APPENDIX E

THE ASSESSMENT PHASE OF THE DATA LIFE CYCLE

The assessment phase of the Data Life Cycle includes verification and validation of the survey data and assessment of quality of the data. Data verification is used to ensure that the requirements stated in the planning documents are implemented as prescribed. Data validation is used to ensure that the results of the data collection activities support the objectives of the survey as documented in the Quality Assurance Project Plan (QAPP), or permit a determination that these objectives should be modified. Data Quality Assessment (DQA) is the scientific and statistical evaluation of data to determine if the data are of the right type, quality, and quantity to support their intended use (EPA 1996a). DQA helps complete the Data Life Cycle by providing the assessment needed to determine that the planning objectives are achieved. Figure E.1 illustrates where data verification, data validation and DQA fit into the Assessment Phase of the Data Life Cycle.

There are five steps in the DQA Process:

- Review the Data Quality Objectives (DQOs) and Survey Design
- Conduct a Preliminary Data Review
- Select the Statistical Test
- Verify the Assumptions of the Statistical Test
- Draw Conclusions from the Data

These five steps are presented in a linear sequence, but the DQA process is applied in an iterative fashion much like the DQO process. The strength of the DQA process is that it is designed to promote an understanding of how well the data will meet their intended use by progressing in a logical and efficient manner.

E.1 Review DQOs and Survey Design

The DQA process begins by reviewing the key outputs from the Planning phase of the Data Life Cycle that are recorded in the planning documents (*e.g.*, the QAPP). The DQOs provide the context for understanding the purpose of the data collection effort. They also establish qualitative and quantitative criteria for assessing the quality of the data set for the intended use. The survey design (documented in the QAPP) provides important information about how to interpret the data.

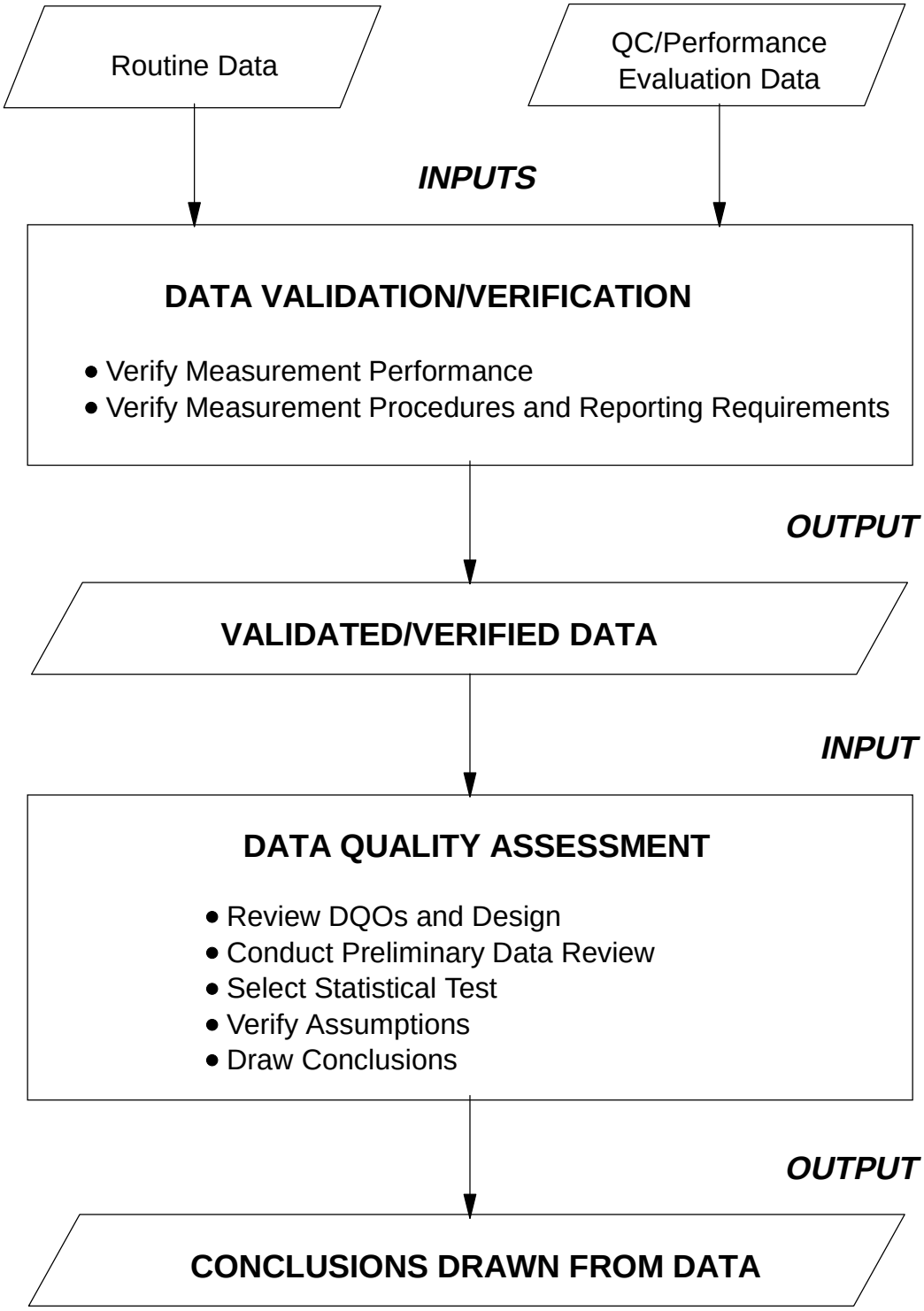

Figure E.1 The Assessment Phase of the Data Life Cycle (EPA 1996a)

There are three activities associated with this step in the DQA process:

- Translating the data user's objectives into a statement of the hypotheses to be tested using environmental data. These objectives should be documented as part of the DQO Process, and this activity is reduced to translating these objectives into the statement of hypotheses. If DQOs have not been developed, which may be the case for historical data, review Appendix D for assistance in developing these objectives.

- Translating the objectives into limits on the probability of committing Type I or Type II decision errors. Appendix D, Section D.6 provides guidance on specifying limits on decision errors as part of the DQO process.

- Reviewing the survey design and noting any special features or potential problems. The goal of this activity is to familiarize the analyst with the main features of the survey design used to generate the environmental data. Review the survey design documentation (*e.g.*, the QAPP) with the data user's objectives in mind. Look for design features that support or contradict these objectives.

For the final status survey, this step would consist of a review of the DQOs developed using Appendix D and the QAPP developed in Chapter 9.

E.2 Conduct a Preliminary Data Review

In this step of the DQA process, the analyst conducts a preliminary evaluation of the data set, calculating some basic statistical quantities and looking at the data through graphical representations. By reviewing the data both numerically and graphically, the analyst can learn the "structure" of the data and thereby identify appropriate approaches and limitations for their use.

This step includes three activities:

- reviewing quality assurance reports
- calculating statistical quantities (*e.g.*, relative standing, central tendency, dispersion, shape, and association)
- graphing the data (*e.g.*, histograms, scatter plots, confidence intervals, ranked data plots, quantile plots, stem-and-leaf diagrams, spatial or temporal plots)

Chapter 8 discusses the application of these activities to a final status survey.

E.3 Select the Statistical Test

The statistical tests presented in Chapter 8 are applicable for most sites contaminated with radioactive material. Chapter 2 discusses the rationale for selecting the statistical methods recommended for the final status survey in more detail. Additional guidance on selecting alternate statistical methods can be found in Section 2.6 and in EPA's DQA guidance document (EPA 1995).

E.4 Verify the Assumptions of the Statistical Test

In this step, the analyst assesses the validity of the statistical test by examining the underlying assumptions in light of the environmental data. The key questions to be resolved are: "Do the data support the underlying assumptions of the test?", and: "Do the data suggest that modifications to the statistical analysis are warranted?"

The underlying assumptions for the statistical tests are discussed in Section 2.5. Graphical representations of the data, such as those described in Section 8.2 and Appendix I, can provide important qualitative information about the validity of the assumptions. Documentation of this step is always important, especially when professional judgement plays a role in accepting the results of the analysis.

There are three activities included in this step:

- Determining the approach for verifying assumptions. For this activity, determine how the assumptions of the hypothesis test will be verified, including assumptions about distributional form, independence, dispersion, type, and quantity of data. Chapter 8 discusses methods for verifying assumptions for the final status survey statistical test during the preliminary data review.

- Performing tests of the assumptions. Perform the calculations selected in the previous activity for the statistical tests. Guidance on performing the tests recommended for the final status survey are included in Chapter 8.

- Determining corrective actions (if any). Sometimes the assumptions underlying the hypothesis test will not be satisfied and some type of corrective action should be performed before proceeding. In some cases, the data for verifying some key assumption may not be available and existing data may not support the assumption. In this situation, it may be necessary to collect new data, transform the data to correct a problem with the distributional assumptions, or select an alternate hypothesis test. Section 9.3 discusses potential corrective actions.

E.5 Draw Conclusions from the Data

The final step of the DQA process is performing the statistical test and drawing conclusions that address the data user's objectives. The procedure for implementing the statistical test is included in Chapter 8.

There are three activities associated with this final step:

- Performing the calculations for the statistical hypothesis test (see Chapter 8).

- Evaluating the statistical test results and drawing the study conclusions. The results of the statistical test will be either accept the null hypothesis, or reject the null hypothesis.

- Evaluating the performance of the survey design if the design is to be used again. If the survey design is to be used again, either in a later phase of the current study or in a similar study, the analyst will be interested in evaluating the overall performance of the design. To evaluate the survey design, the analyst performs a statistical power analysis that describes the estimated power of the test over the full range of possible parameter values. This helps the analyst evaluate the adequacy of the sampling design when the true parameter value lies in the vicinity of the action level (which may not have been the outcome of the current study). It is recommended that a statistician be consulted when evaluating the performance of a survey design for future use.