

Revised Draft - What do we need a procedure to do? 4/26/06

Introduction

At its December 8-9, 2005 meeting, the Federal Advisory Committee on Detection and Quantitation Approaches and Uses in Clean Water Act Programs (FACDQ, committee) agreed by consensus that its recommendations concerning analytical procedures for detection and quantitation in Clean Water Act programs should be based on what members of the advisory committee need procedures to do.

Members of the committee discussed what they needed procedures to do in the ensuing months in Policy Work Group meetings, with additional input from the Technical Work Group. At its March 29-30, 2006 FACDQ, the committee reviewed a draft document, "What do we need a procedure to do," which identified 13 objectives. After discussion, the committee agreed to the 13 objectives and added a 14th. The committee also agreed that the objectives would apply to long-term committee recommendations, but that the setting of any numeric objectives (i.e. false positive, false negative, precision and accuracy) would apply only to the pilot study.

Individual caucuses then reviewed the draft document, including how each objective would be evaluated, and provided comments. The facilitators consolidated the comments into a revised document. The committee created a subgroup, consisting of Bob Avery, Richard Burrows, Michael Murray, John Phillips and Jim Pletl, and asked it to consider the caucus comments and to refine the 14 objectives and ways to measure them as input into the pilot study design.

The subgroup held a two-hour call on Monday, April 24, to review the objectives and to revise the document in light of the comments. In carrying out this assignment, the subgroup noted the following:

- The objectives defined in the document are intended to be used to evaluate procedures tested in the pilot study. The committee does not expect that procedures will meet all of these objectives. After receiving the pilot study results, the FACDQ may decide to revisit the objectives or it may seek to revise the procedures so they better meet the objectives.
- The committee acknowledged that cost and contracting restraints are factors that will affect the pilot study. To the maximum extent possible, the pilot will be conducted using a wide range of labs and methods.
- The committee agreed to specific measurement quality objectives (MQOs) for false positives, false negatives, and precision to be used in the pilot study. For accuracy (bias), the committee assigned the Technical Work Group and Pilot Design Team to establish values based on the specific analytical methods accuracy levels and existing data. The detailed recommendations are included at the end of this document and need to be considered in evaluating the procedures.

The remainder of this document identifies the 14 objectives for testing procedures and suggests how each objective can be evaluated as part of the pilot test. The term “limit” is used generally to refer to detection and quantitation limits since the FACDQ has not yet defined them. Examples of how to measure specific objectives are sometimes written broadly and may not apply in every case (L_C , L_D , L_Q , other).

The procedure(s) will:

1. **provide an explicit estimate of bias at L_Q for limits that must be verifiable by labs at those limits.**

To be evaluated by:

- a. reviewing procedure(s) and specifically identifying the quantitative limit for bias at L_Q that is tested in the pilot study.
- b. requiring labs to analyze samples (spikes, blind or otherwise as appropriate) and comparing observed bias to that cited by the procedure(s).

See Appendix for specific MQOs adopted by the committee for the pilot study

2. **provide an explicit estimate of precision at L_Q for limits that must be verifiable by labs at those limits.**

To be evaluated by:

- a. reviewing procedure(s) and specifically identifying the quantitative limit for precision at L_Q that is tested in the pilot study
- b. requiring labs to analyze samples (spikes, blind or otherwise as appropriate) and comparing observed precision to that cited by the procedure(s).

See Appendix for specific MQOs adopted by the committee for the pilot study

3. **provide an explicit false positive rate for L_C .**

To be evaluated by:

- a. reviewing procedure(s) and specifically identifying the false positive error rate predicted for each limit that is tested in the pilot study.
- b. comparing the false positive rate of lab blanks at the estimated levels of L_C to those predicted by the procedure(s).

Note: The intent is to look at long term performance, however for the pilot study the number of samples may be relatively small.

See Appendix for specific MQOs adopted by the committee for the pilot study

4. **provide an explicit false negative rate at L_C for the true value at L_D or L_Q that must be observed in labs at L_C for the estimated values of L_D or L_Q .**

To be evaluated by:

- a. reviewing procedure(s) and specifically identifying the false negative error rate predicted for L_D/L_Q that is tested in the pilot study.
- b. comparing the false negative rate of results obtained by analyzing samples spiked at the L_D/L_Q concentration to those predicted by the procedure(s).

Note: The intent is to look at long term performance, however for the pilot study the number of samples may be relatively small.

See Appendix for specific MQOs adopted by the committee for the pilot study

5. **provide that qualitative identification criteria defined in the analytical method are met at the determined detection and quantitation limits.**

To be evaluated by:

- a. Requiring that all method qualitative identification criteria be satisfied in order for detection to occur.
- b. Requiring modification of L_Q or L_D if all spikes at L_Q or L_D are not detected.

6. **adequately represent routine variability in lab performance.**

To be evaluated by determining whether the procedures:

- a. use data to calculate limits that are collected over enough time to capture variability in performance relative to MQOs.
- b. recalculate limits at a frequency that captures variability in performance relative to MQOs.
- c. incorporate variability due to the use of multiple instruments per lab.
- d. incorporate variability due to use of multiple analysts per lab.
- e. incorporate variability occurring across laboratories (not for single lab. procedure).
- f. adjust or account for recovery.
- g. provide recommendations or limit choices for outlier tests.
- h. address varying numbers of different concentrations (spikes) that can be used between laboratories (may only apply to multi/inter lab procedures).
- i. address varying numbers of replicates per concentration (spike) that can be used between laboratories (may only apply to multi/inter lab procedures).
- j. address varying combinations of concentrations (spikes) that can be used between laboratories (may only apply to multi/inter lab procedures).

- k. adequately accommodate different models of instruments used per analyte and technology to calculate limits.
7. **be capable of calculating limits using matrices other than lab reagent grade water.**
To be evaluated by:
- a. reviewing procedures and determining that there is nothing precluding the use of matrices other than reagent grade water to calculate limits.
 - b. reviewing procedures to determine if they incorporate steps to verify when limits adopted for an analytical method can or cannot be met in a matrix other than lab reagent grade water.
 - c. reviewing procedures to determine if they provide instructions on preparing an analyte-free matrix that approximates the matrix in question.
8. **use only data that results from test methods conducted in their entirety.**
To be evaluated by determining whether the procedure(s):
- a. require that samples used to calculate detection and quantitation limits undergo all routine steps outlined in an analytical method as specified in the laboratory's SOP (prep method, extraction, etc.).
 - b. reviewing procedures to determine if they incorporate steps to verify when limits adopted for an analytical method can or cannot be met when a sequence of non-routine steps are used.
9. **explicitly adjust or account for situations where method blanks always return a non-zero result/response.**
To be evaluated by:
- a. reviewing the procedure(s) and determining if they include a process to address occasions when method blanks always return a non-zero result.
 - b. reviewing the procedure(s) and determining if they require calculation of statistics regarding non-zero results/responses.
 - c. reviewing the procedure(s) and determining if they mathematically adjust limits for non-zero results/responses.
10. **explicitly adjust or account for situations where method blanks are intermittently contaminated.**
To be evaluated by:

- a. reviewing the procedure(s) and determining if they define intermittent contamination and provide explicit instructions to deal with this situation.
- b. reviewing the procedure(s) and determining if they mathematically adjust limits for non-zero results/responses.

11. be clearly written with enough detail so that most users can understand and implement them.

To be evaluated by:

- a. asking users to interpret data prior to the after-procedure calculations are carried out. Examples include: What is the resulting detection limit? What is the resulting quantitation limit? and What is the blank bias?
- b. asking users questions about the procedure characteristics and using the matrix as a point of reference. Examples include: Do the procedures address recovery? How often is a limit calculated by the user? and How often is data generated to calculate limits for a given procedure?
- c. asking users to perform calculations or run software and interpret results.
- d. asking users to select spikes for given circumstances.
- e. reviewing the procedure(s) and determining which ones minimize the amount of data required to calculate analytical limits beyond that normally generated by analytical methods.
- f. determining that the procedure(s) do not require skills of users in addition to those that are normally required by laboratories.

12. be cost effective.

To be evaluated by:

- a. reviewing the procedure(s) and determining which ones minimize the amount of data required to calculate analytical limits beyond that normally generated by analytical methods.
- b. determining whether the procedure(s) require the purchase of software or equipment in addition to that which is normally required by laboratories.
- c. determining that the procedure(s) do not require skills of users in addition to those that are normally required by laboratories.

13. assess multi- and inter-laboratory variability when data from more than one lab is used.

To be evaluated by:

- a. comparing results from multi-, inter-, and single lab studies.

- b. Calculating intra-lab, inter-lab and pooled or multi-lab variability and the associated variance error components.

14. be applicable to all users and test methods.

To be evaluated by:

- a. testing procedures against objectives 1-13 among a representative sample of labs (states, EPA, commercial, municipal, small, medium and large, etc.).
- b. testing procedures against objectives 1-13 among a representative sample of analytical test methods (different technologies and analytes).

APPENDIX

Excerpt from Decisions Made at the March 29-30, 2006 FACDQ Meeting Relative to Measurement Quality Objectives (MQOs)

Measurement Quality Objectives

Alpha (False Positive)

For purposes of pilot testing, the committee agreed, by consensus, to set the false positive rate equal to or less than 1%.

Beta (False Negative)

A. Reporting

The committee agreed, by consensus, that if or when data are reported below L_Q , then the data points that fall between L_C and L_Q would be reported, for example, as detected but not quantified (e.g., DNQ). For purposes of pilot testing, numerical data could be used in the calculations. Associated with that value would be a lower bound of L_C and an upper bound of L_Q with some probability. A number with a flag would not be reported.

B. Consideration of L_D

The committee agreed, by consensus, that determination of L_D was not a requirement for purposes of pilot testing, so long as data between L_C and L_Q is reported as detected but not quantified.

C. False negative rate

For purposes of pilot testing, the committee agreed, by consensus, to set the false negative rate equal to or less than 1% measured at L_C for the true value at L_Q or L_D .

Precision

The committee agreed, by consensus, that the goal for the pilot test of 20% relative standard deviation (RSD) is based on the mean recovery, understanding that there will be instances where this %RSD may show conflicts with accuracy (that is, set precision targets may inherently define accuracy targets). This may not be applied universally after the pilot study is complete. The study design team will consider higher precision targets (higher %RSD) if the goal cannot be met.

Accuracy

The committee agreed, by consensus, that, for the pilot, the study design team will ask participating laboratories to use accuracy based on mean accuracy and that the Technical Work Group study design team should make decisions on specific goals for accuracy based on an evaluation of existing data. The study design team will ensure that the batch-by-batch data is available for the FACDQ to have analyzed.