approval of an optimal design for sampling and analysis to support a decision or when  
70 available historical data are sufficient to support a decision.

71 In practice, project planning teams often do a cursory job on the first four steps, wanting to get  
72 into technical design issues immediately. Without carefully defining the problem and the desired  
73 result, the project planning team may develop a design that is technically sound but answers the  
74 wrong question, or answers the questions only after the collection of significant quantities of  
75 unnecessary data. Time spent on the first four steps is time well spent. Extra effort must be given  
76 to assure that Steps 1 to 4 are adequately addressed.

77 When applying the DQO process, or any planning approach, it is important to document the  
78 outputs of each step to assure that all participants understand and approve the interim products,  
79 and that they have a clear record of their progress. It is sometimes useful to circulate an approval  
80 copy with signature page to ensure agreement of the stakeholders.

### 81 **B3.0 The Seven Steps of the DQO Process**

82 Each step of the DQO process will be discussed in the following sections. Not all items will be  
83 applicable to every project. The project planning team should apply the concepts that are  
84 appropriate to the problem.

#### 85 **B3.1 DQO Process Step 1: State the Problem**

86 The first step is to define the problem clearly. The members of the project planning team present  
87 their concerns, identify regulatory issues and threshold levels, and review the site history. The  
88 project planning team should develop a concise description of the problem. Some elements to  
89 include in the description might be the study objectives, regulatory context, groups who have an  
90 interest in the study, funding and other resources available, previous study results, and any  
91 obvious sampling design constraints. The more facts, perceptions and concerns of the key  
92 stakeholders—including important social, economic, or political issues—that are identified  
93 during this step, the better the chances are that the issues driving the decisions and actions will be  
94 identified.

95 The primary decision maker should be identified. The resources and relevant deadlines to address  
96 the problem are also defined at this time. If possible, a “site conceptual model” should be  
97 developed. This will help structure and package the diverse facts into an understandable picture  
98 of what the various issues are and how those issues can be focused into a specific problem.  
99 The expected outputs of Step 1 are:

- 100 • A conceptual model that packages all the existing information into an understandable picture  
101 of the problem;
- 102 • A list of the project planning team members and identification of the decision maker;
- 103 • A concise description of the problem; and
- 104 • A summary of available resources and relevant deadlines for the study.

### 105 **B3.2 DQO Process Step 2: Identify the Decision**

106 During Step 2 of the DQO Process, the project planning team defines what decision must be  
107 made or what question the project will attempt to resolve. The decision (or question) could be  
108 simple, like whether a particular discharge is or is not in compliance, or the decision could be  
109 complex, such as determining if observed adverse health is being caused by a non-point source  
110 discharge. Linking the problem and the decision focuses the project planning team on seeking  
111 only that information essential for decision making, saving valuable resources (time and money).

112 The result may be a comprehensive decision for a straightforward problem, or a sequence of  
113 decisions for a complex problem. For complex problems with multiple concerns, these concerns  
114 should be prioritized in order of importance. Often a complex concern is associated with a series  
115 of decisions that need to be made. Once these decisions have been identified, they should be  
116 sequenced in a logical order so the answer to one decision provides input in answering the next  
117 decision. It may be helpful to develop a logic flow diagram (decision framework), arraying each  
118 element of the issue in its proper sequence along with its associated decision that requires an  
119 answer.

120 The term “action level” is used in this document to denote the numerical value that will cause the  
121 decision maker to choose one of the alternative actions. The action level may be a derived  
122 concentration guideline level, background level, release criteria, regulatory decision limit, etc.  
123 The action level is often associated with the type of media, analyte and concentration limit. Some  
124 action levels, such as the release criteria for license termination, are expressed in terms of dose or

125 risk. The release criterion typically is based on the total effective dose equivalent (TEDE), the  
126 committed effective dose equivalent (CEDE), risk of cancer incidence (morbidity) or risk of  
127 cancer death (mortality) and generally can not be measured directly. A radionuclide-specific  
128 predicted concentration or surface area concentration of specific nuclides that can result in a dose  
129 (TEDE or CEDE) or specific risk equal to the release criterion is called the “derived concentra-  
130 tion guideline level” (DCGL). A direct comparison can be made between the project’s analytical  
131 measurements and the DCGL (MARSSIM, 1997).

132 The project planning team should define the possible actions that may be taken to solve the  
133 problem. Consideration should be given to the option of taking no action. A decision statement  
134 can then be developed by combining the decisions and the alternative actions. The decision rule  
135 and the related hypothesis test will be more fully developed in the DQO process at Steps 5 and 6.

136 By defining the problem and its associated decision clearly, the project planning team has also  
137 begun to define the inputs and boundaries (DQO process Steps 3 and 4). At the end of Step 2, the  
138 project planning team has:

- 139 • Identified the principal decisions or questions;
- 140 • Defined alternative actions that could be taken to solve the problem based on possible  
141 answers to the principal decisions and questions;
- 142 • Combined the principal decisions and questions and the alternative actions into decision  
143 statements that expresses a choice among alternative actions; and
- 144 • Organized multiple decisions.

### 145 **B3.3 DQO Process Step 3: Identify Inputs to the Decision**

146 During Step 3, the project planning team makes a formal list of the specific information required  
147 for decision making. The project planning team should determine what information is needed and  
148 how it can be acquired. The project planning team should specify if new measurements are  
149 required for the listed data requirements. The data required are based on outcomes of discussion  
150 during the previous two steps. The project planning team should define the basis for setting the  
151 action level. Depending on the level of detail of the discussion during the previous steps, then  
152 efforts associated with Step 3 may be primarily to capture that information. If the first two steps  
153 have not defined the inputs with enough specificity, then those inputs should be defined here.

154 However, before going further, the output should be reviewed to assure that the problem, the  
155 decision steps and the input are compatible in complete agreement.

156 An important activity during Step 3 is to determine if the existing data or information, when  
157 compared with the desired information, has significant gaps. If no gaps exist, then the existing  
158 data or information may be sufficient to resolve the problem and make the decision. (Although  
159 there may be no gaps in the data, the data may not have enough statistical power to resolve the  
160 action level. See Step 6 for more discussion.) In order to optimize the use of resources, the  
161 project planning team should maximize the use of historical information. If new data are  
162 required, then this step establishes what new data (inputs) are needed. The specific environmental  
163 variable or characteristic to be measured should be identified. The DQO Process clearly links  
164 sampling and analysis efforts to an action and a decision. This linkage allows the project  
165 planning team to determine when enough data have been collected.

166 If the project planning team determines that collection of additional data is needed, the analytical  
167 laboratory acquisition strategy options should be considered at this stage. Identifying suitable  
168 contracting options should be based on the scope, schedule, and budget of the project, and the  
169 capability and availability of laboratory resources during the life of the project, and other  
170 technical considerations of the project. If an ongoing contract with a laboratory is in place, it is  
171 advisable to involve them with the radioanalytical specialists as early as possible.

172 The project planning team should ensure that there are analytical protocols available to provide  
173 acceptable measurements. If analytical methods do not exist, the project planning team will need  
174 to consider the resources needed to develop a new method, reconsider the approach for providing  
175 input data, or perhaps reformulate the decision statement.

176 The expected outputs of Step 3 are:

- 177 • A list of information needed for decision making;
- 178 • Determination of whether data exists and are sufficient to resolve the problem;
- 179 • Determination of what new data, if any, are required;
- 180 • Defined the characteristics that define the population and domain of interest;
- 181 • Defined the basis for the action level;
- 182 • Confirmation that appropriate analytical protocols exist to provide the necessary data; and
- 183 • A review of the planning output to assure the problem, decision and inputs are fully linked.

184 **B3.4 DQO Process Step 4: Define the Study Boundaries**

185 In Step 4, the project planning team should define clearly the geographic area within which the  
186 decisions will apply. The project planning team specifies the spatial and temporal boundaries  
187 covered by the decision statement. The spatial boundaries define the physical aspects to be  
188 studied in terms of geographic area, media, and any appropriate subpopulations (e.g., an entire  
189 plant, entire river basin, one discharge, metropolitan air, emissions from a power plant). When  
190 appropriate, divide the population into strata that have relatively homogeneous characteristics.  
191 The temporal boundaries describe the time frame the study data will represent (e.g., possible  
192 exposure to local residents over a 30-year period) and when samples should be taken (e.g.,  
193 instantaneous samples, hourly samples, annual average based on monthly samples, samples after  
194 rain events). Changing conditions that could impact the success of sampling and analysis and  
195 interpretation need to be considered. These factors include weather, temperature, humidity, or  
196 amount of sunlight and wind.

197 The scale of decision is also defined during this step. The scale of decision selected should be the  
198 smallest, most appropriate subset of the population for which decisions will be made based on  
199 the spatial or temporal boundaries. During Step 4, the project planning team also should identify  
200 practical constraints on sampling and analysis that could interfere with full implementation of the  
201 data collection design. These include time, personnel, equipment, and seasonal or meteorological  
202 conditions when sampling is not possible or may bias the data.

203 In practice, the study boundaries are discussed when the decision makers agree on the problem  
204 and its associated decision. For instance, a land area that may be contaminated or a collection of  
205 waste containers would be identified as part of the problem and decision definition in Steps 1 and  
206 2. The boundaries also would be considered when determining inputs to the decision in Step 3. If  
207 the study boundaries had not been addressed before Step 4 or if new issues were raised during  
208 Step 4, then Steps 1, 2, and 3 should be revisited to determine how Step 4 results are now  
209 influencing the three previous steps.

210 The outputs of Step 4 are:

- 211 • A detailed description of the spatial and temporal boundaries of the problem; and
- 212 • Any practical constraints that may interfere with the sampling and analysis activities.

213 **B3.5 Outputs of DQO Process Steps 1 to 4 Lead Into Steps 5 to 7**

214 At this stage in the DQO process, the project planning team has defined with a substantial degree  
215 of detail the problem, its associated decision, and the inputs and boundaries for addressing that  
216 problem. The project planning team knows whether it needs new data to fill specific gaps and  
217 what that data should be. The remaining three steps are highly technical and lead to the selection  
218 of the sampling and analysis design. Even when new data is not required (i.e., a data collection  
219 design is not needed), the project planning team should continue with Steps 5 and 6 of the DQO  
220 Process. By establishing the formal decision rule and the quantitative estimates of tolerable  
221 decision error rates, the project planning team is assured that consensus has been reached on the  
222 actions to be taken and information to establish criteria for DQA process.

223 It is important to emphasize that every effort must be made to assure that Steps 1 to 4 are  
224 adequately addressed. If the necessary time is taken in addressing carefully the first four steps  
225 and assuring consensus among the project planning team, then the three remaining steps are less  
226 difficult.

227 **B3.6 DQO Process Step 5: Develop a Decision Rule**

228 In Step 5, the project planning team determines the appropriate statistical parameter that  
229 characterizes the population, specifies the action level, and integrates previous DQO process  
230 outputs into a single “if ..., then ...” statement (called a “decision rule”) that describes a logical  
231 basis for choosing among alternative actions. (The statistical parameters are discussed in more  
232 detail in Chapter 19, *Measurement Statistics*.)

233 The four main elements to the decision rule are:

- 234 1. THE PARAMETER OF INTEREST. A descriptive measure (e.g., mean, median, or proportion) that  
235 specifies the characteristic or attribute that the decision maker would like to know and that  
236 the data will estimate. The characteristics that define the population and domain of interest  
237 was established in Step 3.
- 238 2. THE SCALE OF DECISION MAKING. The smallest, most appropriate subset for which decisions  
239 will be made. The scale of decision making was previously defined in Step 4.
- 240 3. THE ACTION LEVEL. A threshold value of the parameter of interest that provides the criterion  
241 for choosing among alternatives. Action levels may be based on regulatory standards or they

242 may be derived from site- and analyte-specific criteria such as dose or risk analysis. The basis  
243 for the action level was determined in Step 3.

244 4. THE ALTERNATIVE ACTIONS. The actions the decision maker would take, depending on the  
245 “true value” of the parameter of interest. The alternative actions were determined in Step 2.

246 The decision rule is a logical, sequential set of steps to be taken to resolve the problem. For  
247 example, “If one or more conditions exists then take action 1, otherwise take action 2.”

248 The outputs of Step 5 are:

- 249 • The action level;
- 250 • The statistical parameter of interest; and
- 251 • An “if ..., then ...” statement that defines the conditions that would cause the decision maker  
252 to choose among alternative courses of action.

### 253 **B3.7 DQO Process Step 6: Specify the Limits on Decision Errors**

254 In Step 6 of the DQO process, the project planning team assesses the potential consequences of  
255 making a wrong decision and establishes a tolerable level for making a decision error. The  
256 project planning team defines the types of decision errors (Type I and II) and the tolerable limits  
257 on the decision error rates. In general, a Type I error is deciding against the default assumption  
258 (the null hypothesis) when it is actually true; a Type II error is not deciding against the null  
259 hypothesis when it is actually false (see Attachment B1 and Appendix C for detailed  
260 discussions). The limits on the decision errors will be used to establish measurement  
261 performance criteria for the data collection design.

262 Traditionally, the principles of statistical hypothesis testing (see Chapter 19) have been used to  
263 determine tolerable levels of decision error rates. Other approaches applying decision theory have  
264 been applied (Bottrell, et al., 1996a,b). Based on an understanding of the possible consequences  
265 of making a wrong decision in taking alternative actions, the project planning team chooses the  
266 null hypotheses and judges what decision error rates are tolerable for making a Type I or Type II  
267 decision error.

268 The project planning team also specifies a range of possible values where the consequences of  
269 decision errors are relatively minor (the gray region). Specifying a gray region is necessary  
270 because variability in the population and imprecision in the measurement system combine to  
271 produce variability in the data such that the decision may be “too close to call” when the true

272 value is very near the action level. The gray region establishes the minimum distance from the  
273 action level where it is most important that the project planning team control Type II errors. (For  
274 additional information on the gray region, hypothesis testing, and decision errors, see EPA  
275 (2000), NRC (1998), and Chapter 19, *Measurement Statistics*.)

276 The tolerable decision error rates are used to establish performance goals for the data collection  
277 design. Overall variability in the result can be attributed to several sources, including sample  
278 location, collection, and handling; laboratory handling and analysis; and data handling and  
279 analysis. In many environmental cases, sampling is a much larger source of uncertainty than  
280 laboratory analyses. The goal is to develop a sampling and analysis design that reduces the  
281 chance of making a wrong decision. The greater certainty demanded by the decision makers, the  
282 more comprehensive and expensive the data collection process is likely to be. In this step, the  
283 project planning team has to come to an agreement on how to determine acceptable analytical  
284 uncertainty and how good the overall data results are required to be. The team has to reach a  
285 consensus on the trade-off between the cost of more information and the increased certainty in  
286 the resulting decision.

287 Often the project planning team does not feel comfortable with the concepts and terminology of  
288 hypothesis testing (Type I and Type II errors, gray zone, critical region, tolerable decision error  
289 rates). As a result the project planning team may have difficulty (or want to skip) this step of the  
290 directed planning process. If these steps are skipped or insufficiently addressed, it is more likely  
291 that the data will not be of the quality needed for the project. Attachment B1 is provided to give  
292 some additional guidance on these concepts. MARLAP recommends that for each radionuclide  
293 of concern an action level, gray region and limits on decision error rates be established during a  
294 directed planning process.

295 Figure B2 summarizes the outputs of the decisions made by the project planning team in a  
296 Decision Performance Goal Diagram (EPA, 2000). The horizontal axis represents the (unknown)  
297 true value of the parameter being estimated. The vertical axis represents the decision maker's  
298 desired probability of concluding that the parameter exceeds an action limit. The "gray region"  
299 (bounded on one side by the action level) defines an area where the consequences of decision  
300 error are relatively minor (in other words, it defines how big a divergence from the action level  
301 we wish to distinguish). The gray region is related to the desired precision of the measurements.  
302 The height of the indicated straight lines to the right and left of the gray region depict the  
303 decision maker's tolerance for Type I and Type II errors.

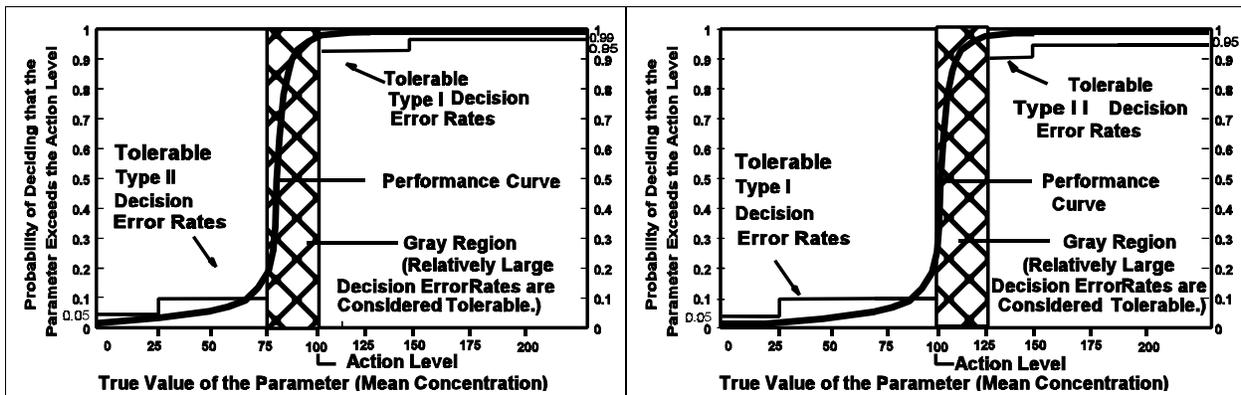


Figure B2(a)—Decision performance goal diagram null hypothesis: the parameter exceeds the action level.

Figure B2(b)—Decision performance goal diagram null hypothesis: the parameter is less than the action level.

304 For purposes of this example, the default assumption (null hypothesis) was established as the  
 305 measured concentration exceeded the action level (Figure B2a). The Type I error (5 percent at  
 306 true concentration between 100 and 150; 1percent at >150 units) making a decision NOT to take  
 307 action to solve an environmental problem (e.g., remediate) when that action was in fact required  
 308 (e.g., analyte concentrations are really above an action level). The Type II error (5 percent at true  
 309 concentrations <25 units; 10 percent between 25 and 75 units) is understood as taking an action  
 310 when in fact that action is not required (e.g., analyte concentrations are really below the action  
 311 level).

312 In Figure B2(b), the default assumption (null hypothesis) was established as the measured  
 313 concentration is less than the action level. The Type I error (5 percent at true concentrations <25  
 314 units; 10 percent between 25 and 100 units) is understood as taking an action when in fact that  
 315 action is NOT required (e.g., analyte concentrations are really below the action level). The Type  
 316 II error (10 percent at true concentration between 100 and 150; 5 percent at >150 units) is  
 317 understood as making a decision not to take action to solve an environmental problem (e.g.,  
 318 remediate) when that action was in fact required (e.g., analyte concentrations are really above an  
 319 action level).

320 The output of Step 6 is:

- 321 • The project planning team’s quantitative measure of tolerable decision error rates based
- 322 on consideration of project resources.

323 **B3.8 DQO Process Step 7: Optimize the Design for Obtaining Data**

324 By the start of Step 7, the project planning team has established their priority of concerns, the  
325 definition of the problem, the decision or outcome to address the posed problem, the inputs and  
326 boundaries, and the tolerable decision error rates. They have also agreed on decision rules that  
327 incorporate all this information into a logic statement about what action to take in response to the  
328 decision. During Step 7, the hard decisions are made between the planning team's desire to have  
329 measurements with greater certainty and the reality of the associated resource needs (time, cost,  
330 etc.) for obtaining that certainty.

331 During Step 7, the project planning team optimize the sampling and analytical design and  
332 established the measurement quality objectives (MQOs) so the resulting data will meet all the  
333 established constraints in the most resource-effective manner. The goal is to determine the most  
334 efficient design (combination of sample type, sample number and analytical procedures) to meet  
335 all the constraints established in the previous steps. Once the technical specialists and the rest of  
336 the project planning team come to agreement about the sampling and analysis design, the  
337 operational details and theoretical assumptions of the selected design should be documented.

338 If a proposed design cannot be developed to meet the limits on decision error rates within budget  
339 or other constraints, then the project planning team will have to consider relaxing the error  
340 tolerance, adjusting the width of the gray region, redefining the scale of decision, or committing  
341 more funding. There is always a trade off between quality, cost and time. The project planning  
342 team will need to develop a consensus on how to balance resources and data quality. If the  
343 proposed design requires analysis using analytical protocols not readily available, the project  
344 planning team must consider the resources (time and cost) required to develop and validate a  
345 method, generate method detection limits relevant to media of concern, and develop appropriate  
346 QA/QC procedures and criteria (Chapter 6, *Selection and Application of an Analytical Method*).

347 If the project entails a preliminary investigation of a site or material for which little is known, the  
348 planners may choose to employ MQOs and requirements that typically are achieved by the  
349 selected sampling and analytical procedures. At this early point in the project, the lack of detailed  
350 knowledge of the site or material may postpone the need for the extra cost of more expensive  
351 sampling and analytical procedures and large numbers of samples, until more site or material  
352 knowledge is acquired. The less-demanding MQOs, however, should be adequate to further  
353 define the site or material. For situations when the measured values are distant from an action  
354 level the MQO-compliant data could also be sufficient to support the project decision.

355 The planning of data collection activities is typically undertaken to determine if a characteristic  
356 of an area or item does or does not exist above an action level. Since the area of interest (popula-  
357 tion) is usually too large to be submitted to analyses, in its entirety, these data collection activities  
358 generally include sampling. If sampling is done correctly, the field sample or set of field samples  
359 will represent the characteristics of interest and, if analyzed properly, the information gleaned  
360 from the samples can be used to make decisions about the larger area. However, if errors occur  
361 during implementation of the project, the samples and associated data may not accurately reflect  
362 the material from which the samples were collected and incorrect decisions could be made.

363 The planning team attempts to anticipate, quantify, and minimize the uncertainty in decisions  
364 resulting from imprecision, bias, and blunders—or in other words, attempts to manage uncer-  
365 tainty by managing its sources. The effort expended in managing uncertainty is project dependent  
366 and depends upon what constitutes an acceptable level of decision uncertainty and the proximity  
367 of the data to a decision point. For example, Figure B3(a) presents a situation where the data  
368 have significant variability. Yet the variability of the data does not materially add to the  
369 uncertainty of the decision since the measurements are so far removed from the action level.  
370 More resources could be expended to control the variability. However, the additional expenditure  
371 would be unnecessary, since they would not alter the decision or measurably increase confidence  
372 in the decision.

373 In contrast, Figure B3(b) depicts data with  
374 relatively little variability, yet this level of  
375 variability is significant since the measured  
376 data are adjacent to the action level, which  
377 results in increased uncertainty in the  
378 decision. Depending upon the consequences  
379 of an incorrect decision, it may be advisable  
380 to expend more resources with the intention  
381 of increasing confidence in the decision.

382 The output of Step 7 is:

- 383 • The most resource-effective design for  
384 sampling and analysis that will obtain  
385 the specific amount and quality of data  
386 needed to resolve the problem within  
387 the defined constraints; and

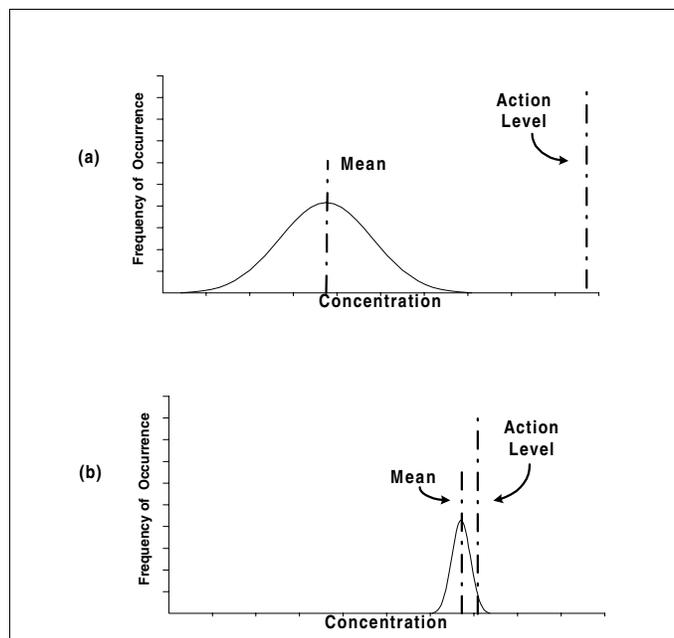


Figure B3 — How Proximity to the action level determines what is an acceptable level of uncertainty.

- 388       • Detailed plans and criteria for data assessment.

389       **B3.9 References**

- 390       American National Standards Institute (ANSI) N42.23. *American National Standard*  
391       *Measurement and Associated Instrument Quality Assurance for Radioassay Laboratories.*  
392       1996.
- 393       American Society for Testing and Materials (ASTM) D5283. *Standard Practice for Generation*  
394       *of Environmental Data Related to Waste Management Activities: Quality Assurance and*  
395       *Quality Control Planning and Implementation.* 1992.
- 396       American Society for Testing and Materials (ASTM) D5792. *Standard Practice for Generation*  
397       *of Environmental Data Related to Waste Management Activities: Development of Data*  
398       *Quality Objectives,* 1995.
- 399       American Society for Testing and Materials (ASTM) D6051. *Standard Guide for Composite*  
400       *Sampling and Field Subsampling for Environmental Waste Management Activities.* 1996.
- 401       Bottrell, D., S. Blacker and D. Goodman. 1996a. “Application of Decision Theory Methods to  
402       the Data Quality Objectives Process.” *In*, Proceedings of the Computing in Environmental  
403       Resource Management Conference, Air and Waste Management Association.
- 404       Bottrell, D., N. Wentworth, S. Blacker and D. Goodman. 1996b. “Improvements to Specifying  
405       Limits on Decision Errors in the Data Quality Objectives Process.” *In*, Proceedings of the  
406       Computing in Environmental Resource Management Conference, Air and Waste  
407       Management Association.
- 408       MARSSIM. 1997. Multi-Agency Radiation Survey and Site Investigation Manual. NUREG-  
409       1575, EPA 402-R-97-016.
- 410       U.S. Environmental Protection Agency (EPA). 2000. *Guidance for the Data Quality Objective*  
411       *Process* (EPA QA/G-4). EPA/600/R-96/055, Washington, DC. available from [www.epa.gov/](http://www.epa.gov/quality1/qa_docs.html)  
412       [quality1/qa\\_docs.html](http://www.epa.gov/quality1/qa_docs.html).
- 413       U.S. Environmental Protection Agency (EPA). 1986. *Development of Data Quality Objectives,*  
414       *Description of Stages I and II.* Washington, DC.

415 U.S. Nuclear Regulatory Commission (NRC). 1998. *A Nonparametric Statistical Methodology*  
416 *for the Design and Analysis of Final Status Decommissioning Surveys*. NUREG-1505, Rev.  
417 1.

# ATTACHMENT B-1 DECISION ERROR RATES AND THE GRAY REGION

## B-1.1 Introduction

This attachment is provided to present some additional discussion on decision error rates and the gray region. The project planning team will need to specify a range of possible values where the consequences of decision errors are relatively minor—the “gray region.” Specifying a gray region is necessary because variability in the population and imprecision in the measurement system combine to produce variability in the data such that the decision may be “too close to call” when the true value is very near the action level. The gray region establishes the minimum distance from the action level, where it is most important that the project planning team control Type II errors.

## B-1.2 The Region of Interest

The first step in constructing the gray region is setting the range of concentrations that is a region of interest (a range of possible values). Usually there is an action level (such as the derived concentration guideline level, a regulatory limit) that should not be exceeded. If the project planning team wants a method to measure sample concentrations around this level, they would not select one that worked at concentrations at 10 to 100 times the action level, nor would they select one that worked from zero to half the action level. They would want a method that worked well around the action level—perhaps from 0.1 to 10 times the action level, or from one-half to two times the action level. For the purpose of the example in this attachment, the action level is 1.0 and the project planning team selected a region of interest that is zero to twice the action level (0-2), as shown on the x-axis in Figure B-1.1.

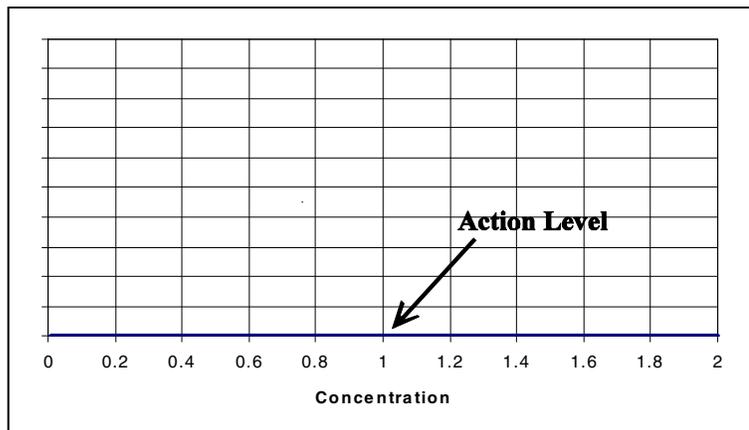


FIGURE B-1.1

449

### B-1.3 Measurement Uncertainty at the Action Level

450  
451  
452  
453  
454  
455

The action level marks the concentration level that the project planning team must be able to distinguish. The project planning team wants to be able to tell if the measured concentration is above or below the action level. Does this mean that the project planning team needs to be able to distinguish 0.9999 times the action level from 1.0001 times the action level? Sometimes, but not usually. This is fortunate, because current measurement techniques are probably not good enough to distinguish that small a difference in concentrations.

456 How close to the action level can  
457 the project planning team plan to  
458 measure? For this example, we will  
459 assume that the standard uncertainty  
460 (1 sigma,  $\sigma$ ) of the measured  
461 concentration is 10 percent of the  
462 action level. With that kind of  
463 measurement “precision,” can the  
464 project planning team tell the  
465 difference between a sample with  
466 0.9 times the action level from one  
467 right at the action level? Not  
468 always. Figure B-1.2 shows the  
469 distribution of the concentration that  
470 is measured (assuming a normal distribution). This means that about 16 percent of the time, the  
471 measured concentration (in the shaded area) will appear to be 0.9 times the action level or less,  
472 even though the true concentration is exactly equal to the action level.

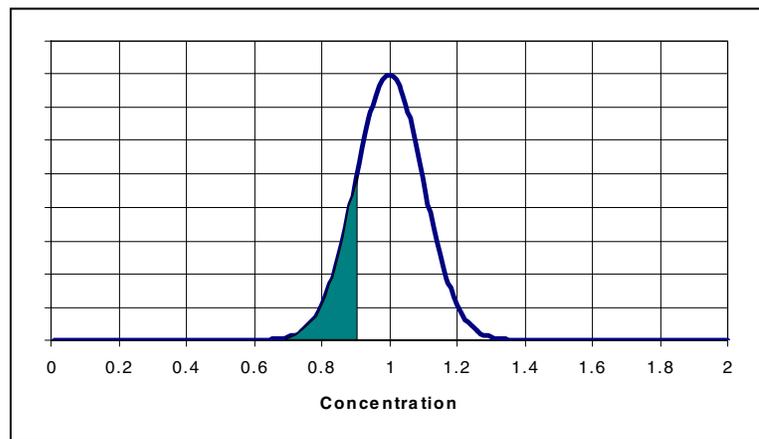


FIGURE B-1.2

473 Similarly, about 16 percent of the  
474 time, the measured concentration  
475 will appear to be at or above the  
476 action level (as shown in the shaded  
477 area in Figure B-1.3), even though  
478 the true concentration is only 0.9  
479 times the action level.

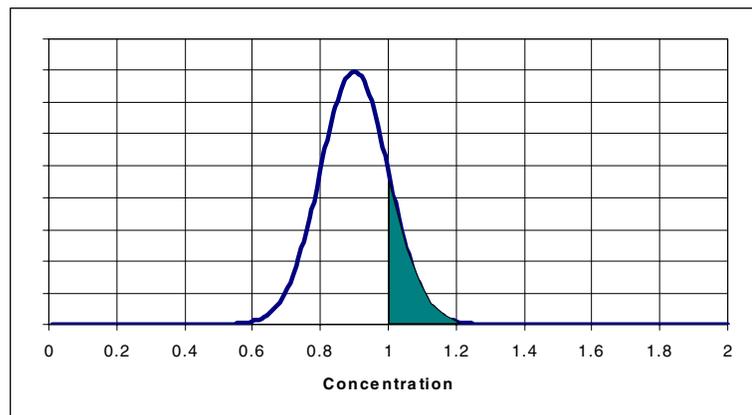


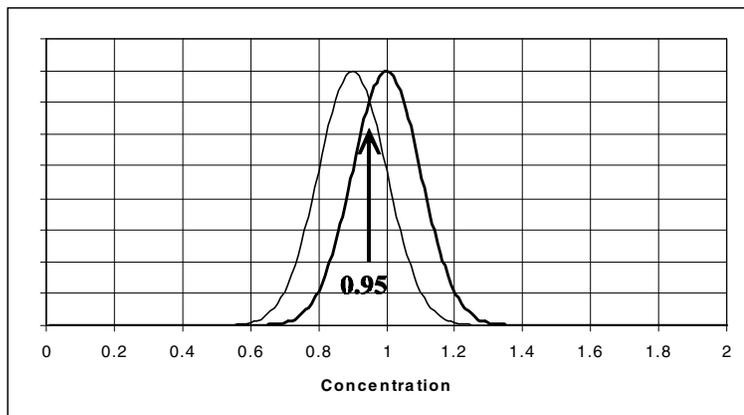
FIGURE B-1.3

480 The problem is, when there is only  
481 the measurement result to go by, the  
482 project planning team cannot tell the

483 difference with confidence. If the measured concentration is 0.9, it is more likely that the true  
484 concentration is 0.9 than it is 1.0, but there remains a chance that it is really 1.0.

### 485 **B-1.4 The Null Hypothesis**

486 If the measured concentration is 0.95,  
487 it is equally likely that the true  
488 concentration is 0.9 as it is 1.0 (see  
489 Figure B-1.4). How does the project  
490 planning team decide what is the true  
491 concentration? The project planning  
492 team starts by asking:



493 “Which mistake is worse: (1) saying  
494 the true concentration is 0.9 when it  
495 is 1.0 or more? or (2) saying the true  
496 concentration is 1.0 when it is 0.9 or  
497 less?”

**FIGURE B-1.4**

498 What does the project planning team mean by “worse”? The project planning team really does  
499 not want to make a mistake that is likely to remain undiscovered or will be difficult or expensive  
500 to correct.

#### 501 **Case 1: Assume The True Concentration is Over 1.0**

502 If a true concentration of 1.0 or more is over a regulatory limit, the project planning team will not  
503 want to make mistake (1) above. If the project planning team decides the true concentration is  
504 less than 1.0, the project planning team is not likely to look at the sample again. That would  
505 mean that the mistake would probably not be discovered until much later, if at all. On the other  
506 hand, if the project planning team decides that the true concentration is over 1.0 when it really is  
507 not, the project planning team will discover the mistake while they are trying to figure out how to  
508 “correct” the high reading. So the project planning team will make a rule: Assume the true  
509 concentration is over 1.0 unless they are really sure it is under. This is the default assumption, the  
510 “null hypothesis.”

511 How sure does the project planning team need to be? For this example, we will assume that the  
512 project planning team would like to be 95 percent sure. To be 95 percent sure, they would have to

513 stay with their assumption that the  
 514 true concentration is over 1.0 unless  
 515 the measured concentration is 0.84  
 516 or less (Figure B-1.5). The project  
 517 planning team knows that this will  
 518 only happen about 5 percent of the time  
 519 when the true concentration is  
 520 really 1.0. That is, the measurement  
 521 has to be less than 0.84 to be 95  
 522 percent sure the true concentration  
 523 is less than 1.0.

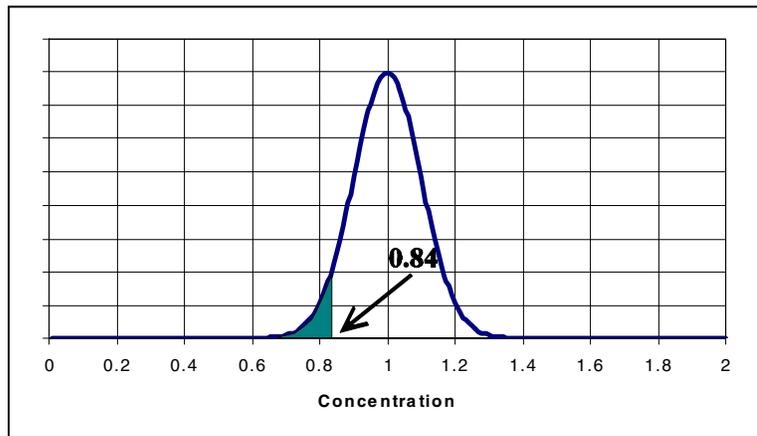


FIGURE B-1.5

524 But what if the true concentration is  
 525 0.9 or less—mistake (2) above?

526 Under the new rule (default assumption or null hypothesis), how often will the project planning  
 527 team say that the true concentration is over 1.0 when it is really only 0.84? As seen in Figure B-  
 528 1.6, there is only a 50-50 chance of making the right decision when the true concentration really  
 529 is 0.84. That is the price of being sure they are not over the action level.

530 How low does the true concentration  
 531 have to be in order to have a pretty  
 532 good chance of deciding that the  
 533 true concentration is below the  
 534 limit? To be 95 percent sure, the  
 535 true concentration needs to be twice  
 536 as far below the action level as the  
 537 decision point, namely at about 0.68.  
 538 That is, the project planning team  
 539 will need a concentration of 0.68 or  
 540 less to be 95 percent sure that they  
 541 will be able to decide the true  
 542 concentration is less than 1.0 (see  
 543 the unshaded portion in Figure B-  
 544 1.7). In other words, it is only when the true concentration is 0.68 or less that the project planning  
 545 team can be pretty sure that they will decide the true concentration is less than 1.0. (Note how  
 546 similar this looks to an MDC in reverse.)

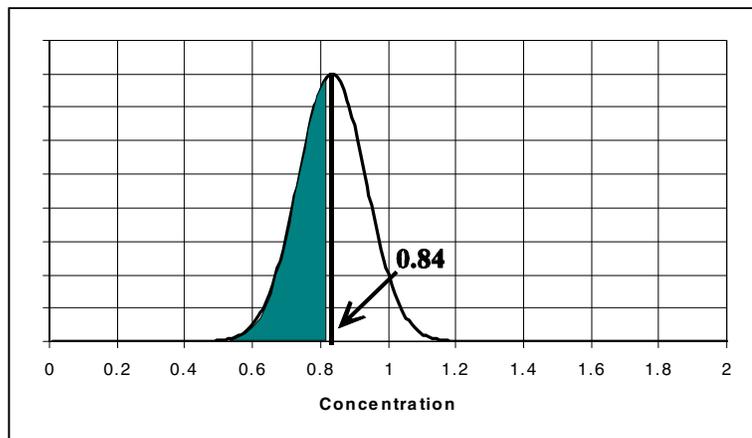


FIGURE B-1.6

547 **Case 2: Assume The True**  
548 **Concentration is 0.9**

549 As stated previously, the mistake  
550 that is most serious determines the  
551 null hypothesis. Suppose that the  
552 project planning team determined  
553 that it is worse to decide that the true  
554 concentration is over 1.0 when it is  
555 0.9 (than it is to decide it is 0.9  
556 when it is 1.0). Then, the default  
557 assumption (the null hypothesis)  
558 would be that the true concentration  
559 is 0.9, unless the measured  
560 concentration is large enough to convince the planning team otherwise. Only when the measured  
561 concentration reaches 1.06 does the planning team decide the true concentration is over 1.0  
562 (Figure B-1.8). The team will have to have a true concentration of 1.22 or more to be 95 percent  
563 sure that they will be able to decide the true concentration is over 1.0.

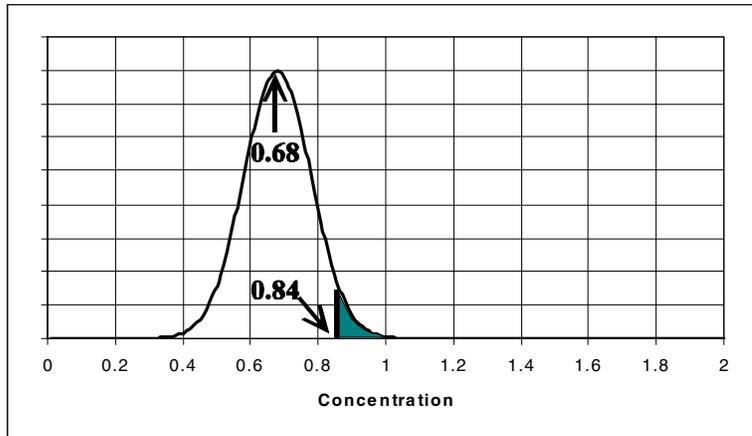


FIGURE B-1.7

564 **B-1.5 The Critical Region**

565 The mistake that is “worse” defines  
566 the null hypothesis and also defines  
567 a “Type I” error. The probability of a  
568 Type I error happening is called the  
569 “Type I error rate,” and is denoted  
570 by alpha ( $\alpha$ ). Under the original null  
571 hypothesis (Case 1: Assume the true  
572 concentration is over 1.0), a Type I  
573 error would be deciding that the  
574 concentration was less than 1.0  
575 when it really was not. In general, a  
576 Type I error is deciding against the null hypothesis when it is actually true. (A Type I error is also  
577 called a “false positive.” This can be confusing when the null hypothesis appears to be a  
578 “positive” statement. Therefore, MARLAP uses the neutral terminology.)

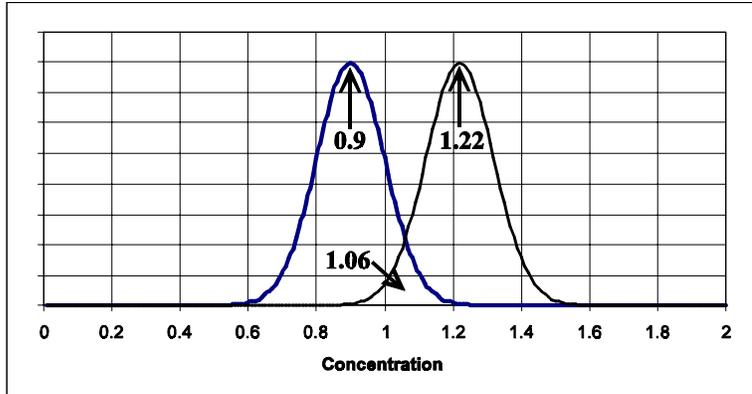


FIGURE B-1.8

579 The “less serious” mistake is called a Type II error, and the probability of it happening is the  
580 “Type II error rate,” denoted by beta ( $\beta$ ). Under the original null hypothesis that the concentration

581 was 1.0 or more, a Type II error would be deciding that the concentration was more than 1.0  
582 when it really was not. In general, a Type II error is not deciding against the null hypothesis when  
583 it is actually false.

584 In both Case 1 and Case 2, the probability of both Type I errors and Type II errors were set to 5  
585 percent. The probabilities were calculated at multiples of the standard deviation, assuming a  
586 normal distribution. This will not always be the case. However, the probability of a Type I error  
587 is always calculated as the probability that the project planning team will decide to reject the null  
588 hypothesis when it is actually true. This is simple enough, as long as there is a clear boundary for  
589 the parameter of interest.

590 The parameter of interest in both Case 1 and Case 2 was the true concentration. The true  
591 concentration had a limit of 1.0. Therefore, all the project planning team had to do was calculate  
592 the probability that they would get a measured concentration that would cause them to decide  
593 that the true concentration was less than 1.0, even though it was equal to 1.0. In the example, the  
594 project planning team actually started with the probability (5 percent) and worked out the critical  
595 value. The “critical value” (or decision point) is the measured value that divides the measurement  
596 results into two different sets: (1) those values that will cause us to reject the null hypothesis and  
597 (2) those values that will cause us to leave the null hypothesis as the default. Set (1) is called the  
598 “critical region.”

599 The Type I and Type II error rates,  $\alpha$  and  $\beta$ , often are both set at 5 percent. This is only by  
600 tradition. They do not have to be equal. Neither error rate needs to be set at 5 percent. The way  
601 the project planning team should set the value is by examining the consequences of making a  
602 Type I or a Type II error. What consequences will happen as a result of making each type of  
603 error? This is a little different than the criterion that was used to define the null hypothesis. It  
604 may be that in some circumstances, a Type II error is riskier than a Type I error. In that case,  
605 consider making  $\alpha$  bigger than  $\beta$

## 606 **B-1.6 The Gray Region**

607 In the previous sections (B-1.1 to B-1.4) the project planning team:

- 608 • Set the region of interest for the measured concentrations between zero and about twice the  
609 action level;
- 610 • Assumed that the true concentration exceeds 1.0, unless they measure “significantly” below  
611 that, the default assumption (null hypothesis);

- 612 • Defined “significantly below” to mean a concentration that would be observed less than 5  
613 percent of the time, when the true concentration is actually 1.0. To describe their uncertainty,  
614 the project planning team used the normal distribution, with a relative standard deviation of  
615 10 percent at the action level, as a model;
- 616 • Developed an operational decision rule: If the measured concentration is less than 0.84, then  
617 decide the true concentration is less than 1.0. Otherwise, decide there is not enough reason to  
618 change the default assumption (null hypothesis); and
- 619 • Found using this operational decision rule that they were pretty sure (95 percent) of deciding  
620 that the true concentration is less than 1.0 only when the true concentration is actually 0.68 or  
621 less.

622 If the true concentration is between 0.68 and 1.0, all the project planning team really can say is  
623 that the probability of deciding that the true concentration is less than 1.0 will be between 5  
624 percent (when the true concentration is 1.0) and 95 percent (when the true concentration is 0.68).  
625 Conversely, when the true concentration is in this range, the probability of deciding that the true  
626 concentration is not less than 1.0 (i.e., the probability of a Type II error) will be between 5  
627 percent (when the true concentration is 0.68) and 95 percent (when the true concentration is just  
628 under 1.0). This range of concentrations is called the “gray region.”

629 When the null hypothesis is that the true concentration exceeds the action level (1.0), the gray  
630 region is bounded from above by the action level. This is where  $\alpha$  is set. It is bounded from  
631 below at the concentration where  $\beta$  is set. There is some flexibility in setting the lower boundary  
632 of the gray region (LBGR). If the project planning team specifies a concentration, they can  
633 calculate the probability  $\beta$ . If they specify  $\beta$ , they can calculate the value of the true concentration  
634 that will be correctly detected as being below 1.0 with probability  $1-\beta$ .

635 In our example, the project planning team found that they needed the true concentration to be  
636 0.68 or less to be at least 95 percent sure that they will correctly decide (by observing a measured  
637 value of 0.84 or less) that the true concentration is less than 1.0. If the project planning team  
638 doesn’t like that, the project planning team can find that a true concentration of 0.71 will be  
639 correctly detected 90 percent of the time (also by observing a measured value of 0.84 or less).  
640 The critical value, or decision point, is determined by  $\alpha$ , not  $\beta$ .

641 If the project planning team decides to raise the LBGR (i.e., narrow the gray region) the Type II  
642 error rate at the LBGR goes up. If they lower the LBGR (i.e., widen the gray region) the Type II

643 error rate at the LBGR goes down. Nothing substantive is really happening. The project planning  
 644 team is merely specifying the ability to detect that the null hypothesis is false.

645 If the project planning team wants to make a substantive change, they need to change the  
 646 probability that an error is made. That is, they need to change the uncertainty (standard deviation)  
 647 of the measurements. Suppose the relative standard deviation of the measurements at the action  
 648 level is 5 percent instead of 10 percent. Then the value of the true concentration that will be  
 649 correctly detected to be below the action level (by observing a measured value of 0.92 or less) 95

650 percent of the time, is 0.84. Cutting  
 651 the standard deviation of the  
 652 measurement in half has cut the  
 653 (absolute) width of the gray region  
 654 in half, but left the width of the gray  
 655 region in standard deviations  
 656 unchanged. Previously, with  $\sigma = 10$   
 657 percent, the width of the gray region  
 658 was  $1.0 - 0.68 = 0.32 = 3.2 (0.10) =$   
 659  $3.2\sigma$ . As Figure B-1.9 illustrates,  
 660 with  $\sigma = 5$  percent, the width of the  
 661 gray region is  $1.0 - 0.84 = 0.16 = 3.2$   
 662  $(0.05) = 3.2\sigma$ .

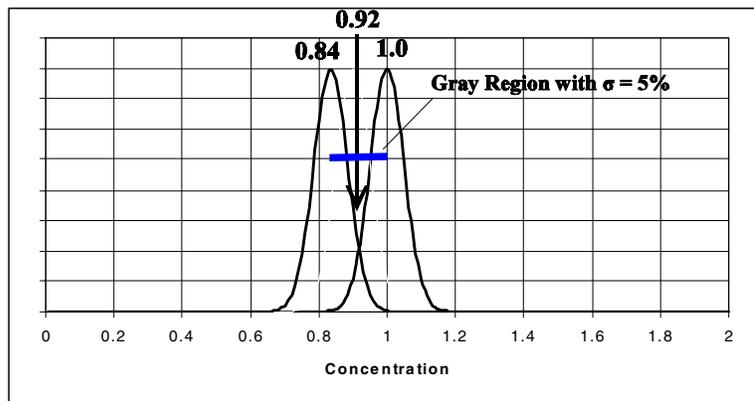


FIGURE B-1.9

663 What is important is the width of the gray region in standard deviations; not the width of the gray  
 664 region in concentration. In order to achieve the same specified Type II error rate at the LBGR, the  
 665 action level and the LBGR must be separated by the same number of standard deviations. The  
 666 width of the gray region (action level minus LBGR) will be denoted by delta ( $\Delta$ ), the “shift.”  $\Delta/\sigma$   
 667 is how many standard deviations wide the gray region is.  $\Delta/\sigma$  is called the “relative shift.”

668 If the gray region is less than one standard deviation wide, the Type II error rate may be high at  
 669 the LBGR. The only way to improve the situation would be to decrease the standard deviation  
 670 (i.e., increase the relative shift,  $\Delta/\sigma$ ). This can be done by employing a more precise measurement  
 671 method or by averaging several measurements. When the width of the gray region is larger than  
 672 about three standard deviations (i.e.,  $\Delta/\sigma$  exceeds 3), it is overkill. It may be possible to use a  
 673 simpler, less expensive measurement method or take fewer samples.