NCER's Guidance for Quality Assurance Project Plans (QAPPs)

General Guidance for Writing and Reviewing QAPPs
For EPA/NCER and the STAR Grant Program

General Information about QAPPs –
The purpose for writing a QAPP is to ensure that the activities associated with the collection, generation, use, and/or reporting of data will provide information suitable for answering a question or making a decision, such as:

- a hypothesis is correct, or not;
- a method works, or not;
- a new technology is beneficial, and it is "this much" better than the old standard;
- this model can tell us "XYZ", but with these limitations.

Basically, state what you are going to do, how you are going to do it, and how you will know that you have done it "right". In the QAPP, *you* will define what "right" is for *your* research. For example, when you provide specific criteria for some measure of accuracy, then also explain the reasons for needing that level of accuracy (or inaccuracy, as the case may be). The goal is to be specific enough about how you will conduct the research that it could be reproduced by another team of researchers with reasonably similar results. Assume that the reader has already read your proposal, and refer to specific sections of it as necessary.

The QAPP must cover all areas / phases of your research, so ensure that the plan is complete. If necessary, break it into logical sections for different activities. If the details of later phases depend on information gleaned in the early ones, provide the information you have now, and provide your EPA PO with an addendum with the remaining information at a later date. However, all research activities must be covered under a QAPP approved by your EPA PO before you can begin working on those activities. So ensure timely submission of your plans for the next phase. Failure to obtain EPA approval on the requisite documentation in a timely manner may result in a freeze on your funding.

QAPPs must be written in active language indicating exactly what will be done during the course of the project and not in terms of "may", "should", or "could" as might have been done during the proposal stage of the research. Now that the research has been funded, write in concrete terms of what will be done with the money. The QAPP must be significantly more detailed than the Quality Assurance Statement that was submitted with the proposal. There is no page limit; but short, to the point, yet complete is preferred.

This summary guidance provides information on the general content to be provided in QAPPs, but **it is not a requirement to follow the format described below**. It is far more important to cover all of the required details in a format that makes sense for the type of research project in question. The EPA QA/R-5 document (referenced below) provides a generally accepted, formalized structure that many people are already familiar with and may wish to use.

This information is a short summary of EPA's "Guidance for Quality Assurance Project Plans (EPA QA/G-5)" (http://www.epa.gov/quality/qs-docs/g5-final.pdf) and "EPA Requirements for Quality Assurance Project Plans (EPA QA/R-5)" (http://www.epa.gov/quality/qs-docs/r5-final.pdf). If necessary, additional details can be found within these online documents.

The first part of this guide present general guidance that would be applicable to almost any type of research. Then the following section covers additional issues that are specific to certain types of data and/or specific types of research (e.g. modeling, technology development, analytical method development, surveys, etc.). One or more of the issues discussed may apply to any given project, but it is not necessary to read the ones that obviously do not apply.

Parts of a QAPP are:

Background/purpose: Lay out the question to be answered and/or decision to be made.

Experimental Design: Describe what will be done.

Data Gathering Methods: Connect what will be done with how it will be done.

Data Quality: Show how we can be sure it will be done well enough (acceptance criteria).

Data Reduction: Discuss the procedures that will make the data meaningful.

Interaction of the Players: Show the responsibilities of each organization / researcher.

Each of these segments is discussed on the following pages.

NOTE: ***It is not necessary to repeat information that is in the Research Plan (RP) that was submitted with the proposal. Simply make reference to pages in the RP as necessary.***

Background/Purpose - Lay out the question to be answered and/or decision to be made.

- State explicitly the *question to be answered and/or the decision to be made*. (This must be clear if we are to evaluate the usefulness of methodologies.) Indicate the reasons why this research is important.

- Describe *how the research group will obtain the data necessary to provide the information needed* to answer the question or make the decision, so that the Project objectives are connected with the Data objectives. (i.e., how will the methods be used to get you the data you need?) Show what general kind of information is needed (e.g. sampling, monitoring, analysis, compilation of secondary data, etc.)

Design - Describe what will be done.

- Provide a summary of *all* technical work to be performed and/or products to be produced.

- Describe the connections between your stated purpose and what will be done to successfully achieve it. [*AND* define the criteria for "success" for the research].

- Describe the scheme for different types of data collection (or compilation) activities to be used (include conditions and assumptions). Explain the reasons for designing the project in such a way, and **provide explanation for any quality control (QC) criteria you set in this section.**

- Show how this scheme will produce the group of results needed to answer the question, etc. from above. (This is essentially the Data Quality Objectives (DQO) Process, which is a QA system name for a set of steps that mirror the scientific method. The bigger the project, the more important it is to utilize true DQOs in developing the QAPP. For additional information see http://www.epa.gov/quality/qs-docs/g4-final.pdf.)

Data Gathering Methods - Connect what will be done with how it will be done.
(Includes analytical work, secondary data, meta-data, surveys, modeling efforts, etc.)

- Describe what specific types of data need to be gathered, how they will be gathered or produced, and how each one connects with and/or supports the stated purpose and criteria for success of the research.

- Describe the specific methods and equipment to be used. If established methods are to be used cite them and attach copies if they are not readily available. (Use of tables may be helpful.)

- Include operating procedures for major analytical work and/or unique research procedures, and provide a table listing all other repetitive operating procedures that are followed in the process of the research. (It is important to show that the research is repeatable.)

Data Quality - Show how we can be sure it will be done well enough (acceptance criteria).

- **Data quality encompasses both Quality Control (QC) and Quality Assurance (QA) practices.** QC is the system of technical activities used in data gathering activities/methods to ensure the quality of each individual data point (as applicable: precision, accuracy, representativeness, completeness, & comparability); and Quality Assurance (QA) is the overall program that includes planning, QC, documentation, and Assessments or audits.

- List the QC, QA, and Assessment activities that will be part of the project, to show how the required data quality will be achieved. (What will the numbers represent and what will we do with them?)

- Explain the reasons for the research design, and provide explanation for any quality control (QC) criteria you set in this section.

- *For each activity*, state appropriate acceptance limits/criteria. Include all QC parameters for each type of laboratory analysis, secondary data sources, model validation, statistical procedures, etc. Indicate how each will be documented and what corrective actions will be taken if QC criteria are not met. (Again, use of tables may be helpful.)

- *For each QC parameter* (Precision, Accuracy, Representativeness, Completeness, Comparability), state the how conformance with the stated limits will be determined, and *the consequences of failure* --such as for precision, what is the acceptable range, how will it be determined (average of 3 lab duplicates prior to the start, etc.), if fail, recalibrate and repeat until satisfactory, etc. Who makes the decisions if some parameters fail?

- *For each assessment type* (field audit, lab audit, data quality assessment, etc.), state the type and acceptance level, if any, and the consequences of failure.

- Of all of the data quality parameters (Precision, Accuracy, Representativeness, Completeness, Comparability), *the most important is Representativeness*. Describe the representativeness of the sample(s) in the plan. How is the statistical network designed? (If the research does not use representative data, it doesn't matter how good it is, or how good the use of it is. This is the main connection between the science and the politics: making sure that the right samples will be collected to validly address the issue at hand.)

Description of Data Reduction Methods and Procedures - Discuss the procedures that will make the data meaningful.

- Describe all data reduction methods and procedures.

- Identify the specific descriptive statistical methods (for example, regression analyses, analysis of variance, or multivariate analyses) that will be used to present results. Discuss how raw and processed data that are used in statistical analyses will be verified after statistical analyses have been completed. Also include a complete citation of software programs that will be used for these statistical analyses and to present results.

- When using secondary data, include a discussion of the quality of these data and how the data will be transferred into computer files for various analyses and how they are verified through these processes. Discuss how the quality of these secondary data affect the results being reported. Indicate that a complete citation of these data sources will be complied during data collection so that these data can be reviewed later if necessary.

- Discuss how original or raw data measurements will be verified after they have been transferred from instrument data recording devices and/or floppy diskettes or laboratory notebooks and processed by computer or manually.

Interaction of the Players - Show the responsibilities of each organization / researcher.

- Need clear *organizational charts* for the project, showing who is responsible for what, and who reports to whom. This must include *every organization* involved in the project: management, project management including co-PIs, field groups, lab workers, QA, contractors, consultants. Indicate how each party will interact to achieve the end result.

- Plan must include statements of who has the authority to change it or the operating procedures.

- It should be *signed* by a representative of each organization *to confirm their involvement*, understanding, and acceptance of the stated roles, responsibilities, and authorities. (A common flaw is that one group writes and signs the Plan and sends it to EPA without the concurrence of all parties.)

Special Topics:

GIS/Remote Sensing data: Discuss how the following elements will be addressed: positional accuracy; attribute accuracy; logical consistency; time; lineage; resolution accuracy; and the completeness of coverage, classification, and verification.

Conducting surveys: Discuss the justification for the size of the proposed sample for both the overall project and all subsamples for specific treatments or tests. Identify and explain the rational for the proposed statistical techniques (e.g., evaluation of statistical power).

Development or refinement of models:

(i) Discuss the scope and purpose of the model, key assumptions to be made during development/refinement, requirements for code development, and how the model will be documented.
(ii) Discuss verification techniques to ensure the source code implements the model correctly.
(iii) Discuss validation techniques to determine that the model (assumptions and algorithms) captures the essential phenomena with adequate fidelity.
(iv) Discuss plans for long-term maintenance of the model and associated data.

Development or operation of environmental technology:

(i) Describe the overall purpose and anticipated impact of the technology.
(ii) Describe the technical and quality specifications of each technology component or process that is to be designed, fabricated, constructed, and/or operated.
(iii) Discuss the procedure to be used for documenting and controlling design changes.
(iv) Discuss the procedure to be used for documenting the acceptability of processes and components, and discuss how the technology will be benchmarked and its effectiveness determined.
(v) Discuss the documentation requirements for operating instructions/guides for maintenance and use of the system(s) and/or process(s).

Secondary Data Collection:

(Secondary data is data that will be used for purposes other than those for which they were originally collected. They may be obtained from many sources, including literature, industry surveys, compilations from computerized databases and information systems, and computerized or mathematical models of environmental processes. This includes purchased data sets, too.)

(i) Identify the types of secondary data needed to satisfy the project objectives. Specify requirements relating to the type of data, the age of data, geographical representation, temporal representation, and technological representation, as applicable.
(ii) Specify the source(s) of the secondary data discuss the rational for selection.
(iii) Establish a plan to identify the sources of the secondary data in all deliverables/products.
(iv) Specify quality requirements and discuss the appropriateness for their intended use. Accuracy, precision, representativeness, completeness, and comparability need to be addressed, if applicable.
(v) Describe the procedures for determining the quality of the secondary data.
(vi) If no quality documentation exists or if the quality of the secondary data will not be evaluated under the grant, the QAPP shall require that a disclaimer be added to all deliverables/products (including models and other decision support tools). The wording for the disclaimers shall be defined. (Use of disclaimers must receive prior approval from the EPA Project Officer.)
(vii) Describe all data reduction procedures including calculations, equations, scripts, and statistical analysis. At some point, relate these back to satisfying the project objectives.
(viii) Include a plan for data management / integrity.