



Tennessee Valley Authority, Post Office Box 2000, Soddy-Daisy, Tennessee 37384-2000

March 18, 2004

State of Tennessee
Department of Environment and Conservation
Division of Water Pollution Control
Enforcement & Compliance Section
6th Floor, L & C Annex
401 Church Street
Nashville, Tennessee 37243-1534

Gentlemen:

SEQUOYAH NUCLEAR PLANT – TRC STUDIES, EPA INFORMATION, BMP
PLAN, BIODETERGENT 73551

Please find enclosed:

Sequoyah and Watts Bar Total Residual Chlorine (TRC) Method Detection Limit (MDL) studies

In a joint letter between Sequoyah (SQN) and Watts Bar (WBN) dated September 15, 2003, SQN and WBN decided to report a Minimum Level of Quantification (ML) of 0.08 mg/L analyzing grab samples at the discharge outfall using a colorimeter. If at any time, both the sample analysis and the established permit limit is less than the ML of 0.08 mg/L, the reporting facility (SQN or WBN) will report < 0.08 mg/L on the DMRs. For purposes of evaluating compliance and calculating averages, these values will be considered equal to 0 mg/L. Enclosed in the BMP plan is when SQN proposes to take grab samples versus using the calculation. The SQN and WBN TRC MDL studies are provided per your request.

EPA's lawsuit information

Reference material provided per your request.

Sequoyah's Best Management Practices (BMP) Plan and Corrosion Control Tables

Sequoyah revised the BMP Plan to incorporate changes in the corrosion control program (oxidizing and non-oxidizing biocide treatment), to introduce Biodetergent 73551 use in June 2004, addition of liquid radwaste release points, and various other changes. Enclosed are the revised BMP Plan and Corrosion Control Tables.

JE25

State of Tennessee
Page 2
March 18, 2004

Sequoyah's chemical change information on Biodetergent 73551

Sequoyah currently uses CL-363 dimethylamide (DMAD). DMAD penetrates and softens deposits and biofilm. It is injected before Towerbrom-960 treatments to enhance effectiveness of chlorination. CL-363 and PCL-222/401 help avoid buildup of hard manganese-iron deposits. CL-363 is injected into both the Essential Raw Cooling Water (ERCW) Train A & B and Raw Cooling Water (RCW) systems two to three times per week for 30 minutes in duration at a target concentration of 0.5 ppm. CL-363 is to be phased out of production by June 2004 and replaced by Biodetergent 73551. Biodetergent 73551 is used to remove and disperse "soft foulant" (mud, silt and clay) deposits in cooling water systems. It is injected before Towerbrom-960 treatments to enhance effectiveness of chlorination. 73551 will be injected into the ERCW Train A & B and RCW systems two to three times per week for 30 minutes in duration at a target concentration of 2.0 ppm. The limit at the Diffuser (Outfall 101) discharge will be 0.100 mg/L daily max. Sequoyah will begin injecting 73551 in June 2004. Enclosed are the calculated 73551 injection rate record sheet and Material Safety Data Sheet (MSDS).

Please contact me at (423) 843-6700 if you have any questions or comments.

I certify under penalty of law that this document and all attachments were prepared under my direction or supervision in accordance with a system designed to assure that qualified personnel properly gather and evaluate the information submitted. Based on my inquiry of the person or persons who manage the system, or those persons directly responsible for gathering the information, the information submitted is, to the best of my knowledge and belief, true, accurate, and complete. I am aware that there are significant penalties for submitting false information, including the possibility of fine and imprisonment for knowing violations.

Sincerely,



Stephanie A. Howard
Principal Environmental Engineer
Signatory Authority for
Richard T. Purcell
Site Vice President
Sequoyah Nuclear Plant

Enclosures
cc: See page 3

State of Tennessee
Page 3
March 18, 2004

cc (Enclosures):

Chattanooga Environmental Assistance Center
Division of Water Pollution Control
State Office Building, Suite 550
540 McCallie Avenue
Chattanooga, Tennessee 37402-2013

U.S. Nuclear Regulatory Commission
ATTN: Document Control Desk
Washington, D.C. 20555

SQN TRC MDL Study

All analyses involved in this study used a HACH DL 850 Colorimeter. All standards and river matrix spikes were analyzed using manganese interference detection and removal. Residual chlorine standards were obtained from a RTC standard with a TRO concentration of 1040 (after 1000:1 dilution, 1.04 ppm). A working standard was prepared in a DI water matrix to a concentration of 100 ppm. From this 100 ppm standard all additional standards and spikes were made. Using the Hach DR 850 colorimeter procedures manual (1st edition revision 3 4/99 page 44) the Method Detection Limit (MDL) was performed for the DR 850 colorimeter. A Method Detection Limit (MDL) determination was performed for river matrix only. A tabulation of the data determination is presented in Table 5. Samples used to determine the estimated detection limit of 0.016 ppm was prepared using river water matrix that had been adjusted (approximately 0.75 ppm) to account for the chlorine demand and were analyzed 9 times using the HACH DR 850 colorimeter. The student t-test used the first 7 values obtained of the analysis.

A 0.1 ppm chlorine standard was prepared in DI water matrix and analyzed to determine if previous dilutions were accurate. Table 1.

Table 1
DI Water spike to 0.100 ppm TRC

Theoretical	Observed ppm	Absorbance	Deviation
0.100	0.11	0.063	0.010
0.100	0.10	0.061	0.000
0.100	0.11	0.066	0.010
0.100	0.10	0.061	0.000
0.100	0.11	0.062	0.010
0.100	0.10	0.061	0.000
0.100	0.11	0.064	0.010
0.100	0.11	0.063	0.010
0.100	0.11	0.062	0.010
Averages	0.107	0.063	0.007

A river sample was spiked with standard to obtain the theoretical value of 0.100 ppm chlorine. As expected the analyses of the river matrix showed a demand of the matrix for chlorine. Table 2.

Table 2
River Matrix spike to 0.100 ppm TRC

Theoretical	Observed ppm	Absorbance	Deviation
0.100	0.040	0.025	-0.06
0.100	0.040	0.024	-0.06
0.100	0.050	0.027	-0.05
0.100	0.050	0.028	-0.05
0.100	0.040	0.023	-0.06
0.100	0.040	0.023	-0.06
0.100	0.050	0.027	-0.05
0.100	0.040	0.023	-0.06
0.100	0.050	0.026	-0.05
Averages	0.04444	0.02511	- 0.05556
Standard Deviation	0.00527		

A 0.075 ppm chlorine standard was prepared in DI water matrix and analyzed to determine if dilutions were accurate. Table 3.

Table 3
DI Water Spike to 0.75 ppm TRC

Theoretical	Observed ppm	Absorbance	Deviation
0.075	0.08	0.044	0.005
0.075	0.07	0.041	-0.005
0.075	0.07	0.040	-0.005
0.075	0.07	0.041	-0.005
0.075	0.08	0.043	0.005
0.075	0.08	0.042	0.005
0.075	0.08	0.044	0.005
0.075	0.07	0.040	-0.005
0.075	0.07	0.041	-0.005
Averages	0.07444	0.04178	-0.00056
Standard Deviation	0.0053		

A river sample was spiked with sufficient standard to obtain a 0.020 ppm above the demand of the river matrix. Since the demand of the river matrix was determined to be 0.055ppm, the theoretical spike value was 0.075 ppm. When analyzed the expected results would be approximately 0.020ppm.

Table 4
River Spike for 0.02 ppm TRC residual.

Theoretical	Observed ppm	Absorbance	Deviation
0.075	0.02	0.013	-0.055
0.075	0.03	0.020	-0.045
0.075	0.02	0.013	-0.055
0.075	0.03	0.016	-0.045
0.075	0.03	0.015	-0.045
0.075	0.03	0.018	-0.045
0.075	0.02	0.014	-0.055
0.075	0.03	0.015	-0.045
0.075	0.03	0.014	-0.045
Averages	0.02666	0.01533	- 0.04833
Standard Deviation	0.0050		

TABLE 5
t Test for Hach DR 850 in river water (adjusted to compensate for chlorine demand)

Actual Chlorine Concentration (ppm)	Hach DR 850 Chlorine Concentration (ppm)	Hach DR 850 Absorbance reading
0.075	0.02	0.013
0.075	0.03	0.020
0.075	0.02	0.013
0.075	0.03	0.016
0.075	0.03	0.015
0.075	0.03	0.018
0.075	0.02	0.014
Standard Deviation	0.0050	

The standard deviation for the DR 850 data in Table 5 is 5.00E-3. Multiplying the standard deviation by the student t value of 3.143 the Method Detection Limit (MDL) is calculated to be 0.016 ppm.

The Limit of Quantification (LOQ) is calculated by multiplying 10 by the Standard Deviation 5.00E-3 giving a (LOQ) value of 0.05 ppm.

WBN TRC MDL Study

Residual chlorine standards were obtained from Environmental Resource Associates with a TRO concentration of 1860 (after 1000:1 dilution, 1.86 ppm). Standards were prepared in a DI water matrix and a Tennessee River matrix. Standards were prepared that ranged from 0.02 ppm to 2.20 ppm. Both colorimeter methods could detect 0.02 ppm Total Residual Chlorine in both DI water and the Tennessee River matrix; however, the Tennessee River water matrix presented a chlorine demand that ranged from 0.03-0.09 ppm as is seen in Table 1. After this demand is satisfied in a sample of river water, a 0.02 ppm spike can be readily detected. It must be remembered that each time a chlorine sample is analyzed, a blank using the same water will zero out any interference from manganese, turbidity, and color.¹ Table 2 presents data from samples prepared in demineralized water which does not exhibit any demand for chlorine.

TABLE 1¹
DPD Chlorine standards in Tennessee River Matrix

Actual Chlorine Conc. (ppm)	Hach DR 2000 Chlorine Conc. (ppm)	Hach DR 2000 Abs reading	Apparent chlorine demand Conc. using river water & 2000 (ppm)	Hach DR 850 Chlorine Conc. (ppm)	Hach DR 850 Abs. reading	Apparent chlorine demand Conc. using river water & 850 (ppm)
0.044	0.00	0.001	0.044	0.00	-0.008	0.044
0.066	0.02	0.006	0.046	0.02	0.011	0.046
0.088	0.06	0.031	0.028	0.06	0.034	0.028
0.264	0.19	0.098	0.074	0.17	0.100	0.094
0.56	0.53	0.223	0.03	0.50	0.294	0.06

TABLE 2
DPD Chlorine standards in DI Water Matrix

Actual Chlorine Concentration (ppm)	Hach DR 2000 Chlorine Concentration (ppm)	Hach DR 2000 Absorbance reading	Hach DR 850 Chlorine Concentration (ppm)	Hach DR 850 Absorbance reading
0.022	0.02	0.008	0.02	0.009
0.044	0.04	0.026	0.04	0.022
0.088	0.08	0.044	0.08	0.048
0.132	0.14	0.075	0.13	0.076
0.264	0.24	0.150	0.24	0.138
0.56	0.56	0.291	0.57	0.329
1.55	1.55	0.786	1.55	0.900
2.20	2.20	1.116	2.17	1.264

¹ A comparison of DI water with chemical added and used as a blank and then treating river water with the chemicals and running the river water as a sample gave an absorbance reading of 0.02 (corresponding to a 0.04 ppm chlorine concentration) using the Hach DR 2000 colorimeter and a 0.023 absorbance reading of 0.02 (which also read 0.04 ppm chlorine concentration) using the Hach DR 850 colorimeter.

Using the Hach DR 850 colorimeter procedures manual (1st edition revision 3 4/99 page 44) the Method Detection Limit (MDL) was performed for both the DR 2000 and DR 850. A tabulation of the data determination is presented in Table 4 for the DR 2000 in river water and the data obtained for the DR 850 is presented in Table 3. A standard that is 5 times the estimated detection limit of 0.02 ppm (approximately 0.1 ppm) was prepared using river water matrix that had been adjusted to remove chlorine demand ² and analyzed seven times using the Hach DR 850 and the DR 2000 colorimeters.

TABLE 3

t Test for Hach DR 850 in river water (adjusted to compensate for chlorine demand)²

Actual Chlorine Concentration (ppm)	Hach DR 850 Chlorine Concentration (ppm)	Hach DR 850 Absorbance reading
0.09	0.08	0.048
0.09	0.08	0.048
0.09	0.09	0.051
0.09	0.09	0.051
0.09	0.08	0.048
0.09	0.10	0.058
0.09	0.09	0.053

TABLE 4

t Test for Hach DR 2000 in river water (adjusted to compensate for chlorine demand)²

Actual Chlorine Concentration (ppm)	Hach DR 850 Chlorine Concentration (ppm)	Hach DR 850 Absorbance reading
0.09	0.10	0.049
0.09	0.08	0.044
0.09	0.10	0.049
0.09	0.10	0.052
0.09	0.09	0.045
0.09	0.10	0.053
0.09	0.09	0.045

The standard deviation for the DR 850 data in Table 3 is 7.559E-3 and the standard deviation for the DR 2000 data in Table 4 is 7.868E-3. If you take the Student t value of 3.143 given for 7 test portions and multiply by the standard deviation, the DR 850 gives a method detection limit of 0.024 ppm and the DR 2000 method detection limit is 0.025 ppm. It must be remembered that river water exhibits a chlorine demand that ranges from 0.03-0.09 ppm and this had to be adjusted prior to making the chlorine standard spike. ²

² 4 ml of a 1.86 ppm standard had to be added to 200 ml of river water to satisfy the chlorine demand (which calculated was 0.037 ppm) prior to the addition of a 10 ml spike of 1.86 ppm to 200 ppm to give a 0.09 ppm standard for analysis.

Standards were prepared in demineralized water by taking 10 ml of 1.86 ppm chlorine standard and diluting to 200 ml to prepare a 0.09 ppm standard. The results of 7 replicate analyses on the DR 850 and DR 2000 are presented in Table 5 and 6 respectively

TABLE 5
t Test for Hach DR 850 in demineralized water

Actual Chlorine Concentration (ppm)	Hach DR 850 Chlorine Concentration (ppm)	Hach DR 850 Absorbance reading
0.09	0.09	0.043
0.09	0.08	0.037
0.09	0.09	0.043
0.09	0.09	0.043
0.09	0.09	0.043
0.09	0.08	0.038
0.09	0.09	0.043

TABLE 6
t Test for Hach DR 2000 in demineralized water

Actual Chlorine Concentration (ppm)	Hach DR 2000 Chlorine Concentration (ppm)	Hach DR 2000 Absorbance reading
0.09	0.09	0.046
0.09	0.08	0.040
0.09	0.08	0.040
0.09	0.08	0.042
0.09	0.09	0.046
0.09	0.09	0.046
0.09	0.09	0.046

The standard deviation of the data in Table 5 for the DR 850 in demineralized water is 4.880E-3 and the standard deviation for the data in Table 6 for the DR 2000 is 5.345E-3. If we multiply each standard deviation by the student t value of 3.143 for seven test runs, we obtain a method detection limit of 0.015 ppm for the DR 850 and a method detection limit of 0.017 for the DR 2000 in demineralized



**Technical Support Document
for the Assessment of
Detection and Quantitation
Approaches**

February 2003

U.S. Environmental Protection Agency
Office of Water (4303T)
1200 Pennsylvania Avenue, NW
Washington, DC 20460

EPA-821-R-03-005

Technical Support Document
for the Assessment of
Detection and Quantitation Approaches

Engineering and Analysis Division
Office of Science and Technology
U.S. Environmental Protection Agency
Washington, DC 20460

February 2003

Acknowledgments and Disclaimer

This document was prepared by Maria Gomez-Taylor, Henry D. Kahn, William A. Telliard, Khouane Ditthavong, and Leonid Kopylev, of the Engineering and Analysis Division in EPA's Office of Water. Harry McCarty, Lynn Riddick, Ken Miller, and Joan Cuddeback, with DynCorp I&ET, and Dale Rushneck with Interface, Inc. provided assistance under EPA Contract No. 68-C-01-091. Sidina Dedah and Kathleen Stralka with Science Applications International Corporation provided assistance under EPA Contract No. 68-C-99-233.

Questions or comments about general aspects of this assessment should be addressed to:

William A. Telliard
USEPA (4303T)
1200 Pennsylvania Avenue, NW
Washington, DC 20460
telliard.william@epa.gov

Questions or comments about statistical issues related to this assessment should be addressed to:

Henry Kahn
USEPA (4303T)
1200 Pennsylvania Avenue, NW
Washington, DC 20460
kahn.henry@epa.gov

This document has been reviewed and approved for publication by the Engineering and Analysis Division, Office of Science and Technology. Neither the United States Government nor any of its employees, contractors, or their employees make any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use of or the results of such use of any information, apparatus, product, or process discussed in this report, or represents that its use by such party would not infringe on privately owned rights.

February 2003

Table of Contents

Acknowledgments and Disclaimer	i
Table of Contents	iii
Chapter 1 Introduction	1-1
1.1 Background	1-1
1.2 Clause 6 Settlement Agreement Requirements	1-1
1.2.1 Clause 6a	1-1
1.2.2 Clause 6b	1-2
1.2.3 Clause 6d	1-2
1.2.4 Clause 6e	1-2
1.2.5 Clause 6f	1-2
1.3 EPA's Approach to Conducting this Assessment	1-3
1.3.1 Study Plan	1-4
1.3.2 Material and Data used in the Assessment	1-4
1.4 Peer Review of the Agency's Assessment	1-9
1.5 Terminology used in this Document	1-10
Chapter 2 Overview and History of Detection and Quantitation Limit Approaches	2-1
2.1 Currie's Call for Standardization	2-1
2.2 Development of the MDL and ML as Practical Embodiments of Currie's Proposal ...	2-3
2.2.1 Method Detection Limit	2-3
2.2.2 Minimum Level of Quantitation	2-4
2.3 Approaches Advanced by Other Organizations	2-5
2.3.1 EPA Approaches	2-5
2.3.2 Industry-supported Approaches	2-7
2.3.3 Approaches Advocated by the Laboratory Community and Voluntary Consensus Standards Bodies	2-8
2.3.4 Approaches Advocated by Other U. S. Government Agencies and Other Governments	2-9
Chapter 3 Issues Pertaining to Detection and Quantitation	3-1
3.1 Analytical Chemistry Approaches to Detection and Quantitation	3-2
3.1.1 Blank versus Zero Concentration	3-2
3.1.2 Lack of Instrument Response	3-2
3.1.3 Matrix Effects	3-4
3.1.4 Recovery Correction	3-6
3.1.5 Measurement Quality over the Life of a Method	3-8
3.2 CWA Regulatory Issues Affecting Detection and Quantitation	3-9
3.2.1 Detection and Quantitation Limit Applications Under CWA	3-9
3.2.2 Descriptive versus Prescriptive Uses of Lower Limits to Measurement	3-13
3.2.3 Compliance Evaluation Thresholds	3-14
3.2.4 Accepting the Procedures of Voluntary Consensus Standards Bodies	3-14

3.2.5	National versus Local Standards for Measurement	3-16
3.2.6	Cost and Implementation Issues	3-16
3.2.7	Use of a pair of related detection and quantitation procedures in all Clean Water Act applications.	3-17
3.2.8	Alternative Procedures	3-18
3.3	Statistical Issues	3-18
3.3.1	Sources of Variability	3-19
3.3.2	Censoring Measurement Results	3-20
3.3.3	Outliers	3-22
3.3.4	Criteria for the Selection and Appropriate Use of Statistical Models	3-24
3.3.5	Methodology for Parameter Estimation	3-28
3.3.6	False Positives and False Negatives	3-28
3.3.7	Statistical Prediction and Tolerance	3-30
3.3.8	Design of Detection and Quantitation Studies	3-33
Chapter 4	Evaluation Criteria	4-1
4.1	Criterion 1	4-1
4.2	Criterion 2	4-1
4.3	Criterion 3	4-2
4.4	Criterion 4	4-4
4.5	Criterion 5	4-4
4.6	Criterion 6	4-6
Chapter 5	Assessment	5-1
5.1	Detection Limit Approaches	5-1
5.1.1	Evaluation of the MDL	5-1
5.1.2	Evaluation of the ASTM International Interlaboratory Detection Estimate (IDE)	5-7
5.1.3	Evaluation of the ACS Limit of Detection	5-11
5.1.4	Evaluation of the IUPAC/ISO Critical Value (CRV)	5-13
5.1.5	Evaluation of the IUPAC/ISO Detection Limit	5-15
5.2	Quantitation Limit Approaches	5-17
5.2.1	Assessment of the EPA Minimum level of Quantitation (ML)	5-17
5.2.2	Assessment of the IQE	5-20
5.2.3	Assessment of the ACS Limit of Quantitation	5-24
5.2.4	Assessment of the IUPAC/ISO Limit of Quantitation	5-26
Chapter 6	Conclusions	6-1
References	6-1
Appendix A	Literature Search Regarding Detection and Quantitation Limit Approaches	
Appendix B	Characterizing Measurement Variability as a Function of Analyte Concentration for a Variety of Analytical Techniques	
Appendix C	Computation of Detection and Quantitation Limits	

1.1 Background

On June 8, 1999 (64 FR 30417), EPA promulgated (i.e., published in a final rule) Method 1631B: *Mercury in Water by Oxidation, Purge and Trap, and Cold Vapor Atomic Fluorescence Spectro-metry* (the "method") for use in EPA's Clean Water Act programs. The method was developed specifically to measure mercury at ambient water quality criteria levels and includes a method detection limit (MDL; see 40 CFR part 136, Appendix B) of 0.2 nanograms per liter (ng/L).

Following promulgation, a lawsuit was filed challenging EPA on the validity of the method. The basis of the challenge included several specific aspects of Method 1631 as well as the general procedures used to establish the MDL and minimum level of quantitation (ML) published in the method. In order to settle the lawsuit, EPA entered into a settlement agreement (the "Settlement Agreement") with the Alliance of Automobile Manufacturers, Inc., the Chemical Manufacturers Association, and the Utility Water Act Group (collectively the "Petitioners") and the American Forest and Paper Association ("Intervenor") on October 19, 2000. Under Clause 6 of the Settlement Agreement, EPA agreed to perform an assessment of detection and quantitation limit concepts. The complete text of Clause 6 is provided in Exhibit 1-1 of this chapter. A summary of Clause 6 is provided in Section 1.2. The summary is followed by a description of EPA's approach to the assessment, including the material and data evaluated (Section 1.3), the use of an independent peer review to evaluate the Agency's assessment (Section 1.4), and a brief discussion of the terminology used in this document.

1.2 Clause 6 Settlement Agreement Requirements

Clause 6 of the Settlement Agreement is titled *Reassessment of Method Detection Limit and Minimum Level Procedures*. Clause 6 consists of five subclauses, a - b and d - f. (There is no subclause c).

1.2.1 Clause 6a

Clause 6a broadly defines the scope of the assessment and provides a schedule for completing the initial phase. Specifically, Clause 6a requires EPA to:

- Sign and forward to the Office of Federal Register (OFR) a notice inviting public comment on a reassessment of existing EPA procedures for determining the detection and quantitation limits of contaminants in aqueous samples.
- Forward the notice to the OFR on or before February 28, 2003.
- Provide a period of at least 120 days for public comment on the notice.
- At a minimum, include the MDL procedure published at 40 CFR part 136, Appendix B, and the ML procedure described in Section 17.8 of Method 1631B, in the reassessment of detection and quantitation limits.
- Invite comment on one or more alternative procedures for determining and describing test sensitivity.

Clause 6a also provides EPA with the option of proposing modifications to the existing procedures.

1.2.2 Clause 6b

Clause 6b requires that EPA obtain a peer review of its reassessment, and describes six specific topics that must be included in the charge to the peer reviewers. Specifically, Clause 6b requires EPA to:

- Submit the reassessment of existing procedures (including any proposed modifications thereof) and any evaluation of alternatives for peer review by experts in the field of analytical chemistry and the statistical aspects of analytical data interpretation.
- Conduct the peer review in accordance with EPA's peer review policies.
- Prepare a charge to the peer review panel that requests the peer reviewers to consider:
 - Criteria for selection and appropriate use of statistical models
 - Methodology for parameter estimation
 - Statistical tolerance and prediction
 - Criteria for design of detection and quantitation studies, including selection of concentration levels ("spiking levels")
 - Interlaboratory variability, and
 - Incorporation of elements of probability design.

1.2.3 Clause 6d

Clause 6d requires EPA to provide the Petitioners and Intervenor (the "litigants") with an opportunity for review of the Agency's assessment concurrent with the Clause 6b peer review.

1.2.4 Clause 6e

Clause 6e requires EPA to provide the litigants with:

- An opportunity to meet periodically (i.e., every six months) to discuss the Agency's progress during development of the assessment,
- A plan for performing the assessment on or before the second of these meetings, and
- Copies of relevant documents, where appropriate, in advance of these meetings.

1.2.5 Clause 6f

Clause 6f establishes a schedule and requirements concerning final action on the notice described in Clause 6a. Specifically:

- On or before September 30, 2004, EPA is to sign and forward to the OFR a notice taking final action on the notice described in Clause 6a, and
- Coincident with publication of this notice of final action, EPA is to provide the litigants with an opportunity to meet and discuss the implications of the final notice and/or the need for any subsequent EPA action in light of the final notice.

Exhibit 1-1. Full Text of Clause 6 of the Settlement Agreement

6. Reassessment of Method Detection Limit and Minimum Level Procedures

- a. On or before February 28, 2003, EPA shall sign and forward to the Office of the Federal Register for prompt publication a notice inviting public comment on a reassessment of the existing Agency procedures for determination of sensitivity of analytic test methods for aqueous samples, specifically, EPA procedures for determining the detection limits and levels of quantitation of contaminants in aqueous samples, including, at a minimum, the "Definition and Procedure for Determination of the Method Detection Limit" published at 40 C.F.R. Part 136, Appendix B, as well as the "minimum level" procedures, which is described in section 17.8 of Method 1631B. The notice shall invite comment on EPA's evaluation of one or more alternative procedures for determining and describing test sensitivity. The notice also may propose modifications to the existing procedures. The notice shall invite public comment for a period of no less than one hundred twenty (120) days.
- b. Prior to publishing the notice inviting public comment on EPA procedures for determining test sensitivity, EPA shall submit its reassessment of existing procedures (including any proposed modifications thereof) and its evaluation of alternatives for peer review by experts in the field of analytical chemistry and the statistical aspects of analytical data interpretation. In its charge to the peer review panel, EPA shall request that the peer review consider: criteria for selection and appropriate use of statistical models; methodology for parameter estimation; statistical tolerance and prediction; criteria for design of detection and quantitation studies, including selection of concentration levels ("spiking levels"); interlaboratory variability; and incorporation of elements of probability design. EPA (or its authorized representative) shall conduct the peer review in accordance with EPA's current peer review policies in the January 1998 Science Policy Council Handbook (EPA 100-B-98-00) [sic], including any subsequently-developed EPA peer review documents that may revise or amend that Handbook.

[Note - the correct document number for the Science Policy Council Handbook is EPA 100-B-98-001]

[c. Note - there is no clause "6.c" in the Settlement Agreement]

- d. During the peer review period, EPA shall also provide an opportunity for concurrent review and comment by the Petitioners and Intervenor.
- e. In the development of the reassessment/assessment of alternatives, EPA shall provide the Petitioners and Intervenor with a periodic opportunity to meet (i.e., every six (6) months) on the Agency's progress. EPA shall prepare and present the Petitioners and Intervenor with the Agency's "plan" for conducting the reassessment/assessment of alternatives on or before the second such periodic meeting. Where appropriate, EPA shall provide the Petitioners and Intervenor with copies of relevant documents in advance of such meetings.
- f. On or before September 30, 2004, EPA shall sign and forward to the Office of the Federal Register for prompt publication a notice taking final action on the notice described in subparagraph 6.a. Coincident with publication of the notice of final action, EPA shall provide Petitioners and Intervenor an opportunity to meet to discuss the implications of the final notice and/or the need for any subsequent EPA action in light of the final notice.

1.3 EPA's Approach to Conducting this Assessment

This document details the Agency's assessment of methodology for the determination of method sensitivity, specifically: detection and quantitation limits. This assessment is being conducted in accordance with a plan summarized in Section 1.3.1 and is based, in part, on an assessment of the data described in Section 1.3.2.

1.3.1 Study Plan

EPA developed a technical approach for 1) conducting the assessment, and 2) complying with all applicable requirements of the Settlement Agreement. The approach was documented in a draft study plan that has since formed the general framework for the assessment described in this Assessment Document. EPA also conducted a literature search to identify and review issues and concepts that should be considered when developing the plan. A summary of this literature review is provided in Appendix A to this Assessment Document.

The study plan described roles and responsibilities for implementing the plan, provided a background discussion of detection and quantitation limit concepts, including the MDL and ML, and outlined a series of 11 events associated with the Agency's assessment of detection and quantitation limit approaches. The relationship between those planned events and this Assessment Document is summarized in Exhibit 1-2 at the end of this chapter.

Although the Settlement Agreement did not require that EPA seek formal peer review on its draft plan, the Agency chose to conduct a peer review of the draft plan. The peer review was initiated in December 2001, conducted in accordance with EPA's peer-review policies, and performed by two statisticians and two chemists. EPA reviewed the comments and recommendations offered by these reviewers, and where appropriate, revised the plan to reflect the peer-review comments. EPA also reviewed, and where appropriate, revised the plan to reflect comments provided by the litigants following their concurrent review.

1.3.2 Material and Data used in the Assessment

In order to perform the assessment described in this document, EPA sought to collect documentation describing existing detection and quantitation limit concepts and procedures and data that could be used to evaluate these concepts and procedures.

Documentation concerning the existing concepts and procedures was obtained by performing a literature search as described in Appendix A to this Assessment Document, and where appropriate, by purchasing copies of documents describing concepts or procedures from the organizations that published them.

In performing this assessment, EPA hoped to identify a substantial amount of data containing results of direct relevance to the determination of detection and low-level measurement capability. That is, measurement results in the low concentration region. To date, EPA has been able to identify only six data sets that were of use in fully evaluating variability in the range of analytical detection and quantitation. Three of the six were developed by EPA for the express purpose of studying the relationship between measurement variation and concentration across a wide variety of measurement techniques and analytes. EPA refers to these data sets as "EPA's ICP/MS Study of Variability as a Function of Concentration," "EPA's Multi-technique Variability Study" (also referred to as the "Episode 6000 study"), and "EPA's GC/MS Threshold Study" (also referred to as "the Episode 6184 study"). In all three cases, replicate measurement results from each combination of analyte and measurement technique were produced by a single laboratory over a wide range and large number of concentrations. The fourth data set was developed by the American Automobile Manufacturer's Association (AAMA) for the purpose of estimating one particular kind of quantitation value. That quantitation value is called an alternative minimum level (AML; see Gibbons *et al.*, 1997). In the AAMA study, replicate results were measured at a limited number of concentrations by multiple laboratories using EPA Method 245.2 (cold vapor atomic absorption; CVAA) for mercury and EPA Method 200.7 (inductively coupled plasma/atomic emission spectroscopy; ICP/AES) for twelve other metals. The final two data sets were

jointly gathered by EPA and the Electric Power Research Institute (EPRI) to support interlaboratory validation of EPA Methods 1631 and 1638.

The studies from which these six data sets were obtained are summarized in sections 1.3.2.1 - 1.3.2.6 below. Additional information about these studies can be found in Appendices B and C to this Assessment Document.

Although the litigants offered specific suggestions for other data sets that they believed should be considered in this assessment, EPA found that these data sets did not include a sufficient number of results in the region of detection and quantitation to yield information for the assessment, overlapped with data already used in the assessment, or exhibited signs of significant contamination that made the data inappropriate for inclusion in the assessment. These data, and EPA's decisions regarding the data, are discussed in Section 1.3.2.7 below.

1.3.2.1 EPA's ICP/MS Study of Variability as a Function of Concentration

The objective of the ICP/MS study was to characterize variability as a function of concentration using EPA's draft Method 1638 for determination of nine metals by inductively coupled plasma with mass spectroscopy (ICP/MS). The nine metals were silver, cadmium, copper, nickel, lead, antimony, selenium, thallium, and zinc. The ICP/MS instrument used in this study averages triplicate scans to produce a single measurement of each element at each concentration. Such averaging is typical of ICP/MS design and use.

In preparation for the study, the ICP/MS was calibrated using triplicate scans averaged to produce a single measurement of 100, 1,000, 5,000, 10,000, and 25,000 nanograms per liter (ng/L) for each element. Originally, the instrument was calibrated using unweighted least squares estimates under the assumption of linearity. Subsequently, the analytical results were adjusted with weighted least squares estimates. Weighted least squares estimates are based on the knowledge that variability (expressed as the standard deviation) increases with increasing analyte concentration.

Although the instrumentation has the capability to provide intensity results for each of the three scans at each concentration, averaging the three scans to produce a single measurement is the normal operating mode, and the average was used to produce the measurements in this study. Draft Method 1638 specifies the use of average response factors rather than least squares estimation of a linear calibration, although it does allow for the use of such procedures.

All nine metals were spiked into reagent water to produce solutions at concentrations of: 0, 10, 20, 50, 100, 200, 500, 1,000, 2,000, 5,000, 10,000, and 25,000 ng/L. Each solution was divided into seven replicate aliquots for subsequent analysis. The aliquots were analyzed beginning with the blank (zero concentration) followed by analyses from the highest to the lowest concentration. This sequence was chosen to minimize carry-over effects and to allow the analyst to stop at the concentration that returned zero results. Carry-over is caused by residual sample remaining in the inlet system of the instrument, in this case, the ICP/MS. Carry-over can occur when analysis of a high-concentration sample is followed by analysis of a relatively low-concentration sample, as could occur if the replicates were analyzed in random order. Use of the highest to lowest analytical sequence ensured that each successive concentration analyzed was close enough to the previous concentration that any effects of carryover would be negligible and, therefore, would not compromise study results. (A more in-depth discussion of the randomized design and the effects of carry-over issues is provided in Chapter 3, Section 3.3.8.2).

Results at multiple mass-to-charge ratios, or m/z 's, were reported for each metal, although draft Method 1638 specifies only one m/z for eight of the nine metals. For lead, m/z 's 206, 207, and 208 are

specified. Only data associated with m/z's specified in draft Method 1638 were used in the ICP/MS study.

1.3.2.2 EPA's Multi-technique Variability Study (the "Episode 6000 Study")

In 1997 and 1998, EPA conducted a study of variability vs. concentration for a number of analytical methods. Five laboratories were employed for the analyses; each analyte and method combination was tested by one of these laboratories. Details of the study design are described in EPA's *Study Plan for Characterizing Variability as a Function of Concentration for a Variety of Analytical Techniques* (July 1998). Based on the sampling episode number assigned to the study by the EPA Sample Control Center, the study and results have become known as the Episode 6000 study and data. The analytes and analytical techniques studied were:

- Total suspended solids (TSS) by gravimetry
- Metals by graphite furnace atomic absorption spectroscopy (GFAA)
- Metals by inductively-coupled plasma atomic emission spectrometry (ICP/AES)
- Hardness by ethylene diamine tetraacetic acid (EDTA) titration
- Phosphorus by colorimetry
- Ammonia by ion-selective electrode
- Volatile organic compounds by purge-and-trap capillary column gas chromatography with a photoionization detector (GC/PID) and electrolytic conductivity detector (GC/ELCD) in series
- Volatile organic compounds by gas chromatography with a mass spectrometer (GC/MS)
- Available cyanide by flow-injection/ligand exchange/amperometric detection
- Metals by inductively-coupled plasma spectrometry with a mass spectrometer (ICP/MS)

In this study, an initial (range finding) MDL was determined for each combination of analyte and analytical technique using minor modifications to the MDL procedure at 40 CFR part 136. Specifically, the modifications made the optional iterative step 7 of the MDL procedure mandatory and required the spike concentration to be no more than a factor of three times the determined MDL (instead of a factor of five times). During the study, however, two of the laboratories found that the reduction in the allowable spike range necessitated an unreasonably large number of iterations. In continuing the study, EPA returned to the spike-to-MDL ratio of five published in the 40 CFR part 136, Appendix B procedure.

After determining the initial MDL, each laboratory analyzed 7 replicate samples spiked at concentrations that were 100, 50, 20, 10, 7.5, 5.0, 3.5, 2.0, 1.5, 1.0, 0.75, 0.50, 0.35, 0.20, 0.15, and 0.10 times the initial MDL. In a few instances, laboratories analyzed more than 7 replicates. As often as possible, the replicate analyses at each concentration level were produced using the same calibration that was used in determining the initial MDL. Where laboratory reports indicated that multiple calibrations were conducted, each result was associated with its calibration in the data analysis.

Spiked aqueous solutions were analyzed in order from the highest concentration (100 times the MDL) to the concentration at which 3 or more non-detects (zeros) were encountered among the 7 replicates, or the lowest concentration specified (0.1 times the MDL), whichever occurred first. This analysis order (1) minimized carryover that could occur in some methods if a low-concentration sample had followed a high-concentration sample (as may happen when samples are analyzed in random order), and (2) prevented collection of a large number of zeros if the signal disappeared.

For methods that do not produce a signal for a blank, the signal will disappear somewhere below the MDL, i.e., a zero will be reported. Laboratories were instructed that when three nondetects (out of seven measurements) were reported, it was not necessary to move to the next lower concentration,

because it would be of no practical value to have laboratories measure seven zeros, move to a lower level, measure seven zeros, etc.

A variant of the iterative procedure for determining the MDL was used for organic compounds determined by chromatographic methods. Methods for organics normally list many (15 to 100) analytes, and the response for each analyte is different. Therefore, to determine an MDL for each analyte, the concentration of the spike would need to be inversely proportional to the response. Making a spiking solution with 15 to 100 different concentrations is cumbersome and error prone. The approach used in the study was to run seven replicates at decreasing concentrations until signal extinction, then select the concentration(s) appropriate for the determining the MDL for each analyte according to the MDL procedure. In some cases, the laboratories selected the concentrations, in others cases, EPA did. This approach was generally applied for organics analysis. However, laboratories also had the option of using some combination of the monotonically decreasing concentrations described above and a few selected concentrations to achieve the desired spiking levels.

1.3.2.3 EPA's GC/MS Threshold Study (the "Episode 6184 Study")

Data from the Episode 6184 study of variability vs. concentration were used to evaluate the effect of GC/MS thresholds on the ability to identify semivolatile organic compounds at low concentrations. Details of the design of this study are described in EPA's *Study Plan for Characterizing Error as a Function of Concentration for Determination of Semivolatiles by Gas Chromatography/Mass Spectrometry* (December 1998). Data were generated for 82 semivolatile organic compounds using EPA Method 1625C (semivolatile organic compounds by GC/MS). MDLs were not determined for these compounds. Instead, solutions of the analytes were prepared and analyzed at concentrations of 50.0, 20.0, 10.0, 7.50, 5.00, 3.50, 2.00, 1.50, 1.00, 0.75, 0.50, 0.35, 0.20, 0.15, 0.10, 0.075 and 0.050 ng/ μ L (or μ g/mL). Each solution was injected into the GC/MS in triplicate with the mass spectrometer threshold set to 0, and again in triplicate with the mass spectrometer threshold set to a level typical of that used in routine environmental analyses. As with the ICP/MS study and the Episode 6000 study, and for the same reasons described in Section 1.3.2.1, samples were analyzed in order from the highest to the lowest concentration.

1.3.2.4 AAMA Metals Study of Methods 200.7 and 245.2

The American Automobile Manufacturer's Association conducted an interlaboratory study of EPA Method 200.7 (metals by ICP/AES) and Method 245.2 (mercury by CVAA). The study was designed to estimate a quantitation value based on a concept termed the alternative minimum level (AML) that had been described in the literature (Gibbons *et al.*, 1997). Nine laboratories participated in the study, and each reported data for the following 13 metals: aluminum, arsenic, cadmium, chromium, copper, lead, manganese, mercury, molybdenum, nickel, selenium, silver and zinc. Study samples were analyzed by EPA Method 200.7 for 12 of the metals. Mercury was determined by EPA Method 245.2.

As part of the study design, the nine laboratories were randomized prior to the start of the study. Five sample matrices (including reagent water) were studied, including four wastewater matrices that are representative of the automotive industry. Starting from a blank, or unspiked sample, all target analytes were spiked at four concentrations to yield a total of five concentrations per matrix. Concentrations ranged from 0.01 to 10 μ g/L for mercury and selenium on the low end, and from 2.0 and 1000 μ g/L for mercury and selenium on the high end. In addition, the concentrations were matrix-dependent. The same concentration ranges for each metal by matrix combination were used for all five weeks of the study.

Matrix A (reagent water) was analyzed in all nine laboratories, and three laboratories analyzed each of the other four matrices. All analyses were repeated weekly over a five-week period. As a result,

a total of 6,825 observations were obtained, which includes 2,925 observations for matrix A (9 labs × 13 metals × 5 spike concentrations × 5 weeks), and 975 observations (3 labs × 13 metals × 5 spike concentrations × 5 weeks) for each of the other four matrices (6,825 = 2,925 + (975 × 4)). There were two missing values for chromium in matrix A from laboratories 1 and 9.

1.3.2.5 Method 1631 Interlaboratory Validation Study

The Method 1631 interlaboratory validation study was conducted by EPA to evaluate performance of the method and to gather data to evaluate existing performance specifications, including detection and quantitation limits. To accommodate stakeholder interests and expand the scope of the study, the Electric Power Research Institute (EPRI) funded the distribution of additional samples to study participants.

This jointly funded study involved an international community of twelve participating laboratories and one referee laboratory. Each participating laboratory analyzed four different matrices, each containing mercury at a concentration selected to allow for characterization of method performance across the measurement range of the method. Each of the 12 participating laboratories was provided with 13 sample pairs (a total of 26 blind samples). These included 1 filtered effluent pair, 1 unfiltered effluent pair, 4 filtered freshwater pairs, 1 filtered marine water pair, 1 unfiltered marine water pair, and 5 spiked reagent water pairs. All 12 laboratories received and analyzed the same sample pairs (a total of 312 analyses). To measure the recovery and precision of the analytical system, and to monitor matrix interferences, the laboratories were instructed to analyze matrix spike and matrix spike duplicate samples on specified field samples for each filtered and unfiltered matrix, spiked at 1-5 times the background concentration of mercury determined by analysis of an unspiked aliquot of the sample. The laboratories were instructed to perform all other QC tests described in Method 1631, including the analysis of blanks, and to conduct MDL studies in reagent water following the procedure at 40 CFR part 136.

1.3.2.6 Method 1638 Interlaboratory Validation Study

The Method 1638 interlaboratory validation study was conducted by EPA to evaluate performance of the method and to gather data that would allow revision of existing performance specifications, including detection and quantitation limits. To accommodate stakeholder interests and expand the scope of the study, the Electric Power Research Institute funded the distribution of additional samples to study participants.

A total of eight laboratories (and a referee laboratory) participated in the study. The study was designed so that each participating laboratory would analyze sample pairs of each matrix of interest at concentrations that would span the analytical range of the method. Each laboratory was provided with 11 sample pairs (a total of 22 blind samples). These included 1 filtered effluent pair, 1 unfiltered effluent pair, 4 filtered freshwater pairs, and 5 spiked reagent water pairs. All eight laboratories received and analyzed the same sample pairs (a total of 176 analyses). To measure the recovery and precision of the analysis, and to monitor matrix interferences, the laboratories were instructed to analyze a matrix spike and matrix spike duplicate of specified field samples in each filtered and unfiltered matrix, spiked at 1-5 times the background concentration of the analytes determined by analysis of an unspiked aliquot of the sample. The laboratories were instructed to perform all other QC tests described in Method 1638, including the analysis of blanks, and to conduct MDL studies in reagent water following the procedure at 40 CFR part 136.

1.3.2.7 Data Considered but not Used in this Assessment

The Petitioners and Intervenor to the Settlement Agreement suggested ten specific data sets that EPA should consider in its assessment of detection and quantitation limits. EPA evaluated each of these data sets to determine if the design of the study, including the concentrations targeted in the study, would provide sufficient data for evaluating measurement variability in the region of interest (i.e., at concentrations below, at, and above the region of detection and quantitation). If such data were determined to be present, EPA further evaluated the data set to ensure that it was of sufficient quality to support the Agency's assessment. Four of the ten data sets met these requirements and were used in EPA's assessment. Table 1 identifies each of the data sets suggested by the petitioners along with EPA's rationale for using or excluding the data from this assessment.

Table 1. Data Sets Suggested by Petitioners

Dataset Source and Year	Analytes and technology	EPA Decision Regarding Use
AAMA 1996-1997	Metals by ICP/AES (200.7)	Used in this assessment and described in Section 1.3.2.4
AAMA 1996-1997	Mercury by CVAA (245.2)	Used in this assessment and described in Section 1.3.2.4
EPA/EPRI 1997-1998	Mercury by CVAF (1631)	Used in this assessment and described in Section 1.3.2.5
EPA/EPRI 1997-1998	Metals by ICP/MS (1638)	Used in this assessment and described in Section 1.3.2.6
EPRI 1987	Metals by GFAA (EPA 200)	Not used in this assessment because of insufficient low-level data
EPRI 1990	Metals by ICP/AES (EPA 200.7)	Not used in this assessment because of insufficient low-level data
EPRI 1994	Al, Be, Ti by GFAA (EPA 200)	Not used in this assessment because of overlap with EPA's Episode 6000 Study, which provides data on the same analytes but covers a larger number of concentrations in the region of interest
AAMA 1996-1997	PCBs by GC/ECD (608.2)	Not used in this assessment because of overlap with EPA's Episode 6000 Study, which provides data on the same analytes but covers a larger number of concentrations in the region of interest
EPRI 1996	Cd, As, Cr by GFAA (EPA 200)	Not used in this assessment because of overlap with EPA's Episode 6000 Study, which provides data on the same analytes but covers a larger number of concentrations in the region of interest
MMA 2000-2001	Aroclors 1016 and 1260 by GC/ECD	Not used in this assessment. Although the study examined the region of detection and quantitation, samples spiked with low levels of Aroclors exhibited average recoveries >500%, with RSDs >200% across 10 laboratories, indicating contamination of the samples from an unknown source.

1.4 Peer Review of the Agency's Assessment

In August 2002, EPA conducted a formal peer review of the Agency's assessment. This peer review, which satisfied requirements in Clause 6b of the Settlement Agreement, was conducted in accordance with EPA's peer review policies described in the Science Policy Council Handbook (EPA 100-B-00-001). The review was performed by two experts in the field of analytical chemistry and two experts in the statistical aspects of analytical data interpretation. Each reviewer was provided with a draft

version of this Assessment Document, which documented the Agency's approach to the assessment and the Agency's preliminary findings and conclusions. Reviewers also were provided with copies of all data evaluated in the assessment, statistical programs used to analyze the data, and copies of the detection and quantitation concepts and procedures evaluated by EPA. In accordance with the Agency's peer review policies, the reviewers also were provided with a written 'charge' intended to ensure the evaluation would meet EPA needs.

In its charge to the peer reviewers, EPA requested a written evaluation of whether the assessment approach described by EPA is valid and conceptually sound. Reviewers also were asked to consider and address eight specific questions pertaining to the adequacy of the concepts and issues considered, the evaluation criteria developed by EPA, EPA's assessment and conclusions, the data used to perform the assessment, suggested improvements to the procedures discussed, and EPA's consideration of interlaboratory vs. intralaboratory issues. Comments from peer reviewers were generally supportive of EPA's assessment and its presentation of the assessment in the draft Assessment Document. Where appropriate, EPA revised the Assessment Document to reflect specific suggestions and comments offered by the peer reviewers. This version of the Assessment Document reflects those revisions. Copies of all materials associated with the peer review, including the peer review charge, the materials provided to the peer reviewers for review, complete copies of the peer reviewers' comments, and detailed EPA responses to each of the comments are provided in the public docket supporting the Agency's assessment.

1.5 Terminology used in this Document

We use the term "quantitation" in this document because of its common usage among analytical chemists, even though we recognize that the term "quantification" (i.e., the act of quantifying) is the term listed in most dictionaries. Also, when referring to detection and quantitation, we use the words "approach" or "concept" to refer, generically, to the procedures used to establish detection and quantitation limits or the theories on which those procedures are based. We use the word "limit" rather than "level" to indicate that the detection and quantitation concepts are directed at the lowest concentration or amount at which an analyte is determined to be present (detection) or may be measured (quantitation). In choosing the word 'limit' we do not mean to imply any sense of permanence. We recognize that measurement capabilities generally improve over time, and that detection or quantitation 'limits' established today may be superseded by future developments in analytical chemistry.

Exhibit 1-2. Relationship of Assessment Document to Assessment of Detection and Quantitation Limit Approaches

Event 1, Develop a detailed plan for responding to Clause 6 the Settlement Agreement: This event was completed in April 2002 when the draft plan was revised to reflect peer review and Litigant comments.

Event 2, Identify and explore issues to be considered: The Settlement Agreement identified six specific issues that should be considered during the assessment of detection and quantitation limit concepts, and subjected to formal peer review. During development of the technical approach, EPA identified a number of other issues that should be considered during the assessment. EPA listed and described each of these issues in the study plan and noted that identification of issues is likely to be a dynamic process, in that as a suite of issues is identified and discussed, other issues may surface. Finally, EPA stated its intent to prepare an "issue paper" that fully explained and discussed each of the identified issues. Chapter 3 of this Assessment Document serves the function of the issue paper described in the plan.

Event 3, Develop criteria against which concepts can be evaluated: After fully considering all relevant issues, EPA developed a suite of criteria that could be used to evaluate the suitability of various detection and quantitation procedures for use in CWA programs. Chapter 4 of this Assessment Document provides and describes the criteria selected by EPA after its consideration of all pertinent issues.

Event 4, Evaluate existing procedures for establishing detection and quantitation levels: EPA evaluated existing detection and quantitation limit concepts used or advanced 1) by voluntary consensus standards bodies (VCSBs), 2) in the published literature, 3) by EPA. As per the terms of the Settlement Agreement, the MDL and ML were explicitly targeted for inclusion. EPA committed to evaluating concepts published by ASTM International and ISO and to consider approaches and procedures offered by other organizations such as the American Chemical Society (ACS) and the International Union of Pure and Applied Chemistry (IUPAC), as well as other approaches that have been adopted by EPA for use in other programs or that were identified during EPA's review of the published literature. Chapter 2 describes the concepts that EPA evaluated in the assessment. Where appropriate, these approaches also are discussed in context to the issues that are identified and discussed in Chapter 3. Chapter 5 presents the results of EPA's assessment of each approach against the evaluation criteria established in Chapter 4. Appendices B and C of this document present additional details of EPA's assessment of each approach, using the data described in Chapter 1, Section 1.3.

Event 5, Develop and evaluate alternative procedures: EPA planned to develop and evaluate alternative procedures and modifications to existing procedures only if the Agency's assessment of existing procedures suggested that modifications or alternatives to the existing procedures were needed. EPA noted that its primary objective in developing such alternatives (or modifications) would be to address deficiencies noted in Event 4 and improve the performance of the procedures that best meet the criteria established in Event 3. In accordance with the plan and with EPA's findings during the assessment, this Assessment Document includes suggested modifications to the existing MDL and ML procedures.

Event 6, Conduct peer review of the Agency's assessment: EPA documented results of the Agency's assessment in a draft Assessment Document that was completed in August, 2002. EPA conducted a formal peer review of the assessment in accordance with the Agency's peer-review policies and guidance. The peer review was performed by two experts in the field of analytical chemistry and two experts in the statistical aspects of analytical data interpretation.

Events 7 - 11, Actions taken following peer review. After considering peer review comments, EPA revised its assessment and the draft Assessment Document to reflect peer review comments. EPA also finalized its strategy regarding the FR notices to be published per the terms of Settlement Agreement Clause 6a and took the actions necessary to ensure publication of those notices.

Chapter 2 Overview and History of Detection and Quantitation Limit Approaches

It is not possible to measure the concentration of a substance in water all the way down to zero. As an analogy, consider the following example: imagine measuring an object less than 1/16th of an inch in length with a ruler marked in 1/16th-inch increments. How well can the length of the object be measured using only the ruler? Similar issues arise as chemists try to measure ever smaller concentrations of substances in water. In response to the challenges associated with measuring low concentrations, chemists have defined numerical values that provide points of reference for reporting and using measurement results. These values are usually referred to as detection and quantitation limits. This chapter provides an overview of detection and quantitation approaches and procedures in analytical chemistry and their use in Clean Water Act applications.

2.1 Currie's Call for Standardization

Since 1968, most of the literature regarding detection and quantitation has referenced the work of Dr. Lloyd Currie, recently retired from the National Institutes of Science and Technology (NIST, formerly the National Bureau of Standards). In 1968, Currie published a paper in which he reviewed the then current state of the art regarding detection and quantitation, presented a three-tiered concept, and demonstrated his concept with operational equations for a single laboratory. In his paper, Currie reviewed eight existing definitions for the concept of detection, and reported that when these eight operational definitions were applied to the same data, they resulted in numerical values that differed by nearly three orders of magnitude. These results made it impossible to compare the detection capabilities of measurement methods using available publications. Currie proposed standardizing the terminology using theoretical definitions that he called the *critical value*, the *detection limit*, and the *determination limit*. (In 1995, writing on behalf of International Union of Pure and Applied Chemistry (IUPAC), Currie used the term "quantification limit" instead of his original term "determination limit." Substantial agreement with the International Organization for Standardization (also known as "ISO") on the meaning and language of detection and quantitation was achieved later, although some "subtle differences in perspective" remain [Currie, 2000]). His purpose for these definitions was to create a system in which the standard documentation of any measurement method would include a statement of capabilities that were directly comparable to any other method for measuring the same substance.

Currie used terms from statistical decision theory as the basis for his three-tiered system. In 1968 and 1995, Currie defined the *critical value* as the measured value at which there is a small chance that the concentration in the sample is zero. Consequently, any measured result greater than or equal to the critical value is considered evidence that the sample contains the substance of interest. Currie was careful to emphasize that the decision as to whether the substance has been detected is made by comparing the measurement result to the critical value. Figure 2-1 shows a critical value selected such that measurements greater than the critical value have less than a 1% chance of being associated with a sample that does not contain the substance of interest. The area under the curve to the right of the critical value represents the probability that a measured value will exceed the critical value. The area

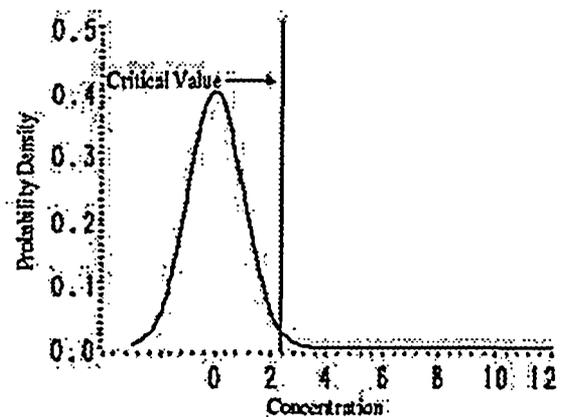


Figure 2-1

under the curve to the left of the critical value represents the (much greater) probability of observing a value that is less than the critical value when the true concentration is zero.

Currie (1968 and 1995) used the term *detection limit* to refer to a true concentration that has a high probability of generating measured values greater than the critical value. That is, measurements on samples that contain concentrations equal to the *detection limit* have a high probability of exceeding the *critical value* and are, therefore, unlikely to result in a decision that the substance is not detected in the sample. In Currie's concept, the *critical value* and the *detection limit* are related and functionally dependent, but it is clear that the detection decision is made on the basis of comparing sample by sample measurements to the *critical value*. While Currie's terminology is consistent with standard statistical decision theory, it is in all likelihood responsible for a great deal of confusion among chemists and others who may associate the term 'limit' with some sort of decision point. Currie (1995) states: "The single, most important application of the detection limit is for planning. It allows one to judge whether the CMP (Chemical Measurement Process) under consideration is adequate for the detection requirements." Figure 2-2 shows a detection limit selected such that 99% of the measurements on a sample containing this concentration are expected to be above the critical value. The bell-shaped curve centered at the detection limit illustrates how likely various measurement responses are when the concentration of the substance in a sample is equal to the detection limit. That is, the figure shows the probability density of values measured in a sample with a true concentration equal to the detection limit. The area under the curve to the left of the critical value is equal to 1% of the total area, while the area to the right is equal to 99%.

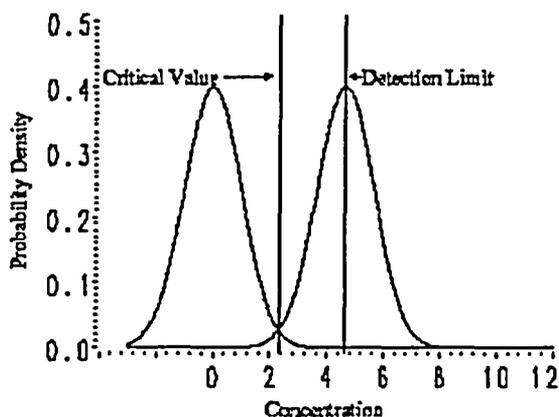


Figure 2-2

Currie (1968, 1995) defined the *determination limit*, later renamed the *quantification limit*, as (quoting Currie, 1995) "performance characteristics that mark the ability of a CMP to adequately 'quantify' an analyte." Quantification limits "serve as benchmarks that indicate whether the CMP can adequately meet the measurement needs. The ability to quantify is generally expressed in terms of the signal or analyte (true) value that will produce estimates having a specified relative standard deviation (RSD) commonly 10%." This translates into a quantification limit equal to a multiplier of 10 times the standard deviation (a measure of measurement variability) at the limit. The multiplier of 10 (equal to the inverse of the 10% RSD) is arbitrary, but has been used widely. IUPAC selected 10 as a "default value" (Currie, 1995), implying other values are possible. In papers published in 1980 and 1983, the American Chemical Society's Committee on Environmental Improvement also recommended the use of a multiplier of 10 for determining quantitation limits (see MacDougall, *et al.*, 1980 and Keith, *et al.*, 1983). Measured concentrations greater than the quantitation limit are considered to be reliable by chemists, although from a statistical perspective, any measured value, along with knowledge of the precision of the measurement, is useful.

Currie's goal of having method developers publish directly comparable descriptions of detection and quantitation capability remains elusive more than thirty years after publication of his first paper on this topic. Even if Currie's three-tiered concept were used, the treatment of related issues causes difficulty in comparing methods. Some of these issues include interlaboratory variability, selection of appropriate statistical models, design of detection and quantitation capability studies, and statistical prediction and tolerance. These and other issues are discussed in Chapter 3 of this document.

2.2 Development of the MDL and ML as Practical Embodiments of Currie's Proposal

In 1981, staff at EPA's Environmental Monitoring and Support Laboratory in Cincinnati, Ohio, published a procedure for determining what they referred to as a method detection limit (MDL) (Glaser *et al.*, 1981). The MDL functions as a practical, general purpose version of Currie's *critical value*. The MDL was subsequently promulgated for use in CWA programs on October 26, 1984 (49 FR 43234) at 40 CFR part 136, Appendix B. Prior to formal development of the MDL in 1981, the EPA Office of Water had included the term "minimum level" (ML) or "minimum level of quantitation" in some methods for analysis of organic pollutants. These methods were proposed on December 3, 1979 and subsequently promulgated on October 26, 1984, along with the MDL. Additional information about the MDL and ML is provided below in Sections 2.2.1 and 2.2.2.

2.2.1 Method Detection Limit

Conscious of the definitions provided by Currie and others, Glaser *et al.* (1981) stated "[t]he fundamental difference between our approach to detection limit and former efforts is the emphasis on the operational characteristics of the definition. [The] MDL is considered operationally meaningful only when the method is truly in the detection mode, i.e., [the] analyte (the substance of interest) must be present." Expanding on this reasoning, Glaser *et al.* (1981) developed MDL estimates for methods that produce a result of zero for blanks, such as EPA Methods 624 and 625 for determination of organic pollutants by gas chromatography/mass spectrometry (GC/MS). Blank variability exists, whether or not it can be detected by measurement processes. Failure to detect this variability may be attributed to insufficient sensitivity of the measurement process or, as is the case with some measurement processes, thresholds that are built into equipment which censor measurements below certain levels. Currie's critical value is dependent on the ability to estimate measurement variability of blank samples. In cases where the substance is not detected in direct measurements on blanks, an alternative approach to estimating blank variability must be used. One option is to estimate measurement variability at concentrations that represent the lowest possible levels where a signal can be detected. This is the basic approach of the MDL, which provides a general purpose, straightforward, operational procedure for estimating a quantity analogous to the Currie critical value when measurement processes applied to blank samples do not produce detectable signals. More complex statistical procedures for estimating blank variability are possible and may be preferable from a rigorous statistical perspective, but the MDL has been found to be satisfactory by chemists in a wide range of applications.

In 1984, the MDL became a regulatory option for wastewater discharge permits authorized under the Clean Water Act. To determine the MDL, at least seven replicate samples with a concentration of the pollutant of interest near the estimated detection capabilities of the method are analyzed. The standard deviation among the replicate measurements is determined and multiplied by the *t*-distribution for $n-1$ degrees of freedom (in the case of 7 replicates, the multiplier is 3.143, which is the value for 6 degrees of freedom). The decision to base the MDL on a minimum of seven replicates reflected a consensus among EPA chemists and statisticians that a requirement of seven replicates is not overly burdensome for laboratories and that laboratories could reasonably be expected to perform the analyses in a single batch.

Both the MDL concept and the specific definition at part 136 have been used within EPA by the Office of Ground Water and Drinking Water (OGWDW), the Office of Solid Waste (OSW), the Office of Emergency and Remedial Response (OERR), and others. The MDL also has been used outside of EPA in *Standard Methods for the Examination of Water and Wastewater*, published jointly by the American Public Health Association (APHA), the American Water Works Association (AWWA), and the Water Environment Federation (WEF), and in methods published by the ASTM International, and elsewhere.

Despite such widespread use, some members of regulated industry and others have claimed that the MDL is a less than ideal concept for detection. Specifically, critics have faulted the MDL because:

- There are some inconsistencies between the definition and the procedure
- It does not account explicitly for false negatives
- It does not account for bias
- A prediction or tolerance limit adjustment is not provided, and
- It does not account for interlaboratory variability

These issues are discussed later in this document.

2.2.2 Minimum Level of Quantitation

The minimum level of quantitation (ML) was originally proposed on December 5, 1979 (44 FR 69463) in footnotes to Table 2 of EPA Method 624 and to Tables 4 and 5 of EPA Method 625. The ML was defined as the "level at which the entire analytical system must give recognizable mass spectra and acceptable calibration points" (in the footnote to Table 2 in Method 624) and as the "*level at which the entire analytical system must give mass spectral confirmation*" (in the footnotes to Tables 4 and 5 in EPA Method 625).

Between 1980 and 1984, EPA also developed Methods 1624 and 1625 and promulgated these methods along with the final versions of EPA Methods 624 and 625 on October 26, 1984 (49 FR 43234). The definitions of the ML in the promulgated versions of EPA Methods 1624 and 1625 were the "*level at which the analytical system shall give recognizable mass spectra (background corrected) and acceptable calibration points*" (in footnote 2 to Table 2 in Method 1624) and as the "*level at which the entire GC/MS system must give recognizable mass spectra (background corrected) and acceptable calibration points*" (in footnotes 2 to Tables 3 and 4 in Method 1625).

As EPA developed additional methods over the next decade, the definition of the ML was generalized to "*the lowest level at which the entire analytical system must give a recognizable signal and acceptable calibration point for the analyte*" (see, e.g., Section 24.2 of EPA Method 1613 at 40 CFR part 136, Appendix A). In generating actual numerical values for MLs, the lowest calibration point was estimated from method development studies and included in the methods, although a specific calculation algorithm was not used. EPA methods that include the ML generally specify the number of calibration standards to be used and the concentrations of those standards. As a result, laboratories using those methods calibrate their analytical systems with a multi-point calibration (i.e., calibrate using a series of standards at different concentrations over the range of the instrument) that includes a standard at the lowest calibration point listed in the method (i.e., the ML).

In response to a need to establish a compliance evaluation threshold when the water quality-based permit limit is below the detection limit of the most sensitive analytical method published at 40 CFR part 136, EPA refined the definition of the ML in 1994 as 10 times the same standard deviation used to calculate the MDL¹. Because the MDL is commonly determined as 3.14 times the standard deviation of seven replicate measurements, the ML was commonly calculated as 3.18 times the MDL. (The figure of 3.18 was derived by dividing 10 by 3.14; if more than 7 replicates were used to determine the MDL,

¹The refined definition of the ML first appeared in EPA's 1994 draft *National Guidance for the Permitting, Monitoring, and Enforcement of Water Quality-based Effluent Limitations Set Below Analytical Detection/Quantitation Levels*. The draft guidance was very controversial and never finalized. However, the refined definition of the ML has remained in use on newer analytical methods.

both the MDL and the ML multipliers are adjusted accordingly, based on values from the *t*-distribution.) This calculation makes the ML analogous to Currie's quantification limit and the American Chemical Society's limit of quantitation (LOQ), which is defined as ten times the standard deviation of replicate or low concentration measurements (MacDougall, *et al.*, 1980 and Keith, *et al.*, 1983).

To simplify implementation of the ML, the definition also was expanded to state that the calculated ML is rounded to the whole number nearest to (1, 2, or 5), times 10^n , where *n* is an integer. The reason for this simplification is that calibration of an analytical system at some exact number (e.g., 6.27) is difficult and prone to error, whereas rounding to the whole number nearest to (1, 2, or 5) $\times 10^n$ provides a practicable value. The most recent definition of the ML is "*the lowest level at which the entire analytical system must give a recognizable signal and acceptable calibration point for the analyte. It is equivalent to the concentration of the lowest calibration standard, assuming that all method-specified sample weights, volumes, and cleanup procedures have been employed. The ML is calculated by multiplying the MDL by 3.18 and rounding the result to the number nearest to (1, 2, or 5) $\times 10^n$, where *n* is an integer,*" and this definition was contained in the version of EPA Method 1631 that was promulgated on June 8, 1999 (64 FR 30417) (see Section 17.8 of EPA Method 1631 Revision B).

The ML will generally be somewhat lower than Currie's quantitation limit, even when similar sample sizes and estimation procedures are used. This is because the standard deviation used to calculate the ML will generally be smaller than the standard deviation at the lowest concentration at which the relative standard deviation is 10%. This is due to the fact that, in almost all cases, standard deviation is non-decreasing with increasing concentration, e.g., it generally tends to increase as concentration increases.

Although the ML has been used successfully in EPA methods for more than 20 years, some members of the regulated industry and others have claimed that the ML is less than an ideal concept for quantitation because it:

- Does not account for interlaboratory variability, and
- Is based on a multiple of the standard deviation rather than a fitted model

These concerns are discussed later in this document.

2.3 Approaches Advanced by Other Organizations

To expand somewhat on Currie (1968), standardizing the operational definitions of detection and quantitation would benefit society by making it easier to compare and select measurement methods based on low-level measurement capability and requirements in particular applications. Unfortunately, in spite of agreement on general principles and definitions advanced by Currie and his supporters, consensus on procedures that would result in comparable detection and quantitation estimates has been elusive. Sections 2.3.1 - 2.3.3, which are by no means an exhaustive list of the various approaches advanced to date, highlight approaches that have been most widely advanced for environmental applications.

2.3.1 EPA Approaches

Over the years, a number of detection and quantitation limit approaches have been developed, suggested, or used by EPA among the various organizations charged with responding to differing program mandates. In part, this situation reflects actual differences in the mandates, and in part, it reflects the fact that no concept advanced to date has emerged as a clear 'winner' that meets all needs for all people. Approaches that have been used or suggested by EPA include the:

- MDL and ML (described in Sections 2.2.1 and 2.2.2)
- Instrument detection limit (IDL)
- Practical quantitation limit (PQL)
- Estimate quantitation limit (EQL)
- Contract-required detection limit (CRDL) and contract-required quantitation limit (CRQL)

Instrument Detection Limit: EPA methods for analysis of metals have historically included an instrument detection limit, or IDL. Functionally, the IDL is similar to the MDL except that the IDL includes temporal variability (it is determined on 3 non-consecutive days) and does not include all sample processing steps (the IDL characterizes the detection capabilities of the instrument as opposed to the method). Because IDLs do not reflect the entire measurement process and, for the most part, have been used only for measurement of metals, EPA did not consider the IDL as a potential alternate to the MDL when conducting the assessment described in this Assessment Document.

Practical Quantitation Limit: The practical quantitation limit, or PQL, was established in the 1980s by EPA's drinking water program as the lowest concentration at which reliable measurements can be made. The PQL is defined as "the lowest concentration of an analyte that can be reliably measured within specified limits of precision and accuracy during routine laboratory operation conditions" (52 FR 25690, July 8, 1987). The PQL is a means of integrating information on the performance of approved analytical methods into the development of a drinking water regulation. The PQL incorporates the following:

- Quantitation,
- Precision and bias,
- Normal operations of a laboratory, and
- The programmatic need to have a sufficient number of laboratories available to conduct compliance monitoring analyses of drinking water samples.

EPA uses two main approaches to determine a PQL for an analyte under the Safe Drinking Water Act (SDWA). One approach is to use the data from Water Supply (WS) studies (e.g., laboratory performance evaluation studies conducted by the Agency as part of the certification process for drinking water laboratories). The PQL is established at the concentration at which at least 75% of the laboratories in the study, or the subset representing EPA Regional laboratories and state laboratories, obtain results within some predetermined percentage of the true value of the test samples (e.g., $\pm 30\%$). This approach is used in most cases when WS data are available to calculate a PQL. The WS data approach was used to determine the PQLs for Phase V inorganic chemicals such as antimony, beryllium, cyanide, nickel and thallium (July 17, 1992; 57 FR 31776), as well as many other contaminants regulated under the SDWA.

In the absence of WS data, the second approach that EPA uses is the multiplier method. In this approach, the PQL is calculated by multiplying the EPA-derived MDL by a factor between 5 and 10. The exact multiplier varies and sometimes depends on the degree of concern about the specific contaminant (i.e., based on a human health risk assessment for consumption of drinking water).

Application of the PQL has been traditionally limited to drinking water. Furthermore, the PQL may not be related to the lowest quantitation limit because 1) the PQL is associated with the analyte and may have been determined irrespective of a specific analytical method (e.g., using data from a variety of methods approved for that analyte at 40 CFR part 141), 2) the performance evaluation (PE) samples from which it is derived contain pollutant concentrations that may be well above the true limit of quantitation, 3) the multiplier used to calculate a PQL when PE data are not available is somewhat dependent on concerns about risks from human exposure to contaminants in drinking water, and 4) the resulting PQLs may be too high for purposes other than the Safe Drinking Water Act (e.g., other EPA programs). In addition, because EPA has privatized the performance evaluation program for drinking water laboratory

certification, it is not yet clear that appropriate data will be available in the future. Based on these facts, EPA did not conduct an assessment of the PQL for CWA applications.

In the late 1980s, EPA's Office of Solid Waste (OSW) adopted a different version of the PQL as a quantitation limit. No procedure for establishing the limits was given; instead values were extrapolated from the Contract Laboratory Program CRQLs (see below). Since 1994, OSW has actively removed the term "PQL" from its revised methods, replacing it with the term "estimated quantitation limit" (EQL). The term PQL and the original numerical values remain in a few older OSW guidance documents.

Estimated Quantitation Limit: EPA's Office of Solid Waste has defined the EQL as:

"The lowest concentration that can be reliably achieved within specified limits of precision and accuracy during routine laboratory operating conditions. The EQL is generally 5 to 10 times the MDL. However, it may be nominally chosen within these guidelines to simplify data reporting. For many analytes the EQL analyte concentration is selected as the lowest non-zero standard in the calibration curve. Sample EQLs are highly matrix dependent. The EQLs in SW-846 are provided for guidance and may not always be achievable." (see SW-846, Chapter 1).

As noted in most newer SW-846 methods, the EQLs are provided for guidance and may not always be achievable. Because the EQL is not rigorously defined and is guidance, because the EQL may be based on the MDL, and because the EQL can be the lowest calibration point and would, therefore, overlap the ML, EPA did not consider the EQL further in its assessment of detection and quantitation approaches.

Contract-Required Detection and Quantitation Limits: EPA's Superfund program has adopted the use of contractually-required limits that are based on consensus among analytical chemists about levels that can realistically be achieved in commercial laboratories using a contractually-specified method. Laboratories that participate in the Superfund Contract Laboratory Program (CLP) are required to demonstrate that they can achieve the specified CRDLs and CRQLs. The CRDLs are consensus values that apply to the analyses of metals using CLP methods. The CRQLs apply to organic analytes and are based on the concentration of the lowest non-zero calibration standard specified in the CLP methods, in a fashion analogous to the original derivation of the ML. Because few CWA applications involve the use of the CLP methods, EPA did not consider the CRDL or the CRQL as viable alternatives to the MDL and ML when conducting the assessment described in this document.

2.3.2 Industry-supported Approaches

The regulated community has demonstrated an interest in detection limit approaches since EPA first promulgated the MDL and ML for use in CWA programs in 1984 (49 FR 43234). As part of that rule, EPA promulgated Methods 601 through 613, 624, 625, 1624, and 1625 for organic compounds at 40 CFR part 136, Appendix A and EPA Method 200.7 for metals by inductively coupled plasma spectrometry (ICP) at 40 CFR part 136, Appendix C. EPA also promulgated the MDL procedure at 40 CFR part 136, Appendix B. The Virginia Electric Power Company (VEPCO) brought suit against EPA, challenging the Agency's use of the MDL in the promulgated methods. In a settlement, EPA agreed that the MDL would be applicable only to the 600-series organic methods, as these methods already contained MDL values; i.e., it would not be applicable to EPA Method 200.7. The settlement agreement did not preclude future use of the MDL by EPA or the right of VEPCO to bring suit in such future use.

After the VEPCO settlement, the regulated community, mainly through efforts of the Electric Power Research Institute (EPRI), remained involved in detection and quantitation approaches to be used under EPA's CWA programs. The first approaches that industry advanced were the compliance

monitoring detection level (CMDL) and compliance monitoring quantitation level (CMQL) (Maddalone, *et al.*, 1993). The CMDL/CMQL were variants of EPA's MDL/ML that attempted to adjust for interlaboratory variability.

The regulated community continued its efforts to develop alternative detection and quantitation approaches with development of the "alternate minimum level" (AML) in the mid-1990s (Gibbons *et al.*, 1997). The AML is based on statistical modeling of standard deviation versus concentration, which requires large amounts of data.

Most recently, the regulated community has funded development of the interlaboratory detection estimate (IDE) and interlaboratory quantitation estimate (IQE). The IDE and IQE have been balloted and approved by ASTM's Committee D-19 for water as Standard Practices D-6091 and D-6512, respectively. These approaches take into account nearly all sources of variability to arrive at detection and quantitation limits that are higher, on average, than the limits produced by other approaches (see Appendix C of this Assessment Document). Because the regulated community has shifted support from the CMDL/CMQL and the AML to the IDE and IQE, and because EPA is not aware of other organizations that currently advocate the earlier approaches, EPA did not consider industry approaches other than the IDE/IQE in its assessment of possible alternatives to the MDL and ML.

As with all other approaches advocated to date, the IDE and IQE have fallen short of being ideal approaches for detection and quantitation for all organizations and applications. To date, EPA is not aware of a demonstrated implementation of the IDE or IQE in the development of an analytical method. Specific concerns that have been raised about the IDE and IQE are that:

- They contain an allowance for false negatives that may be inappropriate,
- The IDE and IQE are based on the use of prediction and/or tolerance intervals, which may be inappropriate,
- The IDE and IQE require a large amount of data in order to be able to model variability versus concentration, including data generated in multiple laboratories, and
- The complex statistical procedures involved in calculating an IDE and IQE would place a heavy burden on the analytical chemists that typically develop, modify, and use methods.

These concerns are discussed in detail later in this document.

2.3.3 Approaches Advocated by the Laboratory Community and Voluntary Consensus Standards Bodies

In 1980 (MacDougall *et al.*, 1980) and 1983 (Keith *et al.*, 1983), the American Chemical Society's Committee on Environmental Improvement (CEI) advanced approaches for the Limit of Detection (LOD) and Limit of Quantitation (LOQ). The ACS LOD is defined as the lowest concentration level that can be determined to be statistically different from a blank. The recommended value for the LOD is three times the standard deviation of replicate measurements of a blank or low-level sample. The LOD is roughly equivalent to the MDL in numerical terms and conceptually equivalent to Currie's critical value.

The ACS LOQ is defined as the level above which quantitative results may be obtained with a specified degree of confidence. The recommended value for the LOQ is 10 times the standard deviation of replicate measurements of blanks or low-level samples. Because the LOD and LOQ are still used by the analytical community, they have been included in EPA's reassessment of detection and quantitation approaches.

In the mid-1980s, the ACS CEI introduced the concept of the Reliable Detection Limit (RDL) and the Reliable Quantitation Limit (RQL). The RDL and RQL were attempts at simplification of the LOD and LOQ. Both the RDL and the RQL involved applying a multiplier to the standard deviation derived from replicate measurements of a low-level sample. Neither concept received acceptance by the analytical community. Because the RDL and RQL are no longer being advanced by ACS, they were not considered for evaluation in EPA's assessment of detection and quantitation approaches.

In 1999 (Currie, 1999a and 1999b), IUPAC and ISO reached substantial agreement on the terminology and approaches documented by Currie (1995), although "subtle differences in perspective" of the organizations remain (Currie, 2000). IUPAC and ISO have not, to date, published methods that include limits reflecting these standards. Similarly, although ASTM International adopted the IDE in 1997 and the IQE in 2000, ASTM International has not included any IDE or IQE values in methods approved through the ASTM ballot process. On the other hand, ISO and ASTM International have published methods that employ the MDL. Because IUPAC and ISO have approved the critical value, detection limit, and quantification limit, and because ASTM International has approved through ballot the IDE and IQE, EPA has included these approaches in its assessment of detection and quantitation approaches.

At the ACS Annual Meeting held in August, 2002, CEI members discussed the issue of detection and quantitation, with the objective of determining if the LOD and LOQ approaches should be re-visited. At that meeting, several members suggested that the committee consider adopting a sample-specific detection limit approach in which the ratio of instrument signal to background noise is used to estimate a detection limit for each analyte in each sample analyzed. EPA did not include the signal-to-noise ratio concept in this assessment because its application is limited to specific types of measurement techniques, such as gas chromatography/mass spectrometry. Limitations of this concept for use in general environmental chemistry are best illustrated by the fact that it would not apply to any of the techniques traditionally used to determine the "conventional pollutants" cited in the Clean Water Act (the only pollutants cited by name in the Act), i.e., biochemical oxygen demand (BOD), total suspended solids (TSS), fecal coliforms, and pH.

2.3.4 Approaches Advocated by Other U. S. Government Agencies and Other Governments

Within the U.S., EPA found that other Federal agencies tend to rely on the detection and quantitation limit approaches described above or on variants of those procedures. For example, the USGS National Water Quality Laboratory (NWQL) began using the EPA MDL procedure in 1992. USGS has since developed a variant of the MDL called the long-term MDL (LT-MDL) that employs at least 24 spiked samples prepared and analyzed by multiple analysts on multiple instruments over a 6- to 12-month period. (Unlike EPA programs that rely on hundreds of commercial, Federal, state, and local laboratories for sample analysis, nearly all samples analyzed for USGS programs are analyzed by the USGS National Water Quality Laboratory.) As described by USGS, the long-term MDL is based on many of the same fundamental assumptions as the MDL, namely:

1. Normal data distribution,
2. Constant standard deviation, and
3. Best-case detection condition (because LT-MDLs typically are determined by spiking the analyte in a clean matrix, e.g., reagent water).

The primary differences between the EPA MDL and the USGS LT-MDL are the longer time period and mixing of instruments and analysts (Oblinger Childress, *et al.*, 1999). Because the MDL and LT-MDL approaches otherwise are so similar, EPA did not evaluate the long-term MDL approach in this

assessment. Instead, EPA considered the underlying differences between the two approaches (namely the effects of temporal, instrument, and analyst variability) in its assessment of issues (see Chapter 3).

Outside the U.S., EPA found that the European Union (EU) relies on the terminology and conventions developed by Currie, IUPAC, and others (Eurachem, 2000). The EU advocates reporting all results along with an estimate of the uncertainty associated with each value. In its discussion of the issue, the EU indicates that use of the term 'limit of detection' only implies a level at which detection becomes problematic and is not associated with any specific definition. Instead, the EU focuses its attention on ways to estimate uncertainty, basing its approach on the ISO *Guide to the Expression of Uncertainty in Measurement* (1993). However, the EU also notes that the use of uncertainty estimates in compliance statements and the expression and use of uncertainty at low levels may require additional guidance. The United Kingdom's Valid Analytical Measurement Programme (VAM) has adopted a similar approach that is based on both the ISO and the Eurachem guidance (Barwick and Ellison, 2000). Because these approaches are focused on estimating uncertainty rather than at establishing or defining limits for detection and quantitation, EPA did not consider the European approaches in this assessment.

Chapter 3

Issues Pertaining to Detection and Quantitation

As part of the Settlement Agreement concerning EPA's reassessment of detection and quantitation limit approaches, the Agency agreed to consider several specific issues pertaining to these approaches. These issues included:

- Criteria for selection and appropriate use of statistical models,
- Methodology for parameter estimation,
- Statistical tolerance and prediction,
- Criteria for design of detection and quantitation studies, including selection of concentration levels ("spiking levels"),
- Interlaboratory variability, and
- Incorporation of elements of probability design.

In developing its plan for conducting this assessment, EPA identified a number of other issues that should be considered. These issues include:

- Concepts of the lower limit of measurement,
- The need for approaches that can support CWA programs, including:
 - method performance verification at a laboratory,
 - method development and promulgation,
 - National Pollutant Discharge Elimination System (NPDES) applications,
 - non-regulatory studies and monitoring,
 - descriptive versus prescriptive uses of lower limits to measurement, and
 - use of a pair of related detection and quantitation procedures in all OW applications
- Censoring of measurement results,
- Sources of variability (including, but not limited to interlaboratory variability),
- False positives and false negatives,
- Measurement quality over the life of a method,
- Matrix effects,
- Background contamination,
- Outliers,
- Instrument non-response,
- Accepting the procedures of voluntary consensus standards bodies (VCSBs),
- National versus local standards for measurement,
- Ease of use (i.e., ability of study managers, bench chemists, and statisticians to do what is required by a detection or quantitation limit procedure),
- Cost to implement the procedures, and
- Laboratory-specific applications.

Approaches to establishing the lower limits of measurement were discussed in Chapter 2. For clarity and brevity, EPA has organized the remaining issues into three subsections that follow. Section 3.1 discusses the issues that are primarily driven by analytical chemistry concerns, Section 3.2 discusses the issues that are primarily driven by CWA regulatory considerations, and Section 3.3 discusses issues that are primarily driven by statistical concerns. Table 3-1, at the end of this chapter, provides a summary of each issue discussed in Sections 3.1 - 3.3.

3.1 Analytical Chemistry Approaches to Detection and Quantitation

3.1.1 Blank versus Zero Concentration

Analytical chemists rarely, if ever, say that a sample contains zero concentration of a substance of interest. Even when the sample is created in a laboratory for the purpose of containing as little substance of interest as possible (a blank), analytical chemists recognize the possible contribution of the blank to the final measurement result. The ability of a laboratory to reduce the concentration of a substance in the blank is often the limiting factor in attempts to make measurements at ever lower levels.

A classic example of the potential problem is illustrated by the seminal works of Patterson in the late 1960s and 1970s (e.g., Patterson and Settle, 1976). Patterson demonstrated that the majority of concentrations of lead reported in the literature for such diverse matrices as urban dust, open ocean waters, and biological tissues were in error by several orders of magnitude. The source of the "gross positive errors" was contamination introduced during sample collection, handling, and analysis. Interlaboratory studies of the day designed to determine consensus values for reference materials were, in fact, determining the consensus values for background contamination across laboratories. Patterson recognized the value in running blank samples (samples thought not to contain the substance of interest) to demonstrate that the sample collection, handling, and analysis processes were not introducing contamination. Patterson subsequently developed the techniques for "evaluating and controlling the extent and sources of industrial lead contamination introduced during sample collecting, handling, and analysis" that form the basis of the "clean techniques" used for metals analysis today, and that are incorporated in EPA Method 1631 among others.

The most common analytes for which contamination problems are encountered in environmental measurements are metals, primarily zinc because of its ubiquity in the environment. On the other hand, it is rare to find contamination in the measurement of organic compounds, except for methylene chloride, acetone, and a few other volatile organic compounds used as solvents in analytical laboratories. Therefore, for determination of metals, a blank is usually included or compensated in the calibration whereas, for organics, except for the solvents, the concentration in the blank is generally assumed to be zero and there is no compensation of the calibration.

Measurement methods designed to determine substances at very low concentrations may include requirements for the preparation and analysis of a variety of blanks that are designed to identify the extent and the sources of contamination. Analysts understand that "blank" does not mean zero concentration, but that through careful control and evaluation, it is possible make measurements for which the blank contribution is sufficiently small to be considered negligible.

Useful detection and quantitation limit approaches should address the potential contribution of the blank through both the design of the study that generates the detection and quantitation limit estimates and the evaluation of the study results.

3.1.2 Lack of Instrument Response

Instruments do not always produce a result from an appropriately prepared sample. Sometimes this is attributable to uncontrollable instrument limitations, sometimes it is attributable to controllable instrument settings (thresholds) established by the manufacturer or the laboratory, and sometimes it occurs randomly. As an example, gas chromatograph/mass spectrometer (GC/MS) instruments often contain thresholds below which no instrument signal is reported. With no instrument signal reported, no measurement result can be reported, and the instrument will report zero to indicate the lack of a signal.

To understand how instrument thresholds are used, it may be helpful to think of background static heard on a citizen-band (CB) radio or a walkie-talkie. The static is present, but it has no meaning. Turning the "squelch" knob to the point at which the static is filtered out also may make it impossible to hear the caller. In the context of detection, increasing the instrument threshold may cause the instrument to miss a substance of interest at a low level.

In 1997, EPA conducted study of 82 semivolatile (acid and base/neutral) organic compounds measured by EPA Method 1625 in order to observe the performance of a GC/MS instrument both with and without an instrument threshold (see Chapter 1, Section 1.3.2.3). In the study, solutions at up to 17 concentration levels were analyzed with the threshold on (i.e., low level signals are automatically suppressed) and with the threshold off (i.e., there is no suppression of signals). Samples were analyzed at decreasing concentrations, including a blank, with triplicate determinations at each concentration. For measurement results obtained with the threshold turned on, all of the measurements made on the blank were reported as zero. This was not a surprising result, given the purpose of the instrument threshold. For measurements obtained without the threshold, 27 of 230 measurements on the blank (11%) were reported as 0.000 ng/mL and no measurement results were reported lower. Instrument non-response at a low concentration has both direct and indirect impacts on estimating detection and quantitation limits.

The main direct impact of non-response at low concentrations is that it is not possible to estimate the standard deviation of measurements at zero concentration. By definition, however, this standard deviation is required to calculate the Currie critical value. The EPA MDL procedure was constructed to deal with this problem by providing for a means estimation of a standard deviation at a low concentration. The MDL procedure includes step-by-step instructions for determination of a concentration as close to zero as is possible that will generate a measurement.

In order to meet the requirements of the MDL definition, it is necessary to find the concentration at which the measurement method ceases to generate measurement results, and many laboratories have run repeat measurements in order to find this concentration. This problem manifested itself in EPA's variability versus concentration (Episode 6000) studies. The 40 CFR part 136, Appendix B procedure suggests iteration until the calculated MDL is within a factor of 5 of the spike level. For the Episode 6000 studies, EPA instructed laboratories to use a factor of 3 instead of 5 in an attempt to more narrowly define the lowest spike level at which measurements could be made.

This change to a factor of 3 also was suggested by one of the peer reviewers charged with evaluating EPA's assessment of detection and quantitation limits, who noted:

"However, the use of as much as five times the critical level for the spike concentrations could be problematic. The inflation of the MDL by using a spike at the critical level is only 25% for a method with a high-level CV of 20% (this and other calculations here are done with the Rocke and Lorenzato 1995 variance function assuming a sample size of 7). A spike concentration of 3 times the critical level inflates the MDL to a value 140% higher, which even there may be tolerable. Use of a value 5 times the critical level gives an inflation of over 280%. ..."

Following some theoretical example calculations that are not reproduced here, the reviewer's comment continues with:

"Thus, I would recommend that the procedure be altered to use concentrations that are no more than 3 times the detection limit, and perhaps to permit concentrations lower than the critical level, including possibly blanks" (Rocke, 2002).

The reviewer's calculations suggest that MDL may be strongly inflated for a spike level of 5 times the MDL, but only moderately inflated at a spike level of 3 times the MDL. However, during the Episode 6000 studies, several laboratories asked for relief from the requirement, and EPA relented after learning of the difficulties in attempting to achieve the factor of 3. If the reviewer's example calculations are correct and a practical procedure for determining the MDL using the factor of 3 could be implemented, it could exacerbate the concern from the regulated community that MDL values are too low.

Given the competing theoretical and practical considerations, one conclusion that can be drawn is that detection limits are somewhat variable and not easy to define. Further details are in the results of the studies given in Appendices B and C to this Assessment document.

In summary, both Currie's approach and EPA's approach have theoretical problems with addressing instrument non-response. Any operational approach to detection or quantitation should take this issue into account.

3.1.3 Matrix Effects

"Sample matrix" is a term used to describe all of the substances, other than the substance(s) of interest, present in an environmental sample. In the case of a wastewater sample, this would include the water itself, as well as any other dissolved or suspended materials. For any given measurement, some of the substances may interfere with the measurement, while others may be substances that have no effect on the measurement. Interferences in the sample may act either positively (i.e., increasing the measured result), negatively (i.e., decreasing the measured result), or even preventing the measurement from being made.

"Matrix effect" is a term used to describe a situation in which a substance or combination of substances in the sample (other than the substance[s] of interest) influence the results of the measurement. Positive interferences may inflate the results for the substance or make it difficult to distinguish one substance from another. However, unless the positive bias is consistent and predictable, the measurement result may be unreliable. Negative interferences may suppress the results for the substance to the point that the results cannot be distinguished from background instrument noise.

In some cases, finding a matrix effect indicates that the analyst should select a more appropriate method. For example, a colorimetric method for the measurement of sulfide may be a poor choice for the analysis of a sample that is very cloudy or darkly colored. In other cases, characteristics of the sample such as its pH may destroy the substance of interest, effectively preventing analysis for that substance.

Nearly all of the newer analytical methods approved at 40 CFR part 136 describe the preparation and analysis of quality control samples that are designed to indicate the presence of matrix effects (e.g., matrix spike and/or matrix spike duplicate samples). Many of these methods also contain techniques for addressing matrix effects. Further, EPA has developed guidance documents that amplify the discussions in those methods (e.g., *Guidance on Evaluation, Resolution, and Documentation of Analytical Problems Associated with Compliance Monitoring*, June 1993, EPA 821-B-93-001). For determination of mercury by EPA Method 1631 that is the subject of the Settlement Agreement, additional guidance on resolving matrix interferences to achieve specified detection and quantitation limits is provided in EPA's *Guidance for Implementation and Use of EPA Method 1631 for the Determination of Low-Level Mercury* (March 2001, EPA 821-R-01-023). Following the techniques in the methods and guidance will usually reduce adverse effects of the sample matrix on detection/quantitation limits and measurement results.

3.1.3.1 Allowance for Matrix Effects in Detection and Quantitation Limits

There are those who believe that detection and quantitation limits should be determined in "real-world" matrices, rather than in reference matrices intended to simulate method performance in a particular matrix type. Problems with such an approach, however, are that:

- Many "real-world" matrices contain the target pollutant at levels well above the detection or quantitation limit, making it impossible to characterize what can and cannot be detected at low levels. Diluting the sample to dilute the target pollutant concentration is an option. However, this also has the potential to dilute any interferences that might be present, thereby defeating the purpose of using the real-world matrix.
- It is not possible to anticipate and obtain samples of every possible matrix on which a method might be used when the method is being developed and detection/quantitation limits are being established.
- Although use of a reference matrix to establish detection and quantitation limits allows the results to be reproduced (i.e., confirmed) by an independent party, such a confirmation may not be possible with many real world matrices that may be subject to seasonal, diurnal, or other types of variability.
- The cost of determining detection and quantitation limits in every possible matrix would be prohibitive.

Given these difficulties, EPA believes that a reference matrix or reference matrices should be used to establish method detection and quantitation limits, but that the procedures for defining these limits should allow for evaluation of data collected in particular matrices of concern. EPA also believes that such matrix-specific determinations should only be used when all efforts to resolve matrix interferences have been exhausted.

3.1.3.2 Repository of Reference Matrices

Two of the four peer reviewers charged with evaluating EPA's assessment of detection and quantitation limit approaches suggested that EPA create a repository of reference matrices, similar to those developed by NIST, and that these reference matrices be used to challenge a test method and to establish detection and quantitation limits (Cooke, 2002 and Wait, 2002). EPA has considered such a repository from time to time and again in response to this suggestion, but has been unable to resolve all of the issues surrounding such a repository. Some of these issues are:

- The stability of aqueous samples,
- The holding times necessary to assure stability,
- The argument that no matrix from a given industrial discharge in industrial category or subcategory reflects the characteristics of another discharge in that or other industrial categories or subcategories,
- The cost of maintaining such a repository, and
- The potential conflict with NIST and with non-governmental organizations that provide reference matrices.

Given these issues, EPA believes that the development and maintenance of standard reference materials (SRMs) and certified reference materials (CRMs) are best left to NIST and the commercial marketplace. EPA agrees that such reference materials are a useful means of challenging a test method and has suggested in recent methods that reference matrices be analyzed, when available, as an additional QC measure. For example, when EPA developed an appendix to Method 1631 for application to matrices other than water, EPA specified use of a quality control sample (QCS) with the statement that "many certified reference materials (CRMs) are available for total mercury in plants, animals, fish, sediments, soils, and sludge" and the requirement that "recovery and precision for at least one QCS per batch of samples must meet the performance specifications provided by the supplier."

Although EPA agrees that SRMs and CRMs could be useful in establishing detection and quantitation limits, EPA believes that practical considerations are likely to preclude their use for this purpose in most situations. This is because the materials would need to contain the analytes of interest at levels that are near the detection limit (e.g., within 1 to 5 times the concentration of a determined MDL). Such concentrations are unlikely to occur in an SRM produced by NIST or a CRM produced by a vendor, and diluting the CRM/SRM would diminish matrix effects, as indicated in Section 3.1.3.1.

As an alternative to using standard reference materials, EPA commonly tests its analytical methods on a variety of real-world matrices, and allows for this variability in the QC acceptance criteria for the matrix spike (MS) and matrix spike duplicate (MSD) samples. For example, EPA published performance data in Table 3 of EPA Method 1631B for reagent water, fresh water, unfiltered and filtered marine water, and unfiltered and filtered secondary effluent, and allowed for the variability among these matrices in the QC acceptance criteria for the MS/MSD in the method. ASTM Committee D 19 allows this approach in development of QC acceptance criteria for methods (see Section 6.5.1.1 of ASTM D 5847: *Standard Practice for Writing Quality Control Specifications for Standard Test Methods for Water Analysis*.)

3.1.4 Recovery Correction

This section addresses correction for recovery in detection and quantitation limits. The purpose of a recovery correction is to adjust a measured concentration in a sample for the amount by which the measured concentration differs from the true concentration (if known).

To illustrate the potential need for recovery correction, consider the case of certain compounds such as organic bases (e.g., benzidine) and acids (e.g., phenols) that are either not totally (100%) recovered in the extraction process, or are adsorbed on the surface of a GC column at low (nanogram) levels. As a result, the measured concentration of such compounds is always less than the true concentration in the water sample. These incomplete recoveries have led some developers of detection and quantitation limit approaches to believe that these limits should be recovery corrected (i.e., that the detection or quantitation limit should be adjusted inversely proportional to the recovery). For example, if an analyte is recovered at 50%, the detection and/or quantitation limit should be doubled. EPA believes that recovery correction may be appropriate if (1) the recovery is consistent across laboratories, matrices, and conditions, and (2) the relative variability (as relative standard deviation) remains constant as the recovery decreases. These two requirements are rarely met; therefore, recovery correction would be appropriate only in rare circumstances.

The first requirement (consistent recovery) would need to be tested under a variety of conditions because, if the recovery varies among laboratories, matrices, and analytical conditions, then a detection and/or quantitation limit would need to be developed for each of these conditions. EPA's experience is that poor recovery is rarely consistent; i.e., if one laboratory measures a recovery of 40%, another laboratory may measure 20%, or 60%, but not exactly 40%.

The normal condition in environmental analytical measurements is that the variability (as standard deviation) remains approximately constant as the recovery decreases (i.e., the relative precision [as RSD] is poorer at low recovery). For example, if the RSD is 10% at 100% recovery, the RSD may be 50% at 50% recovery, and may be 100% at 10% recovery. This increase in relative variability is not the result of measurements being made at lower levels, as is the normal case, but as a result of variability in the extraction (partitioning) process. For examples of the effect of poor recovery on precision, see the quality control (QC) acceptance criteria for the semivolatiles organic compounds in Table 8 of EPA Method 1625 (see 40 CFR part 136, Appendix A). Because nearly all detection and quantitation limits

are based on precision (as standard deviation), including a recovery correction for an analyte with poor precision at low recovery is, in effect, a double counting for poor precision.

A third concern with the issue of recovery correction is "where does it stop?" It makes little sense to recovery correct those measurements made in the region of detection and quantitation if similar recovery-correction steps are not employed for measurements at higher concentrations. EPA has traditionally viewed recovery correction with great caution, and has preferred to require that laboratories analyze quality control samples to demonstrate that analytes are recovered within an acceptable level. For example, EPA's Office of Water methods require that laboratories prepare and analyze both a reference matrix and a sample matrix that have been spiked with the analytes of interest, and that these analytes be recovered within method-specified acceptance criteria. If the recovery criteria are met, then samples analyzed in the batch are considered to be reliable within the overall level of error associated with the method, and results are reported without correcting for the recovery. EPA believes that it would be inconsistent to correct for recovery in measurements made at the detection or quantitation limit, if such corrections are not made to results obtained at higher concentrations (e.g., during the routine analysis of samples).

EPA acknowledges that recovery-correction techniques are employed in some Agency methods. Most notably are those methods that employ isotope dilution techniques, in which a stable, isotopically labeled analog of each target analyte is spiked into each sample. Because of their structural similarity to the analytes of interest, the labeled analogs are assumed to behave exactly like their unlabeled analogs (the target analytes). Because the recovery of the labeled analog will be similar to that of the target analyte, the technique allows for recovery correction of each target analyte and is particularly useful in highly complex matrices. In these methods, recovery correction techniques are specified as part of the procedures for calculating and reporting results and are dependent on the one-to-one relationship of the target analyte and the labeled analog. Inclusion of a further procedure for recovery-correction in a detection and quantitation limit approach could result in double-counting of bias.

Few of the "traditional" approaches to establishing detection and quantitation limits include procedures for recovery correction. For example, the issue was not addressed by Currie in his original proposal of a critical value or quantitation limit. Similarly, neither EPA's MDL and ML nor the American Chemical Society's LOD and LOQ, all of which are based on the approaches advanced by Currie, include a mechanism for recovery correction. When Currie introduced his critical value, he defined it as "the minimum significant value of an estimated net signal or concentration, applied as a discriminator against background noise" (Currie, 1995). Because the critical value is defined as a *measured* concentration rather than a *true* concentration, a recovery correction is not included.

The use of recovery correction has been included in several of the most recently developed approaches for detection and quantitation. For example, the minimum detectable value (MDV) recently adopted by ISO and IUPAC, and the interlaboratory detection estimate (IDE) and interlaboratory quantitation estimate (IQE) adopted by ASTM include procedures for recovery correction. The IQE also contains a further correction that we have termed a "bias" correction.

In the MDV approach, recovery is treated as a linear function versus concentration, and an extrapolation is used to estimate the recovery at zero concentration. EPA has found that this projection of the regression line to zero concentration can lead to errors because, depending on the intercept (in concentration units), the recovery at zero concentration can be positive, zero, or negative, resulting in an inflated MDV, an MDV very close to zero, or a negative MDV. For further details, see the section titled "Negative detection limits for the ISO/IUPAC MDV" in Appendix C to this Assessment Document, and the data in Table 2 of that appendix.

The IDE and IQE fit recovery versus concentration in a way analogous to the fitting in the MDV. The difference between the treatment of recovery in the MDV and the IDE/IQE is that an unweighted model is used in the MDV, whereas the linear model in the IDE and IQE is weighted as determined by the model of standard deviation versus concentration that is used in calculating the IDE and IQE. (If this model is the constant model, the weighting is the same as for the MDV.)

The IQE, but not the IDE, includes an additional correction for the bias associated with an estimate of the true standard deviation at each concentration as compared to the measured standard deviation at each concentration. In this context (a "bias" correction to the IQE), the word "bias" means the amount by which the estimated sample standard deviation differs from the true population standard deviation, and should not be confused with common use of the word "bias" in analytical chemistry measurements (the deviation of a result from the true value, usually expressed as percent).

The effect of these corrections on detection and quantitation limits was calculated using data generated in EPA's Multi-technique Variability Study (the "Episode 6000 Study"). Details of these effects are discussed in Appendix C.

3.1.5 Measurement Quality over the Life of a Method

We have all heard the expression "Practice makes perfect." Although there is no such thing as a "perfect" measurement, the idea that results get better with practice applies to the quality of measurements made with a given method over time. We can demonstrate it using simple techniques like laboratory control charts. The improvements are a result of experience, as well as improvements in equipment over time. EPA expects changes in performance when new staff are trained. For this reason, many EPA methods specify that "start up tests" be repeated each time new staff arrive. It is not unusual to see slight increases in measurement variability as new staff are trained. However, when new staff become as good as the existing staff, control charts should show it.

The use of quality control (QC) charts as a means of tracking method and laboratory improvement as a function of time is described in EPA's *Handbook for Analytical Quality Control in Water and Wastewater Laboratories* (referenced in the 40 CFR part 136, Appendix A methods). Although these charts are instructive in tracking improvement, they have two significant drawbacks: (1) they do not establish an absolute limit within which an analysis must be operated and (2) continued improvement can lead to unusually stringent limits that, eventually, will not be met. As long as absolute QC acceptance criteria (limits), such as those found in EPA methods, are established for the determination, and as long as there is a recognition that stringent limits may be an artifact of improvement beyond what is routinely achievable, QC charts can be instructive in identifying statistically significant losses of, or improvements in, analyte responses in the region of interest. ASTM Committee D 19 adopted the philosophy of establishing absolute limits for analytical methods in approving Standard Practice D 5847.

As with most other areas of technology, measurement instruments continue to improve. Instrument manufacturers and laboratories are increasing data processing power, speed of analysis, and the reduction of chemical or electronic "noise." Any of these instrument improvements can be expected to improve the measurement method in determining the concentrations of environmental pollutants. This process can be illustrated for a variety of EPA methods. A case in point is EPA Method 1613 for determination of polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzofurans. Development of this method began in 1988. At the time, high resolution mass spectrometer systems that were commercially available were able to achieve a detection limit of approximately 4 pg/L and an ML of 10 pg/L. By the time that EPA proposed the method in 1991, the Canadian government published its own version that included a quantitation limit 5 pg/L. By the time EPA officially promulgated Method 1613

in 1997, many laboratories performing the analysis had replaced or supplemented their old instruments with newer models. As a result, many laboratories performing analyses using Method 1613 routinely measure sample results at levels 10 times lower than those analyzed routinely only 10 years earlier.

Given that measurement capabilities tend to improve over time, EPA believes that a detection and quantitation limit approach should be supported by procedures that will allow individual laboratories and other organizations to affordably characterize such improvements.

3.2 CWA Regulatory Issues Affecting Detection and Quantitation

Section 3.2.1 below provides a brief overview and a discussion of Clean Water Act activities that involve chemical measurements and are, therefore, directly impacted by detection and quantitation limit approaches. Specific issues that must be considered in the context of these CWA applications and EPA's regulatory obligations are discussed in Sections 3.2.2 - 3.2.6.

3.2.1 Detection and Quantitation Limit Applications Under CWA

The Clean Water Act directs EPA, States, and local governments to conduct a variety of data gathering, permitting, and compliance monitoring, and enforcement activities. Many of these activities depend directly on environmental measurements and, therefore, are affected by detection and quantitation limit approaches as discussed in the subsections that follow.

3.2.1.1 Method Development and Promulgation

Section 304(h) of the Clean Water Act (CWA; the "Act") requires EPA to promulgate test procedures (analytical methods) to be used for data gathering to support certification, permitting, and monitoring under the Act. These methods are promulgated at 40 CFR part 136, and include methods developed by EPA as well as those developed by other organizations, such as the publishers of *Standard Methods for the Examination of Water and Wastewater*, as well as AOAC-International, ASTM International, the U.S. Geological Survey, instrument manufacturers, and others. Upon request by a laboratory, permittee, instrument manufacturer, or other interested party, EPA considers alternate testing procedures (ATPs). If EPA deems these ATPs to be acceptable for nationwide use, they too, may be published at 40 CFR part 136. A primary objective in promulgating methods developed by EPA and by other organizations is to provide the regulatory community, permittees, and laboratories with multiple options so that they may choose the method that yields the best performance at the lowest cost for the application.

In recent years, EPA has focused on developing methods for promulgation at 40 CFR part 136 where no other methods are available that meet an immediate or anticipated regulatory need. The National Technology Transfer and Advancement Act of 1995 (NTTAA) urges government agencies to consider methods published by voluntary consensus standards bodies (VCSBs), such as Standard Methods and ASTM International, when VCSB methods are available. EPA accepts that many of these methods have been through a sufficient level of testing, peer review, and scientific acceptance to warrant proposal if they meet EPA's regulatory needs. When an individual laboratory, permittee, or other organization submits a request for approval of an alternate test procedure, however, EPA generally requires that the procedure be subjected to a level of testing that demonstrates that the method provides sensitivity, accuracy, and other measures of performance comparable to an approved method.

The lack of widespread consensus on detection limits has obvious impacts on EPA's responsibility to promulgate methods under CWA. Most organizations that develop methods use

different approaches, and many organizations have changed approaches over the years. The result is that a number of different approaches for detection and quantitation are embodied in the methods approved at 40 CFR part 136. The vast majority of the approved methods include the MDL which, as noted in Section 2.2.1, has been used by several EPA Offices, *Standard Methods*, AOAC, ASTM, and others. Other approaches embodied in the methods at 40 CFR part 136 include, but are not limited to: 1) a method "range" that is usually not defined, but is often interpreted as the lower end of the range in which pollutants either can be identified or quantified, 2) an "instrument detection limit" that has been defined by a variety of procedures, but is intended to capture instrument sensitivity only, 3) an "estimated detection limit" that may be based on best professional judgement, single laboratory data, or some other source of information, 4) a "practical quantitation limit," that has typically been determined according to one of the scenarios described in Section 2.3.1, and 5) "sensitivity" that is an undefined concept similar in result to the MDL.

The most obvious solution to this problem would be for the Office of Water to force all methods promulgated at 40 CFR part 136 to contain uniform approaches for detection and quantitation. Unfortunately, taking such action would confound methods promulgation. Problems with this solution are that:

- To date, no single detection and quantitation limit approach has emerged to meet the needs of all organizations for all applications.
- If the Office of Water were to select an approach that differs from those of other organizations, those organizations would be required to conform their method to accommodate OW's approach. Doing so would mean that these organizations would have to invest additional laboratory resources to develop detection and quantitation limits that conformed to OW definitions.
- If outside organizations decided against conforming their approaches to that of OW, fewer methods would be promulgated at 40 CFR part 136. This would result in fewer options for the regulatory, permittee, and laboratory communities.
- If EPA selected an approach that has burdensome procedures for developing detection and quantitation limits, it could discourage development of innovative technology or method modifications.

Given these issues, and EPA's desire to 1) encourage the development of improved measurement techniques, and 2) provide the stakeholder community with a variety of measurement options whenever possible, EPA believes it would be impractical to force standardization on a single detection or quantitation limit approach on method developers and promulgate only those methods that contain this approach. The Agency also believes, however, that there are real benefits to standardization, and that 1) all new methods developed by EPA for promulgation at 40 CFR part 136 should reflect such standardization, and 2) EPA should strongly encourage outside organizations to include these standardized approaches in their methods.

3.2.1.2 Method Performance Verification at a Laboratory

Just as sensitivity is important for evaluating measurement method performance, it is important to verify that a laboratory using a method can achieve acceptable levels of sensitivity for making measurements. Such demonstrations can take many forms and should be viewed in the context of the decision to be made. The analytical methods published at 40 CFR part 136 are designed for monitoring compliance with CWA permits. Most pollutants in permits have a numeric limit, and compliance with this limit is determined by laboratory analysis of samples from the waste stream or water body regulated by the limit. The laboratory that conducts such analyses must be able to demonstrate that its detection or quantitation limits are low enough to assure reliable measurements.

Thus, even where a method describes the sensitivity measured or estimated by the developer or the organization that published the method, some means are needed to demonstrate that a given laboratory can achieve sufficient sensitivity to satisfy the regulatory decision (e.g., monitoring compliance).

The EPA MDL procedure provides a means for verifying laboratory performance and has long been used in this fashion by EPA and various other Federal and state agencies. Other procedures may be employed, including analysis of reference materials containing the analytes of interest at concentrations that are at or below the regulatory limits of interest, spiked samples that are similarly prepared (e.g., matrix spikes), or performance evaluation (PE) samples such as those used in laboratory accreditation studies.

The IDE and IQE were advanced by the regulated industry and subsequently approved by ASTM International as a means of characterizing the performance of a method in laboratories that participate in an interlaboratory study. The idea in developing these approaches was to establish detection and quantitation limits that could be met by any laboratory that participated in the study. An advantage of this approach is that individual laboratories do not have to demonstrate sensitivity. However, potential disadvantages also exist. For example, it may not be possible to develop a realistic IDE or IQE for a new method involving a highly innovative technique because there may not be a sufficient number of laboratories practicing the technique to allow development of an IDE/IQE. Also, establishing detection and quantitation limits that can be met by all laboratories that practice methods that are in widespread use can potentially lead to worst-case limits that are significantly higher than limits that can be achieved by many commercial laboratories.

Developers of the IDE/IQE have recognized that an analogous approach is desirable for single-laboratory application and have begun work on a within-laboratory detection estimate (WDE), to be followed by a within-laboratory quantitation estimate (WQE). As with the IDE/IQE, these approaches will capture a wide range of sources of variability such as temporal variability, and will include a prediction or tolerance limit (or both), but will not include interlaboratory variability. EPA would consider such single laboratory approaches if and when they are adopted by ASTM International.

3.2.1.3 National Pollutant Discharge Elimination System

The National Pollutant Discharge Elimination System (NPDES) serves as the primary means by which EPA, States, and Tribes control point source releases into the nation's waters. Under this system, individual facilities are issued NPDES permits that provide limitations on the type, concentration, and volume of pollutants that may be legally discharged. Typically, these pollutant controls are based on technology-based standards. If, however, these technology-based controls are not adequate to protect the water-quality standard designated for the facility's receiving water, stricter controls are warranted. In such cases, NPDES permits generally contain water quality-based controls.

Development and Implementation of Technology-based Controls (Effluent Guidelines)

EPA promulgates national effluent limitations guidelines and standards under the authority of Clean Water Act Sections 301, 304, 306, 307, 308, and 501. The regulations allow the discharge of pollutants from normal industrial processes when the discharges have been treated using various levels of available treatment technologies that are affordable. Functionally, these industry-specific guidelines establish standards for the quality of wastewater discharges to waters of the United States. They are generally stated in the form of concentration-based limits for selected substances that are not to be exceeded. For example, the maximum oil concentration in wastewater separated from oil pumped out of an offshore well and discharged on any single day shall not exceed 42 milligrams per liter (mg/L). This form is called a numeric effluent guideline limit or numeric limit.

Development and Implementation of Water Quality-based Controls

States designate water-quality standards for various bodies of water within their boundaries. Each standard consists of a designated use, criteria to support that designated use, and an anti-degradation policy. Examples of designated uses include public water supply, recreation, and propagation of fish and wildlife. When the water-quality standard is not met, waste-load allocations are developed to indicate the maximum amount of a substance that can be discharged to a particular water body without impairing the designated use. EPA and authorized states calculate water quality-based effluent limits based on the waste-load allocation and the variability of the substance in the wastewater discharge. The concept is to prohibit discharge of a substance beyond the level at which a designated use would be impaired. Water quality-based permits generally specify the use of measurement methods promulgated at 40 CFR part 136 under the Clean Water Act Section 304(h).

A special case occurs when the water quality-based effluent limit is less than the detection limit of the most sensitive analytical method. This case is addressed in Section 3.2.3 below, on compliance evaluation thresholds.

Permit Compliance Monitoring

Under Clean Water Act Sections 318, 402, and 405, NPDES permits are issued to owners of facilities that discharge wastewater to waters of the United States (coastal areas, lakes, rivers, streams, certain wetlands, etc.). Specific discharge limits are established either for individual facilities or for classes of facilities. Individual permits are established for industries with many site-specific issues that determine the substances discharged, such as the pharmaceutical industry in which the specific drugs produced could influence the water quality. General permits are issued when the substances discharged do not vary widely among facilities (e.g., coastal oil and gas extraction industry facilities). The permit limits are typically established using technology-based effluent guidelines, unless the facility is discharging into a water body that does not meet its designated use or that will not meet the designated use if a technology-based limit is permitted.

Detection plays a role in compliance monitoring because of concerns with measurement results at the low end of any measurement method. All measurement results are variable. At the low end of most measurement methods, there comes a point at which a particular measurement result is unacceptably likely (a policy decision) to have come from a sample in which the substance of interest is absent (zero concentration). Such a measurement result would be below the critical value defined by Currie (1995) and in common usage it would be called below detection. In practice, the reporting limit may be set equal to a critical value, detection limit, or quantitation limit. Assuming that the reporting limit is a detection limit of 1 mg/L of oil and grease, the measurement result would be reported as "less than 1 mg/L of oil and grease."

3.2.1.4 Non-Regulatory Studies and Monitoring

EPA conducts a variety of non-regulatory studies and monitoring activities to support its Clean Water Act programs. These activities range from long term surveys, such as the Great Lakes Water Quality Surveys that are conducted each spring and summer to monitor trends in water quality against established baselines, to short-term studies that are used to establish baselines, model pollutant cycles, and guide direction for future study and policy. Examples of such studies include the National Study of Chemical Residues in Fish that was conducted in the late 1980s (a follow-up to that study is currently underway), and the Lake Michigan Mass Balance Study conducted in the early 1990s.

When designing a study or monitoring program, EPA uses information about detection and quantitation limits, along with information on the risks associated with the pollutant(s) of interest and the cost of measurement, to select an appropriate method for measuring the pollutant. Accepting all positively valued measurement results and selecting a measurement method with a detection limit lower than the level of concern for the substance being measured would provide some assurance that measurement results associated with that concentration would be positively valued. Selecting a measurement method with a quantitation limit lower than the level of concern for the substance being measured would generate measurement results that are easier to explain to the data user and the general public.

3.2.2 Descriptive versus Prescriptive Uses of Lower Limits to Measurement

The literature on detection and quantitation generally assumes that these procedures are descriptive, as opposed to prescriptive. In other words, detection and quantitation studies are described as characterizing the current performance of a laboratory or laboratories using a method to measure a substance. Two possible reasons for this treatment are: (1) the intended audience includes laboratory staff and measurement methods developers who wish to make new methods useable to as many laboratories as possible, and (2) the author may have an institutional reason for not attempting to control variability and thus lower detection and quantitation limits. On the other hand, the technology-based and water quality-based effluent limitations programs administered by EPA's Office of Water have an institutional goal of protecting human health and the environment. Providing this protection requires that the Agency be able to measure pollutants at ever lower concentrations. Establishing stringent standards and a compliance scheme for laboratories is one way to more rapidly develop the ability to measure at these concentrations. A prescriptive strategy concerning detection and quantitation limits would be to:

- Determine the detection and quantitation limits at multiple laboratories.
- Establish a detection limit and a quantitation limit for the method that is based on some performance of these laboratories. These limits could be established as the limits reported by the mean or median laboratory, or by some other criterion, such as the pooled value of the limits achieved by all laboratories, or the limits that are met by a certain percentage of the laboratories.
- Use the established detection and quantitation limits as performance standards that must be demonstrated by laboratories that practice the method.

Such an approach is consistent with other performance standards included in EPA methods, such as standards for instrument calibration, recovery of spiked reference and matrix samples, etc.

The use of such an approach would help ensure that prescriptive detection and quantitation limits (i.e., performance standards) reflect the capabilities of multiple laboratories, rather than a single state-of-the-art research laboratory. Of course, it is possible that even when multiple laboratories are used to establish performance standards for detection and quantitation, some laboratories may not be able to achieve these standards using their current operations. However, most laboratories facing this problem should be able to achieve these standards by investing in staff training, improved equipment, a stronger quality assurance program, or higher quality maintenance and operations.

There is of course, a risk that some members of the laboratory community will not be able to meet the standard, either because they are not willing to invest the resources necessary to do so, or for other reasons. That risk should be considered when using a prescriptive approach to detection and quantitation (i.e., establishing limits that act as performance standards). Conversely, the risk of using a descriptive approach is that it can result in detection and quantitation limits that reflect a broad community of laboratories, including those that have made little if any effort to control variability at these levels, thus raising detection and quantitation limits to a level that is higher than desired.

3.2.3 Compliance Evaluation Thresholds

3.2.3.1 Compliance Evaluation Thresholds

When technology-based effluent limitations are developed, the limits are expressed as being at or above the quantitative measurement capabilities (e.g., the ML) of one or more analytical methods that are available to support compliance monitoring. Therefore, it is possible to monitor and evaluate permit compliance at concentrations with an accepted degree of measurement certainty.

A situation that arises frequently in addressing water quality-based limits is the setting of the permit limit below the detection or quantitation limit of the most sensitive, approved analytical method. This subject was addressed in EPA's draft *National Guidance for the Permitting, Monitoring, and Enforcement of Water Quality-based Effluent Limitations Set Below Analytical Detection/Quantitation Levels* (WQBEL guidance). The WQBEL guidance suggested use of the minimum level of quantitation (ML) as the compliance evaluation threshold (CET) when the water quality-based effluent limit (WQBEL) is below the detection or quantitation limit of the most sensitive, approved analytical method. In comments on the WQBEL guidance, the regulated industry objected to the CET, claiming that it did not include interlaboratory variability and other sources of variability. States objected to the CET, claiming that it would not allow them to be as protective as if the detection limit were used. (This 1994 draft guidance document was never finalized due to the controversy.)

From a technical standpoint, a one-sided limit that addresses false positives only, such as Currie's critical value or EPA's MDL, is the most appropriate approach for producing a CET for the situation in which the WQBEL is less than detection limit in the most sensitive analytical method because the one-sided limit allows measurement to the lowest possible level while protecting a discharger from the risk of a false violation. For example, consider the situation in which 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (dioxin) is to be evaluated against the ambient water quality criterion of 13 parts-per-quintillion (ppqt). The most sensitive analytical method approved at 40 CFR part 136 is EPA Method 1613, with an MDL of 4 parts-per-quadrillion (ppq) and an ML of 10 ppq. The MDL is more than 300 times greater than the ambient criterion. Therefore, if dioxin is detected in the receiving water as a result of a discharge (i.e., the measurement result is greater than the MDL of 4 ppq), there must have been an exceedance of the ambient criterion. In the WQBEL guidance, EPA suggested use of the ML because it was the point at which the measurement could be considered reliable. However, from a purely technical standpoint, the MDL (or Currie's critical value) is most appropriate if the goal is to protect the receiving water. It is important to note, however, that because individual states are responsible for implementation and enforcement of NPDES permits, use of the MDL and ML as regulatory reporting and compliance evaluation thresholds varies among the states.

Detection and quantitation limits have been used to establish CETs and permit limits. For example, see application of the ML to establishment of permit limits in Procedure 8 of Appendix F of Water Quality Guidance for the Great Lakes System at 40 CFR part 132. However, EPA believes that the decision to establish CETs based on a detection or quantitation limit is a separate issue from the question of which detection and quantitation limit approach is most valid. The objective of the assessment described in this document is to evaluate the merits of each approach.

3.2.4 Accepting the Procedures of Voluntary Consensus Standards Bodies

In February 1996, Congress enacted Public Law 104-113 (15 USC 3701), the National Technology Transfer and Advancement Act (NTTAA). This act directs "*federal agencies to focus upon increasing their use of (voluntary consensus) standards whenever possible, thus reducing federal procurement and operating costs.*" The Act gives Federal agencies discretion to use other standards

where the use of voluntary consensus standards would be *"inconsistent with applicable law or otherwise impractical."*

The NTTAA is implemented by Federal agencies based on the policies described in Circular A-119 from the Office of Management and Budget (OMB). The current version of this OMB circular was published in the *Federal Register* on February 19, 1998 (63 FR 8546).

Neither the NTTAA nor Circular A-119 require that agencies replace existing government standards with standards from a voluntary consensus standard body (VCSB). In other words, if EPA already has standards in place for detection and quantitation approaches, EPA is not obligated by NTTAA to replace these with VCSB standards.

Circular A-119 also discusses the effect of the policy on the regulatory authorities and responsibilities of Federal agencies. The circular states that:

"This policy does not preempt or restrict agencies' authorities and responsibilities to make regulatory decisions authorized by statute. Such regulatory authorities and responsibilities include determining the level of acceptable risk; setting the level of protection; and balancing risk, cost, and availability of technology in establishing regulatory standards. However, to determine whether established regulatory limits or targets have been met, agencies should use voluntary consensus standards for test methods, sampling procedures, or protocols."

Thus, EPA is responsible for establishing the levels of risk and protection, not only for the regulatory limits applied to discharges, but also to the risks of decision errors (e.g., false negatives or false positives) in the detection and quantitation approaches applicable under the Clean Water Act.

Finally, Circular A-119 describes two types of technical standards: performance standards and prescriptive standards. A performance standard is defined as:

"a standard ... that states requirements in terms of required results with criteria for verifying compliance but without stating the methods for achieving required results." In contrast, a prescriptive standard is one "which may specify design requirements, such as materials to be used, how a requirement is to be achieved, or how an item is to be fabricated or constructed."

Neither the NTTAA nor Circular A-119 direct agencies to favor performance standards over prescriptive standards, or vice versa. EPA believes that the current MDL procedure is a prescriptive standard, in that it specifies both the design of the MDL study and how the requirement to establish method sensitivity be achieved. There is some obvious flexibility or opportunity for judgement in employing the MDL procedure, and much of the historical debate over the utility of the MDL procedure would suggest that it may not be prescriptive enough. The alternative approaches for establishing detection and quantitation are also prescriptive standards, rather than performance standards.

One option that EPA may consider is to employ a performance-based approach to establishing detection and quantitation limits in which method developers, laboratories, and others would be free to use any one of a variety of approaches to establishing these limits, including the existing MDL procedure or a VCSB standard. Thus, establishing method sensitivity could be considered a performance standard under NTTAA and Circular A-119, rather than a prescriptive standard. The fact that different approaches (prescriptive standards) yield different answers would be immaterial if EPA evaluates the answers relative

to a specific decision. That evaluation should not be divorced from knowledge of the decision to be made (e.g., the regulatory limit for a given pollutant).

3.2.5 National versus Local Standards for Measurement

In accordance with the Settlement Agreement, EPA is re-examining the approaches of detection and quantitation used with methods approved for use at 40 CFR part 136. The Clean Water Act authorizes states and local governments to implement permits, with the requirement that they be at least as protective (stringent) as the national standards established by EPA. Thus, EPA must take into account the impact of any revised or new detection/quantitation limit approaches and procedures on state and local governments, as well as on those affected by state and local requirements. EPA also is aware that some states have implemented approaches to detection and quantitation that are either specific to that state, result in lower numeric limits in discharge permits, or both. Given the ability of state and local governments to use more stringent approaches, any decision by EPA with regard to this re-evaluation of detection and quantitation approaches may not have an effect on those states and local governments.

3.2.6 Cost and Implementation Issues

Detection and quantitation limit procedures are typically employed by organizations that develop methods and by laboratories that use the methods. Method developers typically include governmental organizations such as EPA, NOAA, USGS, and DOE, or voluntary consensus standards bodies (VCSBs) such as the American Public Health Association (APHA), ASTM International, AOAC-International, and ISO/IUPAC. Method developers also may include manufacturers of instruments or supplies used in testing. Methods users generally are the laboratories performing tests to assess and assure product quality, to support regulatory compliance monitoring, or to support scientific studies.

Method development requires a more diverse set of skills than method use because such development generally demands an understanding of quality systems, statistics, and analytical technologies. Staff working for the method developer will usually include a project manager, measurement analysts, and statisticians. Method use requires a focus on obtaining reliable results in the analysis of a given sample. Staff working for laboratory typically include the manager and measurement analysts.

3.2.6.1 Implementation of a Detection/Quantitation Limit Procedure by a Method Developer

The basic resources available to the method developer are time, money, and the technical skills of its staff. The fundamental decision for implementing a detection or quantitation procedure is whether that procedure is intended to characterize the performance of the method at a well-performing laboratory or if it is intended to characterize the performance of the method across a group of laboratories. If the procedure is intended to characterize the performance of the method across a group of laboratories, it is also necessary to decide if there will be some way to compare the performance of individual laboratories to the group performance standard. There are serious time, cost, and skill issues with each of these decisions. Ordering these decisions from the least resource intensive to the most, they are characterizing the performance of the method: (1) at a well-performing laboratory, (2) at a group of laboratories, or (3) at a group of laboratories with comparisons of individual laboratories. Other costs for the method developer could include planning, data management, reference laboratory services, and whether laboratories are willing to volunteer for the study or if their services must be purchased.

An independent decision is whether to assume a simple model for measurement variability and limit the number of test concentrations, iterate assuming a simple model, or to design a study of the relationship between measurement variation and the concentrations of the substances measured by the

method. This decision will greatly influence the number of samples measured in the study. If the laboratories do not volunteer for the study, then the direct cost for measuring these samples or blanks ranges from a few dollars per sample to more than \$1,000 per sample for some analytes. Until such time as the relationship between measurement results and standard concentrations becomes well known, such studies will require the active participation of professional statisticians in design, implementation, and analysis.

3.2.6.2 *Implementation of a Detection/Quantitation Limit Procedure by a Laboratory*

A laboratory may implement detection or quantitation procedures for its own quality control purposes, because of regulatory requirements, or as part of the study of a method by some other organization. When participating in the study of another organization, the laboratory may voluntarily accept some cost of the study for marketing purposes, professional development, or to benchmark the performance of the laboratory.

In each case, a detection or quantitation limit approach will be of little utility if it is not capable of being implemented by the laboratory. An advantage of straightforward approaches such as the EPA MDL, the ACS limit of detection, and the ISO/IUPAC critical value is that they can, in principle, be understood by analysts expected to use the approach. Likewise, the procedures described for implementing the MDL approach are straightforward and have been implemented by thousands of laboratories. In contrast, the ASTM IDE and IQE procedures are highly complex and, as a consequence, are beyond the capability of nearly all environmental testing laboratories.

Another disadvantage of highly complex procedures is that they are usually more costly to implement than simple procedures. As noted in Section 3.1.5, method performance generally improves over time, and EPA believes that a detection and quantitation limit approach should be supported by procedures that will allow individual laboratories and other organizations to affordably characterize such improvement. A significant shortcoming of interlaboratory procedures is that small laboratories that develop new techniques or modify existing techniques to achieve improved measurement sensitivity would have to rely on, and perhaps even pay, other laboratories to demonstrate the sensitivity of their procedures. Such a limitation has the effect of hindering method development and improvement.

3.2.7 **Use of a pair of related detection and quantitation procedures in all Clean Water Act applications.**

In Section 3.2.1, we discussed several different applications for detection and quantitation limits under the Clean Water Act. To review, these applications are:

- Method development and promulgation,
- Method performance verification at a laboratory,
- Technology-based effluent guidelines development,
- Water quality-based effluent limits development,
- Permit compliance monitoring, and
- Non-regulatory studies and monitoring.

Although EPA could develop a separate detection and quantitation approach for each of these applications and attempt to define and evaluate each of these approaches in our re-examination of detection and quantitation approaches, the resulting matrix of applications and approaches would cause confusion for regulators, permittees, and the laboratory community. Further, when proposed, each element of the matrix of approaches and applications would, individually, be subject to contention and second-guessing, and it is likely that the outcome would be nearly the same as if a single pair of

approaches is selected. EPA prefers to avoid this confusion by using a single pair of related detection and quantitation procedures to meet any the needs of all Clean Water Act applications.

3.2.8 Alternative Procedures

One of the peer reviewers who evaluated a draft version of this assessment document noted that:

"EPA has stated in the TSD that one primary procedure is needed for clarity and to avoid confusion among stakeholders. If alternate procedures are needed, the EPA Clean Air Act system of reference and equivalent methods has worked well, and could be a model for EPA to follow under the Clean Water Act." (Cooke, 2002)

The system of reference methods used under the Clean Air Act is similar to the existing "alternate test procedure" (ATP) program for analytical methods currently used within the Office of Water. The difference between the ATP program and the case of the procedures for establishing detection and quantitation limits is that in an ATP program, the goal is clear and agreed upon, whereas there remain fundamental theoretical issues surrounding detection and quantitation.

For example, when a test procedure is developed for use in the Clean Air Act or Clean Water Act programs, the reference method is designed to measure Analyte X, in Matrix Y, at some concentration related to a regulatory need (i.e., a compliance limit). Alternative procedures may be capable of making measurements of Analyte X in Matrix Y, at the level of concern, using completely different instrumentation. Thus, the demonstration of equivalency between the reference method and a possible alternative method is judged using a metric that consists of Analyte X, Matrix Y, and the level of concern, as well as other aspects of method performance.

In contrast, the primary differences between the EPA MDL/ML approaches and potential alternatives such as the ASTM IDE and IQE are related to the philosophical differences of how detection and quantitation limits should be derived and applied. These differences are described at length in this chapter and the rest of the Assessment Document. Therefore, EPA does not believe that a variant of existing ATP programs is likely to be an effective model for assessing other detection and quantitation procedures.

What EPA would be willing to consider is that an analytical method from a VCSB or other source may be acceptable for approval at 40 CFR Part 136 and use in Clean Water Act programs even if it employs an alternative procedure for establishing method sensitivity. For example, consider the theoretical situation of an ASTM method for the determination of an analyte regulated under the NPDES program that uses the IDE or IQE to describe method sensitivity *and* for which the value of the IDE or IQE was below the relevant regulatory limit. EPA would evaluate the overall performance of such a method for approval at 40 CFR Part 136, despite the fact that the method did not contain an MDL determined using the Appendix B procedure.

3.3 Statistical Issues

The goal of this section is to provide a brief explanation of the key statistical issues involved in the development of detection and quantitation limits.

3.3.1 Sources of Variability

Various known and unknown sources of variability will influence measurements made by a laboratory using a specific method. These sources may include random measurement error, differences in analysts, variations between different equipment manufacturers and models, variations in analytical standards, routine fluctuations in equipment performance, and variations in facility conditions (e.g., varying levels of background contributions).

There are a number of ways in which variability can be controlled. One is a strong quality assurance (QA) program that includes use of: 1) trained and qualified staff, 2) properly maintained equipment, 3) standards that are fresh and properly prepared and stored, 4) written standard operating procedures and methods for all sample handling, analysis, and data reduction/reporting activities, 5) procedures for monitoring ongoing laboratory performance and 6) quality control (QC) samples and QC acceptance criteria to ensure that the laboratory systems are in control. The EPA methods promulgated at 40 CFR part 136 require the use of qualified staff, appropriately cleaned and calibrated equipment, and properly prepared standards. Each method also provides detailed steps for performing all sampling handling and analysis activities.

Even when prescribed EPA requirements are implemented, however, it is not possible to completely eliminate all variability within or between laboratories. The potential effects of sources of variability should be considered when establishing detection and quantitation limits. Even with procedures in place to control quality and reduce variability, it should be recognized that some laboratories may achieve lower detection and quantitation limits than others. Ultimately, some laboratories may not be capable of meeting low-level measurement requirements without some effort to improve operations.

One of the peer reviewers asked to evaluate EPA's reassessment pointed out that EPA should consider added sources of variability, such as matrix interferences, when developing detection and quantitation limits (Wait, 2002). Such sources of variability are considered in establishing detection and quantitation limits for analytical methods under EPA's Clean Water Act programs because these detection and quantitation limits are established in a single-laboratory study, verified in multiple single-laboratories, and, where necessary, further verified in an interlaboratory study. If matrix effects surface in these studies, either procedures are put in place to overcome matrix interferences, or the detection and quantitation limits are adjusted upward to account for these matrix effects. However, it has been EPA's experience that concern over matrix effects may be somewhat overblown. For example, in EPA's interlaboratory validation studies of the 600-series wastewater methods, the recoveries of some organic analytes from some real-world matrices were closer to 100% than from a reagent water matrix. This effect is thought to be attributable to dissolved solids in the real-world matrix that, in effect, "salt out" the organic compounds.

What EPA has not done, and does not believe is appropriate, is to aggressively pursue matrix effects in method development or in establishing detection and quantitation limits (i.e., EPA has not attempted to find worst-case matrices in order to maximally exacerbate matrix effects). EPA has not pursued this approach because, for any method, it is possible to contrive and synthesize a sample matrix that would render the method unusable (e.g., by saturating an sample in which organic analytes are to be determined with a wide range of polymeric materials, both water soluble and water insoluble). Rather, EPA considers the type of matrix that would be encountered in a wastewater discharge and that would be regulated under the Clean Water Act (e.g., the effluents that are discharged from properly designed and operated secondary treatment plants).

A source of variability that is *not* considered in any of the detection and quantitation limits is the variability that is associated with the sample itself. Detection and quantitation limits focus exclusively on the capabilities of the measurement process. However, measurements are only as reliable as the sample being measured. If the sample is not truly representative of the population from which it was collected, then the variability associated with measurements made in the region of detection or quantitation may be immaterial.

For example, EPA's Technology Innovation Office conducted a study to characterize the effects of sampling variability on measured results. In that study, results from seven discrete samples collected within a two-foot distance of one another were evaluated. Each sample was analyzed for the explosive TNT on-site, using a colorimetric test kit, and in a traditional laboratory using EPA SW-846 Method 8330 (high-performance liquid chromatography). Analysis of the results from these measurements indicated that 95% of the total variability was due to sampling location and only 5% was due to differences between the analytical methods. Put another way, differences in sampling location caused 19 times more uncertainty in the data results than did the choice of analytical method, over a distance of only 2 feet (Crumbling, 2002). While EPA does not wish to diminish the importance of understanding measurement error in the region of detection and quantitation, EPA believes it is equally important to understand it in the context of the overall sampling and analysis error.

3.3.2 Censoring Measurement Results

Measurement results are often reported as less than some detection, quantitation, or reporting limit (see Section 3.2.1.3, Permit Compliance Monitoring) without providing a single best estimate for the numeric result. For example, if a direct reading of the measurement results would indicate a concentration of 3 mg/L and the reporting limit for the substance is 5 mg/L, the laboratory may only report that the measurement result is less than 5 mg/L. Statisticians call this process of suppression of results less than a specified amount "censoring." Reasons for the practice of censoring relate directly to issues surrounding the development of detection and quantitation limits (i.e., the premise that measurement results below certain low levels may not useable for certain purposes).

In order to evaluate low-level variability, EPA conducted a comprehensive study of results from 1/10th of the MDL to concentrations into the usual quantitation range. Ten different analytical techniques were evaluated in the study (see Appendix B, Characterizing Measurement Variability as a Function of Analyte Concentration for a Variety of Analytical Techniques). Data from this study indicate that measurement results may be generated at low concentrations which are quite variable in relation to the true concentration in the sample. While this observation has not been demonstrated with every substance measured, it is suggested by plots of data from most of the measurement techniques observed in the study. An example is Ammonia as Nitrogen by Method 350.3. Plotting measurement results versus concentrations spiked into reagent water samples (Figure 3-1), we see the strong relationship between measurement results and spike concentrations that would be expected. However, it is difficult to see what is going on at low concentrations in a graphic that covers measurement results over several orders of magnitude.

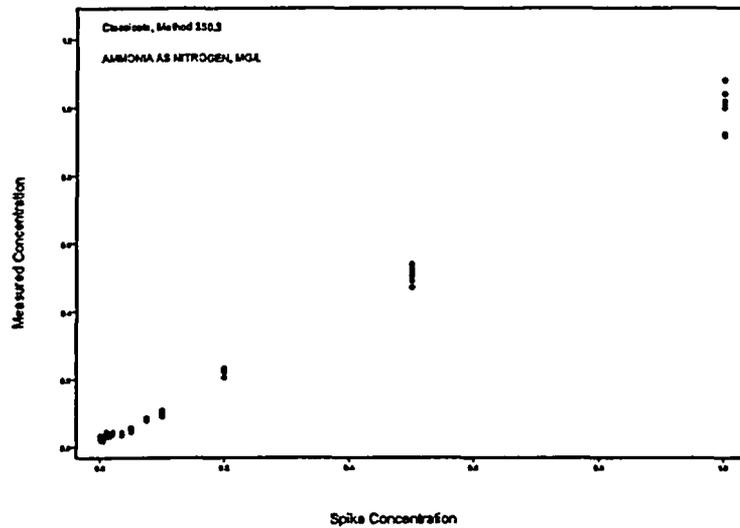


Figure 3-1

By plotting the measurement results versus the spike concentrations on log scales (Figure 3-2), we would expect to see an expansion of variability at low concentrations, a contraction of variability at high concentrations and points mostly plotted along a 45° angle to indicate that measurement results are approximately equal to the spike concentrations. However, we only see the 45° angle at higher concentrations. Measurement results in the lowest order of magnitude appear to have reached a plateau below which they do not go. ASTM Committee E-1 has termed the model that describes the general pattern displayed by these data as the "General Analytical Error Model."

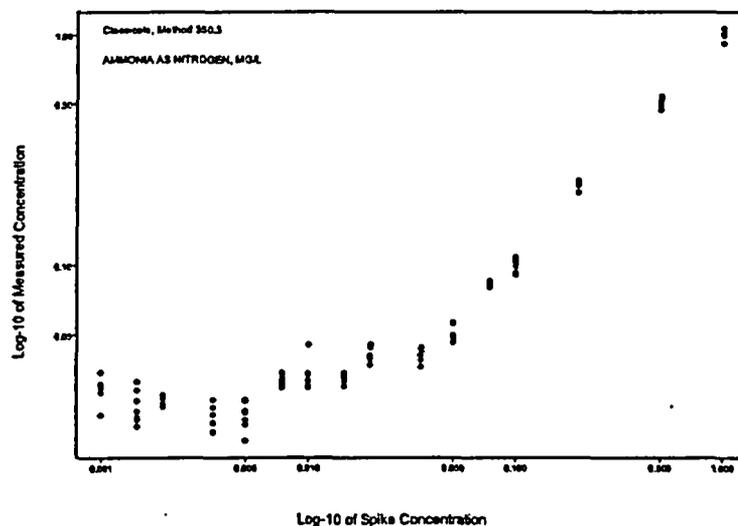


Figure 3-2

Despite such evidence that results at low concentrations can be quite variable, some data users prefer to use the actual measurement results (even if they are negative values), rather than to censor the results at a reporting or detection limit, because censoring data at such a limit can introduce bias into the data set. If all low values are eliminated, then the mean of the remaining data would have a positive bias. In other words, while negative or extremely low values may be considered problematic by some, they are of value to statisticians and modelers, especially when dealing with analytes that are highly toxic and/or environmentally significant.

Some programs, such as EPA's Superfund Contract Laboratory Program, require laboratories to report the measurement result obtained from the analysis in conjunction with a qualifier that the result is below a specified detection, quantitation, or reporting level. Going back to the example in the first paragraph, the laboratory might report both a measured value of 3 mg/L and a reporting limit of 5 mg/L. Under certain assumptions, measurement results below the specified level could then be used to calculate averages and statistical estimates that would be superior to estimates calculated using censored data.

EPA believes that such an approach provides the greatest degree of flexibility for data users, but also believes that it should be used with care. First, data users who choose to use values reported below a detection or quantitation limit need to have a firm understanding of the limitations of those data. Second, and as noted in Section 3.2.1.3, Permit Compliance Monitoring, reporting data below a detection or quantitation limit can lead to misinterpretation.

One of the peer reviewers that evaluated EPA's assessment of detection and quantitation limit approaches noted that European Union (EU) has adopted another variant for reporting or censoring data.

"In this case, the EU has adopted EPA Method 1613B (for analysis of dioxins and furans) as well as EPA's MDL approach. However, the EU has further specified that the MDL be used as an Upper Bound reporting limit where all non-detects are found in the analysis of human or animal foodstuff. This forces laboratories to achieve levels available with modern instrumentation, otherwise, the Upper Bound reporting level is above the regulatory compliance level, and the data (or foodstuffs) are rejected" (Cooke, 2002).

EPA agrees that this approach, which yields a "worst-case" (or highest possible) estimate of the pollutant concentration, can serve as a useful regulatory tool for encouraging the analytical and regulated community to pursue measurements at the lowest levels necessary to protect human and ecological health. However, EPA also cautions that this approach also should be recognized as a regulatory strategy that effectively censors measurements made below the MDL.

EPA believes that while the issue of censoring is important, it should not be a consideration when selecting a detection and quantitation limit approach. The decision to censor data is a data reporting and data use issue, rather than a detection and quantitation issue. This issue will apply regardless of what detection or quantitation limit approach is used. The EU approach reflects a similar point of view, in that it relies on the MDL as a detection approach, and also establishes this limit as the reporting level for non-detects in order to encourage development of lower MDLs.

3.3.3 Outliers

Outliers are extreme or aberrant measurement values that, on inspection, do not follow the characteristics of a set of data. Outliers may be generated by a number of causes, such as errors in following an analytical procedure, errors in recording numerical results, or the result of extreme random variation in a properly operating process. For example, if a new measurement method is being tested but

the laboratory fails to follow the procedure correctly with some samples, the associated measurement results may stand out as outliers. A graphic example is provided in Figure 3-3, which shows measurement results for aluminum, determined using EPA Method 1620, versus concentration. At a spike concentration of 250 $\mu\text{g/L}$, one of the measured values is about 750 $\mu\text{g/L}$, and visually stands out from the rest of the values. This result may turn out to be an outlier.

A common process for identifying potential outliers is to apply one or more statistical procedures for identifying values far from the mean (average) of the data. An example of such a procedure is ASTM Practice D-2777. Because extreme values can be expected to occur on occasion, it is not necessarily appropriate to exclude them from the measurement results used to develop detection or quantitation values. As recommended in the ASTM procedure, a review of the analyst's records associated with the measurement may establish whether the extreme value was caused by failure to follow the method or by some rare event associated with the method. If the method under study was not followed, it is appropriate to exclude the measurement result from the detection or quantitation analysis. If the measurement result is a rare event associated with the method under study it may also be appropriate to exclude the measurement result from the results in the study.

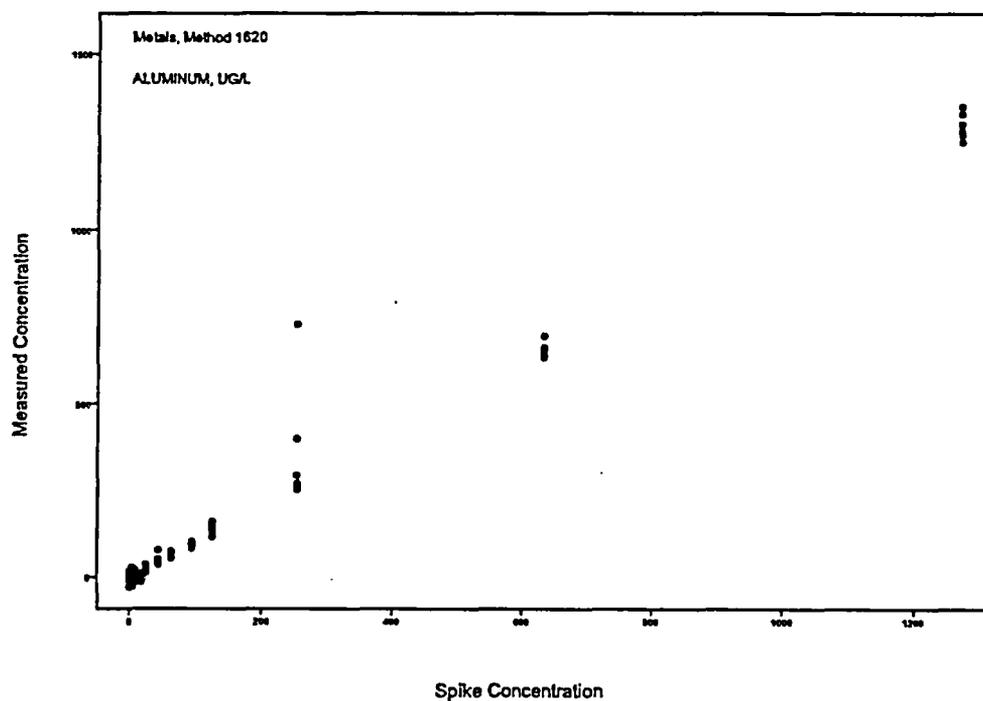


Figure 3-3

Influential early work in using ranking procedures to help identify outlying laboratories in studies was conducted by Youden (Youden, W. J. and E. H. Steiner, *Statistical Manual of the Association of Official Analytical Chemists*, 1975).

3.3.4 Criteria for the Selection and Appropriate Use of Statistical Models

Detection and quantitation limits may be based on statistical models of the relationship between measurement variation and the concentration of a substance in the sample. Results are produced by adding varying known amounts of the substance to the sample ("spiking"), making replicate measurements at each concentration, and modeling the variability of the results as a function of concentration. This section summarizes the history of modeling variability versus concentration, considers criteria for selecting models, and discusses current practices with regard to available data.

3.3.4.1 Short History of Modeling Measurement Results

Over time, a number of different models have been used to estimate measurement variation. Currie (1968) modeled variation in radiochemical measurement methods using a procedure associated with counting large numbers of distinct objects which are appropriately modeled with the Poisson distribution. However, he relied on large sample sizes and standard normal distributions to describe all other types of measurement methods. Hubaux and Vos (1970) developed a procedure based on an estimated calibration relationship that uses smaller sample sizes to estimate Currie's detection and quantitation limits. Again, measurement results were assumed to follow standard normal distributions, but it was also assumed that measurement variation was constant throughout the range of interest. Similarly, Glaser *et al.* (1981) suggested that measurement variation increases linearly with concentration, but they did not provide estimators under this theory because they believed that measurement variation is usually approximately constant in the range of detection. Glaser *et al.* (1981) did suggest that, when appropriate data were available, a linear regression analysis of the relationship over the analytical range be performed. Clayton *et al.* (1987) discussed transforming the measurement results (using logarithms or square root functions). Gibbons *et al.* (1991) suggested that measurement variability may be proportional to concentration. Rocke and Lorenzato (1995) proposed a model motivated by physical characteristics of measurement processes, in which measurement variability is approximately constant at low concentrations, but changes in a continuous mathematical manner to a relationship where variability increases as concentration increases.

Figure 3-4 illustrates the fundamental analytical measurement models in linear and logarithmic domains. The models are applicable to nearly all analytical measurements; we will not deal with the exceptions because they represent a small percentage of cases. As can be seen from the top two graphs, response is a linear function of concentration in both the linear and log domains. The middle two graphs and the bottom two graphs are those most pertinent to the discussion detection and quantitation.

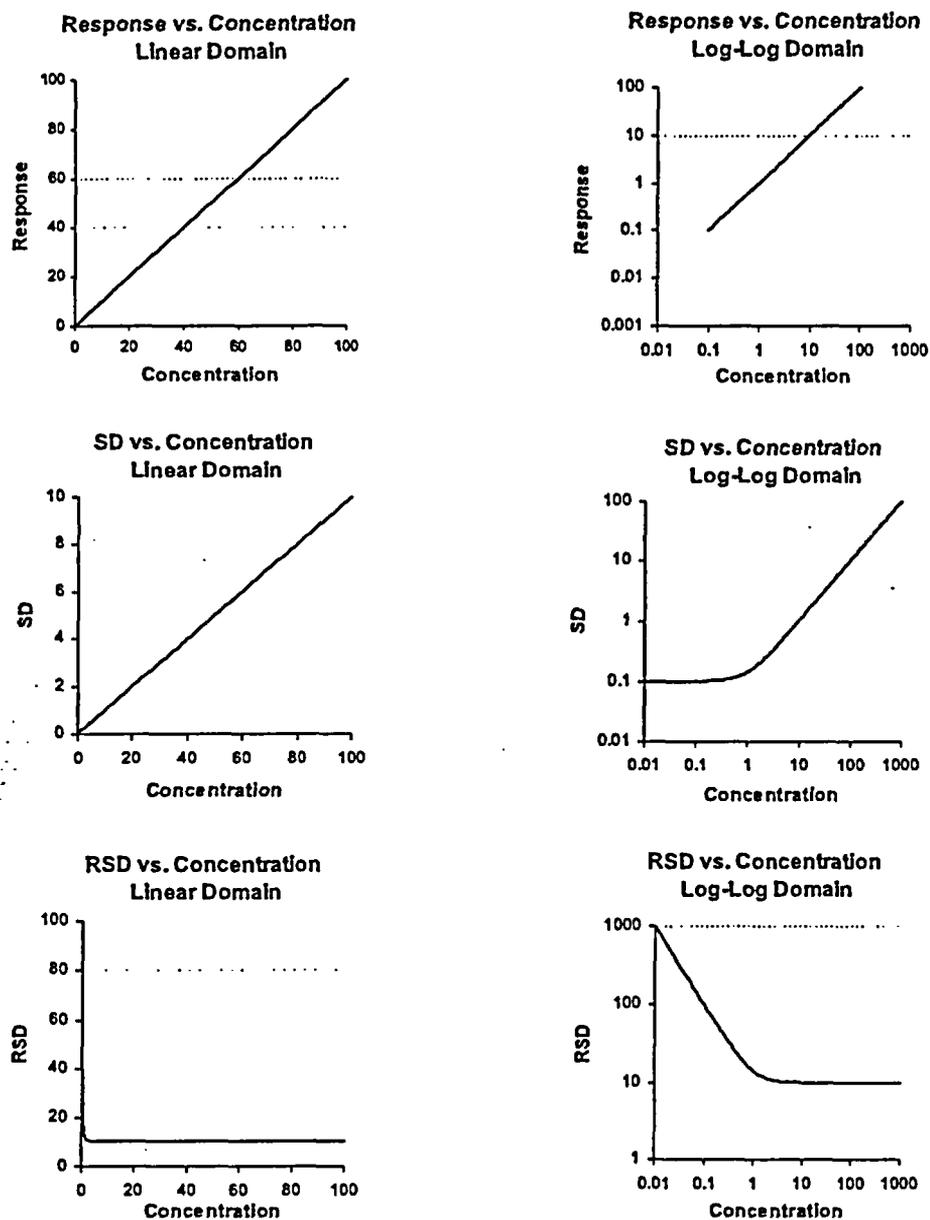


Figure 3-4

3.3.4.1.1 *Detection Limits Using Variability at Low Concentrations*

The middle two graphs in Figure 3-4 show variability versus concentration and show the model postulated by Rocke and Lorenzato. The flat (constant) portion of the graph in the linear domain is difficult to see because it occurs near the origin, but it can be seen easily in the log domain. Most detection approaches (e.g., Currie's critical value and detection limit; EPA's MDL; the ACS LOD) are constructed assuming that the flat (constant) region of the variability versus concentration relationship in the graph holds true, although the graph is rarely displayed (a horizontal line would be singularly uninteresting). Detection approaches such as Currie's critical value, detection limit, LOD, and MDL are constructed by multiplying the standard deviation in the flat region by some constant.

Contention and differences of opinion occur in determining how to arrive at an "appropriate" standard deviation and what to do with the standard deviation when you have it. Currie's critical value and EPA's MDL use a multiple of the standard deviation in a similar manner (a *t*-statistic adjusted for the number of replicates used for Currie's critical value; 3.14 for 7 replicates in EPA's MDL). The IDE uses an additional upward adjustment based on a statistical tolerance limit calculation.

3.3.4.1.2 Quantitation Limits Using Standard Deviation Multiples and Models of Standard Deviation versus Concentration and RSD versus Concentration

The limit of quantitation (LOQ) advanced by Currie and the American Chemical Society's Committee on Environmental Improvement, and EPA's minimum level of quantitation (ML) result from multiplication of the standard deviation by a factor of 10, again assuming a flat portion of the variability versus concentration graph. This factor of 10 is directed at achieving a relative standard deviation (RSD) of 10 percent. An advantage of this approach is that a quantitation limit is produced, regardless of what the RSD turns out to be.

For example, it is known that the determination of 2,4-dinitrophenol by EPA Method 625 produces highly variable results and that 10 percent RSD cannot be achieved for this compound. Multiplying the standard deviation of replicate measurements of low-level samples results in a quantitation limit that is considerably higher than the quantitation limits for other compounds analyzed by Method 625. The RSD at this quantitation limit could be 30, 50, or 70 percent. Arbitrarily limiting the quantitation limit to some value (e.g., 30%, as with the ASTM IQE) could prohibit the use of EPA Method 625 for determination of 2,4-dinitrophenol. If 2,4-dinitrophenol were present at high concentration in a discharge, it would not be reported. Although it could be argued that a more precise method should be used for determination of 2,4-dinitrophenol, determination of pollutants by a large suite of different methods would be quite costly with little meaningful benefit. Increasing precision (i.e., decreasing measurement error) would be critical only if the concentration at issue was near a compliance limit.

Another means of arriving at a limiting RSD is to graph RSD versus concentration, as shown in the bottom two graphs of Figure 3-3. This approach is used by the ASTM IQE. It has the advantage that a model is fit to data, rather than using a point estimate such as the Currie and ACS LOD or the EPA ML. However, this approach requires considerably more data than are necessary for approaches based on point estimates. In addition, how a model is selected can play a major role in the outcome.

3.3.4.2 Criteria for Selecting Models

Both statistical and graphical procedures have been proposed for selecting between models for measurement results versus spike concentrations.

Statistical Criteria

While statistical criteria are available for choosing between models of similar types, the currently available criteria are not satisfactory for choosing between the wide variety of models considered for the relationship between measurement variation and spike concentration, based on EPA's studies. More technically, statistical criteria include using: (1) the simplest model to obtain statistical significance, (2) the model with the smallest estimated variability, and (3) the model with the smallest likelihood ratio. Given the wide variety of models considered for detection and quantitation, there are problems associated with each of these procedures. Data that obviously do not follow the model may produce statistically significant results, variability may be estimated with weights that make the various estimates incomparable, and the likelihood function may not be comparable between models.

Graphical Criteria

Graphical criteria may be susceptible to some subjectivity in their application, but they are currently the best available method for choosing between models. At the most basic level, the primary graphical criteria is for the form of the model to be suggested by the available data. To consider the quality of the graphical analysis, it is useful to see if some small number of data are overly influential in determining if a model does or does not fit. Given the ability of the human eye to discern deviations from a straight line rather than from a curved line, a useful technique is to plot the data so that they will indicate a straight line if they follow the model of interest.

3.3.4.3 *Assessment of Current Models*

EPA graphed variability versus concentration data with regard to how real data from measurement methods used under the Clean Water Act would conform to a number of different models. For details of how data sets were selected and how data were collected within the data sets, see Appendix B, *Characterizing Measurement Variability as a Function of Analyte Concentration for a Variety of Analytical Techniques*. Four sets of composite scatter plots for all combinations of analytical technique by analyte by study were produced. These sets include:

1. Measurement versus Spike Concentration,
2. Log Measurement versus Log Spike Concentration,
3. Observed Standard Deviation versus Spike Concentration,
4. Log Standard Deviation versus Log Spike Concentration, and
5. Relative Standard Deviation (RSD) versus Log Spike Concentration.

There are hundreds of scatter plots in each set, sorted by the source, measurement technique, and study. The first set of scatter plots can be used to evaluate how well measurement results match the spiked concentration in the water. If the assumed straight line model is true, then the relationship outlined by the plotted data will be approximately linear. These relationships are plotted using log-log plots so that small deviations from the line can be easily visualized. All the graphs are contained in attachments to Appendix B of this Assessment Document.

The plot of observed standard deviations versus spike concentrations can be used to evaluate the reasonableness of the constant variation and/or linearly increasing variability models (Currie, 1968, Hubaux and Vos, 1970, and Glaser *et al.*, 1981). If the constant model for standard deviation is true, there would be no apparent relationship between the standard deviation and spike concentration. If the straight-line model for standard deviation is true, plots are expected to indicate an approximately linear relationship. Analogously, the standard deviation/spike concentration versus spike concentration is expected to show a straight-line relationship when variability is proportional to the spike concentration (Gibbons *et al.*, 1991). The log-log plots of standard deviation versus spike concentration are expected to indicate if log or square root transformations may be appropriate (Clayton *et al.*, 1987) or to display a shape that approximates a "hockey stick" when it is appropriate to use the model proposed by Rocke and Lorenzato (1995). With the Rocke and Lorenzato model, variability near zero will be approximately constant, but variability will increase proportionally with concentration in the higher concentration range.

The large number of plots make it difficult to draw general conclusions. For the most part, conclusions must be considered on a case-by-case basis. One somewhat general observation is that measurement variability over low concentrations does not appear to fit a particular curvilinear shape, and thus may be considered to be approximately constant in this range for a large number of analytical techniques.

3.3.5 Methodology for Parameter Estimation

Along with various approaches of detection and quantitation and models for measurement, a number of specific procedures have been suggested for estimating model parameters. Maximum likelihood and least squares are two generally applicable statistical methods that can be used in estimating model parameters. There are advantages and disadvantages to both that must be weighed in particular cases. A standard statistical practice for evaluating the quality of an estimation procedure is to calculate the precision and bias, usually best understood by examining a plot of residuals from a fit to a function. All else being equal, the estimation procedure with the greatest precision and least bias is preferred. In some cases, precision and bias can be calculated based on the assumptions behind the estimation procedure. In other cases, it is either necessary or convenient to estimate precision and bias using simulations. From a general theoretical perspective, the maximum likelihood estimation methodology is preferable because it generates estimates that are generally best with regard to properties of precision and bias (especially for larger sample sizes), while also being approximately normally distributed. Unfortunately, maximum likelihood methodology sometimes can be problematic because the method requires the solution of complex equations. Least squares estimation is generally more tractable, and thus is more generally applicable, although the estimates that result may not be as desirable from a theoretical statistical perspective.

What can sometimes be overlooked in considering estimating for model fitting is that direct measurement of variation of the blank or low level concentration may be the most cost-effective and least difficult method to implement. The loss in statistical efficiency in comparison to more elaborate estimation and model fitting methodology would be offset by the relative ease and lower cost.

3.3.6 False Positives and False Negatives

In this section, we discuss the impact of detection, quantitation, and reporting levels on false positive measurement results and false negative measurement results. The definitions of false positives and false negatives are directly related to the concepts of critical value and detection limit used by Currie (1995). These terms were adapted from statistical decision theory to establish the framework for decision making with regard to detection of analytes. The critical value, as defined by Currie, is the point at which the detection decision is made. That is, measured values that are less than the critical value are judged to be "not detected." Measured values that exceed the critical value are judged to be "detected."

The critical value is defined such that when the analyte is not present in a sample, there is a small possibility that a measurement will exceed the critical value. A measurement that indicates the critical value has been exceeded is, therefore, the result of one of two circumstances: (i) the analyte is present in the sample; or (ii) the analyte is not present in the sample and, by chance, the measurement has exceeded the critical value. The occurrence of (ii) is defined as the "false positive" situation. A measurement that is less than the critical value occurs when: (iii) the analyte is not present in the sample; or (iv) the analyte is contained in the sample but the measurement procedure fails to indicate its presence. The occurrence of (iv) is defined as the "false negative" situation.

The following table summarizes the two possible outcomes for each decision: "detected" or "not detected."

<i>Decision</i>	<i>State of the Sample</i>	
	<i>Analyte Present</i>	<i>Analyte Not Present</i>
<i>Detected</i>	Correct (i)	False Positive (ii)
<i>Not Detected</i>	False Negative (iv)	Correct (iii)

As formulated by Currie, the Detection Limit is a value greater than the Critical Value that is used to evaluate the capabilities of analytical procedures. In the terminology of statistical decision theory, the Detection Limit corresponds to a true value referred to as the "Alternative" (see, e.g. *Introduction to Mathematical Statistics*, by Hogg and Craig, 5th edition, [1995]). The Detection Limit is *not* a part of the detection decision process that is applied to individual sample results. The Detection Limit is defined such that when the analyte is present in the sample at a value equal to the Detection Limit, there is a small probability that a measured value will be less than the Critical Value, and thereby result in the false negative decision of "Not Detected."

A common error in many published discussions of false negatives in relation to detection and quantitation (such as the ASTM IDE) is the claim that using Currie's detection limit as a reporting limit or action level will somehow "control" false negatives. That claim is both false and counter-productive. To illustrate the problem with this error, consider the scenario in which the true concentration for the substance of interest in a sample is equal to the critical value. Also assume that measurement variability is approximately normal in distribution throughout the region of concern. The critical level (alpha level) is set to 0.01 or 1% throughout the remainder of this discussion.

If the reporting limit is set equal to the critical value, then given a large number of measurements on the sample, about half of the results will be reported as being measured above the reporting limit, and about half of the measurement results will be reported as being measured below the reporting limit. Measurement results below the reporting limit are treated as if there is no analyte in the sample. These are false negative measurements, and the false negative rate is 50%.

Now set the reporting limit to Currie's *detection limit*. Recall that the true concentration in the sample is equal to the *critical value*. Given a large number of measurements on the sample, about 1% of the measurement results will be reported as being measured above the reporting limit, and 99% of the measurement results will be reported as being measured below the reporting limit. This is antithetical to Currie's formulation of detection.

To illustrate the intent of Currie's detection limit, set the reporting limit equal to Currie's *critical value*, and create a sample with a true concentration equal to Currie's *detection limit*. Given a large number of measurements on this sample, about 99% of the measurement results will be reported as being measured above the reporting limit, and 1% of the measurement results will be reported as being measured below the reporting limit. Knowledge of Currie's *detection limit* can be used to determine if the measurement method meets the needs of a study. For instance, a study concerned with a wastewater treatment technology that is not expected to be effective at concentrations below 10 mg/L may call for a relatively inexpensive measurement method capable of detecting the analyte at 10 mg/L, rather than a more expensive measurement method capable of measuring a hundred times lower.

3.3.7 Statistical Prediction and Tolerance

When we define a critical value, detection limit, or quantitation limit, different descriptive terminology will dictate differences in the numerical value of the limit. We will use a critical value as an example, but the questions motivating detection and quantitation limits can be phrased in similar fashion. Do we want a critical value that tells us how likely it is that:

1. A measurement result was produced by measuring a blank sample,
2. The next measurement result will be produced by measuring a blank sample, or
3. The next [pick any number] of measurement results will be produced by measuring a blank sample?

In statistical terms, these three objectives may be addressed, respectively, by application of methodology for determining:

1. Percentiles;
2. Prediction intervals; and
3. Tolerance intervals.

Percentiles are fairly straight forward to interpret, i.e., they specify the percentage of a distribution that falls below a given percentile value. Prediction and tolerance intervals are, in effect, confidence intervals on percentiles and can be somewhat more difficult to understand and apply. There are many excellent textbook and literature references that present the theory and application of prediction and tolerance intervals such as Hahn and Meeker, *Statistical Intervals*, Wiley, 1991, and Pratt and Gibbons, *Concepts of Non-parametric Theory*, Springer-Verlag, 1981. Hahn and Meeker describe at length the different statistical intervals including their properties, applications, methodology for constructing the intervals. Pratt and Gibbons have an excellent discussion of tolerance intervals that is general in application due to the non-parametric perspective, i.e., no distributional assumptions are required for the results to be valid.

One of the peer reviewers of EPA's reassessment states:

"Tolerance intervals are inappropriate for environmental monitoring. The main issues here are 1) is the true concentration greater than some specified safe of action level, with sufficient confidence, and 2) what interval of possible concentrations is consistent with one or a series of measurements, with a specified degree of confidence? Both are statements about a given sample or series of samples, and not about the hypothetical variability of future estimates. Suppose that one has a sample of 10 observations with mean concentration of 1 ppb and standard deviation of 0.5 ppb. Then the estimated 99% critical level is $(2.326)(0.5) = 1.2$ ppb. One may choose to use a t-score instead of a normal score so that the chance that a future observation will exceed this level is in fact 99%. In this case, the critical level estimate would be $(3.250)(0.5) = 1.6$ ppb. This does actually correspond to a prediction interval for future observations from a zero concentration sample.

"If one asked instead for a 95% confidence interval for the .99 percentage point of the true distribution of measurements (assuming normality) when the true quantity is zero, this can be calculated approximately using a chi-squared distribution and covers the interval (0.9 ppb, 2.4 ppb). It does not, however, make sense to use 2.4 ppb as a threshold, since the chance of a future observation exceeding 2.4 ppb when the true mean concentration is 0 is about .0005, far smaller than the intended false-positive limit of .01." (Rocke, 2002)

Another of the peer reviewers of EPA's reassessment states:

"the operational definition as taken from pp. 5-2/5-3 of

$$MDL = t_{0.99} (df) S$$

does not correspond to a confidence statement that I can interpret.... This should be replaced, although I agree that a number of statistical quantities could be used; this is where the "fray" seems to be most boisterous. (By the way, the TSD, and I, should be more careful in the use of statistical terminology. We both refer often to confidence "intervals," when in fact the quantity of interest is a confidence limit — or tolerance limit, etc. — on some underlying parametric quantity.)...

"If we accept the TSD's argument on p. 3-25 that the practical value of tolerance limits is limited, then the MDL should be viewed as a prediction limit. And if so, it must contain an additional term as per Gibbons (1994, p. 98):

$$t_{0.99} = (df) S \sqrt{1 + \frac{1}{n}}$$

"Also, to reemphasize, the single most problematic issue when developing a detection limit is correction for false negatives. I took from the TSD (in §3.3.6) an implicit emphasis on LC-type values such as the MDL [when correctly calculated, as in (1)], as motivated by an underlying sort of practical/environmental conservatism that essentially removes false negatives from the estimator's development. I am willing to accept this interpretation. I suspect the fray will continue, however, since there seems to be a fair amount of confusion on the issue in the analytical chemistry literature. The bottom line from my reading of the TSD is that, in effect, we are calculating an LC, but using terminology that makes some readers think it's an LD. I can accept the argument that false negative errors are not the critical issue here, and hence that the approach is reasonable (once correct calculations are undertaken). But, the Agency should put forth an effort to overcome this confusion in terminology. (I expect they will ask me how, and in reply I'd suggest emphasizing that an LC calculation is a form of decision limit, not a detection limit. But here I suspect many users will still confuse the terms, or reverse their meaning, or not see the difference, or who knows what else? I don't know how winnable this battle is...)

"One caveat: although I think the prediction limit argument is acceptable, if the use of tolerance limits rather than prediction limits is in fact desired, then Gibbons' (1994, p. 99) presentation or an equivalent approach should be used instead to correct the MDL calculation." (Piegorsch, 2002)

Similarly, Hahn and Meeker describe situations in which the various intervals or limits are appropriate to use. (As noted by the peer reviewer, the terms "intervals" and "limits" are sometimes used interchangeably). They also give examples of the sort of applications that are suitable for each type of limit although the decision to use a particular type of limit in a given application is not determined strictly by theoretical considerations but is also a matter of judgment.

Prediction intervals contain results of future samples from a previously sampled population with a specified level of confidence. Prediction limits are not estimators of parameters such as means or percentiles. For example, a prediction interval may be constructed to contain future sampling results expressed as a mean or standard deviation of a future sample or all of a certain number of individual future sampling results.

Therefore, EPA agrees with the first peer reviewer that the use of tolerance intervals in environmental monitoring is inappropriate. EPA also agrees with the second peer reviewer that there is considerable confusion in the terminology.

3.3.7.1 Tolerance Intervals

Tolerance intervals contain a specified proportion of a population of measured values with a given statistical confidence level. For example, we say that a proportion, P , of a population is contained within the intervals (L_1, L_2) with $(1-\alpha)100\%$ confidence. The lower and upper ends of the interval, L_1 and L_2 , respectively, are referred to as tolerance intervals. A tolerance interval is therefore the endpoint of an interval of random length that is determined on the basis of having a specified probability of $1-\alpha$ that its coverage of the population is at least equal to a specified value P . The quantity $1-\alpha$ is referred to as the confidence level for the interval and P is the minimum proportion of the population contained in the interval. Tolerance intervals are not estimators of values such as a mean or a percentile but rather values that are always guaranteed to be either greater than or less than the desired value at some level of statistical confidence. Pratt and Gibbons discuss this and other properties that affect the utility of tolerance intervals and create difficulties in the interpretation and application of tolerance intervals.

In effect, the determination of what, if any, interval to use is a policy decision. The choice of which kind of interval to use should consider how easy it is to estimate the interval you want under the conditions that exist. As Pratt and Gibbons point out, the interpretation of tolerance intervals (and analogously, prediction intervals) can be problematic, especially when issues of sample size and the choice of confidence level come into play. Pratt and Gibbons cite examples where the interplay of sample size and high percentile and confidence level make tolerance intervals useless.

3.3.7.2 Use of Tolerance and Prediction in Setting Detection and Quantitation Limits

Statistical intervals can, and have by a number of authors, be adapted for use in setting detection and quantitation limits. The basic approach requires a functional definition of detection or quantitation that includes a statistical term or terms. An interval could then be constructed about the statistical term which could be used to assess the detection or quantitation limit, or make an adjustment to a calculated value that would result in the detection or quantitation limit. For example, most detection limit estimators are functionally dependent on an estimate of standard deviation of measurement error. A statistical interval could be constructed about the standard deviation and the length of the interval could be used to assess the detection limit. The end points of the interval could be used as the basis for an adjustment (upward or downward) in the calculated limit.

However, the use of prediction and/or tolerance limits in setting detection and quantitation limits is not an absolute requirement and should be evaluated in the context of specific applications and policy considerations. In practice, the effect of adjustment of detection and quantitation limits by use of prediction and tolerance intervals can be quite large, depending on the amount of data available and the choices of percentiles and confidence levels.

3.3.8 Design of Detection and Quantitation Studies

The issues associated with the design of detection and quantitation studies include: how well a selection of spike concentrations can be used to differentiate between different models for the relationship between measurement results and spike concentrations, how the distance between spike concentrations can impact estimates of detection and quantitation limits, how to reduce the influence of uncontrollable factors in the measurement process (probability design), how complete to make the design factors in terms of the physical measurement process, and how flexible to make the design factors in terms of the physical measurement process.

3.3.8.1 Spike Concentrations and Modeling

If a model under consideration cannot be described by the number of spike concentrations in the design, then it is not possible to tell if the model is appropriate. To take the simplest example, it is not possible to describe the slope of a line associated with linearly increasing variation from a single spike concentration. Two well-spaced spike concentrations would allow you to estimate a slope, but provide you with no idea of the variability of the estimate. Three well-spaced spike concentrations represent the minimum requirement for estimating the linear relationship and the variability of that relationship.

Clayton *et al.* (1987) describe the relationship between the spread of the spike concentrations, the number of spike concentrations, and the number of replicate measurements with regard to estimated variability when a linear model is used. While the specific equation used in their paper does not apply to all models, it indicates principles that do apply. Increasing the number of replicate measurements, increasing the number of spike concentrations, and reducing the spread of the spike concentrations are all expected to reduce estimated variability along with the associated detection and quantitation limits. However, one of the components of variability associated with detection and quantitation is that associated with estimating the calibration relationship. To account for this source of variation, it may be appropriate to cover the entire calibration range. On the other hand, many replicates at a high concentration may improperly weight the data in favor of high detection and quantitation estimates.

3.3.8.2 Probability Design

The process known as randomization is an important statistical consideration in the design and interpretation of experimental studies. Randomization involves the allocation of experimental units to factors and treatments under study according a design determined by probability. Randomization avoids bias and systematic errors that can occur in studies where randomization is not used. Randomization is discussed in classic texts such as *Statistics for Experimenters*, by Box, Hunter, and Hunter, Wiley, 1978.

In studies of measurement methods, randomization can be used in the process of creating spike concentration solutions and the ordering of analyses. However, randomization has practical drawbacks, particularly with regard to studies designed to establish detection or quantitation limits. For example, consider a simple study involving the analyses of samples spiked at five concentrations of the analyte of interest, with five replicates of each sample analyzed. A total of 25 analyses are required for the study, and the analyses of the samples can be organized in a 5 by 5 matrix. A random number is assigned to each block in the matrix, as a means of randomizing the order of the replicates at each concentration.

By virtue of this randomized design, a sample with a high concentration of the analyte of interest may end up being analyzed immediately prior to a sample with a very low concentration of the analyte. Unfortunately, this can lead to problems that result from the "carry-over" of analyte within the instrumentation from one analysis to the next. When carry-over occurs, the apparent concentration of the low-concentration sample can be inflated because some of the high-concentration sample may be carried

into the low-concentration sample 2. In the context of a study designed to establish "how low you can go" (i.e., establishing a detection limit), carry-over of the analyte into a low-concentration sample may compromise the results by inflating the result for low-concentration sample 2, but not inflating the results for other low-concentration samples because the randomized design did not cause them to be analyzed immediately following a high-concentration sample.

Analysts are aware of the potential for carry-over and generally take steps during routine analyses to minimize the chance that it will occur. Examples of steps that can minimize carry-over problems include analyzing "cleaner" samples before "dirtier" samples, and interspersing "blanks" between samples when possible. Obviously, the intentional segregation of low and high concentration samples defeats the purpose of the randomized design. Interspersing blanks between the samples can be effective, as well as blocking similar concentrations together and randomizing blocks. But in order to ensure that the blanks do not have other effects on the results, blanks would be needed between each sample or block analysis, and this would greatly increase the cost of the study (e.g., 25 samples and 24 blanks would be required in case of pure randomization). Therefore, despite the statistical benefits, in practice, randomization of the sample analysis sequence can be difficult to apply in detection and quantitation limit studies.

In the Agency's studies of variability as a function of concentration discussed in Sections 1.3.2.1 - 1.3.2.3 of this document, EPA chose to use a non-random design to avoid carry-over problems and to limit the potential difficulties with measurements at very low concentrations. For example, if there was no instrument response at concentration X , then it would be unlikely that there would be a response at a concentration of $X/2$. In the non-random design, EPA permitted the analyst to stop analyses of ever-lower concentrations, whereas a randomized design would have required that all the samples be analyzed, even when there was no instrumental response for many of those samples.

One of the peer reviewers evaluating EPA's draft version of this Assessment Document commented that the effects of carry-over could have been mitigated by studying variability around the calibration line rather than the mean of the replicates. However, carry-over affects subsequent samples differently. The effect of the carry-over cannot be mitigated, regardless of whether variability is studied around the calibration line or the mean of the replicates, unless the amount of carry-over is known and can be subtracted from the affected (low-concentration) sample. This subtraction has limitations because of error accumulation and because the amount of carry-over cannot be determined precisely without extensive studies at multiple concentrations.

3.3.8.3 *Completeness*

The physical measurement process can be studied using rough approximations or it can be studied more rigorously. A rough approximation could use the available components of a method as applied to convenient samples. A more rigorous study would use a complete, specific, and well-defined measurement method with all sample processing steps. The appropriate level of study will probably depend on the purpose of the study.

Measurement procedures (methods) may be more or less strictly designed. Variability in what is allowed in the procedures may add to variability in the measurement results. To the extent that permutations of a method's procedures are not expected to be used in a particular detection or quantitation study, EPA recommends that this information be included in the report on the study results. While there may be physical/chemical reasons for extrapolating the results of a variability study on one set of procedures to permutations of those procedures, there is no statistical basis for making such an extrapolation. Statistical theory by itself is only able to describe conditions that have been observed. On the other hand, a knowledge of the underlying physics of the measurement process can guide the completeness of the modeling process when statistical procedures fail. For example, the Rocke and

Lorenzato model in the linear or log-log domain may be the best general characterization of a physical measurement process. Therefore, this model can be applied to data to produce a complete answer when statistical procedures fail to deduce the "correct" model.

Table 3-1. Summary of Issues Considered

Issue Considered	Description	Assessment/Relevance to Detection and Quantitation Level Approaches
Blank versus zero concentration	It is not possible to measure 'zero' concentration of a substance. As chemists push detection capabilities closer to 'zero', analysis of blanks is needed establish if the result measured is a true concentration or background contamination.	Useful approaches should address the potential contribution of the blank, both through the design of studies that generate detection and quantitation limits, and through evaluation of the study results.
Lack of instrument response	Instruments do not always produce a response (i.e., sometimes they are not capable of yielding a result at the known concentration of a pollutant.)	Operational systems for detection and quantitation need to take instrument non-response into account.
Matrix effects	If not properly controlled, substances in a particular matrix can interfere with measurements of pollutants, resulting in high or low bias in the measurement.	Procedures for determining detection and quantitation limits should allow for evaluation of data in specific matrices of concern when all efforts to resolve the matrix interferences have been exhausted.
Bias (recovery) correction	Some detection and quantitation approaches apply a correction for bias based on the slope of recovery versus concentration line.	The typical relationships between recovery, precision, and concentration in analytical chemistry show that precision worsens as recovery falls. Therefore, increasing detection and quantitation limits to allow for lowered recovery may, in fact, provide a double correction. Preliminary tests have shown that negative detection limits can be produced when a recovery model has a negative slope.
Measurement quality over the life of a method	Measurements made by a given method may improve in quality over time as analysts gain more experience and/or as improvements are made in the underlying analytical technology.	Detection and quantitation limit approaches should be supported by procedures that will allow individual laboratories and other organizations to affordably characterize the improvements.
Method development and promulgation	CWA Section 304(h) requires EPA to promulgate test methods that are used in certain certification and permitting activities under the Act. If necessary, EPA must develop these methods. EPA also can promulgate methods developed by other organizations if they meet EPA needs. However, organizations presently use a variety of differing approaches for detection and quantitation.	There are real benefits to standardization of detection and quantitation limit approaches, but it is impractical to force outside organizations that develop methods to accept standardization adopted by EPA.
Laboratory-specific performance verification	Laboratories that use a particular method must be capable of demonstrating they can achieve the required detection limit.	This objective can be achieved by establishing a reasonable detection limit based on tests in an experienced laboratory and requiring individual laboratories to confirm they achieve this detection limit.

Table 3-1. Summary of Issues Considered

Issue Considered	Description	Assessment/Relevance to Detection and Quantitation Level Approaches
Effluent guideline development	CWA requires EPA to develop technology-based guidelines and standards concerning the treatment and discharge of pollutants into US waters.	Effluent guidelines and standards are based on existing technological capabilities. Because pollutant limits are derived from statistical analysis of data that captures all sources of variability in the industrial treatment process, including analytical variability, facilities employing well-designed and operated treatment should be capable of achieving technology-based limits.
Development of water quality-based controls	CWA requires EPA and authorized States to implement water quality-based controls when technology-based controls will not sufficiently protect designated uses for a water body. Detection and quantitation limit capabilities are not a consideration when water quality criteria are established, meaning that water quality criteria may be lower than levels that can be reliably measured using available technology.	Ideally, detection and quantitation limit approaches will include procedures that encourage the lowering of detection and quantitation limits so that human health and the environment are protected.
Compliance monitoring	CWA requires EPA and authorized States to issue NPDES permits to facilities that discharge into waters of the U.S., and requires municipalities to issue pre-treatment permits to facilities that discharge into publicly operated treatment systems.	Legal issues must be considered when establishing permit limits. Compliance limits must reflect realistic measurement capabilities.
Non-regulatory studies and monitoring	EPA conducts a variety of non-regulatory studies and monitoring activities the goals of the CWA.	Ideally, detection and quantitation limits should allow reliable detection and measurement of pollutants at levels that could become of environmental concern.
Descriptive versus prescriptive uses of detection and quantitation limits	Descriptive approaches characterize the current performance of a laboratory or laboratories that might use a method to analyze pollutants. Prescriptive approaches define a limit that laboratories must demonstrate they can achieve before practicing a method. The prescriptive approach to detection and quantitation is consistent with EPA's use of other prescriptive performance standards that laboratories must achieve.	The prescriptive approach requires specification of reasonably attainable detection limits. Even so, a prescriptive approach could result in exclusion of laboratories that cannot achieve the limit. A descriptive approach can reflect capabilities of poor laboratories and can, therefore, have the effect of raising detection and quantitation limits to levels that are higher than desired and not consistent with the performance of laboratories that make reasonable efforts to control variability.
Compliance evaluation thresholds	Water quality-based permitting can suggest permit limits that are below the detection or quantitation limits of the most sensitive, approved method.	EPA's draft guidance suggested that the permit limit be established from the water quality standard but that compliance be evaluated at the quantitation limit of the most sensitive method. Permit writers should have the flexibility to use the detection limit, quantitation limit, or other limit as the compliance evaluation threshold so that the environment is protected.

Table 3-1. Summary of Issues Considered

Issue Considered	Description	Assessment/Relevance to Detection and Quantitation Level Approaches
Accepting the procedures of VCSBs	The National Technology Transfer Advancement Act directs EPA to focus on increasing their use of standards published by VCSBs such as ASTM, ISO, etc., when it is consistent with the Agency's mission.	VCSBs use differing approaches to detection and quantitation needs to be considered. EPA would continue to accept analytical methods from VCSBs that may use other detection and quantitation approaches when the methods meet EPA's regulatory needs.
National versus local standards for measurement	CWA authorizes State and local governments to implement provisions of the Act, as long as they do so in a way that is at least as protective of the environment as the national standards established by EPA.	This assessment must recognize the impact of any new or revised detection or quantitation approaches on State and local governments and requirements. Given that state and local governments may use more stringent approaches, adoption of new or modified approaches by EPA as a result of this assessment may have no practical impact if states choose to continue use of their existing approaches.
Cost and implementation issues	The financial and technical resources required to determine detection limit approaches vary widely according the complexity of the procedures involved. Organizations that develop methods typically have greater resources available for determining limits than do organizations that use the methods.	EPA must be sensitive the capabilities of the organizations that develop and that use methods. Data from EPA studies indicate that the true detection/quantitation limits at a given point in time can only be arrived at by running hundreds of replicates. A better alternative would be to identify a simple procedure that yields a reproducible estimate and to allow laboratory-specific adjustment based on actual conditions in the laboratory.
Use of multiple approaches or a single pair of approaches	Analytical methods are used to support multiple applications under CWA, including development of regulatory requirements, compliance monitoring and enforcement, and non-regulatory studies and monitoring. EPA could adopt multiple detection and quantitation level approaches, choosing for each application an approach that is best suited to that need.	Use of a single pair of approaches that meets the needs of all CWA programs is preferable to the adoption of multiple, application-specific approaches. Selection of multiple approaches would likely yield a matrix of approaches and applications that would cause confusion and frustration among regulators, permittees, and the laboratory community. It is also likely that the suitability of each approach to each designated application would be subject to the same level of contention that has been applied to nearly all existing approaches.
Sources of variability	Various known and unknown sources of variability can impact laboratory results. Steps can be taken to control known sources of variability. When such steps are taken, some variability will still exist, and it can be expected that interlaboratory variability will be greater than intralaboratory variability.	The potential impacts of interlaboratory variability must be considered when selecting detection and quantitation limit approaches. Even if prescriptive measures are used to control variability, it should be recognized that some laboratories may achieve lower detection and quantitation limits than others.

Table 3-1. Summary of Issues Considered

Issue Considered	Description	Assessment/Relevance to Detection and Quantitation Level Approaches
Censoring measurement results	Measurement results are often reported as less than some detection, quantitation, or reporting limit (i.e., they are censored below a designated reporting threshold). The primary reason for censoring is to avoid reporting results with a higher degree of uncertainty, based on a policy decision. Although such results may not have the desired level of certainty for most applications, they may be of value to statisticians and modelers who handle large volumes of data.	Although the issue of censoring is important, it should not be a consideration when selecting a detection and quantitation limit approach. The decision to censor data is a data reporting and data use issue, and it will apply regardless of how the detection or quantitation limit is established.
Outliers	Outliers are measurement results that are inconsistent with the vast majority of results. They may arise from random variation or some deviation in the measurement process.	Data sets used for development of detection and quantitation limits, for QC acceptance criteria, and for other purposes, should continue to be screened for outliers.
Selection of statistical models	Method sensitivity is usually established based on measurement variation. Nearly all analytical techniques produce results that can generally be classified according to one of three basic models.	Approaches that rely on collection of large data sets and selection of appropriate models based on graphical analysis of the results have been proposed as a more accurate means of determining detection and quantitation limit approaches than simple models based on the collection of small data sets. Regardless of which model is selected, it should yield a reasonable and reproducible estimates of detection and quantitation. Differences in the cost and complexity of these approaches must be weighed against differences in observed outcome using either approach.
Methodology for parameter estimation	Methods for estimating the goodness of fit of a set of data to a particular model include use of statistical estimation procedures for precision and bias and graphical displays. However, the exact tests to be used may not be detailed well enough in the detection/quantitation limit approach.	Details of estimating goodness of fit must be specified and must be based on actual tests of real-world variability versus concentration data.
False positives and false negatives	If a pollutant is present in a sample, but not measured, the reported result is a "false negative." If a pollutant is not present in a sample, but a positive result is measured and reported, the reported result is a "false positive." The issue of allowing for false negatives is contentious.	A common error in discussions of detection and quantitation limit approaches is that reporting limits can be used to control both the false positive and false negative rates. As long as the reporting limit is the only tool for controlling false positive and false negatives, setting the reporting limit higher only reduces the probability of a false positive at the expense of the false negative rate.

Table 3-1. Summary of Issues Considered

Issue Considered	Description	Assessment/Relevance to Detection and Quantitation Level Approaches
Statistical prediction and tolerance	Currie's original approaches are based on confidence limits. More recently, others have suggested that detection and quantitation approaches should be based on statistical prediction or tolerance intervals because compliance measurements are future samples, whereas the limits are based on available measurements.	Statistical intervals can be, and have been, adapted for use in setting detection and quantitation limits by a number of authors. However, the use of prediction and/or tolerance limits in setting detection and quantitation limits is not a requirement and should be evaluated in the context of specific applications and policy considerations.
Design of detection and quantitation studies	Studies designed to characterize detection and quantitation limits can be affected by the selection of concentrations studied, how well uncontrollable factors in the measurement process are reduced, the degree to which entire measurement process is studied, and the flexibility of the design factors in terms of the physical measurement process.	Resources may be insufficient to support detection/quantitation limit approaches that model variability versus concentration because the selection of concentrations may require iteration when results do not meet their respective criteria.

Chapter 4 Evaluation Criteria

This chapter presents the criteria developed by EPA as a means for selecting acceptable detection and quantitation limit approaches for use in Clean Water Act (CWA) programs. These criteria reflect EPA's careful consideration of the issues identified and discussed in Chapter 3. A total of six criteria were established, and are discussed in Sections 4.1 - 4.6. Table 4-1 at the end of this chapter summarizes the relationship between each issue discussed in Chapter 3 and the criteria discussed in Sections 4.1 - 4.6.

4.1 Criterion 1

Criterion 1: The detection and quantitation limit approaches should be scientifically valid.

The concept of scientific validity is widely accepted but loosely defined. For the purposes of this evaluation, a detection/quantitation approach or methodology will be considered scientifically valid if it meets the following conditions:

- It can be (and has been) tested,
- It has been subjected to peer review and publication,
- The error rate associated with the approach or methodology is either known or can be estimated,
- Standards exist and can be maintained to control its operation (i.e., it is supported by well-defined procedures for use), and
- It has attracted (i.e., achieved) widespread acceptance within a relevant scientific community.

While EPA acknowledges that other measures could be established to demonstrate scientific validity, EPA has adopted the conditions cited because they reflect those discussed by the U.S. Supreme Court as considerations pertaining to assessments of scientific validity when considering the admissibility of expert scientific testimony¹. EPA believes that considerations discussed by the Court as necessary to demonstrate the scientific validity of an expert's reasoning or methodology are equally valid for demonstrating the scientific validity of a detection/quantitation approach.

4.2 Criterion 2

Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

As discussed in Chapter 3 of this Assessment Document, the detection and quantitation limit(s) for an analyte in an analytical method can be established from a single-laboratory study, multiple single-laboratory studies, or an interlaboratory study. Historical methods developed by EPA under Clean Water Act programs, and nearly all methods developed by EPA under Safe Drinking Water Act programs, were developed by EPA's research laboratory in Cincinnati, Ohio. In the course of method development, this single laboratory established detection and quantitation limits. In many instances, these detection and quantitation limits were found to be unrealistic, in that they could not be achieved in many non-research laboratories. However, with time laboratory and method performance, as well as analytical instrumentation improved, making detection and quantitation limits more easily achievable in nearly all laboratories. Therefore, the difficulty created was in initial application of the research methods.

¹*Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579 (1993) and *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999)

In recent years, EPA's Office of Science and Technology has used single-laboratory studies to develop an initial estimate of the detection and quantitation limit for a new or modified method, and has verified these limits in interlaboratory studies or by conducting additional single-laboratory studies in other laboratories.

Voluntary consensus standards bodies (VCSBs) such as ASTM International have historically used interlaboratory studies to establish method performance. Over the past 5 to 10 years, ASTM International has been developing interlaboratory and single-laboratory approaches for detection and quantitation. Whereas the single-laboratory studies at EPA's research laboratory in Cincinnati produce the lowest detection and quantitation limits, approaches such as those published by ASTM International gather all sources of variability to produce the highest detection and quantitation limits. A realistic expectation of method and laboratory performance likely lies somewhere in between.

As noted in Section 3.2.2 of this Assessment Document, laboratory and method performance can be affected by the use of performance criteria that serve as prediction or tolerance limits. Examples of such criteria include measures to demonstrate that a laboratory is producing accurate results at a concentration of interest (i.e., analysis of reference standards or spiked samples), measures to demonstrate that results are not biased by contamination (i.e., analysis of blanks), and measures to demonstrate that the laboratory can achieve the sensitivity required to reliably detect pollutants at low concentrations (i.e., at the detection limit). It is likely that laboratory performance will improve (and variability will be lower) when laboratories are required to meet specified performance criteria in order to report results.

A further consideration concerning routine variability is the means for rejection of outliers. True outliers can occur in laboratory data and some means of resolving outlier issues must be included. Statistical procedures are available for the identification of candidate outlier values. Once a candidate outlier has been identified, evaluation of the value from a chemical analytical perspective (e.g., some procedural error or quality control error has occurred) should be the basis of exclusion of the value from a data set. In cases where no cause for the outlier has been identified, it may be reasonable to reject an outlier on statistical grounds, but every effort should be made to justify the exclusion on technical grounds.

In examining each approach against this criterion, EPA will evaluate whether the approach can be used to provide a realistic expectation of laboratory performance. As part of this assessment, EPA will examine the sources of variability captured by the approach, and the degree to which the statistics that underlie the approach realistically reflect these sources of variability.

4.3 Criterion 3

Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.

Any approach or procedure should be simple, complete, and cost-effective to implement (i.e., it should be reliable and "laboratory-friendly"). The laboratories that can be expected to use detection or quantitation procedures will range from large laboratories and laboratory chains with a wide range of technical capabilities, to "mom and pop" laboratories operated by one or a few people with a limited set of statistical skills. If a procedure is complicated, it will be prone to error in its use. Similarly, if a procedure requires investment of extensive resources that cannot be billed to the client, laboratories will have a disincentive to use the procedure. Therefore, if the Agency wishes to encourage the development and use of innovative techniques that improve measurement performance or lower measurement costs, the Agency must consider practicality and affordability as significant, if not equal, considerations to scientific validity.

After evaluating each of the issues discussed in Chapter 3 of this document, EPA concluded that successful implementation of CWA programs depends on the ability of laboratories to easily and affordably:

1. demonstrate that a method works in a particular matrix at the levels of concern,
2. characterize improvements in measurement capabilities in terms of measurement sensitivity, and
3. characterize the sensitivity of new methods.

A matrix effect is an interference in a measurement that is caused by substances or materials in the sample other than the analyte of interest that are not removed using the procedures in the method or other commonly applied procedures. In the context of detection and quantitation, matrix effects may manifest themselves by precluding measurements at levels as low as could be measured were the interference not present. From a practical perspective, it is not possible to test the sensitivity of each new method in every possible matrix in which it may be used. At a minimum, it is unlikely that EPA or any other organization could possibly identify and obtain samples of every matrix to which the method might be applied, and even if such a feat were possible, the cost and logistics of doing so would be prohibitive. The situation for characterizing matrix effects on analytical sensitivity is similar to the situation for characterizing matrix effects on measurement performance at higher concentration levels. In the latter case, EPA typically uses one or more spiked reference matrices (e.g., reagent water, sand, diatomaceous earth) to establish QC acceptance criteria for real-world matrix samples that are spiked with the analyte of interest at a mid-to-high concentration. Each analytical method includes QC acceptance criteria for such matrix spikes, along with a suite of quality control requirements designed to verify that failures are attributable to the matrix rather than to an analytical system that is out of control. EPA prefers to identify a similar concept that allows for characterization of measurement sensitivity in representative matrices and that is supported by a simple, cost-effective procedure that would allow individual laboratories to evaluate, on an as-needed basis, the effects of specific matrices on measurement sensitivity. Because methods approved at 40 CFR part 136 already contain a suite of quality control procedures and QC acceptance criteria, EPA believes that it is not necessary to verify method sensitivity in each and every batch of each and every matrix analyzed. Rather, such testing could be done only on an as-needed basis when it is suspected that matrix interferences may preclude reliable measurements at low levels.

Another consideration is that measurement capabilities generally improve over time. This is attributable to a variety of factors, including:

1. increased staff experience with a given technique,
 2. technological upgrades or improvements in the instrumentation used for analysis, and
 3. development of new instrumentation or techniques that improves sensitivity, precision, or bias.
- In each case, the improvements may not be observed across the entire laboratory community. In the case of increased staff experience, for example, it is obvious that a laboratory that specializes in one type of analysis, such as low-level mercury measurements, will develop greater experience than a laboratory that rarely performs this measurement. Likewise, it is easy to see how one or a few laboratories that concentrate their business on a particular type of analysis might be willing to invest significant resources in new or upgraded equipment to improve performance, whereas laboratories that rarely perform such analyses would not find such upgrades to be cost-effective.

Improvements in measurement capability, including the development of new methods, may create a dynamic decision-making process, in that measurements at lower levels may allow EPA and states to identify previously undetected pollutants. Such situations offer a means for monitoring and controlling (i.e., regulating) the discharge of previously unregulated, but harmful, pollutants. Therefore, it is in the best interest of the environment for EPA to encourage the development and use of improved environmental analysis procedures and equipment.

In evaluating this criterion, EPA will favor affordable and easy-to-use approaches and procedures that allow analysts in a single laboratory to 1) determine matrix-specific variations when necessary, based on realistic data, and 2) demonstrate lower detection and quantitation limits associated with improvements in their measurement capabilities. Procedures for establishing the sensitivity of new methods or improved measurement capabilities must be practical enough to encourage such development. These procedures should specify the spiking level at which measurements are to be made and the corrective action to be taken if the resulting detection or quantitation limit is inconsistent with the data from which it is derived.

4.4 Criterion 4

Criterion 4: The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.

Any approach to establishing levels at which detection decisions are made should be capable of providing regulators, the regulated community, and data users with a high level of confidence that a pollutant reported as being present really is present. Historically, nearly every approach to making detection decisions has set the criterion for detection at 99 percent confidence (i.e., the lowest level at which a pollutant will be detected with a probability of 99 percent). This criterion results in the probability of a false positive (i.e., that a pollutant will be stated as being present when it actually is not [a Type I error]) of one percent.

In evaluating this criterion, EPA will favor approaches and procedures that reflect routine analytical conditions in a well-operated laboratory. That is, the procedure must be capable of generating a detection level when the substance of interest is not present in a blank and/or when instrument thresholds are adjusted for routine operation.

4.5 Criterion 5

Criterion 5: The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.

Measurement capabilities among laboratories vary depending on a number of factors, including, but not limited to, instrumentation, training, and experience. Similarly, measurement capabilities among different analytical methods vary depending on a number of factors, including the techniques and instrumentation employed and the clarity of the method itself.

Historical approaches to recognizing laboratory capabilities in establishing detection and quantitation limits have varied between two extremes of establishing the limit in a state-of-the-art research laboratory to reflect the lowest possible limit that can be achieved, and establishing the limit based on statistical confidence intervals calculated from a large number of laboratories with varying levels of experience, instrumentation and competence. Generally, use of the former has been employed to serve as a goal or performance standard to be met by other laboratories, whereas use of the latter treats the limit, not as a performance standard that needs to be met by each laboratory, but rather as a characterization of the performance of the capabilities of a population of laboratories at the time of method development.

Historical approaches to recognizing method capabilities also have varied between those that allow the error expressed as relative standard deviation, or RSD among low-level measurements to vary, depending on the capabilities of the method, and those that fix this error (RSD) at a specific level.

Initially, Criterion 5 stated that the *"quantitation limit should identify a concentration at which the reliability of the measured result is consistent with the capabilities of the method when a method is performed by experienced staff in a well-operated laboratory."* Reviewers from within EPA questioned the criterion's implication that measurements below a quantitation limit could be considered unreliable. A similar concern was expressed by one of the peer reviewers charged with evaluating EPA's assessment and an earlier draft of this Assessment Document. This reviewer noted that:

"almost all implementations of limits of quantitation have nothing to do with whether the measurements are actually quantitative," and that *"any level at which the instrument can be read, and at which there is a reliably estimated standard deviation is a level at which quantitation is possible"* (Rocke, 2002)

The peer reviewer suggested that Criterion 5 might be rewritten as:

"the quantitation limit should identify a concentration at which the instrument yields a measurable signal at least 99% of the time, and which is no smaller than the detection level. Such a quantitation limit will often be the same as the detection level."

EPA agrees that this is a valid perspective, in that if the pollutant is identified and the analytical system produces a result, quantitation occurs. Although this interpretation of a quantitation limit has validity, implementation of such an approach would require that all values generated by an analytical system be reported, along with an estimate of the uncertainty associated with each value (e.g., the "reliably estimated standard deviation" mentioned by the peer reviewer). As noted in Section 2.3.4, several organizations, including the European Union, are developing procedures for estimating the uncertainty associated with measured results. If successful, such an approach would eliminate many of the data censoring concerns discussed in Section 3.3.2. Given the difficulty in achieving consensus on an appropriate means of establishing a detection limit, however, EPA believes that it would also be difficult, to obtain consensus on an appropriate means for estimating the uncertainty associated with each result measured on each environmental sample. In addition, analytical chemists have used and believe that they understand a quantitation limit to mean the lowest concentration at which an analyte can be identified and determined with some degree of certainty.

Therefore, EPA prefers to monitor developments by the EU and others on this subject, and if appropriate, re-evaluate this issue if and when it becomes widely accepted by the laboratory, regulatory, and regulated communities. In the meantime, EPA believes that the traditional approach of defining a quantitation limit at some level above the detection limit provides a data user with a reasonable degree of confidence in the measured value without requiring that individual estimates of uncertainty be developed and reported. Criterion 5 reflects this belief.

EPA will evaluate various approaches against this criterion by examining the ease of adjustment of the RSD or other performance measures in the context of the measurement capability of the laboratory or the need to adjust the measurement error to allow for environmental decisions. In evaluating the approaches, EPA will give preference to those approaches that strike a reasonable balance between using either state-of-the art laboratories or a highly varied community of laboratories to establish quantitation limits.

4.6 Criterion 6

Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.

The Clean Water Act requires EPA to conduct, implement, and oversee a variety of data gathering programs. As noted in Section 3.2 of this Assessment Document, these programs include, but are not limited to:

- Survey programs to establish baselines and monitor changes in ambient water quality,
- Screening studies to identify emerging concerns and establish the need for more in-depth assessment,
- Effluent guideline studies to establish technology-based standards for the control of pollutants in wastewater discharges,
- Toxicity and environmental assessment studies to establish water quality-based standards for the control of pollutants in wastewater, and
- Risk assessment studies designed to characterize and evaluate human health and environmental risks associated with various water body uses.

In addition, EPA needs to apply a detection limit or quantitation limit approach to permitting, compliance monitoring, and other uses of the 40 CFR part 136 methods. These applications include:

- Permitting,
- Ambient and effluent compliance monitoring under NPDES and the pretreatment program,
- Ambient and effluent compliance monitoring under state and local programs,
- Quality control in analytical laboratories, and
- Method promulgation.

In theory, EPA could evaluate each of these applications independently and identify a detection and quantitation limit approach that is best suited to each application. However, doing so could potentially result in the need for up to 10 different detection and/or quantitation limit approaches. EPA believes that this would increase confusion, increase record keeping burdens, and increase laboratory testing burdens. For these reasons, EPA believes it is desirable to adopt a single pair of related detection and quantitation procedures that can be used to address all Clean Water Act applications.

EPA also believes that 1) it is unrealistic to expect other organizations, such as the U.S. Geological Survey, the Food and Drug Administration, ASTM International, AOAC-International, etc., to adopt and standardize on the approach selected by EPA for its use in CWA programs, and 2) it is desirable to allow use of approaches and methods developed by these and other organizations to be used in CWA programs. The inclusion of such approaches and methods provides the stakeholder community with increased measurement options that may help reduce measurement costs or improve measurement performance for specific situations. This approach is consistent with EPA's movement towards a performance-based measurement system (PBMS) and with the intent of the National Technology Transfer and Advancement Act (NTTAA). Therefore, although EPA prefers to identify and adopt a single pair of detection and quantitation limit approaches that can meet CWA needs, EPA also believes that any approach should be acceptable for use if it meets all of the criteria established above and fulfills the needs of the specific CWA application in which it should be used.

The Clean Water Act authorizes state or local governments to implement specific aspects of the Act, with the proviso that they do so in a way that is at least as protective (i.e., stringent) as the national

standards put forth by EPA. Therefore, this criterion is intended to ensure that any detection and quantitation limit approach adopted by the Office of Water is sufficiently clear and defined that it allows for comparison with approaches adopted by state or local governments. It is important to note that this criterion does not establish the need for an approach or procedure that is less stringent than those already in use by state or local governments.

Finally, it is important to differentiate between detection and quantitation limit approaches and compliance evaluation thresholds. Detection and quantitation limit approaches pertain to measurement process thresholds. More specifically, a detection limit describes the lowest concentration at which it is possible to determine that a substance is present with some stated confidence, and a quantitation limit describes the lowest concentration at which it is possible to quantify the amount of a substance that is present. In contrast, compliance evaluation thresholds are used to support wastewater discharge limits established in National Pollutant Discharge Elimination System (NPDES) or pretreatment program permits. Such limits are usually expressed as either a maximum concentration of pollutant allowed in the discharge or a maximum mass of pollutant allowed to be discharged in a specific time period.

Ideally, analytical methods are available to allow for detection and quantitation of pollutants at concentrations that are lower than the discharge levels needed to protect or restore the quality of the receiving water. When such measurement capability does not exist, permitting authorities must decide how to incorporate detection and quantitation limits into the discharge permit. Historically, EPA has recommended that in such cases, the permitting authority include the water quality-based limit in the permit, but establish the compliance evaluation threshold at the quantitation limit of the most sensitive available method. However, as with other aspects of the Clean Water Act, state and local governments may adopt permitting and compliance evaluation approaches that are at least as stringent as those put forth by EPA, and some states have preferred to use the detection limit as the compliance evaluation threshold.

In examining each approach against this criterion EPA will consider 1) the applicability of various detection/quantitation approaches to the variety of data gathering decisions that must be made under the CWA, including those that do and those that do not involve compliance monitoring, and 2) the ability of the approaches to support state and local obligations for implementing the CWA.

Table 4-1. Relationship of Issues Considered in Chapter 3 to Evaluation Criteria Established in Chapter 4

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
All sections in Chapter 3	Any approach adopted by EPA must be scientifically valid. Although not explicitly discussed in Chapter 3, the need for scientific validity has been an underlying condition throughout EPA's assessment.	<i>Criterion 1:</i> The concept of scientific validity is widely accepted but loosely defined. For the purposes of establishing scientific validity in this evaluation, EPA has adopted conditions discussed by the U.S. Supreme Court as considerations pertaining to assessments of scientific validity when considering the admissibility of expert scientific testimony. These conditions are that it can be (and has been) tested; it has been subjected to peer review and publication; the error rate associated with the approach or methodology is either known or can be estimated; standards exist and can be maintained to control its operation; and it has attracted (i.e., achieved) widespread acceptance within a relevant scientific community.
3.2.2, Descriptive vs. Prescriptive Uses of Lower Limits to Measurement	In order to protect human health and the environment, EPA must measure pollutants at ever lower concentrations. Establishing stringent standards and a compliance scheme for laboratories is one way to more rapidly develop the ability to measure at these concentrations. A prescriptive strategy concerning detection and quantitation limits would be to: determine these limits at one or more well-operated laboratories; use the performance of these laboratories as the basis to establish limits for the method; and use the established limits as a performance standard that must be demonstrated by laboratories that practice the method. The use of such an approach is consistent with EPA's use of other prescriptive laboratory performance standards and would ensure that prescriptive detection and quantitation limits (i.e., performance standards) reflect the capabilities of a well-performing laboratory or laboratories. This is in contrast to a descriptive approach that would base performance on a population of laboratories that may not be representative of the best possible performance.	<i>Criterion 2:</i> "... laboratory and method performance can be affected by the use of performance criteria that serve as prediction or tolerance limits. Examples of such criteria include measures to demonstrate that a laboratory is producing accurate results at a concentration of interest..., measures to demonstrate that results are not biased by contamination..., and measures to demonstrate that the laboratory can achieve the sensitivity required to reliably detect pollutants at low concentrations (i.e., at the detection limit). It is likely that laboratory performance will be better (and variability will be lower) when laboratories are required to meet specified performance criteria in order to report results." <i>Criterion 4:</i> "In evaluating this criterion, EPA will favor procedures that reflect routine analytical conditions in a well-operated laboratory." <i>Criterion 5:</i> "In evaluating the approaches, EPA will give preference to those approaches that strike a reasonable balance between using state-of-the-art laboratories and a highly varied community of laboratories to establish quantitation limits."
3.3.1, Sources of Variability	There are a number of ways in which variability can be controlled. However, it is not possible to completely eliminate all variability within or between laboratories. Even if prescribed quality control and variability control procedures are in place, it should be recognized that some laboratories may achieve lower detection and quantitation limits than others. The potential effects of sources of variability should be considered when establishing detection and quantitation limit approaches.	<i>Criterion 2:</i> "... laboratory and method performance can be affected by the use of performance criteria that serve as prediction or tolerance levels... In examining each approach against this criterion, EPA will evaluate if the approach can be used to provide a realistic expectation of laboratory performance. As part of this assessment, EPA will examine the sources of variability captured by the approach, and the degree to which the statistics that underlie the approaches realistically reflect these sources of variability."

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.3.7, Statistical Prediction and Tolerance	<p>Percentiles and prediction and tolerance intervals are statistical tools for describing how something that already exists (percentiles) and describing a future occurrence (prediction and tolerance limits). Percentiles are fairly straight forward to interpret, i.e., they specify the percentage of a distribution that fall below a given percentile value. Prediction and tolerance limits are, in effect, confidence limits on percentiles and can be somewhat more difficult to apply... Statistical intervals can, and have by a number of authors, be adapted for use in setting detection and quantitation levels... However, the use of prediction and/or tolerance limits in setting detection and quantitation limits is not an absolute requirement and should be evaluated in the context of specific applications and policy considerations. In practice, the effect of adjustment of detection and quantitation limits by use of prediction and tolerance intervals can be quite large, depending on the amount of available data and the choices of percentiles and confidence levels.</p>	
3.3.8, Design of Detection and Quantitation Studies	<p>Studies designed to characterize sensitivity can be affected by the selection of spiking concentrations in studies, how well uncontrollable factors in the measurement process are reduced, the degree to which the entire measurement process is studied, and the flexibility of the design factors in terms of the physical measurement. Resources may be insufficient to support detection/quantitation limit approaches that model variability versus concentration because the selection of concentrations may require iteration when results do not meet their respective criteria.</p>	<p>Criterion 2: "In examining this criterion, EPA will evaluate if the approach can be used to provide a realistic expectation of laboratory performance. As part of this assessment, EPA will examine the sources of variability captured by the approach, and the degree to which the statistics that underlie the approach realistically reflect these sources of variability."</p>

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.3.3, Outliers	<p>One or more statistical procedures may be used to identify extremely large or small measurement values (outliers). Because extreme values are expected to occur, it is not necessarily appropriate to exclude them from measurement results used to develop detection or quantitation values. Ideally, the analyst's records should be reviewed to establish if an extreme value was caused by failure to follow the method or by some rare event associated with the method. In large detection and quantitation studies, it may not be feasible to review all extreme values to determine if they are outliers. In such cases, removing all extreme values as if they were outliers may be acceptable, but study documentation should state this is the case and the percentage of data removed. Removing large percentages of extreme values may cause variability estimates to be understated, indicate that there are systematic problems with following the method, or indicate that there are problems with the procedure for determining the extreme values.</p>	<p><i>Criterion 2:</i> "A further consideration concerning routine variability is the means for rejection of outliers. True outliers can occur in laboratory data and some means of resolving outlier issues must be included. Statistical procedures are available for the identification of candidate outlier values. Once a candidate outlier has been identified, evaluation of the value from a chemical analytical perspective (e.g., some procedural error or quality control error has occurred) should be the basis of exclusion of the value from a data set. In cases where no cause for the outlier has been identified it may reasonable to reject an outlier on statistical grounds but every effort should be made to justify the exclusion on technical grounds."</p>
3.1.3, Matrix Effects	<p>Reference matrices should be used to establish method detection and quantitation limits. The procedures used to define detection and quantitation limits should allow for evaluation of data collected in particular matrices of concern. Matrix-specific determinations should be used only after all efforts to resolve matrix interferences have been exhausted.</p>	<p><i>Criterion 3:</i> "The reality of environmental analysis is that measurement capabilities generally improve over time. This is attributable to a number of factors... In each case, the improvements may not be observed across the entire laboratory community... In evaluating this criterion, EPA will favor affordable and easy-use procedures that allow analysts in a single laboratory to 1) determine matrix -specific variations based on real data and 2) demonstrate that lower detection and quantitation limit approaches associated with improvements in their measurement capabilities."</p>
3.1.4, Measurement Quality over the Life of a Method	<p>Given that measurement capabilities generally improve over time, EPA believes that detection and quantitation limit approaches should be supported by procedures that will allow individual laboratories and other organizations to affordably characterize such improvements.</p>	
3.2.1.2, Method Performance Verification by a Laboratory	<p>Even where a method describes the sensitivity measured or estimated by the developer or the organization that published the method, some means is needed to demonstrate that given laboratory can achieve sufficient sensitivity to satisfy the regulatory decision (e.g., monitoring compliance).</p>	

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.2.6, Cost and Implementation Issues	The financial and technical resources required to determine detection limit approaches vary widely according to the complexity of the procedures involved. Organizations that develop methods typically have greater resources available for determining limits than do organizations that use the methods. EPA must be sensitive to the capabilities of the organizations that develop and use the methods. Data from EPA studies indicate that the true detection/quantitation limits can only be arrived at by running hundreds of replicates. A better alternative would be to identify a simple procedure that yields a reproducible estimate and to allow laboratory-specific adjustment based on actual conditions in the laboratory.	<p><i>Criterion 3:</i> "Any approach or procedure should be simple, complete, and cost effective to implement. The laboratories that can be expected to use detection/quantitation procedures will range from large laboratories and laboratory chains with a wide range of technical capability to "mom and pop" laboratories operated by one or few people with a limited set of statistical skills. If a procedure is complicated it will be error prone in its use... if a procedure requires investment of extensive resources... laboratories will have a disincentive to use the procedure. Therefore, if the Agency wishes to encourage the development and use of innovative techniques that improve measurement performance or lower measurement cost, the Agency must consider practicality and affordability as significant, if not co-equal, considerations to scientific validity."</p>
3.3.4, Criteria for the Selection and Appropriate Use of Statistical Models	What can be sometimes overlooked in considering estimation for model fitting is that direct measurement of variation of the blank or low-level concentration may be the most cost-effective and least difficult method to implement. The loss in statistical efficiency in comparison to more elaborate estimation and model fitting methodology would be offset by the relative ease and lower cost.	
3.3.6, False Positives and False Negatives	A common error in many published discussions of false negatives in relation to detection and quantitation is the claim that using Currie's detection limit (as opposed to the critical level) as a reporting limit or action level will somehow "control" false negatives. That claim is both false and counter-productive... As long as the only tool for setting requirements for false positive and false negative measurement results is the reporting limit, setting the reporting limit higher reduces the probability of a false positive at the expense of increasing the probability of a false negative.	<p><i>Criterion 4:</i> "Any detection limit approach should be capable of providing regulators, the regulated community, and data users with confidence that a pollutant reported as being present really is present. Historically, nearly every detection approach has set the criterion for detection at 99 percent confidence... This criterion results in the probability of a false positive (i.e., that a pollutant will be stated as being present when it actually is not [a Type 1 error]) of one percent."</p>
3.1.1, Blank vs. Zero Concentration	Useful detection and quantitation limit approaches should address the potential contribution of the blank, through both the design of the study that generates the detection and quantitation limit estimates and evaluation of study results.	<p><i>Criterion 4:</i> "In evaluating this criterion, EPA will favor procedures that reflect routine analytical conditions in a well-operated laboratory. For example, the procedure must be capable of arriving a detection limit when the substance of interest is not found in a blank and/or when instrument thresholds are adjusted for routine operation."</p>
3.1.2, Lack of Instrument Response	Procedures for establishing detection or quantitation limits should take into account the impact of instrument non-response.	

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.3.4, Criteria for the Selection and Appropriate Use of Statistical Models	<ul style="list-style-type: none"> • Method sensitivity is usually established based on measurement variation. Nearly all analytical techniques produce results that can generally be classified according to one of three basic models. • The LOQ advanced by Currie and ACS, and EPA's ML result from multiplying the standard deviation of replicate analyses by a factor of 10. This factor of 10 is directed at achieving a relative standard deviation of 10 percent. An advantage of this approach is that a quantitation limit is produced, regardless of what the RSD turns out to be. Another means of arriving at a limiting RSD is to graph RSD versus concentration. This approach is used by the ASTM International IQE. It has the advantage that a model is fit to data, rather than using a point estimate such as the LOQ or ML. However, it requires considerably more data than approaches based on point estimates, and how a model is selected can play a major role in the outcome. 	<p><i>Criterion 5:</i> "Measurement capabilities among laboratories vary depending on a number of factors, including, but not limited to, instrumentation, training, and experience. Similarly, measurement capabilities among different analytical methods vary depending on a number of factors, including the techniques and instrumentation employed and the clarity of the method itself... Historical approaches to recognizing method capabilities also have varied between those that allow the error expressed as relative standard deviation, or RSD among low-level measurements to vary, depending on the capabilities of the method, and those that fix this error (RSD) at a specific level."</p> <p>"EPA will evaluate various approaches against this criterion by examining the ease of adjustment of the RSD or other performance measure in the context of the measurement capability of the laboratory or the need to adjust measurement error to allow for environmental decisions. In evaluating the approaches, EPA will give preference to those approaches that strike a reasonable balance between using state-of-the-art laboratories and a highly varied community of laboratories to establish quantitation limits."</p>
3.2.1.3, NPDES	<p>The NPDES system serves as the primary means by which EPA, states, and Tribes control point source releases into the nation's waters. Under this system individual facilities are issued NPDES permits that provide limitations on the type, concentration, and volume of pollutants that may be legally discharged. Typically, these pollutant controls reflect technology-based standards. If, however, these technology-based controls are not adequate to protect the water quality standard designated for the facility's receiving water, stricter controls are warranted. In such cases, NPDES permits contain water quality-based controls.</p>	<p><i>Criterion 6:</i> "... it is important to differentiate between detection and quantitation limit approaches and compliance evaluation thresholds. Detection and quantitation limit approaches pertain to measurement process thresholds. More specifically, a detection limit describes the lowest concentration at which it is possible to determine that a substance is present with some stated confidence, and a quantitation limit describes the lowest concentration at which it is possible to quantify the amount of a substance that is present. In contrast, compliance evaluation thresholds are used to support wastewater discharge limits established in National Pollutant Discharge Elimination System (NPDES) or pretreatment program permits."</p>
3.2.3, Compliance Evaluation Thresholds	<p>A situation that arises frequently in addressing water quality-based limits is the setting of the permit limit below the detection or quantitation limit of the most sensitive, approved analytical method. Permit writers should have the flexibility to use the detection limit, the quantitation limit, or other limit as the compliance evaluation threshold so that the environment is protected.</p>	

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.2.4, Accepting the Procedures of Voluntary Consensus Standards Bodies	<ul style="list-style-type: none"> The National Technology Transfer and Advancement Act (NTTAA) encourages Federal agencies to focus on increasing their use of voluntary consensus standards whenever possible, and gives Federal agencies discretion to use other standards where use of voluntary consensus standards would be inconsistent with applicable law or otherwise impractical. Two types of technical standards apply to NTTAA; a performance standard and a prescriptive standard. NTTAA does not direct agencies to favor one type of standard over another. One option is for EPA to employ a performance-based approach to establishing detection and quantitation limits, in which method developers, laboratories, and others would be free to use any one of a variety of approaches to establishing these limits, including the existing MDL procedure, or a VCSB. Thus, establishing method sensitivity would be considered a performance standard under NTTAA, rather than a prescriptive standard. The fact that different approaches (prescriptive standards) yield different answers would be immaterial if EPA evaluates the answers (e.g., the detection limit that is determined) relative to a specific decision (e.g., the regulatory limit for a given pollutant). 	<p><i>Criterion 6:</i> "The Clean Water Act requires EPA to conduct, implement, and oversee a variety of data gathering programs... In addition, EPA needs to apply detection to permitting, compliance monitoring, and other uses of the 40 CFR part 136 methods. These applications include: permitting; ambient and effluent compliance monitoring under NPDES and the pretreatment program; ambient and effluent compliance monitoring under state and local programs; quality control in analytical laboratories; and method promulgation...In theory, EPA could evaluate each of these applications independently and identify a detection and quantitation limit approach that is best suited to each application... EPA believes that such an approach would increase confusion, increase record keeping burdens, and increase laboratory testing burdens. For these reasons, EPA believes it is desirable to adopt a single pair of related detection and quantitation procedures that can be used to address all Clean Water Act applications... In examining each approach against this criterion, EPA will consider 1) the applicability of various detection/quantitation approaches to the variety of data gathering decisions that must be made under the CWA, including those that do and those that do not involve compliance monitoring, and 2) the ability of the approaches to support state and local obligations for implementing the CWA."</p>
3.2.1.1, Method Development and Promulgation	<ul style="list-style-type: none"> EPA believes it would be impractical to force standardization on a single detection or quantitation limit approach on method developers and promulgate only those methods that contain the standardized approach. EPA also believes there are real benefits to standardization and that 1) all new methods developed by EPA for promulgation at 40 CFR part 136 should reflect such standardization, and 2) EPA should strongly encourage outside organizations to include these approaches in their methods. 	

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.2.7, Use of a Pair of Related Detection and Quantitation Procedures	<p>Although EPA could develop a separate detection and quantitation limit approach for each application and attempt to define and evaluate each of the separate approaches, the resulting matrix of approaches would cause confusion to regulators, permittees, and the laboratory community. Further, when proposed, each item in the matrix of approaches and applications would, individually, be subject to contention and second-guessing, and it is likely that the outcome would be nearly the same as if a single pair of approaches is selected. To avoid this outcome, EPA believes it is desirable to use a single pair of related detection and quantitation procedures to meet needs where they exist in all CWA applications.</p>	<p><i>Criterion 6:</i> In theory, EPA could evaluate each of these applications independently and identify a detection and quantitation limit approach that is best suited to each application... EPA believes that such an approach would increase confusion, increase record keeping burdens, and increase laboratory testing burdens. For these reasons, EPA believes it is desirable to adopt a single pair of related detection and quantitation procedures that can be used to address all Clean Water Act applications.</p>
3.2.5, National versus Local Standards for Measurement	<p>CWA authorizes states and local governments to implement permits, with the requirement that they be at least as protective (stringent) as the national standards established by EPA. Thus, EPA must take into account the impact of any revised or new detection/quantitation limit approaches and procedures on state and local governments, as well as on those affected by state and local requirements.</p>	<p><i>Criterion 6:</i> "This criterion will be evaluated by studying ... 2) the ability of the approaches to support state and local obligations for implementing the CWA."</p>
3.3.2; Censoring Measurement Results	<p>Measurement results are often reported as less than some detection, quantitation, or reporting limit (i.e., they are censored below a designated limit). The primary reason for censoring is to avoid reporting highly unreliable results. Although such results may have high measurement error in a relative sense, they are of value to statisticians and modelers who are interested in analysis and modeling of measurement processes.</p>	<p>None. Although the issue of censoring is important, it should not be a consideration when selecting a detection and quantitation limit approach. The decision to censor data is a data reporting and data use issue.</p>

This chapter summarizes EPA's assessment of various detection and quantitation limit approaches against the evaluation criteria established in Chapter 4. Assessments of detection limit approaches are presented in Section 5.1 and include an assessment of the:

- EPA method detection limit (MDL; Section 5.1.1),
- ASTM International interlaboratory detection estimate (IDE; Section 5.1.2),
- American Chemical Society (ACS) limit of detection (LOD; Section 5.1.3),
- International Organization for Standardization/International Union of Pure and Applied Chemistry (ISO/IUPAC) critical value (CRV; Section 5.1.4), and
- ISO/IUPAC minimum detectable value (MDV; Section 5.1.5).

Assessments of quantitation limit approaches are presented in Section 5.2 and include an assessment of the:

- EPA minimum level of quantitation (ML; Section 5.2.1),
- ASTM International interlaboratory quantitation estimate (IQE; Section 5.2.2),
- ACS limit of quantitation (LOQ; Section 5.2.3), and
- ISO/IUPAC LOQ (section 5.2.4).

A brief summary of the evaluation is presented in Tables 5-1 (detection limit approaches) and 5-2 (quantitation limit approaches).

EPA limited the assessment to detection and quantitation limit approaches advanced by ASTM International, ACS, ISO/IUPAC, and EPA, for use in EPA's Clean Water Act (CWA) programs, because these approaches are the most widely published and pertinent.

5.1 Detection Limit Approaches

Sections 5.1.1 through 5.1.5 describe EPA's assessment of five detection limit approaches. Each discussion is divided into two major subsections. The first subsection describes the approach and, where applicable, the procedure that supports the approach. The second subsection details EPA's assessment of the approach based on the five criteria established in Chapter 4 for evaluating detection limit approaches.

Note: Six criteria are given in Chapter 4. Four of these pertain to both detection and quantitation limit approaches. Criterion 4 pertains only to detection limit approaches and Criterion 5 pertains only to quantitation limit approaches. Therefore, the discussions of each detection and quantitation limit approach that follow will omit the criterion that does not apply.

5.1.1 Evaluation of the MDL

Section 5.1.1.1 provides an overview of the MDL approach and the procedures used to implement the approach. Section 5.1.1.2 describes EPA's assessment of the MDL against the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-3, and Criteria 4 and 6).

5.1.1.1 Description of the MDL Approach and Procedure

As promulgated at 40 CFR part 136, Appendix B, the MDL is defined as:

“the minimum concentration of a substance that can be measured and reported with 99% confidence that the analyte concentration is greater than zero and is determined from analysis of a sample in a given matrix containing the analyte.”

A six-step procedure is given in Appendix B, with an optional seventh step to verify the reasonableness of the MDL determined in the first six steps. The procedure is intended for use by experienced analytical chemists. A brief summary of the MDL procedure is as follows:

1. The analyst makes an estimate of the detection limit based on one of four options: the instrument signal to noise ratio; three times the standard deviation of replicate blank measurements; a break in the slope of an instrument calibration curve; or known instrument limitations.
2. The analyst prepares a volume of reagent water that is as free of the target analyte as possible (if the MDL is to be determined in reagent water).
3. The analyst prepares a sufficient volume of spiked reagent water (or of an alternate matrix) to yield seven replicate aliquots that have a concentration of the target analyte that is at least equal to or in the same concentration range as the estimated detection limit (it is recommended that the concentration of the replicate aliquots be between 1 and 5 times the estimated detection limit).
4. All of the replicate aliquots are processed through the entire analytical method.
5. The variance (S^2) and standard deviation (S) of the replicate measurements are determined, as follows:

$$S^2 = \frac{1}{n - 1} \left[\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i \right)^2}{n} \right]$$

$$S = \sqrt{S^2}$$

where:

X_i ; $i=1$ to n , = are the analytical results in the final method reporting units obtained from the n sample aliquots and Σ refers to the sum of the X values from $i=1$ to n .

6. The MDL is then determined by multiplying the standard deviation (S) by the Student's t -statistic at a 99% percentile for $n-1$ degrees of freedom. If seven replicates are used, the Student's t -value is 3.143. This information is used to calculate the MDL as follows:

$$MDL = t_{(n-1, 1-\alpha = 0.99)} (S)$$

where:

MDL = the method detection limit

$t_{(n-1, 1-\alpha = .99)}$ = the Student's *t*-value appropriate for a 99% confidence level with *n*-1 degrees of freedom, and

S = the standard deviation of the replicate analyses.

A 95% confidence interval for the determined MDL may be calculated from percentiles of the chi square over degrees of freedom distribution (χ^2/df).

7. The optional iterative procedure to verify the reasonableness of the MDL involves spiking the matrix at the MDL that was determined in Step 6, and analyzing another seven replicates spiked at this level. The F-ratio of the variances (S^2) is determined and compared with the F-ratio found in the table, which is 3.05. If $S^2_A/S^2_B > 3.05$, the analyst is instructed to respike at the most recently calculated MDL and process the samples through the procedure starting with Step 4. If $S^2_A/S^2_B > 3.05$, then the pooled standard deviation is determined. The pooled standard deviation is then used to calculate the final MDL as follows:

$$MDL = 2.681 \times S_{pooled}$$

where 2.681 is equal to $t_{(12, 1-\alpha = .99)}$.

The 95% confidence interval around the final MDL may be determined using the chi squared over degrees of freedom distribution.

The MDL procedure given at 40 CFR part 136, Appendix B is described as being applicable to 1) a wide variety of sample types, ranging from reagent water containing the analyte of interest to wastewater containing the analyte of interest, and 2) a broad variety of physical and chemical measurements. To accomplish this, the procedure was made device- and instrument-independent.

5.1.1.2 Assessment of the MDL Against the Evaluation Criteria

The following five subsections discuss the MDL approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-4, and Criterion 6).

5.1.1.2.1 Criterion 1: The detection and quantitation limit approaches should be scientifically valid.

For the purposes of evaluating scientific validity, EPA is using the conditions discussed by the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* (1993) and *Kumho Tire Co. v. Carmichael*, (1999) (see Chapter 4, Criterion 1).

Condition 1: It can be (and has been) tested. The MDL procedure meets this condition. The MDL has been used experimentally since 1980 and in a regulatory context since 1984. The MDL procedure is the most widely used and, therefore, the most widely tested detection limit procedure in the history of approaches of detection.

Critics of the MDL have noted that the detection limit produced with the MDL procedure can vary depending on the spike levels used. This would suggest, on the surface, that the MDL procedure can be used to obtain results that do not support the MDL approach. This is a misinterpretation of the MDL

based on the mistaken assumption that spike levels may be arbitrarily selected. In fact, step 1) of the MDL procedure specifies a number of criteria, based on chemical analytical considerations, that must be met in selecting the spike levels (see Section 5.1.1.1, Step 1).

In preparation for the assessment of detection and quantitation approaches, EPA exhaustively tested the MDL procedure with 10 different techniques, at decreasing spike concentrations, to evaluate this concern and determine how well the procedure characterized the region of interest. Results of the study suggest that, although the calculated MDL could vary depending on the spike level used, the procedure was capable of reasonably estimating a detection limit when the full iterative procedure was employed. Given these findings, and the previously noted concern that acceptable spike levels have been subject to misunderstanding, EPA believes that Step 1 of the MDL procedure should be revised to improve reader understanding of appropriate spiking levels, and that the iterative procedure in Step 7 of the MDL procedure should be made mandatory for development or revision of an MDL published in an analytical method.

Condition 2: It has been subjected to peer review and publication. The MDL meets this condition. Prior to promulgation by EPA, the MDL approach and supporting procedure was published by Glaser *et al.* in a peer-reviewed journal (Glaser, *et al.*, 1981).

Condition 3: The error rate associated with the procedure is either known or can be estimated. It is possible to estimate error rates associated with the MDL procedure. It is also possible to calculate confidence intervals about estimated MDLs that are expressions of uncertainty in the estimates. Clarification is in order because the promulgated MDL definition may be somewhat confusing in some respects. In particular, the definition is confusing with regard to whether the MDL is a true concentration or a value estimated from measured data. Another source of confusion lies in terminology. Because the MDL employs the term "detection" and is based on the approaches developed by Currie, it has often been incorrectly assumed to be the equivalent of Currie's "detection limit," when in fact, it is the equivalent of Currie's "critical value," which is the point at which the detection decision is made. EPA believes that the approach of MDL can be clarified by slightly revising the definition as follows:

"The method detection limit (MDL) is an estimate of the measured concentration at which there is 99% confidence that a given analyte is present in a given sample matrix. The MDL is the concentration at which a decision is made regarding whether an analyte is detected by a given analytical method. The MDL is calculated from replicate analyses of a matrix containing the analyte and is functionally analogous to the "critical value" described by Currie (1968, 1995) and the Limit of Detection (LOD) described by the American Chemical Society (MacDougall, et al. 1980, and Keith, et al. 1983)."

Condition 4: Standards exist and can be maintained to control its operation. The MDL approach is supported by a clearly defined, published procedure to control its operation. The procedure gives the steps to be followed and instructs the analyst to use the entire measurement process. Hundreds, if not thousands, of laboratories have successfully implemented the MDL procedure since its promulgation in 1984. EPA has found that when laboratories are required to perform MDL studies as part of an interlaboratory study, the results reported by the laboratories are generally consistent (i.e., within the expected variability). EPA has observed similar consistency in use of the MDL by laboratories required to perform the procedure to demonstrate proficiency with a method. Therefore, the MDL meets this condition.

That said, however, EPA believes that additional guidance can be provided to clarify certain aspects of the MDL procedure, particularly with respect to handling outliers, the optional reasonableness step, and multi-analyte test methods. The MDL procedure contains no discussion of outliers. It may be

helpful to clarify that 1) results should be discarded only if the results are associated with a known error that occurred during analysis (e.g., the replicate was spiked twice) or through a statistically accepted analysis of outliers, and 2) that laboratories should not run more than seven replicates and simply pick the best of the seven results. The optional step involves iterative testing to verify that the determined MDL is reasonable; EPA has observed that few organizations bother to perform this step. EPA also has observed that when a method involves a large number of analytes, it can be difficult to get all analytes to pass the iterative test in the same run. The MDL procedure would benefit from the addition of guidance on how and when to address each of these issues.

In addition, EPA notes that the calculation of the 95% confidence interval described in Step 7 is neither routinely performed by laboratories, nor are the results employed by regulatory agencies, including EPA. Therefore, EPA believes that the MDL procedure could be streamlined by deleting this calculation.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The MDL meets this condition. Within EPA, the MDL has been used by the Office of Research and Development, Office of Science and Technology, Office of Ground Water and Drinking Water, Office of Solid Waste, Office of Emergency and Remedial Response, and other offices. The MDL also has been used outside of EPA in methods published by ASTM International, in *Standard Methods for the Examination of Water and Wastewater*, jointly published by the American Public Health Association (APHA), the American Water Works Association (AWWA), and the Water Environment Federation (WEF), and in methods elsewhere. Although the MDL has been criticized by some, EPA believes that it is the most widely used approach of detection within the environmental chemistry community. Many states incorporate the MDL into NPDES permits, for example, and laboratories often advertise MDLs in their sales literature.

5.1.1.2.2 Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

The MDL procedure is designed to demonstrate laboratory performance with a given method, and can be applied to a broad variety of physical and chemical methods. To accomplish this, the procedure was made device- or instrument- independent. The procedure also recognizes the importance of analyst experience, and explicitly directs the analyst to employ all sample processing and computation steps given in the analytical method when determining the MDL. (All of these aspects are addressed in the MDL procedure published at 40 CFR 136, Appendix B).

When the MDL procedure is followed as intended (i.e., the MDL is determined by an experienced analyst on each device or instrument used for a given method), the demonstrated MDL will include routine variability associated with the laboratory and the method. As noted in the previous section, EPA believes the MDL procedure could be improved by describing appropriate means for the identification and treatment of outliers. Such modifications would ensure that laboratories do not inappropriately discard replicate data when calculating MDLs.

EPA recognizes that one laboratory may obtain detection limits that are lower or higher than those in another laboratory. If the MDL is being determined during method development, it is important to determine the MDL at more than one laboratory to ensure the MDL published in the method reflects demonstrated expectations of method performance in a community of laboratories. EPA does not believe that this community should be so broad as to include the entire universe of possible laboratories that might desire to practice the method. Rather, EPA believes this community should include well-operated laboratories that are experienced with the techniques used in the method and that have some familiarity with the method.

In recent years, EPA's Office of Science and Technology has used single-laboratory studies to develop an initial estimate of the MDL for a new or modified method, and has verified these limits in interlaboratory studies or by conducting additional single-laboratory studies in other laboratories. For example, when EPA initially drafted Method 1631 for measurement of mercury, EPA estimated the MDL to be 0.05 ng/L based on results produced by a contract research laboratory. Additional single-laboratory MDL studies conducted in other laboratories suggested that the MDL should be raised to 0.2 ng/L to better reflect existing capabilities of the measurement community. During EPA's interlaboratory study, each laboratory was asked to conduct an MDL study. Every laboratory in the interlaboratory study met the MDL of 0.2 ng/L, the value published in the promulgated version of Method 1631.

EPA believes that 1) the MDL procedure does address demonstrated expectations of laboratory and method performance, including routine variability, and 2) if the MDL procedure is being employed for method development purposes, it should be performed in multiple laboratories to ensure that it adequately demonstrates expectations in a community of qualified laboratories.

5.1.1.2.3 Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.

The MDL is designed for use by a single laboratory. The promulgated version of the MDL procedure can be performed with as few as seven analyses. If the MDL is to be determined in a matrix other than reagent water, additional analyses will be needed.

Use of the optional iterative procedure would increase the number of analyses by seven each time the procedure is implemented. If the procedure is implemented two times in reagent water, a total of 14 analyses are required. If the procedure is implemented two times in an alternative matrix, EPA estimates that 17-20 analyses may be required, given the possible need to determine the background concentration of the analyte in the alternative matrix. In any of these scenarios, the entire MDL determination can be performed in a single analytical batch (most EPA methods specify batch sizes of 20 samples). As a result, EPA believes that the MDL is among the most affordable of the procedures that have been suggested for determining detection limits. In terms of cost, the only approach that compares favorably with the MDL is the instrument detection limit (IDL). Although most versions of the IDL compare favorably in terms of the number of samples analyzed, the requirement to perform the test on three non-consecutive days has the potential to disrupt routine laboratory operations on three days instead of one. In addition, the IDL does not include sample preparation steps and, therefore, does not completely characterize a method.

5.1.1.2.4 Criterion 4: The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.

EPA believes the MDL meets this condition and refers the reader to the discussion of this subject under Section 5.1.1.2.1, Condition 3.

5.1.1.2.5 Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.

The MDL meets this criterion. The MDL has been successfully applied to a variety of decisions under the CWA since 1984. In addition, many states and others have adopted the MDL in their own programs.

5.1.2 Evaluation of the ASTM International Interlaboratory Detection Estimate (IDE)

The interlaboratory detection estimate (IDE) was developed by ASTM International with support from members of the regulated industry in an attempt to provide a scientifically sound, comprehensive detection limit procedure that addresses the concerns of the regulated industry, of statisticians, and of analysts involved in ASTM Committee D 19 on water.

A brief summary of the procedure is given in Section 5.1.2.1 and Section 5.1.2.2 presents EPA's assessment of the IDE against the five criteria established for evaluating detection limit approaches (i.e., Criteria 1-4, and Criterion 6).

5.1.2.1 Description of the IDE Approach and Procedure

ASTM Designation D 6091 is the *Standard Practice for 99 %/95 % Interlaboratory Detection Estimate (IDE) for Analytical Methods with Negligible Calibration Error*. As stated in the practice:

"The IDE is computed to be the lowest concentration at which there is 90 % confidence that a single measurement from a laboratory selected from the population of qualified laboratories represented in an interlaboratory study will have a true detection probability of at least 95 % and a true nondetection probability of at least 99 % (when measuring a blank sample)."

The IDE is determined and verified using a procedure containing 5 major steps with approximately 53 substeps and conditions. The full text of the IDE procedure is available from ASTM International. The five major steps and their functions are given in Section 6 of the IDE procedure and are as follows:

1. Overview of the procedure.
2. IDE Study Plan, Design, and Protocol - in this section, the task manager (study supervisor) chooses the analyte, matrix, and analytical method. Details are given for range finding; the concentrations to be used in the study; the study protocol (ASTM Practice D 2777 is suggested); the allowable sources of variation; and the number of laboratories, analysts, and days over which the study will be conducted.
3. Conduct the IDE Study, Screen the Data, and Choose a Model - after the study data are collected and screened according to ASTM Practice D 2777, interlaboratory standard deviation (ILSD) versus concentration data are tabulated and one of three models is fit to the data. The first attempt is at fitting a constant model. If the attempt fails, a straight-line model is attempted. If the straight-line model fails, an exponential model is fitted. After fitting, the model is evaluated for reasonableness and lack of fit. If the model fails, the study supervisor determines if a subset of the data should be analyzed or if more data are needed.
4. Compute the IDE - the IDE is computed using the ILSD model selected in Step 3 to estimate the interlaboratory standard deviation at a true concentration of zero and at the IDE, using a mean recovery model to transform measured and true concentrations. The IDE is computed as a one-sided 90 % confidence upper statistical tolerance limit.
5. Nontrivial Amount of Censored Data - this section addresses the effect of "non-detects" or "less-thans." Suggestions are given to see if uncensored data can be obtained from the laboratories or if the

study needs to be augmented with additional data. Suggestions are given for fitting a model to data that contain less than 10 % non-detects or less-thans to produce an IDE.

5.1.2.2 Assessment of the IDE Against the Evaluation Criteria

The following five subsections discuss the IDE approach and procedure in the context of the five evaluation criteria that concern detection limit approaches.

5.1.2.2.1 Criterion 1: *The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. EPA is not aware of any organization, including ASTM International, that has conducted a study to test the procedure as written (i.e., designed and implemented an interlaboratory study that involves estimating an initial IDE [IDE₀] and multilaboratory analyses of multiple concentrations of each matrix of interest surrounding IDE₀). Developers of the approach performed limited testing of the approach on 1) simulated data sets and 2) real-world data sets generated for other purposes. However, these real-world data sets are of limited value for testing the IDE because the concentration ranges associated with the data are above the low-level region of interest. As part of this reassessment, EPA tested a variant of the IDE procedure on single-laboratory data sets designed for characterization of an analytical method in the region of detection. Despite the lack of comprehensive testing, EPA believes that the procedure can be tested, and therefore meets part of this condition. Specifically, the IDE meets the condition that it can be tested, but it only partially meets the condition that it has been tested.

Condition 2: It has been subjected to peer review and publication. Although the IDE has not been published in the peer-reviewed scientific literature, the IDE has undergone extensive review and ballot by members of ASTM Committee D 19, many of whom are qualified peer reviewers. Therefore, although the IDE does not meet this condition in the sense of formal peer review and publication, EPA believes it does meet the intent of this condition (i.e., submission to scrutiny of the scientific community). In addition, the IDE was reviewed by four peer reviewers as part of EPA's assessment of detection and quantitation limit approaches.

Condition 3: The error rate associated with the procedure is either known or can be estimated. In theory, expert statisticians could estimate the error rate of the IDE. However, the IDE procedure is extremely complex from an analytical chemistry and statistical perspective. As a result, it is unlikely that the error rate could be estimated by the typical users of the analytical method to which it would be applied, or even by the typical developers of an analytical method. Moreover, EPA found the model selection procedure to be highly subjective, a situation likely to yield different IDEs from the same data set, depending on the staff involved in performing the calculations. In practice, such conditions make it impossible to estimate the actual error associated with the IDE. Therefore, the IDE fails this condition.

One of the four peer reviewers charged with evaluating EPA's assessment of detection and quantitation limit approaches concurred with EPA's assessment of the IDE, specifically stating, "*I agree that the IDE procedure as outlined is so complex as to make simple determination of error rates associated with it untenable.*" (Piegorsch, 2002)

Condition 4: Standards exist and can be maintained to control its operation. The IDE approach and procedure is supported by a published procedure (standard) to control its operation. The procedure gives the steps to be followed in determining the IDE and instructs the study supervisor how to gather the data and compute an IDE.

However, there are several "gray areas" in the published procedure. The most significant gray area is in the description of model selection. The procedure provides insufficient guidance on use of residual plots to evaluate and select models and, as a result, selection of the model may be very subjective, especially if the number of concentrations is low. The discussion of what model to use after rejecting the exponential and linear model is also very vague. The Rocke and Lorenzato (hybrid) model is mentioned, as well as models with more than one coefficient. Much of the data evaluated by EPA have tended to suggest the exponential model, based on the statistical tests discussed. However, those data have almost always shown residual "patterns" when using this model, which would then lead to consideration of other models. In addition, fitting the constant model is never discussed in detail. Most likely, this is done by simply calculating a mean (weighted if necessary) of the variances from the different concentrations; however, such calculations are never explicitly stated.

Another concern with the standard is that it gives procedures that are inconsistent with procedures given in the IQE standard, even though the two approaches should be consistent for a given analyte with a given method. For example, the exponential model figures prominently in the IDE procedure, where it is one of the three main models discussed. The Rocke and Lorenzato model is not discussed in the IDE procedure, but it figures prominently in the IQE procedure. In theory, a single model should support the definition of both the detection and quantitation limits for a given analyte by a given method. As another example, the IDE procedure includes a multiplier to account for bias in estimating the true standard deviation with the sample standard deviation, but the IQE does not.

Finally, the procedure contains statistical errors that, if followed as written, could produce inaccurate IDE values. For example, Table 1 of the procedure contains "Computations to Estimate Straight-Line Model Coefficients by Means of Least Squares— Ordinary and Weighted," but the weighted least squares formulae given in the table are incorrect. The formulae for the weighted means of the spike values and results given in Table 1 of D6091 would only be appropriate if the weighting were done based on the number of replicates per spike level, rather than on the estimated variance calculated using the chosen standard deviation model.

In conclusion, EPA believes that although the IDE is supported by a published procedure, that procedure will not control its operation because of the degree of subjectivity involved implementing the procedure, errors in the procedure, and inconsistencies with its IQE counterpart. Therefore, the IDE fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The IDE fails this condition because it is only familiar to, and has been accepted by, a very narrow segment of the scientific community. Although the IDE has been approved by ASTM for more than 5 years, EPA is not aware of an IDE that has been published in the open literature or in an analytical method, including an ASTM method.

5.1.2.2.2 Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

The IDE procedure is designed to reflect expectations of interlaboratory performance, including routine variability. The procedure contains extensive instructions for dealing with unusual conditions, including sources of variability and outliers. However, EPA studies of a single-laboratory variant of the procedure suggested that the procedure may not always work as intended. For example, model selection based upon hypothesis tests (as described in D6091, Section 6.3.3.2) almost always indicated that the exponential model should be used, even when the data seemed to be show constant or approximately linear error, while examination of residual plot indicated "systematic behavior" (i.e., non-random deviations from the model) for the exponential and linear models. Another concern with the IDE

procedure is that use of the non-mandatory appendices in ASTM D 6512 to determine the fit of a model may produce results that differ from those that would be obtained by using the default procedures for testing model fit that are built into off-the-shelf statistical software. Such observations, along with the concerns described in Section 5.1.2.2.1, condition 4, lead EPA to believe that, while the IDE approach addresses demonstrated expectations of laboratory and method performance, the IDE procedure does not adequately do so. Therefore, the IDE only partially meets this criterion.

5.1.2.2.3 Criterion 3: *The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.*

The IDE procedure is designed for use by an ASTM International study supervisor or task manager and not as a procedure that a single laboratory can use to evaluate method performance. EPA is aware that ASTM Committee D 19 is developing a Within-laboratory Detection Estimate (WDE), but the WDE is presently only in the formative stages. The WDE may meet this criterion, but the IDE does not.

Regarding cost, the IDE procedure would be the most costly of the procedures that EPA has evaluated because of the time it would take to understand and implement the procedure, and requirements for: 1) estimation of IDE₀, 2) interlaboratory data, 3) extensive statistical intervention in determining the correct model, and 4) possible reanalyses if the resulting IDE does not meet the criteria in the procedure.

5.1.2.2.4 Criterion 4: *The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.*

By definition, the IDE is designed to achieve "a true detection probability of at least 95 % and a true nondetection probability of at least 99 %." Although the 99% probability of a "true nondetection" is equivalent to the 99% confidence that the substance is actually present given in Criterion 4, ASTM International also included the simultaneous requirement for a 95% probability of a "true detection." The developers are using the IDE as a means to control the rates of both false positive and false negative results, in essence, making the IDE analogous by definition and formulaic construction to the *detection limit* (DL) defined by Currie (1968). The IDE accomplishes this goal by using a tolerance limit that increases the IDE well above the point at which the detection decision would be made. For a discussion of this issue, see Sections 3.3.6 (false positives and false negatives) and 3.3.7 (prediction and tolerance intervals) in Chapter 3 of this document.

As noted in Section 2.1 of Chapter 2 of this document, Currie (1968) used the term *detection limit* (subsequently termed the *minimum detectable value*) to refer to a true concentration that has a high probability of generating measured values greater than the critical value. That is, measurements on samples that contain concentrations equal to the *detection limit* have a high probability of exceeding the *critical value* and are, therefore, unlikely to result in a decision that the substance is not detected in the sample. However, the *detection decision* is made on the basis of comparing sample measurements to the *critical value*. With regard to his definition of the "detection limit," Currie (1995) states "The single, most important application of the detection limit is for planning."

When the allowance for false negatives and the prediction and tolerance limits are taken into account, the resulting IDE is raised to the point at which the probability of a false positive is less than 0.00000001 (10⁻⁸). This protection against false positive results is excessive and would yield numerical values of little practical value for making the detection decision.

5.1.2.2.5 **Criterion 6:** *Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.*

EPA's comparison of detection limits produced by various detection limit approaches shows that the median IDE is considerably higher than ACS, ISO/IUPAC, and EPA detection limits. Although the IDE could be applied to some decisions to be made under CWA, it may not support decisions when pollutant levels need to be protective of human health and the environment because the IDE is an implementation of Currie *detection level* or *minimum detectable value*, and may be considerably higher than these levels. At best, the IDE only partially meets this criterion.

5.1.3 Evaluation of the ACS Limit of Detection

The limit of detection (LOD) was developed by the Committee on Environmental Improvement (CEI) of the American Chemical Society (ACS). ACS is a professional society for chemists and other scientists and the publisher of a number of scientific journals. It is not a voluntary consensus standards body (VCSB), nor does it develop or publish analytical methods. In 1978, the ACS/CEI began addressing concerns about the lack of useful standards for interlaboratory comparisons. In 1980, the Committee published its *"Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry"* (MacDougall, *et al.*, 1980), which included the approaches of the LOD and the limit of quantitation (LOQ).

5.1.3.1 Description of the ACS LOD

The 1980 *"Guidelines"* define the LOD as:

"... the lowest concentration of an analyte that the analytical process can reliably detect. ... The LOD in most instrumental methods is based on the relationship between the gross analyte signal S_t , the field blank S_b , and the variability in the field blank σ_b ."

and construct the formal relations using the equation:

$$S_t - S_b \geq K_d \sigma$$

where K_d is a constant. ACS recommended a minimal value of 3 for K_d . Thus, the LOD is 3σ above the gross blank signal, S_b . In the 1980 publication, the ACS stated that at $K_d = 3$, there is a 7% risk of false negatives and false positives. Given that the LOD is 3σ above the blank, however, EPA believes that the risk of false positives is somewhat less than 1%.

In 1983, the ACS Committee published *"Principles of Environmental Analysis"* (Keith *et al.*, 1983). That publication occurred after the 1981 paper on the Method Detection Limit (MDL), and ACS/CEI stated that the LOD is numerically equivalent to the MDL as S_b approaches zero. However, neither the 1980 nor 1983 ACS publications provide a specific procedure for estimating the LOD, nor do they provide a minimum number of observations needed to estimate the gross blank signal or the variability term σ_b .

5.1.3.2 Assessment of the LOD Against the Evaluation Criteria

The following five subsections discuss the LOD approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-4, and Criterion 6).

5.1.3.2.1 Criterion 1: The detection and quantitation limit approaches should be scientifically valid.

Condition 1: It can be (and has been) tested. Testing of the ACS LOD is hampered by 1) the lack of a supporting procedure for establishing an LOD, and 2) it's conceptual dependence on the variability associated with measuring blanks. For example, there is no procedure to govern the minimum number of analyses needed to characterize the variability of a blank sample. Because many environmental chemistry techniques yield a zero, or possibly even negative, value when a blank sample is analyzed, and because the LOD approach is based on the standard deviation of these results, directly testing the LOD in such techniques will yield a zero or negative value. One solution for testing is to rely on ACS' 1983 statement that the LOD is conceptually equivalent to the MDL as the blank signal approaches zero, and employ the MDL procedure as a means for indirectly testing the LOD approach. EPA believes that use of the MDL procedure is a viable means for testing the approach; therefore, the LOD meets this condition.

Condition 2: It has been subjected to peer review and publication. The LOD definition was published in the peer-reviewed journal *Analytical Chemistry* in 1980 and 1983. Therefore, the LOD meets this condition.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The error rates can be estimated, so the LOD meets this condition. The error rate for both false positives and false negatives is stated to be 7 % in the 1980 *Analytical Chemistry* article. However, EPA believes that, because the LOD is stated to be 3 times the standard deviation of replicate measurements of a blank, the false positive rate is overstated and is actually somewhat less than 1 % whereas the false negative rate depends on the true concentration in the sample.

Condition 4: Standards exist and can be maintained to control its operation. The LOD lacks a clearly defined procedure for estimating the important terms required to derive it. Although it may be possible to derive LOD values from data used to derive EPA MDL values, there is no procedure giving explicit instructions on the use of replicate blanks, replicate spiked samples, or a minimum recommendation for the number of replicates. Therefore, the LOD fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Because ACS does not develop and publish analytical methods, it is difficult to determine the degree of acceptance of the LOD. EPA has not specifically investigated the numbers of papers published in ACS journals that include LOD values, and EPA's literature search for detection and quantitation approaches did not uncover a large number of citations that promote the LOD in particular. However, ACS LOD values have appeared in the technical literature. Given that ACS is a relevant scientific community, and that use of the LOD has appeared in the technical literature, EPA believes the LOD meets this condition.

5.1.3.2.2 Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

The LOD approach is designed to address demonstrated expectations of laboratory and method performance, including routine variability, and thus appears to meet this criterion. Unfortunately, ACS has not published a procedure to implement the approach. In other words, the LOD addresses demonstrated expectations of laboratory and method performance in theory, but in practice, provides no direct means for performing these demonstrations. Therefore, EPA believes the ACS LOD only partially meets this criterion.

5.1.3.2.3 Criterion 3: *The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.*

The ACS LOD approach is not supported by a clearly defined procedure for establishing the LOD. Therefore, it fails this criterion.

5.1.3.2.4 Criterion 4: *The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.*

The 1983 publication associated the LOD with the "99% confidence level when the difference ($S_t - S_b$) $> 3\sigma$." Therefore, the LOD satisfies this criterion.

5.1.3.2.5 Criterion 6: *The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.*

In the absence of a procedure for determining LOD values, the ACS LOD fails to meet this criterion because it cannot be used in a regulatory context. The LOD passes only if it is assumed to be functionally equivalent to the MDL (i.e., the MDL procedure is used to establish an LOD).

5.1.4 Evaluation of the IUPAC/ISO Critical Value (CRV)

The critical value (CRV) was developed by the International Union of Pure and Applied Chemistry (IUPAC) and the International Organization for Standardization (ISO). IUPAC and ISO are professional societies for chemists and other scientists. ISO develops and publishes analytical methods through its Task Groups. In 1995, Lloyd Currie of the National Institute for Standards and Technology (NIST; formerly the National Bureau of Standards) published a signature discussion of IUPAC approaches for detection and quantitation (*Pure and Appl. Chem.* 67:10, 1699-1722). Although refined during the intervening years (see Currie, L.A., *J. Radiochem. And Nuclear Chem.* 245:1, 145-156, 2000), the CRV approach remains basically as described in 1995.

5.1.4.1 Description of the ISO/IUPAC Critical Value (CRV) Approach and Procedure

The 1995 article states that the critical value (L_c) is:

"... the minimum significant value of an estimated net signal or concentration, applied as a discriminator against background noise. This corresponds to a 1-sided significance test."

For a normal distribution with known variance, L_c reduces to:

$$L_c = z_{(1-\alpha)}\sigma_0$$

where:

$1-\alpha$ is the false positive error rate, recommended at 5 % ($\alpha = 0.05$), and σ_0 is the standard deviation at zero concentration

If σ_0 is estimated by s_0 (replicate measurements of a blank), $z_{(1-\alpha)}$ is replaced by the Student's t -value. For 7 replicates (6 degrees of freedom), the Student's t -value is 1.943, where $\alpha = 0.05$.

5.1.4.2 Assessment of the CRV Against the Evaluation Criteria

The following five subsections discuss the CRV approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-4, and Criterion 6).

5.1.4.2.1 Criterion 1: *The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. The lack of a supporting procedure for establishing the CRV, coupled with its conceptual dependence on the variability of blank measurements makes testing of the approach difficult. For example, if blank measurements fail to produce a response, it is impossible to calculate a CRV because the standard deviation of zero is zero. One solution for testing the approach is to assume that the CRV is functionally equivalent to the MDL as the blank signal approaches zero, and use a slightly modified version of the MDL procedure to test the CRV approach. The slight modification involves selecting a Student's t -value based on $\alpha = 0.05$ instead of $\alpha = 0.01$, for $n-1$ degrees of freedom. EPA believes this is a reasonable assumption, and therefore, that the MDL procedure is a viable means for testing the CRV approach. Therefore, the CRV meets this condition.

Condition 2: It has been subjected to peer review and publication. The IUPAC/ISO definitions meet this criterion. Moreover, it is likely that these definitions have received greater peer review than any of the other approaches.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The error rate is specified by α , with a suggested value of 0.05 (5%). Therefore, the CRV meets this condition.

Condition 4: Standards exist and can be maintained to control its operation. The CRV is defined in the various publications by Currie. However, EPA's search of the literature and the ISO web site found no standard for control of the approach. Therefore, the CRV fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Because IUPAC and ISO are international bodies, it is difficult to determine the degree of acceptance of the CRV in the U.S. and the world community. EPA has not specifically investigated the number of papers in published journals that include CRV values, but EPA's literature search for detection and quantitation approaches did not uncover a large number of citations that promote the CRV in particular. Therefore, it is difficult to determine if the CRV meets this condition.

5.1.4.2.2 Criterion 2: *The approach should address demonstrated expectations of laboratory and method performance, including routine variability.*

The CRV approach is designed to account for the variability of measurements of the blank in the context of a "chemical measurement process" (method). Unfortunately, neither ISO, IUPAC, nor Currie have published a procedure to implement the approach. As a result, the CRV addresses demonstrated expectations of laboratory and method performance in theory, but in practice, provides no direct means for performing these demonstrations. Therefore, EPA believes the CRV partially meets this criterion.

5.1.4.2.3 Criterion 3: *The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance*

The CRV approach is not supported by a clearly defined procedure for establishing a CRV. Therefore, the CRV fails this criterion.

5.1.4.2.4 Criterion 4: *The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.*

Although the CRV suggests $\alpha = 0.05$, resulting in $1-\alpha$ of 0.95 or 95 % probability of detection, the approach allows for the specification of other probabilities. Therefore, the CRV satisfies this criterion.

5.1.4.2.5 Criterion 6: *Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.*

In the absence of a procedure for establishing CRVs, the CRV approach fails to meet this criterion because it cannot be used in a regulatory context. The CRV passes only if it is assumed to be functionally equivalent to an MDL determined with α set at 0.05 instead of 0.01 (i.e., if the MDL procedure, with $\alpha = 0.05$, is used to establish a CRV).

5.1.5 Evaluation of the IUPAC/ISO Detection Limit

The detection limit or minimum detectable value (MDV) was developed by IUPAC/ISO and published in the same papers as the CRV (Section 5.1.4)

5.1.5.1 Description of the IUPAC/ISO Detection Limit Procedure

The 1995 publications define the minimum detectable value (detection limit) as follows:

"The Minimum Detectable Value (MDV) ... [is] ... the net signal (or concentration) of that value (L_D) for which the false negative error is β , given L_C (or α)." (see the CRV for L_C)

For a normal distribution with known variance, L_D reduces to:

$$L_D = z_{(1-\beta)} \sigma_D$$

where:

z is the state variable

$1-\beta$ is the false negative error rate, recommended at 5 % ($\beta = 0.05$), and

σ_D is the standard deviation at the detection limit

Earlier publications refer to the minimum detectable value as the detection limit. To avoid confusion in terminology and to help distinguish the ISO/IUPAC approach from the MDL, LOD, and CRV, EPA will refer to the ISO/IUPAC detection limit as the Minimum Detectable Value, abbreviated as MDV.

5.1.5.2 Assessment of the ISO/IUPAC MDV Against the Evaluation Criteria

The following five subsections discuss the ISO/IUPAC MDV approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-4, and Criterion 6).

5.1.5.2.1 Criterion 1: The detection and quantitation limit approaches should be scientifically valid.

Condition 1: It can be (and has been) tested. The lack of a supporting procedure for establishing the MDV makes testing of the approach difficult. However, EPA believes that the MDV can be tested using data similar to those used to generate MDL values. Therefore, the MDV meets this condition.

Condition 2: It has been subjected to peer review and publication. The IUPAC/ISO definitions meet this condition; moreover, it is likely that this definition has received greater peer review than any of the other approaches.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The error rates are specified by α and β , both with suggested values of 0.05 (5 %). Therefore, the error rate is known.

Condition 4: Standards exist and can be maintained to control its operation. The MDV is defined in the various publications by Currie. However, EPA's search of the literature and the ISO web site found no standard for control of the approach. Therefore, the MDV fails this criterion.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Because IUPAC and ISO are international bodies, it is difficult to determine the degree of acceptance of the MDV in the U.S. and the world community. EPA has not specifically investigated the number of papers in published journals that include MDV values, but EPA's literature search for detection and quantitation approaches did not uncover a large number of citations that promote the MDV in particular. Therefore, it is difficult to determine if the CRV meets this criterion.

5.1.5.2.2 Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

The MDV approach is designed to account for the variability of measurements of the blank in the context of a "chemical measurement process" in the sense that it is used in concert with a critical value that is based on blank measurement variability. The MDV is the true concentration that is used in the planning of method evaluation and development. The actual detection decision is made at the critical value (CRV) which is determined from measured values. The approach of a true concentration MDV and its associated allowance for false negatives is of little practical value in making the actual detection decision. Therefore, the MDV fails this criterion. The allowance for false negatives in a regulatory context is discussed in greater detail in Chapter 3.

5.1.5.2.3 Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance

The MDV approach is not supported by a clearly defined procedure for establishing MDV values. Therefore, the MDV fails this criterion.

5.1.5.2.4 Criterion 4: *The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.*

The allowance for false negatives raises the probability of detection to a value estimated to be greater than 99.999999 % (probability of a false positive less than 10^{-8}). This protection against false positive results is excessive and would yield numerical values of little practical value for making the detection decision. Perhaps more importantly, as noted by Currie (1995) and discussed in Section 5.1.2.2.4 of this document, the *detection decision* is made on the basis of comparing sample measurements to the *critical value*. Therefore, the MDV fails this criterion.

5.1.5.2.5 Criterion 6: *Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government*

In the absence of a procedure for establishing MDV values, the MDV approach fails to meet this criterion because it cannot be used in a regulatory context.

5.2 Quantitation Limit Approaches

Sections 5.2.1 through 5.2.4 describe EPA's assessment of four quantitation limit approaches. Each discussion is divided into two major subsections. The first subsection describes the approach and, where applicable, the procedure that supports the approach, and the second subsection details EPA's assessment of the approach based on the five criteria established in Chapter 4 for evaluating quantitation limit approaches.

Note: Six criteria are given in Chapter 4. Four of these pertain to both detection and quantitation limit approaches. Criterion 4 pertains only to detection limit approaches and Criterion 5 pertains only to quantitation limit approaches. Therefore, the discussions of each detection and quantitation limit approach that follow will omit the criterion that does not apply.

5.2.1 Assessment of the EPA Minimum level of Quantitation (ML)

Section 5.2.2.1 provides an overview of the ML approach and the procedures used to implement the approach. Section 5.2.2.2 contains EPA's assessment of the ML against the five evaluation criteria that concern quantitation limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.1.1 Description of the ML Approach and Procedures

The present definition of the ML includes a statement of the approach and the procedures used to establish the ML. This definition states that the ML is:

"the lowest level at which the entire analytical system must give a recognizable signal and acceptable calibration point for the analyte. It is equivalent to the concentration of the lowest calibration standard, assuming that all method-specified sample weights, volumes, and clean up procedures have been employed. The ML is calculated by multiplying the MDL by 3.18 and rounding the results to the number nearest to $(1, 2, \text{ or } 5) \times 10^n$, where n is an integer."

The ML is designed to provide a practical embodiment of the quantification level proposed by Currie and adopted by IUPAC. It is functionally analogous to Currie's "determination limit" (described in Chapter 2, Section 2.1) and the American Chemical Society's Limit of Quantitation (LOQ). The LOQ is discussed in Section 5.2.3 of this chapter. Chapter 2 (Section 2.2.2) describes the ML approach in additional detail.

The first part of the ML definition (i.e., the lowest level at which the system gives a recognizable signal and acceptable calibration point for the analyte) ties the quantification limit to the capabilities of the measurement system. The second part of the ML definition provides a procedural means for establishing the ML.

The procedural component of the definition is designed to yield an ML value that equals approximately 10 times the standard deviation of replicate analyses used to determine the MDL. (The exact value corresponding to 10 times the standard deviation is rounded to avoid error that would arise from preparation of calibration standards at exact, unrounded concentrations.) The procedure given in the above definition assumes that exactly seven replicates are used to determine the MDL. EPA has observed, however, that laboratories occasionally perform MDL studies with more than the required minimum of seven replicates. When this is done, the Student's *t*-value used to calculate the MDL should be adjusted accordingly. Similarly, the Student's *t*-value would need to be adjusted when a laboratory performs the optional iterative test described in Step 7 of the MDL procedure, or if outlier testing results in the use of less than seven replicates to establish the MDL. Therefore, EPA believes that the ML definition should be revised to eliminate the assumption of seven replicates and clarify its functional equivalence to Currie's critical value and ACS' LOQ. In addition, a detailed procedure should be developed to ensure proper calculation of the ML when more than seven replicates are used to establish the MDL or when iterative testing is used to establish the MDL.

5.2.1.2 Assessment of the ML against the Evaluation Criteria

The following five subsections discuss the ML approach and procedure in the context of the five evaluation criteria that concern quantitation limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.1.2.1 Criterion 1: The detection and quantitation limit approaches should be scientifically valid.

Condition 1: It can be (and has been) tested. The ML meets this condition. The ML has been used experimentally since 1979 and in the regulatory context since 1984. The ML is tested each time a laboratory calibrates an instrument because methods that include the ML require that it be included as the lowest non-zero standard in these calibrations.

Moreover, EPA exhaustively tested the MDL and ML procedure with 10 different techniques at decreasing spike concentrations to evaluate how well the MDL and ML procedures characterized the region of interest in preparation for this reassessment of detection and quantitation limit approaches. Results of the study suggest that 1) although the calculated MDL and ML could vary depending on the spike level used, the procedure was capable of reasonably estimating detection and quantitation limits when the full iterative MDL procedure was employed, and 2) the rounding process employed to determine the ML generally yielded consistent MLs even with slight variations in the calculated MDL.

In other words, if the procedure for establishing an ML is properly implemented for a given method, it will yield an ML value that is consistent with the approach, and this ML value will be verified (tested) by a laboratory each and every time it calibrates the instrument used to analyze samples by the method.

Condition 2: It has been subjected to peer review and publication. The ML has not been published in a peer reviewed journal. However, it was evaluated by four peer reviewers as part of EPA's assessment of detection and quantitation limits. These reviewers noted that:

"The MDL and ML concepts evaluated in Section 5.1.1 and 5.2.1, respectively, are shown in this evaluation to be technically sound and practical." (Wait, 2002)

"With respect to the limit of quantitation concept, the EPA ML is as good as any of the others given..." (Rocke, 2002)

"The MDL and ML have stood the test of time and provide a proven methodology which meets evaluation criteria stated in the TSD." (Cooke, 2002).

In addition, the present definition of the ML describes the approach and the procedures used to establish the ML. This definition is included in EPA Method 1631, which was extensively peer reviewed in accordance with EPA policies on peer review prior to publication and promulgation. Given that EPA's policies on peer review are as stringent as or more stringent than those used by many published journals, EPA believes that the ML has met a high standard of scientific review and scrutiny, and therefore, meets the intent of this condition.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The uncertainty associated with any ML value can be calculated. EPA performed such calculations during this assessment and found that, on average across all techniques tested, the relative standard deviation of replicate measurements at the ML was approximately 7%. Median RSD values calculated for each multi-analyte method tested ranged from 6 to 14 percent. RSD values calculated for each single-analyte method tested ranged from 4 to 16 percent. (See Appendix C to this Assessment Document for a detailed discussion and presentation of results.)

Condition 4: Standards exist and can be maintained to control its operation. The ML meets this criterion. Detailed procedures (i.e., standards) for establishing the ML are given in the definition itself, although, as noted above, EPA believes that a detailed, stand-alone procedure should be created to ensure that the ML is properly calculated when other than seven replicates are used in its determination.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. EPA believes the ML meets this condition. The ML is functionally analogous to the American Chemical Society's LOQ and to the ISO/TUPAC quantification limit, suggesting widespread acceptance.

5.2.1.2.2 Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

The ML procedure is designed to provide a means by which a laboratory can demonstrate performance with a method under routine laboratory operating conditions. All recently developed EPA CWA methods require that a laboratory calibrate its instrument prior to analyzing environmental samples. The ML is defined as the lowest non-zero standard in the laboratory's calibration, and therefore, reflects realistic expectations of laboratory performance with a given method under routine laboratory conditions (i.e., under conditions of routine variability).

Also, the ML is based on the standard deviation of replicate analyses used to establish the MDL. As described in Section 5.1.1.2.2, these analyses are performed to characterize laboratory and method performance, including routine variability, at low concentrations. When a laboratory performs an MDL

study with seven replicates and multiplies the results by 3.18, the laboratory has demonstrated that it can achieve expected levels of performance at the ML.

EPA recognizes that one laboratory may obtain an MDL or ML that is lower or higher than those in another laboratory. If the ML is being established during method development, it is important to determine the ML at more than one laboratory to ensure that the published ML reflects demonstrated expectations of method performance in a community of laboratories. EPA does not believe this community should be so broad as to include the entire universe of possible laboratories that might desire to practice the method. Rather, EPA believes that this community should include well-operated laboratories that are experienced with the techniques used in the method and that have some familiarity with the method. See Section 5.1.1.2.2 for additional discussion of this topic.

5.2.1.2.3 Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.

The ML is designed for use by a single laboratory. The ML can be directly determined from the MDL, which is among the most affordable of procedures for determining detection limits (see discussion in Section 5.1.1.2.3 for additional details regarding affordability). As a result, the ML is among the most affordable of procedures for determining quantitation limits.

5.2.1.2.4 Criterion 5: The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.

The ML meets this criterion. The ML can be verified in a laboratory each time it calibrates an instrument. This calibration is dependent on identifying a recognizable signal for the analyte. In addition, because EPA includes the ML as the low point in the calibration range, that concentration is within the capabilities of the method, as demonstrated by either multiple single-laboratory studies or a multi-laboratory study of the method.

5.2.1.2.5 Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.

The ML meets this criterion. It has been used successfully to support state and local obligations under the Clean Water Act since 1984.

5.2.2 Assessment of the IQE

The Interlaboratory Quantitation Estimate (IQE) was developed by ASTM International with support from members of the regulated industry in an attempt to provide a scientifically sound, comprehensive quantitation limit procedure that addresses the concerns of the regulated industry, statisticians, and analysts involved in ASTM Committee D 19 on water. A brief summary of the procedure for establishing an IQE is given in Section 5.2.2.1. Section 5.2.2.2 presents EPA's assessment of the IQE against the five criteria established for evaluating quantitation limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.2.1 Description of the IQE Approach and Procedure

The ASTM Designation D 6512 is the *Standard Practice Interlaboratory Quantitation Estimate*. As stated in the practice:

"IQE_{Z%} is computed to be the lowest concentration for which a single measurement from a laboratory selected from the population of qualified laboratories represented in an interlaboratory study will have an estimated Z % relative standard deviation (Z % RSD, based on interlaboratory standard deviation), where Z is typically an integer multiple of 10, such as 10, 20, or 30, but Z can be less than 10."

The IQE is determined and verified using a procedure containing 5 major steps with approximately 46 substeps and conditions. The full text of the IQE procedure is available from ASTM International. The 5 major steps and their functions are given in Section 6 of the IQE procedure and are summarized below:

1. Overview of the procedure.
2. IQE Study Plan, Design, and Protocol - in this section, the task manager (study supervisor) chooses the analyte, matrix, and analytical method. Details are given for the appropriate range of study concentrations; the model of recovery vs. concentration; the study protocol (ASTM Practice D 2777 is suggested); the instructions to be given to the participating laboratories, including reporting requirements; the allowable sources of variation; and the number of laboratories, analysts, measurement systems, and days over which the study will be conducted.
3. Conduct the IQE Study, Screen the Data, and Choose a Model - after the study data are collected and screened according to ASTM Practice D 2777, the interlaboratory standard deviation (ILSD) versus concentration data are tabulated and one of three models is fit to the data. The first attempt is at fitting a constant model. If the attempt fails, a straight-line model is attempted. If the straight-line model fails, a hybrid (Rocke/Lorenzato) model is fit. After fitting, the model is evaluated for reasonableness and lack of fit. If the model fails, the study supervisor determines if a subset of the data should be analyzed or if more data are needed.
4. Compute the IQE - the IQE is computed using the ILSD model selected in Step 3 to estimate the relative standard deviation as a function of concentration. The first attempt is at 10% RSD (IQE_{10%}). If this attempt fails, IQE_{20%} is tried, then IQE_{30%}. IQEs greater than 30% are not recommended.
5. Nontrivial Amount of Censored Data - this section of the IQE procedure addresses the effect of "non-detects" or "less-thans." Suggestions are given to see if uncensored data can be obtained from the laboratories or if the study needs to be augmented with additional data. Suggestions are given for fitting a model to data that contain less than 10% non-detects or less-thans to produce an IQE.

5.2.2.2 Assessment of the IQE Against the Evaluation Criteria

The following five subsections discuss the IQE approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.2.2.1 Criterion 1: The detection and quantitation limit approaches should be scientifically valid.

Condition 1: It can be (and has been) tested. EPA is not aware of any organization, including ASTM, that has conducted a study to test the IQE procedure as written (i.e., designed and implemented an interlaboratory study involving multi-laboratory analysis of multiple concentrations of each matrix of interest). It has been tested by its developers using simulated data sets and on interlaboratory data sets that do not adequately characterize the low level region of interest. As part of this reassessment, EPA tested a variant of the IQE procedure on single-laboratory data sets that were designed to characterize an analytical method in the region of detection and quantitation. Despite the lack of comprehensive testing performed to date, however, EPA believes that the IQE procedure can be tested if sufficient resources are invested. In other words, the IQE meets the condition that it "can be" tested, but only partially meets the condition that it "has been" tested.

Condition 2: It has been subjected to peer review and publication. Although the IQE has not been published in the peer-reviewed scientific literature, the IQE has undergone extensive review and ballot by members of ASTM Committee D 19, many of whom are qualified peer reviewers. Therefore, although the IQE does not meet this condition in the sense of formal peer review and publication, EPA believes it does meet the intent of this condition (i.e., submission to scrutiny of the scientific community).

Condition 3: The error rate associated with the procedure is either known or can be estimated. In theory, an expert statistician could estimate the error rate of the IQE. However, the IQE procedure is extremely complex from an analytical chemistry and statistical perspective. As a result, it is unlikely that the error rate could be estimated by the staff of an environmental testing laboratory. Moreover, in attempting to follow the IQE procedure during this reassessment, EPA found the procedure to be highly subjective, particularly with respect to selection of an appropriate model. The subjective nature of the procedure is likely to yield different IQEs from the same data set, depending on the staff involved in analyzing the data and performing the calculations. (The likelihood of this problem is illustrated in Appendix C to this Assessment Document.) EPA believes such conditions make it difficult, if not impossible, to estimate the actual error associated with the IQE. Therefore, the IQE fails this condition.

Condition 4: Standards exist and can be maintained to control its operation. The IQE approach and procedure is supported by a published procedure (standard) to control its operation. The procedure gives the steps to be followed in determining the IQE and instructs the study supervisor how to gather the data and compute an IQE.

However, there are several "gray areas" in the published procedure. The most significant gray area is in model selection. The procedure provides insufficient guidance on the use of residual plots as a basis for selecting models and as a result, selection of the model may be very subjective, especially if the number of concentrations is low. The discussion of what model to use after rejecting the hybrid and linear models also is very vague. The exponential model is mentioned, as well as models with more than one coefficient. Much of the data evaluated by EPA tended to suggest the exponential model, based on the statistical tests discussed. However, those data have almost always shown residual "patterns" when using this model, which would then lead to consideration of other models. In addition, fitting the "constant model" is never discussed in detail. Most likely, this is done by simply calculating a mean (weighted if necessary) of the variances from the different concentrations, however such a calculation never explicitly stated.

As discussed under Condition 4 of Section 5.1.2.2.1 (scientific validity of the IDE procedure), EPA also is concerned about inconsistencies between the IDE and IQE that suggest conceptual problems with these standards. Finally, EPA observed that the IQE contains statistical errors that, if followed as written, could produce inaccurate IQE values. For example, the computations for weighted least squares

given in Table 1 of the procedure are incorrect. The formulae for the weighted means of the spike values and results given in Table 1 of D6512 would only be appropriate if the weighting were done based on the number of replicates per spike level, rather than on the estimated variance calculated using the chosen standard deviation model.

Based on these findings (along with those discussed under Criterion 2 below), EPA believes that, although the IQE is supported by a published procedure, the procedure is not sufficient to control operation of the IQE because of the high degree of subjectivity involved in implementing the procedure, statistical errors in the procedure, and internal inconsistencies with the IDE. Therefore, the IQE fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The IQE fails this condition because it is familiar to, and has been accepted only by, a very narrow segment of the scientific community. Although the IQE has been approved by ASTM for more than 2 years, EPA has not found an IQE in the open literature or in an analytical method, including an ASTM method.

5.2.2.2.2 Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

The IQE procedure is designed to reflect expectations of interlaboratory performance, including routine variability. The procedure contains extensive instructions for dealing with unusual conditions, including sources of variability and outliers. Based on studies of the single-laboratory variant of the procedure in which the model selection proved to be highly subjective, EPA is skeptical about the procedure being able to demonstrate realistic expectations of laboratory and method performance.

The IQE procedure suggests attempting to fit study results to a constant, linear, or hybrid model. If all of these fail, the procedure suggests trying a different model, such as the exponential model. (The exponential model figures more prominently in the IDE procedure, where it is one of the three main models discussed, replacing the Rocke and Lorenzato model.) Although the exponential model may be appropriate for the IDE (which is not tied to a fixed RSD), it yields unacceptable results when applied to the IQE procedure. Under the exponential model, relative variability (standard deviation divided by the true concentration) is a parabolic function (i.e., as concentration increases, relative variability decreases down to a specific percentage, and then begins to increase). This is not realistic of laboratory and method performance. In addition, the exponential model will often result in having two possible values each for $IQE_{10\%}$, $IQE_{20\%}$, and $IQE_{30\%}$.

Another concern with the IQE procedure is that use of the non-mandatory appendices in ASTM D 6512 to determine the fit of a model may produce results that differ from those that would be obtained using the default procedures for testing model fit that are built into off-the-shelf statistical software.

Given the subjectivity and confusion involved in selecting the model, EPA tried using the same data set to calculate a single-laboratory variant of the IQE with each of the available models and found that the calculated IQEs varied widely when different models were used.

Based on the problems described above, EPA believes the IQE fails this criterion.

5.2.2.2.3 Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.

The IQE procedure is neither practical nor affordable in a single-laboratory context. It is designed for use by an ASTM study supervisor or task manager and not as a procedure that a single

laboratory can use to evaluate method performance. EPA is aware that ASTM Committee D 19 is contemplating development of a within-laboratory quantitation estimate (WQE), but the WQE has not been approved through an ASTM ballot and therefore, it cannot be adequately evaluated at this time. The WQE may meet this criterion, but the IQE does not.

Regarding affordability, EPA estimates that the cost of implementing IQE procedure would be more than twice the cost of EPA's present implementation of the ML. The increased cost stems from the additional low-level data required to assure that variability versus concentration is being characterized in the region of detection and quantitation, challenges involved in applying the statistical procedures in the IQE, and because of the anticipated reanalysis and rework required if either the procedure failed to produce an IQE or if the resulting IQE failed to meet the specifications in the IQE procedure.

5.2.2.2.4 Criterion 5: The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.

If the IQE were developed in an interlaboratory study that met the requirements of D 6512, the calculated IQE would likely be achievable by experienced staff in a well-operated laboratory. Therefore, the IQE passes this criterion. However, EPA also notes that although it passes the criterion, based on this assessment, EPA believes that it is very likely that the IQE may not identify the *lowest* concentration at which the signal is recognizable when the method is performed by experienced staff in a well-operated laboratory.

5.2.2.2.5 Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government

Although the IQE could be applied to some decisions to be made under CWA, it may not support decisions when pollutant levels need to be protective of human health and the environment because the IQE may be considerably higher than these levels. At best, the IQE only partially passes this criterion.

5.2.3 Assessment of the ACS Limit of Quantitation

The Limit of Quantitation (LOQ) was developed by the Committee on Environmental Improvement of the American Chemical Society (ACS) and published in the same two papers as the LOD.

5.2.3.1 Description of the ACS LOQ Approach and Procedure

The 1983 "Principles" define the LOQ as:

"... the level above which quantitative results may be obtained with a specified degree of confidence."

The same relationship used to define the LOD is used for the LOQ:

$$S_t - S_b \geq K_d \sigma$$

but the recommended minimal value for K_d be set at 10. Thus, the LOQ is 10σ above the gross blank signal, S_b . According to the 1983 publication, the LOQ corresponds to an uncertainty of $\pm 30\%$ ($10\sigma \pm$

3 σ). This uncertainty statement is based on σ equal to 10% of the LOQ. Other statements of uncertainty are, of course, possible using knowledge of σ and/or the RSD.

Neither the 1980 nor 1983 ACS publications provide a specific procedure for estimating the LOQ, nor do they provide a minimum number of observations needed to estimate the gross blank signal or the variability term σ_b .

5.2.3.2 Assessment of the ACS LOQ Against the Evaluation Criteria

The following five subsections discuss the ACS LOQ approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.3.2.1 Criterion 1: *The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. Testing of the LOQ is hampered by 1) the lack of a supporting procedure for establishing an LOQ, and 2) its conceptual dependence on the variability of blank measurements. If the blank measurements fail to produce a response, it is impossible to calculate an LOQ because the standard deviation of zero is zero. One solution for testing the approach is to assume that the LOQ is functionally equivalent to the ML as the blank signal approaches zero. EPA believes this is a reasonable assumption, and therefore, that the ML procedure is a viable means for testing the LOQ approach. Therefore, the LOQ meets this condition.

Condition 2: It has been subjected to peer review and publication. The ACS LOQ definition was published in the peer-reviewed journal *Analytical Chemistry* in 1980 and 1983. Therefore, the ACS LOQ meets this condition.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The definition of the LOQ specifically estimates the uncertainty associated with a concentration at the LOQ as $\pm 30\%$ based on 10% RSD. Other valid statements in terms of %RSD may be made based on study requirements, policy judgments and/or specific results. For example, the estimate of an uncertainty of $\pm 30\%$ based on 10% RSD is inconsistent with EPA and ISO/IUPAC estimations that place the uncertainty at $\pm 20\%$ (at $\pm 2\sigma$), and is inconsistent with the Episode 6000 data that place the median RSD at 7% and therefore, the $\pm 2\sigma$ uncertainty at approximately $\pm 14\%$.

Condition 4: Standards exist and can be maintained to control its operation. The ACS LOQ lacks a clearly defined procedure for estimating the important terms required to derive it. Although it may be possible to derive ACS LOQ values from data used to derive EPA MDL values, there is no discussion of using replicate blanks, replicate spiked samples, or a minimum recommendation for the number of replicates. Therefore, the ACS LOQ fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Because the ACS does not develop and publish reference analytical methods, it is difficult to determine the degree of acceptance of the LOQ. EPA has not investigated the numbers of papers published in ACS journals that include LOQ values, but EPA's literature search for detection and quantitation approaches did not uncover a large number of citations that promote the LOQ in particular.

5.2.3.2.2 Criterion 2: *The approach should address demonstrated expectations of laboratory and method performance, including routine variability*

The LOQ approach is designed to address demonstrated expectations of laboratory and method performance, including routine variability, and therefore, it appears to meet this criterion. Unfortunately, ACS has not published a procedure to implement the approach. In other words, the LOQ addresses demonstrated expectations of laboratory and method performance in theory, but in practice, provides no direct means for performing these demonstrations. Therefore, EPA believes the ACS LOQ only partially meets this criterion.

5.2.3.2.3 Criterion 3: *The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.*

The ACS LOQ approach is not supported by a clearly defined procedure for establishing the LOQ. Therefore, it fails this criterion.

5.2.3.2.4 Criterion 5: *The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.*

Given the relationship of the ACS LOQ to the ML, EPA believes the LOQ meets this criterion for the reasons outlined in Section 5.2.1.2.4, which discusses EPA's assessment of the ML against Criterion 4 for evaluating quantitation limit approaches.

5.2.3.2.5 Criterion 6: *Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.*

In the absence of a procedure for determining LOQ values, the ACS LOQ fails to meet this criterion because it cannot be used in a regulatory context. The LOQ passes this criterion only if it is assumed to be functionally equivalent to the ML (i.e., the ML procedure is used to establish an LOQ).

5.2.4 Assessment of the IUPAC/ISO Limit of Quantitation

A similar LOQ approach was developed by IUPAC/ISO and published in the same papers as the CRV and MDV (see Sections 5.1.4 and 5.1.5).

5.2.4.1 Description of the ISO/IUPAC LOQ Approach

The 1995 "Recommendations" define the LOQ as:

"... the ability of a CMP [chemical measurement process] to adequately 'quantify' an analyte. The ability to quantify is generally expressed in terms of the signal or analyte (true) value that will produce estimates having a specified relative standard deviation (RSD), commonly 10 %."

The relationship used to define the LOQ is:

$$L_Q = K_Q \times \sigma_Q$$

The recommended value for K_Q is 10. Thus, the LOQ is 10σ above the blank signal, σ_Q .

5.2.4.2 Assessment of the IUPAC/ISO LOQ Against the Evaluation Criteria

The following five subsections discuss the IUPAC/ISO LOQ approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.4.2.1 Criterion 1: *The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. Testing of the IUPAC/ISO LOQ is hampered by 1) the lack of a supporting procedure for establishing and LOQ, and 2) it's conceptual dependence on the variability of blank measurements. If the blank measurements fail to produce a response, it is impossible to calculate an LOQ because the standard deviation of zero is zero. One solution for testing the approach is to assume that the ISO/IUPAC LOQ is functionally equivalent to the ML as the blank signal approaches zero. EPA believes this is a reasonable assumption, and that the ML procedure is a viable means for testing the LOQ approach. Therefore, the ISO/IUPAC LOQ meets this condition.

Condition 2: It has been subjected to peer review and publication. The IUPAC/ISO LOQ definition has been published by Currie in the peer-reviewed journals *Pure and Appl. Chem.* in 1995; in *Anal. Chim. Acta* in 1999, in *Chemometrics and Intelligent Lab Systems* in 1997; and in *J. Radioanal. and Nuclear Chem.* in 2000. Therefore, the IUPAC/ISO LOQ meets this condition.

Condition 3: The error rate associated with the procedure is either known or can be estimated. EPA used data generated in the Episode 6000 study to estimate the error rate associated with the LOQ. The Episode 6000 results show that the median error across all analytes and analytical techniques at 10σ is approximately $\pm 14\%$ with approximately 95% confidence.

Condition 4: Standards exist and can be maintained to control its operation. The IUPAC/ISO LOQ lacks a clearly defined procedure for estimating the important terms required to derive it. Although it may be possible to derive IUPAC/ISO LOQ values from data used to derive EPA MDL values, there is no discussion of using replicate blanks, replicate spiked samples, or a minimum recommendation for the number of replicates. Therefore, EPA believes that the IUPAC/ISO LOQ fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Acceptance by the scientific community is not known. Acceptance would be indicated by use of the LOD in ISO methods. EPA did not perform a search of ISO methods because of copyright restrictions. However, EPA's literature search for detection and quantitation approaches in the open technical literature did not uncover a large number of citations that reference the LOQ. Therefore, it is difficult to determine if the ISO/IUPAC LOQ meets this condition.

5.2.4.2.2 Criterion 2: *The approach should address demonstrated expectations of laboratory and method performance, including routine variability.*

The most recent publication on the IUPAC/ISO LOQ (*J. Radioanal. and Nuclear Chem.*, op. cit.) provides insight into this issue through measurements of ^{14}C by accelerator mass spectrometry. Therefore, EPA believes that the IUPAC/ISO LOQ passes this criterion for at least some measurement techniques.

5.2.4.2.3 Criterion 3: *The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.*

The ISO/IUPAC LOQ approach is not supported by a clearly defined procedure for establishing the LOQ. Therefore, it fails this criterion.

5.2.4.2.4 Criterion 5: *The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.*

Given the relationship of the IUPAC/ISO LOQ to the ML, EPA believes that the LOQ satisfies this criterion for the reasons outlined in Section 5.2.1.2.4, which discusses EPA's assessment of the ML against Criterion 4 for evaluating quantitation limit approaches.

5.2.4.2.5 Criterion 6: *Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government*

In the absence of a procedure for determining LOQ values, the ISO/IUPAC LOQ fails to meet this criterion because it cannot be used in a regulatory context. The ISO/IUPAC LOQ passes only if the ML procedure is used to establish an LOQ.

Table 5-1. Assessment of Detection Limit Approaches Against Evaluation Criteria

Evaluation Criteria	MDL	IDE	ACS LOD	ISO/IUPAC CRV	ISO/IUPAC MDV
<p>The detection limit approach should be scientifically valid:</p> <ul style="list-style-type: none"> • It can be (and has been tested) • Has undergone peer review and publication • Has an error rate that is known or can be estimated • Has standards that can be maintained to control its operation • Has achieved widespread acceptance in a relevant scientific community 	<p>Meets all 5 conditions for scientific validity with slight modifications noted to clarify understanding of error rate.</p>	<p>Meets 1, partially meets 1, and fails 3 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • Can be, but has not been fully tested (partial) • Subjectivity makes calculation of error rate impossible (fails) • Has a standard but, due to the high degree of subjectivity, errors, and conceptual inconsistency, it is unlikely to control its operation (fails) • Is familiar to and accepted by a very narrow segment of the scientific community (fails) 	<p>Meets 4 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • No standards exist to control its operation 	<p>Meets 3 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • No standards exist to control its operation • Degree of acceptance is unclear 	<p>Meets 3 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • No standard exist to control its operation • Degree of acceptance is unclear
<p>The approach should address demonstrated expectations of laboratory and method performance, including routine variability.</p>	<p>Can meet this criterion if properly applied.</p>	<p>Conceptually passes this criterion, but fails in practice due to problems with model selection</p>	<p>Partially meets the criterion. Approach meets the criterion but no procedure for implementing the approach is given. Passes the criterion only if equivalency to the MDL is assumed.</p>	<p>Partially meets this criterion. Approach meets the criterion but no procedure for implementing the approach is given. Passes the criterion only if equivalency to the MDL is assumed.</p>	<p>Could be used in planning method development and evaluation studies as recommended but not in operational detection decision making.</p>
<p>The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.</p>	<p>Meets this criterion. Procedure can be performed by a single laboratory during a single shift, or for method development by multiple labs in a single shift</p>	<p>Fails this criterion. Requires interlaboratory study involving a reference lab or coordinating body, a minimum of 6 complete data sets, and a skilled statistician. The cost of implementing this procedure would exceed most method development budgets.</p>	<p>Fails this criterion. No procedure provided.</p>	<p>Fails this criterion. No procedure provided.</p>	<p>Fails this criterion. No procedure provided.</p>

Table 5-1. Assessment of Detection Limit Approaches Against Evaluation Criteria

Evaluation Criteria	MDL	IDE	ACS LOD	ISO/IUPAC CRV	ISO/IUPAC MDV
<p>The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.</p>	<p>Meets this criterion.</p>	<p>When the allowance for false negatives and for prediction and tolerance are taken into account, the resulting detection limit (IDE) is raised to the point at which detection probability is estimated to be greater than 99.999999%; this yields numerical values that have no practical meaning as a detection standard. Therefore, the IDE fails this criterion.</p>	<p>Meets this criterion.</p>	<p>Meets this criterion.</p>	<p>The MDV is a true concentration value not used in the actual detection decision. Does not meet the criterion.</p>
<p>Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.</p>	<p>Meets this criterion.</p>	<p>At best, only partially passes this criterion. Not likely to meet this criterion in instances in which a compliance limit is close to a detection limit determined by a procedure such as the MDL.</p>	<p>In the absence of a procedure for determining LOD values, fails to meet this criterion.</p>	<p>In the absence of a procedure for determining CRV values, fails to meet this criterion.</p>	<p>In the absence of a procedure for determining MDV values, fails to meet this criterion.</p>

Table 5-2. Assessment of Quantitation Limit Approaches Against Evaluation Criteria

Evaluation Criteria	ML	IQE	ACS LOQ	ISO/IUPAC LOQ
<p>The quantitation limit approach should be scientifically valid.</p> <ul style="list-style-type: none"> • It can be (and has been tested) • Has undergone peer review and publication • Has an error rate that is known or can be estimated • Has standards that can be maintained to control its operation • Has achieved widespread acceptance in a relevant scientific community 	<p>Meets all 5 conditions for scientific validity, though slight modification to the definition is suggested to improve operation when other than 7 replicates are used to estimate the ML.</p>	<p>Meets 1 condition, partially meets 1 condition, and fails 3 conditions.</p> <ul style="list-style-type: none"> • Can be, but has not been fully tested (partial) • Error rate cannot be estimated due to problems with the procedure (fail) • Standards are not likely to control its operation (fail) • Has not achieved widespread acceptance (fail) 	<p>Meets 3 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • Lacks a standard to control its operation • Difficult to determine the degree of acceptance 	<p>Meets 4 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • Lacks a standard to control its operation • Difficult to determine the degree of acceptance
<p>The approach should address demonstrated expectations of laboratory and method performance, including routine variability.</p>	<p>Meets this criterion. Procedure can be performed by a single laboratory during a single shift, or for method development by multiple labs in a single shift.</p>	<p>Fails this criterion due to subjectivity, errors, and theoretical inconsistencies in the procedure.</p>	<p>Partially meets this criterion. The approach is designed to address these expectations but in practice, there is no procedure for performing such demonstrations.</p>	<p>Meets this criterion.</p>
<p>The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.</p>	<p>Meets this criterion.</p>	<p>Fails this criterion. Requires interlaboratory study involving a reference lab or coordinating body, 6 complete data sets, and a highly skilled statistician. The cost of implementing this procedure would exceed most method development budgets.</p>	<p>Fails this criterion.</p>	<p>Fails this criterion.</p>
<p>The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.</p>	<p>Meets this criterion.</p>	<p>Meets this criterion, but is not likely to estimate the <i>lowest</i> level at which reliable measurements can be made by an experienced analyst in a well operated lab</p>	<p>Meets this criterion.</p>	<p>Meets this criterion.</p>

Table 5-2. Assessment of Quantitation Limit Approaches Against Evaluation Criteria

Evaluation Criteria	ML	IQE	ACS LOQ	ISO/IUPAC LOQ
<p>Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.</p>	<p>Meets this criterion.</p>	<p>At best, only partially passes this criterion. Fails for those instances in which the IQE limit is greater than an effluent limit or water quality-based limit.</p>	<p>Fails this criterion. In the absence of a procedure for determining ACS LOQ values, the ACS LOQ cannot be used in a regulatory context.</p>	<p>Fails this criterion. In the absence of a procedure for determining LOQ values, the ISO/IUPAC LOQ cannot be used in a regulatory context.</p>

Chapter 6 Conclusions

This chapter summarizes the results of EPA's assessment of detection and quantitation limit approaches. This assessment, which is detailed in the previous five chapters, was based on:

- Identification of relevant approaches to include in the assessment (Chapter 2),
- Identification of issues that may be relevant to the assessment from an analytical chemistry, statistical, or regulatory perspective (Chapter 3),
- Development of criteria that reflect EPA's views concerning these issues (Chapter 4) and form the primary basis for evaluating the ability of each approach to meet EPA needs under the Clean Water Act,
- Assessment of how well each approach met the evaluation criteria (Chapter 5), and,
- Use of real-world data to evaluate both the theoretical and practical limitations of each approach (Appendices B and C).

EPA evaluated four sets of detection and quantitation limit approaches advanced by EPA, ASTM International, ACS, and both ISO and IUPAC. Each approach was assessed against the suite of criteria described in Chapter 4. The EPA approaches (i.e., the MDL and ML) and the ASTM International approaches (i.e., the IDE and IQE) were supported by clearly defined procedures for implementing the approach. Neither the ACS nor the ISO/IUPAC approaches are supported by detailed procedures for implementation; this lack of supporting procedures was reflected in the outcome of EPA's overall assessment.

After evaluating each approach against each of the evaluation criteria, EPA found that 1) no single pair of detection and quantitation limit approaches perfectly meets EPA's criteria, 2) the MDL and ML are closest to meeting EPA's criteria, and 3) minor revisions and clarifications to the MDL and ML would allow both approaches to fully meet the Agency's needs under the CWA.

EPA also found that, although the IDE and IQE procedures may be acceptable for planning and implementing interlaboratory studies to develop and validate analytical methods, there are a number of difficulties with these procedures that make them unsuitable as the primary means of establishing sensitivity under the Clean Water Act. In particular, the IDE is analogous by definition and formulaic construction to the "*Detection Limit*" defined by Currie (1968, 1995), while it is Currie's "*Critical Value*" approach that is most relevant to Agency needs under the CWA. Currie (1995) states that the decision "detected" or "not detected" is made by comparison of the estimated quantity or measured value with the *critical value*. Currie describes his "*Detection Limit*" as a true concentration that has a high probability of generating measured values that exceed the *critical value*, and states that the single most important application of the detection limit is for planning and evaluation of measurement procedures and that the *detection limit*:

"...allows one to judge whether the CMP (Chemical Measurement Process) under consideration is adequate for detection requirements. This is in sharp contrast to application of the critical value for decision making, given the result of a measurement."

It is important to note that the formulation of the MDL is analogous to the Currie critical value, and as such, is intended to be used to make detection decisions in the manner described by Currie (i.e., the MDL is designed and used to make the decision of "detected" or "not detected"). EPA believes that form of detection decision best supports the use of "detection limits" under CWA programs.

In contrast, although the IDE is intended to be used in a manner analogous to Currie's *critical value* (i.e., to make the decision of "detected" or "not detected"), it is, by definition and design, functionally analogous to Currie's *detection limit* (i.e., it identifies a concentration that will have a high probability of generating measured values that exceed the critical value). (See Chapter 2, Section 2.1 for a discussion of Currie's *critical value* and *detection limit*).

Other drawbacks with the ASTM International approach include the complexities of the IDE and IQE procedures, along with their inability to address individual laboratory performance. Despite these limitations, however, EPA believes the IDE and IQE can be used to establish sensitivity for certain applications. For example, consider the theoretical situation of an ASTM method for the determination of an analyte regulated under the NPDES program that uses the IDE or IQE to describe method sensitivity and for which the value of the IDE or IQE was below the relevant criterion or regulatory limit. EPA would evaluate the overall performance of such a method for approval at 40 CFR part 136, despite the fact that the method did not contain an MDL determined using the procedure described in 40 CFR part 136, Appendix B. (See Chapter 3, Section 3.2.8 for a more in-depth discussion of using alternative procedures to establish sensitivity.)

EPA's assessment of the theoretical and practical applications of each detection and quantitation approach (see Appendices B and C) is summarized in Exhibit 6-1. This exhibit suggests that no approach produces the "right" answer, and that different approaches produce different detection and quantitation limits. Observed differences are largely due to different sources of variability accounted for among the approaches.

As part of this assessment, EPA identified the need for approaches that can support CWA programs, including:

- method performance verification at a laboratory,
- method development and promulgation,
- National Pollutant Discharge Elimination System (NPDES) applications,
- non-regulatory studies and monitoring,
- descriptive versus prescriptive uses of lower limits to measurement, and
- use of a pair of related detection and quantitation procedures in all OW applications

EPA has concluded that the MDL and ML can meet all of these applications and that the addition of a scope and application section to the procedure would help clarify use of the MDL for these applications. However, as noted in Chapter 3, outside organizations use different detection and quantitation approaches that meet their own needs. Given EPA's diverse needs and desire to encourage the development of improved measurement techniques, EPA does not believe it is necessary or appropriate to require the exclusive use of the MDL and ML approaches in CWA programs. As indicated above, EPA would allow use of alternative detection and quantitation procedures to establish detection and quantitation limits in an analytical method, provided that the resulting detection and quantitation limits meet the sensitivity needs for the specific application.

Exhibit 6-1: Theoretical and Practical Application of Each Approach

Finding 1: Each approach yields different values.

Detection Limit Approaches

- The EPA MDL and ACS LOD approaches, which are functionally analogous, produced detection limits that are a median of 1.25 times higher than the limits produced by the CRV advanced by ISO and IUPAC (Appendix C of this document).
- The Minimum Detectable Value (MDV) advanced by ISO and IUPAC produced detection limits that are a median of 1.2 times higher than the limits produced by the MDL and LOD approaches (Appendix C of this document).
- A single-laboratory variant of the IDE (the IDE has been advanced by ASTM International) produced detection limits that are a median of 2.9 times higher than the median limits produced by the MDL and LOD approaches (Appendix C of this document). This result is not surprising given that the IDE is functionally analogous to Currie's *detection level*, while the MDL and LOD are analogous to Currie's *critical value*.

Quantitation Limit Approaches

- The EPA ML and the functionally equivalent ACS LOQ produced quantitation limits that are a median of 1.1 times higher than the limits produced by the LOQ approach advanced by ISO and IUPAC (Appendix C of this document).
- A single-laboratory variant of the IQE (the IQE has been advanced by ASTM International) produced median quantitation limits that are equivalent to the median limits produced by the EPA ML and ACS LOQ approaches (Appendix C of this document).

Finding 2: More than the 5 levels specified by ASTM are required to produce a reliable IDE and IQE

- EPA found that the IDEs produced with a subset of data generated from the minimum of 5 concentrations recommended in the IDE procedure differed widely from the IDEs produced with a larger set of data involving 16 concentrations (which included the subset of 5 concentrations) (Appendix C of this document).
- Findings suggest that more than 5 concentrations are needed to produce a reliable IDE, due to the limited power of the statistical tests for significant model parameters and the difficulty of drawing conclusions based on residual plots with only 5 points (Appendix C of this document).
- Parallel reasoning can be applied to the IQE based on its similarity to the IDE.

Finding 3: The ML procedure yields quantitation limits that are generally in the range of the 10% RSD intended in the ML (and the functionally analogous ACS LOQ) approach.

- EPA calculated the uncertainty associated with replicate measurements made at the ML for a large number of analytes and techniques (Appendix C of this document).
- EPA found that on average, across all techniques tested, the RSD of replicate measurements at the ML was approximately 7%. Median RSDs calculated for each multi-analyte method ranged from 6 - 14%, and RSD values calculated for each single-analyte method ranged from 4 - 16% (Appendix C of this document).

Finding 4: No single model adequately predicts the behavior of all analytes and all methods across the measurement range.

- EPA produced graphs representing hundreds of analyte/method combinations. Selection of an appropriate model based on these graphs is highly subjective, at best, due to the lack of clear patterns and the residuals observed with each model applied (Chapter 3, Section 3.3, and Appendix B of this document).
- The IDE and IQE are the only approaches other than the MDL and ML that are supported by a procedure for their implementation. The IDE and IQE procedures rely heavily on model selection, and the degree of subjectivity involved in selecting these models makes implementation of the IDE and IQE difficult (Chapter 5, Sections, 5.1.2 and 5.2.2, and the third conclusion in Appendix C).

Finding 5: Use of a recovery correction when establishing detection and quantitation limits may not be appropriate.

- EPA found that using a regression to estimate a recovery correction at zero concentration causes great swings in the resulting detection and quantitation limits (Appendix C of this document).
- Use of a recovery-correction procedure also can result in 'double-correcting' for recovery because 1) nearly all methods already contain specifications for acceptable recovery performance, and 2) some methods include recovery correction in the computation of sample results (Chapter 3, Section 3.1.4).

Exhibit 6-2: Summary of Recommended Modifications to the MDL and ML procedures

EPA believes that the following revisions and clarifications to the MDL and ML would allow these procedures to fully meet the Agency's needs under the CWA.

- Refine the definition of the MDL to make it more consistent with the MDL procedure and note the functional analogy of the MDL with the "critical value" described by Currie (1968 and 1995) and with the "limit of detection" (LOD) described by the American Chemical Society in 1980 and 1983 (Chapter 5, Section 5.2.1.1.1).
- Expand the Scope and Application discussion to acknowledge that there are a variety of purposes and analytical methods for which the MDL procedure may be employed and to provide examples of common uses of the MDL procedure (i.e., demonstrating laboratory capability with a particular method; monitoring trends in laboratory performance; characterizing method sensitivity in a particular matrix; and establishing an MDL for a new or revised method for nationwide use).
- Clarify the considerations for estimating the detection limit in Step 1 of the current MDL procedure, and suggest that the method-specified MDL can be used as the initial estimate when performing an MDL study to verify laboratory performance or to demonstrate that the MDL can be achieved in a specific matrix (Chapter 5, Section 5.2.1.1.1).
- Revise the specifications for establishing the test concentration range (i.e., determining the spike levels) in Section 3.1 according to the intended application of the MDL as follows: 1) if verifying a published MDL, the test concentration should be no more than five times the published MDL; 2) if verifying an MDL to support a regulatory objective or the objective of a study or program, the test concentration should be no more than one third the compliance or target limit; 3) if determining an MDL for a new or revised method, the test concentration should be no more than five times the estimated detection limit; and 4) if performing an iteration, the test concentration should be no more than five times the MDL determined in the most recent iteration.
- Delete the calculation of a 95% confidence interval estimate for the MDL from Step 6. EPA has determined that these calculations are neither routinely performed by laboratories, nor are the results employed by regulatory agencies, including EPA.
- Revise Step 7 to 1) require that the iterative procedure be used to verify the reasonableness of the MDL when developing an MDL for a new or revised method or when developing a matrix-specific MDL, but that it remain optional when verifying a method-, matrix-, program-, or study-specific MDL, and 2) provide specific instructions on how to assess the reasonableness of an MDL used to verify laboratory performance (Chapter 5, Section 5.2.1.1.1).
- Add a new Step 8 to the MDL procedure to address the treatment of suspected outliers (Chapter 5, Section 5.2.1.1.1).
- Delete the discussion of analysis and use of blanks included in Section 4(a) of the current MDL procedure. The current discussion applies to methods in which a blank measurement is required to calculate the measured level of an analyte; it requires separate measurements of blank samples for each MDL sample aliquot analyzed and subtraction of the average result of the blank samples from each respective MDL sample measurement. Deletion of this discussion recognizes that subtraction of a single (or average) blank sample result from the result for each MDL sample would not change the standard deviation and thus, would have no effect on the resulting MDL. Although EPA believes laboratories would be prudent to analyze method blanks for assessing potential contamination, EPA also believes that requiring analysis of method blanks or subtraction of method blank results during MDL determinations is unnecessarily burdensome.
- Revise the optional pre-test described in Section 4(b) of the current MDL procedure to provide criteria that allow the analyst to determine if the test samples are the desirable range.
- Improve overall readability and understanding of the MDL procedure through editorial changes to the specific numbering scheme, the addition of clearer titles to some of the steps, and minor clarifications.
- Clarify the ML to emphasize its relationship to Currie's Quantitation Limit and ACS' Limit of Quantitation (LOQ)
- Clarify the ML procedure to address the use of other than seven replicates for determination of the MDL and ML.

Code	Name	Description	Statistical Description
IDL	Instrument detection limit	The lowest concentration of an analyte that, when processed on a specific piece of analytical equipment, produces a signal/response that is statistically distinct from the signal/response arising from instrument "noise" alone	The concentration of analyte in STANDARD SOLUTION that produces an instrument signal/response that is X times the standard deviation above the "EXPECTED" IDL, where X is the student's t-statistic (at 99% confidence; n-1 deg. freedom)
DDL	Daily detection limit	Method detection limit that is calculated on a daily basis using laboratory blanks	The concentration of analyte in a laboratory blank that produces an instrument signal/response that is X times the standard deviation above the "EXPECTED" DDL, where X is the student's t-statistic (at 99% confidence; n-1 deg. freedom)
MDL	Method detection limit	The lowest concentration of an analyte that, when processed through an entire analytical method (including prep and equip), produces a signal/response that is statistically distinct from the signal/response arising from lab reagent blanks (zero values).	The concentration of analyte in SAMPLE MATRIX that produces an instrument signal/response that is X times the standard deviation above the "EXPECTED" MDL, where X is the student's t-statistic (at 99% confidence; n-1 deg. freedom)
SDL	System detection limit	The lowest concentration of an analyte that, when processed through the entire suite of sampling, transport, analysis, and data reduction operations, produces a signal/response statistically distinct from the signal/response arising from field blanks	The concentration of analyte in SAMPLE MATRIX that produces an instrument signal/response that is X times the standard deviation above the FIELD COLLECTION CONTAM. LEVEL, where X is the student's t-statistic (at 99% confidence; n-1 deg. freedom)
UDL	Sample-specific detection limit	Method detection limit that is calculated on a sample basis using lab reagent blanks and adjusting for sample volume.	The conc. of analyte in a lab reagent blank that produces an instrument signal/response that is X times the std. dev. above the "EXPECTED" UDL, divided by the volume of sample, where X is the students t-statistic (at 99% confidence, n-1 deg. freedom)
LOQ-L	Limit of quantification, Low	The lowest concentration of an analyte that produces a signal/response that is sufficiently greater than the signal/response of lab reagent blanks to enable reliable detection (and thus quantification) during routine lab operating conditions	The concentration of analyte in SAMPLE MATRIX that produces an instrument signal/response that is 10 times the standard deviation above the LAB REAGENT BLANK (at 99% confidence; n-1 deg. freedom)
LOQ-H	Limit of quantification, High	Description not available	Statistical description not available
CRDL	Contract required detection limit	The lowest level of detection specified as acceptable under the Statement of Work for the EPA Contract Laboratory Program	Variable, set by contract
CRQL	Contract required quantification limit	The lowest level of quantification specified as acceptable under the Statement of Work for the EPA Contract Laboratory Program	Variable, set by contract

<u>Code</u>	<u>Name</u>	<u>Description</u>	<u>Statistical Description</u>
MCL	Maximum contaminant level	Regulatory concentration for an analyte, set by the EPA in accordance with the Safe Drinking Water Act, above which drinking water is deemed unsafe	Variable, set by regulation
MQL	Method quantification limit	The lowest concentration of an analyte that produces a signal/response that is sufficiently greater than the signal/response of lab reagent blanks to enable reliable detection (and thus quantification) during routine lab operating conditions	The concentration of analyte in SAMPLE MATRIX that produces an instrument signal/response that is 5 times the standard deviation above the LAB REAGENT BLANK (at 99% confidence; n-1 deg. freedom)
RL	Reporting limit	The limit above which a laboratory feels confident in reporting its results. This limit is the level where the laboratory believes results are not subject to laboratory-induced contamination or other sources of bias	Variable, set by laboratory
PQL	Practical quantification limit	The lowest concentration of an analyte that can be reliably measured within specified limits of precision and accuracy during routine laboratory operating conditions	The concentration of analyte in SAMPLE MATRIX that produces an instrument signal/response that is 10 times the Method Detection Limit (MDL)
MDA	Minimum detectable activity	Radiochemistry: The net count rate that must be exceeded before a sample is assumed to contain measurable radioactivity above background levels	Count rate that is X times the standard deviation above the BACKGROUND counting rate, where X is the student's t-statistic (at 95% confidence; n-1 deg. freedom)
LLD	Lower limit of detection	Radiochemistry: The smallest concentration of radioactive materials in a sample that will yield a net count greater than zero	Count rate that is X times the standard deviation above either BACKGROUND or BLANK counting rate, where X is the student's t-statistic (95% confidence; n-1 deg. freedom) corrected for count efficiency, fractional radiochemical yield, and sample mass/volume
LC	Critical level	Radiochemistry: The final instrument measurement of a quantity of analyte at (or above) which a positive amount of analyte is considered present	Count rate that is X times the standard deviation above the ZERO NET counting rate, where X is the student's t-statistic (at 95% confidence; n-1 deg. freedom)
RDL	Required detection limit	New term from the Federal Register ... to be investigated	New term from the Federal Register ... to be investigated
RQL	Required quantification limit	New term from the Federal Register ... to be investigated	New term from the Federal Register ... to be investigated

Name	notes/changes to make
Instrument detection limit	
Daily detection limit	
Method detection limit	check that definition matches or refers to 40 CFR pt. 136, appendix B, <i>revision 1.11</i>
System detection limit	
Sample-specific detection limit	need to add definition - work with Kim to develop definition consistent with other definitions
Limit of quantification, Low	
Limit of quantification, High	
Contract required detection limit	
Contract required quantification limit	

Name	notes/changes to make
Maximum contaminant level	
Method quantification limit	
Reporting limit	
Practical quantification limit	
Minimum detectable activity	
Lower limit of detection	
Critical level	
Required detection limit	
Required quantification limit	

**Best
Management
Practices
Plan**

**Tennessee Valley Authority
Sequoyah Nuclear Plant
NPDES Permit No.
TN0026450**

March 2004

Sequoyah Nuclear Plant's (SQN) Best Management Practices (BMP) plan is divided into two different programs.

1. The non-radiological program is controlled by the State of Tennessee Department of Environment and Conservation (TDEC) and Environmental Protection Agency (EPA).
2. The radiological program is implemented in accordance with 10 CFR 20 which is governed by the Nuclear Regulatory Commission (NRC).

SQN plant procedures and administrative controls are used to properly operate and maintain all plant processing equipment. Plant operating procedures and administrative controls used at SQN are written and maintained considering experience gained in system operation, applicable industry experience, and vendor recommendations. This ensures that SQN's waste management system is operated in as efficient and effective a manner as practical.

Non-Radiological Program

SQN's BMP plan was developed and implemented to minimize the potential for pollutant releases from plant or ancillary activities to surface waters of the State of Tennessee. Pollutants which have the potential for release to waters of the State of Tennessee have been established in SQN's Spill Prevention Control and Countermeasures (SPCC) plan and Storm Water Pollution Prevention Plan (SWPPP) for the Tennessee Multi-Sector Permit for Storm Water TNR050015.

The Spill Prevention Control and Countermeasure (SPCC) Plan

The SPCC plan serves as a management tool to minimize the risk of a Comprehensive Environmental Response Compensation and Liability Act (CERCLA) reportable spill, a reportable spill or release of a hazardous substance as defined in the Federal Clean Water Act (40 CFR 116 and 117), polychlorinated biphenyls (PCB) release or accidental release of other hazardous substances. Consequently, the SPCC plan satisfies the regulatory requirements for Oil Pollution Prevention (40 CFR 112); Resource Conservation and Recovery Act (RCRA) Preparedness and Prevention (40 CFR Part 265 Subpart C), serves as a BMP plan (40 CFR 125 Subpart K), and a RCRA contingency plan (40 CFR Part 265 Subpart D). As such, a copy of this document is kept available for use onsite at the Site Environmental Office at all times and for regulatory review during normal working hours.

Included in the SPCC plan are the identity and quantity of chemicals stored and used at SQN and the types of containment and secondary containments used. Flow paths for discharges from a spill or overflow are diagrammed in the SPCC plan and in the NPDES Permit Renewal Application (submitted June 30, 2003).

Storm Water Pollution Prevention Plan (SWPPP)

The primary purpose of the SQN SWPPP is to provide guidance for meeting storm water control requirements as required in the SQN National Pollutant Discharge Elimination System (NPDES) Permit No. TN0026450 and the Tennessee Storm Water Multi-Sector General Permit for Industrial Activities, Permit No. TNR050015. The goal of these requirements is to prevent storm water pollution that can harm lakes and streams by reducing pollutants contained in storm water discharges.

The SWPPP is applicable to activities at SQN that result in bare areas, loading and unloading, outdoor storage, land disturbance, particulate or dust generation, pipe installation/removal, waste management, and waste disposal. Such activities include demolition, construction, modifications, bulk chemical off-loading and transfers, landfill disposal, open burning, and scrap storage. The SWPPP is an Attachment to the SPCC plan. Runoff potentially contaminated by petroleum products or toxic/hazardous substances, historical documentation of reportable spills, inspection requirements and other elements required by both plans are contained in the text of the SPCC plan. Together these documents provide SQN's program for preventing introduction of pollutants into the environment. The primary focus of the SWPPP is the methodology used to control water generated erosion, since site erosion at SQN is due primarily to rainfall runoff.

National Pollutant Discharge Elimination System (NPDES) Permit

SQN's NPDES Permit No. TN0026450 controls the discharge of process and non-process wastewater through seven outfalls. The permit limits the discharge of various substances such as total suspended solids, oil and grease, chlorine, biocides, etc.

Outfall 101, Diffuser Pond discharge, is the primary discharge point for SQN to the Tennessee River at mile 483.65. Outfall 101 is monitored for flow, temperature, pH, TSS, chlorine, oil and grease, PCBs, and IC25 through procedures to meet all requirements of SQN's NPDES Permit.

Minimization of pollutants that could result from the backwashing activities at Outfalls 116, Backwash for Condenser Circulating Water (CCW), and Outfall 117, Backwash for Essential Raw Cooling Water (ERCW), are procedurally controlled by Operations' procedures: ERCW Strainers and Traveling Screens (0-SO-67-3), and CCW System (0-SO-27-1). Operations staff performs visual inspections for floating debris and oil and grease during backwashes.

SQN has several other permitted outfalls. Procedures are used to ensure that NPDES Permit requirements are met at these outfalls.

Process Chemicals

The types of treatment chemicals (oxidizing biocides, non-oxidizing biocides, slimicides, etc.) used, target concentration feed rates, and treatment schedule has been approved by the Tennessee Department of Environment and Conservative (TDEC). These chemicals are injected per plant procedures designed to ensure NPDES Permit Compliance.

The treatment chemicals currently in use and proposed for near term future use at SQN to control corrosion and biological infestations are described below.

Corrosion Control Program

Towerbrom-960 (oxidizing biocide) and H-130M (non-oxidizing biocide)

SQN Chemistry administers the raw water (Raw Cooling Water (RCW) and Essential Raw Cooling Water (ERCW) Trains A & B) bio-fouling treatment program as follows (Refer to Corrosion Control Tables):

All treatment must be performed in accordance with NPDES permit limitations.

Treatment and controls needed to eliminate macrofouling should be initiated as soon as possible in Spring, when the river intake temperature exceeds 55°F. SQN should treat as late as possible in Fall, until water temperatures are less than 55°F or until there are no veligers present in two consecutive twice per week samples, to eliminate the possibility of a late settlement of veligers that has occurred during the reproductive season. If a low veliger count remains present at temperatures less than 55°F for several measurements, SQN may choose to terminate oxidizing biocide treatment at that time and perform a post spawning period continuous oxidizing biocide treatment of the raw water systems. Treatment shall be conducted to include the following steps as a minimum:

1. Apply continuous oxidizing biocide to maximum target residual allowed by the NPDES Permit (this should be >0.1 ppm free available oxidant (FAO) in the system) from river temperature of approximately 55°F until the river temperature reaches approximately 60°F to 70°F.
2. Terminate oxidizing biocide treatment for approximately 72 hours, and then apply the target residual of non-oxidizing biocide for recommended duration. Note: The target residual for non-oxidizing biocide is that recommended by the chemical product information for 100% mortality during a given duration and temperature.
3. Following a non-oxidizing treatment begin 24 hours of continuous oxidizing biocide to maximum target residual allowed by the NPDES

Permit (this should be >0.1 ppm free available oxidant (FAO) in the system).

4. After the late Spring Asiatic clam and zebra mussel veliger peak is confirmed, apply a second target residual of non-oxidizing biocide treatment for the recommended duration. Caution: Terminate any oxidizing biocide treatment for approximately 72 hours before any non-oxidizing biocide treatment.
5. After the non-oxidizing treatment begin 24 hours of continuous oxidizing biocide to maximum target residual allowed by the NPDES Permit (this should be >0.1 ppm free available oxidant (FAO) in the system).
6. After the Fall Asiatic clam veliger peak is confirmed, apply a third target residual of non-oxidizing biocide treatment for the recommended duration. Caution: Terminate any oxidizing biocide treatment for approximately 72 hours before any non-oxidizing biocide treatment.
7. After the non-oxidizing treatment begin 24 hours of continuous oxidizing biocide to maximum target residual allowed by the NPDES Permit (this should be >0.1 ppm free available oxidant (FAO) in the system).
8. When the river temperature reaches 70°F to 60°F, apply a fourth target residual of non-oxidizing biocide treatment for the recommended duration. Caution: Terminate any oxidizing biocide treatment for approximately 72 hours before any non-oxidizing biocide treatment.
9. After the non-oxidizing treatment begin continuous oxidizing biocide to maximum target residual allowed by the NPDES Permit (this should be >0.1 ppm free available oxidant (FAO) in the system) until the river bottom temperature is <55°F or veliger monitoring indicates zero veliger count for two consecutive measurements OR If veliger count persists at a low level for approximately a three week period, SQN may terminate continuous treatment and perform an approximate three week post spawning period oxidizing biocide treatment to maximum target residual allowed by the NPDES Permit.

When not performing the macrofouling treatment regime described above, perform MIC control oxidizing biocide treatment to maximum target residual allowed by the NPDES Permit. MIC treatment four to eight hours per day, five to seven days per week when river temperature is >65°F and two to five days per week when river temperature is <55°F.

Towerbrom-960

Towerbrom-960, Sodium Dichloroisocyanurate and Sodium Bromide, will continue to be used to mitigate microbiologically induced corrosion and to supplement mitigation of nonnative nuisance aquatic species within plant systems. Pumps for injection of Towerbrom-960 will be calibrated across the range of chemical injection. Daily calibration checks will also be performed. These calibrations will be further verified by a calculated confirmation of pump setting versus chemical used on a daily basis.

During oxidizing biocide treatments of four to eight hours per day the Outfall 101 Total Residual Chlorine (TRC) concentration will be calculated to ensure that the NPDES Permit limits (0.058 mg/L daily max and 0.036 mg/L monthly average) are met. SQN will perform this calculation five per seven days just as the current permit specifies.

During oxidizing biocide treatments exceeding eight hours per day or if the Condenser Circulating Water (CCW) system is chlorinated or if none of the units are discharging flow from the CCW system, then grab samples will be collected at the Diffuser (Outfall 101) discharge and analysis performed once per day. These grab samples will be analyzed using a colorimeter and SQN will report a Minimum Level of Quantification (ML) of 0.08 mg/L.

Currently, the Condenser Circulating Water (CCW) is not treated. This cooling water flow provides dilution of treated cooling water systems. Should the need for CCW treatment be indicated, SQN will update this BMP plan with the details of such treatment.

High Pressure Fire Protection (HPFP) water supply at SQN is potable water, containing only that TRC remaining after piping from Hixson Utility to SQN and storage in HPFP A & B supply tanks. HPFP and additional potable water releases or overflows are taken into consideration in the calculated Outfall 101 TRC.

H-130M

To control mollusks (zebra mussels and Asiatic clams) seen in the plant piping necessary for safe shutdown and reliable operation of the plant, SQN continues to inject H-130M at a frequency of four times per ERCW Train A & B and RCW systems per year. H-130M is a quaternary amine with effectiveness dependent on temperature of the water and concentration of application.

Additionally, SQN proposes to perform whole effluent toxicity (WET) testing of the effluent at Outfall 101 during the spring, summer, and fall non-oxidizing biocide treatment periods if possible based on scheduling availability. Based on

successful demonstration of no toxicity resulting from the treatments using these test results, SQN would then continue to perform toxicity monitoring during at least one H-130M treatment per year if possible based on scheduling availability. SQN will continue to submit veliger monitoring information on a quarterly basis with the Discharge Monitoring Report (DMR) as requested in the April 25, 2003 letter from Saya Qualls to Michael Beavers. SQN will also continue to analyze for H-130M quaternary amine using a low detection level analytical method for all treatments and report these results on the DMR as requested in the April 25, 2003 letter from Saya Qualls to Michael Beavers.

PCL-222/PCL-401

SQN currently uses a combination of two chemicals to provide corrosion protection for plant carbon steel piping. These chemicals are PCL-222, an orthophosphate and hexametaphosphate/copolymer, and PCL-401, a copolymer only. PCL-222 is continuously applied to the RCW system throughout the year. PCL-222 is continuously applied to the ERCW Train A & B system when the ambient temperature is above approximately 40°F. At lower ambient temperatures, PCL-401 is continuously applied to the ERCW Train A & B system because PCL-222 is subject to freezing.

CL-363

SQN currently uses CL-363, dimethylamide (DMAD), however it will be phased out by June 2004 and replaced by Biodetergent 73551. CL-363 is injected into both the ERCW Train A & B and RCW systems. CL-363 penetrates and softens deposits and biofilm. It is injected before Towerbrom-960 treatments to enhance effectiveness of chlorination.

Biodetergent 73551

SQN will begin in June 2004 to use Biodetergent 73551 in place of CL-363. Biodetergent 73551, ethylene oxide – propylene oxide copolymer (EO/PO copolymer), will be injected into both the ERCW Train A & B and RCW systems. Biodetergent 73551 is a non-ionic compound used to remove and disperse "soft foulant" (mud, silt and clay) deposits in cooling water systems. Biodetergent 73551 demonstrates cleaning action through a combination of chemical solubilization and physical scrubbing by entrained air bubbles. Once soft foulant aggregates are dislodged into the bulk fluid, Biodetergent 73551 prevents re-establishment of soft-foulant deposits.

Biodetergent should be complemented with biocide usage so that dislodged microbial populations are killed and not given any chance to re-establish on surfaces. It will be injected before Towerbrom-960 treatments to enhance effectiveness of chlorination.

Radiological Program

Liquid Radwaste Release Points

There are four systems from which liquid effluents are released to the environment. These are the Liquid Radwaste System, the Condensate Demineralizer System, the Turbine Building Sump (TBS), and the Units 1 and 2 Steam Generator Blowdown.

All liquid effluents are ultimately discharged to the Diffuser Pond (Outfall 101) which releases to the Tennessee River (river mile 483.65) where they are evaluated for offsite dose. The Essential Raw Cooling Water (ERCW) routinely provides dilution for liquid effluents at a minimum flow rate of 15,000 gpm. ERCW flow is monitored by radiation monitors 0-RM-90-133, -134, -140, -141.

Liquid Radwaste System

The Liquid Radwaste System processes liquid from the Reactor Building and Auxiliary Building Floor Drains and the laundry/hot shower and chemical drain tanks. The normal release points for liquid radwaste are the Monitor Tank and the Cask Decontamination Collector Tank (CDCT). The Monitor Tank has a capacity of 22,000 gal and is released routinely at a flow rate of 125 gpm. The CDCT has a capacity of 15,000 gal and is also released routinely at a flow rate of 125 gpm. The Monitor Tank and CDCT discharge to the Cooling Tower Blowdown (CTBD) line as a batch release and are monitored by radiation monitor 0-RM-90-122.

Condensate Demineralizer System

The Condensate Demineralizer System processes liquid wastes coming from the High Crud Tanks (HCT-1 and -2), the Neutralization Tank, and the Non-Reclaimable Waste Tank (NRWT). The HCTs have a capacity of 20,000 gal and a maximum discharge flow rate of 245 gpm. The Neutralization Tank has a capacity of 19,000 gal and a maximum discharge flow rate of 245 gpm. The NRWT has a capacity of 11,000 gal and a maximum discharge flow rate of 245 gpm. The Condensate Demineralizer System is routinely released to the CTBD line but can be released to the TBS during periods of low radioactivity levels and is monitored by radiation monitor 0-RM-90-225.

Turbine Building Sump

The Turbine Building Sump (TBS) normally releases to the Low Volume Waste Treatment Pond (LVWTP) but can be released to the Yard Pond (YP). The TBS has a capacity of 30,000 gal and a design discharge release rate of 1,750 gpm per pump. TBS releases are monitored by radiation monitor 0-RM-90-212.

Steam Generator Blowdown

The Steam Generator Blowdown (SGBD) is processed in the Steam Generator Drindown Flash Tanks or SGBD Heat Exchangers. The SGBD discharge has a maximum flow rate of 80 gpm per steam generator. SGBD discharges to the CTBD line are continuous and are monitored by radiation monitors 1,2-RM-90-

120, -121.

Liquid Radwaste Treatment System

The liquid radwaste treatment system described below shall be maintained and operated to keep releases As Low As Reasonably Achievable (ALARA).

The system consists of one reactor coolant drain tank with two pumps and a floor and equipment drain sump inside the containment of each unit and the following shared equipment inside the auxiliary building: one sump tank and pumps, one tritiated drain collector tank with two pumps and one filter, one floor drain collector tank with two pumps and one filter, a waste condensate tank filter, three waste condensate tanks and two pumps, a chemical drain tank and pump, two laundry and hot shower tanks and pump, a spent resin storage tank, a cask decontamination tank with two pumps and two filters, Auxiliary Building floor end equipment drain sump and pumps, and evaporator with two distillate tanks, a Mobile Waste Demineralizer System (if needed) and the associated piping, valves and instrumentation.

CORROSION CONTROL TABLE 1

Treatment Chemicals	Target Concentrations	Treatment Schedule	Treatment Description	Limit at Diffuser Discharge	Comments
<u>Towerbrom-960</u> Sodium Dichloroisocyanurate & Sodium Bromide	0.2 – 1.0 ppm (TRO) <u>Target Feed Rates:</u> <u>ERCW A & B Trains:</u> ~60 gph or ~3785 mL/min each train <u>RCW (4 pumps):</u> ~60 gph or ~3785 mL/min	Temperature and veliger count dependant. (See Corrosion Control Table 2)	Towerbrom-960 is an oxidizing biocide that dissolves in water to produce chlorine and (or) bromine (depending on pH of the river water). Used for MIC, bio-fouling and veliger settling control.	<u>Calculated</u> 0.058 mg/L Daily Max 0.036 mg/L Monthly Average <u>Grabs</u> 0.08 mg/L Daily Max	See Corrosion Control Table 2 and H-130M comments.
<u>H-130M</u> Didecyl-Dimethyl Ammonium Chloride (Quat)	1.5 ppm (Quat) <u>Target Feed Rates:</u> <u>ERCW A & B Train:</u> ~2.5 gph or ~158 mL/min <u>RCW (4 pumps):</u> ~3.8 gph or ~240 mL/min	2/year ~24 hour duration 2/year ~96 hour duration	H-130M is a quaternary amine solution referred to as a non-oxidizing biocide. H-130M will effectively kill adult and juvenile zebra mussels and Asiatic clams.	<u>Calculated</u> 0.050 mg/L Daily Max	The first three H-130M treatments of the year are followed by a 24 hour Towerbrom-960 treatment. The fourth H-130M treatment of the year is followed by a 21 day, 24 hours/day Towerbrom-960 treatment.
<u>PCL-222/PCL-401</u> Copolymer, Orthophosphate & Hexametaphosphate	0.3 ppm (Copolymer) 0.5 ppm (Ortho) 0.5 ppm (Hexameta) <u>Target Feed Rate per 10K gpm:</u> <u>ERCW A & B Train and RCW</u> ~2.0 gph or ~126 mL/min (PCL-222) ~0.6 gph or ~38 mL/min (PCL-401)	Continuous <u>ERCW A & B Trains:</u> PCL-222 is used from April – October. PCL-401 is used from November – March. <u>RCW:</u> PCL-222 used year around	PCL-222 is a copolymer-phosphate blend that serves as a steel corrosion inhibitor, deposit/silt remover and sequestering agent. PCL-401 is a copolymer.	<u>Calculated</u> 0.100 mg/L Daily Max	For ERCW, PCL-401 is used in cold weather months to avoid freezing. PCL-401 contains copolymer only.
<u>CL-363</u> Dimethylamide (DMAD)	0.5 ppm (DMAD) <u>Target Feed Rate per 10K gpm:</u> <u>ERCW A & B Train and RCW</u> ~3.2 gph or ~202 mL/min	2-3/week ~30 minute duration	DMAD penetrates and softens deposits and biofilm. It is injected before Towerbrom-960 treatments to enhance effectiveness of chlorination. CL-363 and PCL-222/401 help avoid buildup of hard manganese-iron deposits.	<u>Calculated</u> 0.100 mg/L Daily Max	To be phased out by June 2004 and replaced by Biodetergent 73551.
<u>Biodetergent 73551</u> Ethylene oxide – propylene oxide copolymer (EO/PO copolymer)	2.0 ppm (EO/PO copolymer) <u>Target Feed Rate per 10K gpm:</u> <u>ERCW A & B Train and RCW</u> ~5.94 gph or ~374 mL/min	2-3/week ~30 minute duration	Biodetergent 73551 is used to remove and disperse "soft foulant" (mud, silt and clay) deposits in cooling water systems. It is injected before Towerbrom-960 treatments to enhance effectiveness of chlorination.	<u>Calculated</u> 0.100 mg/L Daily Max	SQN will begin usage of Biodetergent 73551 in June 2004.

Towerbrom-960 chemical feed duration and frequency is based on river temperature and veliger count as shown in the table below.

CORROSION CONTROL TABLE 2				
Treatment Chemical	Plant System	River Temperature	Target Concentrations	Treatment Schedule
<u>Towerbrom-960</u> Sodium Dichloroisocyanurate & Sodium Bromide	ERCW Train A & B and RCW	< 55°F *	0.2 - 1.0 ppm	2-5/week, 4-8 hours/day
<u>Towerbrom-960</u> Sodium Dichloroisocyanurate & Sodium Bromide	ERCW Train A & B and RCW	> 55°F < 65°F	0.2 - 1.0 ppm	7/week, 24 hours/day
<u>Towerbrom-960</u> Sodium Dichloroisocyanurate & Sodium Bromide	ERCW Train A & B and RCW	> 65°F	0.2 - 1.0 ppm	5-7/week, 4-8 hours/day

* Towerbrom-960 treatment during winter for duration of 21 days, 24 hours/day if intake veliger counts remain above zero at termination of > 55°F < 65°F treatment.

APPENDIX D

Page 1 of 1

NALCO 73551 INJECTION RATE RECORD SHEET

____/____/____ (Procedure Step 6.3[1])
 Month Day Year

NOTE This sheet may be reproduced as needed.

Procedure Step	Parameter	Recording Time	Recorded By	Flow or Injection Rate	Acceptance Criteria
6.3[2]	Diffuser Discharge Flow Rate [Record as (b)]			_____ (b) (cfs)	Flow Rate ≥ 1200 cfs? Yes <input type="checkbox"/> No <input type="checkbox"/>
6.3[4], 6.3[5], 6.3[6]	NALCO 73551 Injection Rates (Below record highest injection rate as (a), provided not treated concurrently with ERCW)			A Train _____ (gph) B Train _____ (gph) RCW _____ (gph)	Each Injection Rate ≤ 11 gph? Yes <input type="checkbox"/> No <input type="checkbox"/>
6.3[7]	NALCO 73551 Injection Times			Date/Time Begin Date/Time End ERCW A _____ _____ ERCW B _____ _____ RCW _____ _____	NO DATA IN THIS SPACE
Calculated Diffuser Discharge DMAD Concentration					
Calculated Diffuser Discharge = $\frac{\text{Highest NALCO 73551 Injection Rate (a), gph}}{\text{Diffuser Discharge Flow Rate (b), cfs}} \times 8.44^A = \text{mg/L EO/PO copolymer}$ Concentration (c)					
6.3[8] 6.3[9] 6.3[10] 6.3[11]	Calculated Diffuser Discharge = _____ (a), gph Concentration (c) _____ (b), cfs	x 8.44 ^A = _____ mg/L EO/PO copolymer			Calculated EO/PO copolymer ≤ 0.100 mg/L? Yes <input type="checkbox"/> No <input type="checkbox"/>
6.3[13]	Performed By: _____ Date/Time _____ / _____				
Comments:					

^A 8.44 = (1hr/3600sec) x (1cf/7.4805gal) x (1.01 NALCO 73551 Sp. Gr.) x (0.225 NALCO 73551 EO/PO copolymer weight fraction) x (1E+6 mg/L).

^B If limit is exceeded, IMMEDIATELY ISOLATE chemical feed AND PROMPTLY NOTIFY Chemistry Shift Supervisor, Shift Manager, and Environmental Supervisor or designee.

**MATERIAL SAFETY DATA SHEET****PRODUCT****73551****EMERGENCY TELEPHONE NUMBER(S)****(800) 424-9300 (24 Hours) CHEMTREC****1. CHEMICAL PRODUCT AND COMPANY IDENTIFICATION****PRODUCT NAME :** 73551**APPLICATION :** DEPOSIT PENETRANT**COMPANY IDENTIFICATION :** Nalco Company
1601 W. Diehl Road
Naperville, Illinois .
60563-1198**EMERGENCY TELEPHONE NUMBER(S) :** (800) 424-9300 (24 Hours) CHEMTREC**NFPA 704M/HMIS RATING****HEALTH: 0/1 FLAMMABILITY: 1/1 INSTABILITY: 0/0 OTHER:**

0 = Insignificant 1 = Slight 2 = Moderate 3 = High 4 = Extreme

2. COMPOSITION/INFORMATION ON INGREDIENTS

Based on our hazard evaluation, none of the substances in this product are hazardous.

3. HAZARDS IDENTIFICATION****EMERGENCY OVERVIEW******CAUTION**

May cause irritation with prolonged contact.

Do not get in eyes, on skin, on clothing. Do not take internally. Wear suitable protective clothing. Keep container tightly closed. Flush affected area with water.

May evolve oxides of carbon (COx) under fire conditions.

PRIMARY ROUTES OF EXPOSURE :

Eye, Skin

HUMAN HEALTH HAZARDS - ACUTE :**EYE CONTACT :**

May cause irritation with prolonged contact.

SKIN CONTACT :

May cause irritation with prolonged contact.

INGESTION :

Not a likely route of exposure. No adverse effects expected.

INHALATION :

Not a likely route of exposure. No adverse effects expected.

Nalco Company 1601 W. Diehl Road • Naperville, Illinois 60563-1198

(630)305-1000

1 / 8



MATERIAL SAFETY DATA SHEET

PRODUCT

73551

EMERGENCY TELEPHONE NUMBER(S)

(800) 424-9300 (24 Hours) CHEMTREC

SYMPTOMS OF EXPOSURE :

Acute :

A review of available data does not identify any symptoms from exposure not previously mentioned.

Chronic :

A review of available data does not identify any symptoms from exposure not previously mentioned.

AGGRAVATION OF EXISTING CONDITIONS :

A review of available data does not identify any worsening of existing conditions.

4. FIRST AID MEASURES

EYE CONTACT :

Flush affected area with water. If symptoms develop, seek medical advice.

SKIN CONTACT :

Flush affected area with water. If symptoms develop, seek medical advice.

INGESTION :

Do not induce vomiting without medical advice. If conscious, washout mouth and give water to drink. If symptoms develop, seek medical advice.

INHALATION :

Remove to fresh air, treat symptomatically. If symptoms develop, seek medical advice.

NOTE TO PHYSICIAN :

Based on the individual reactions of the patient, the physician's judgement should be used to control symptoms and clinical condition.

5. FIRE FIGHTING MEASURES

FLASH POINT : > 400 °F / > 200 °C (COC)

EXTINGUISHING MEDIA :

This product would not be expected to burn unless all the water is boiled away. The remaining organics may be ignitable. Use extinguishing media appropriate for surrounding fire.

FIRE AND EXPLOSION HAZARD :

May evolve oxides of carbon (COx) under fire conditions.

SPECIAL PROTECTIVE EQUIPMENT FOR FIRE FIGHTING :

In case of fire, wear a full face positive-pressure self contained breathing apparatus and protective suit.

6. ACCIDENTAL RELEASE MEASURES

PERSONAL PRECAUTIONS :

Do not touch spilled material. Restrict access to area as appropriate until clean-up operations are complete. Use personal protective equipment recommended in Section 8 (Exposure Controls/Personal Protection). Stop or reduce any leaks if it is safe to do so. Ventilate spill area if possible.

**MATERIAL SAFETY DATA SHEET****PRODUCT****73551****EMERGENCY TELEPHONE NUMBER(S)****(800) 424-9300 (24 Hours) CHEMTREC****METHODS FOR CLEANING UP :**

SMALL SPILLS: Soak up spill with absorbent material. Place residues in a suitable, covered, properly labeled container. Wash affected area. **LARGE SPILLS:** Contain liquid using absorbent material, by digging trenches or by diking. Reclaim into recovery or salvage drums or tank truck for proper disposal. Contact an approved waste hauler for disposal of contaminated recovered material. Dispose of material in compliance with regulations indicated in Section 13 (Disposal Considerations).

ENVIRONMENTAL PRECAUTIONS :

Do not contaminate surface water.

7. HANDLING AND STORAGE**HANDLING :**

Avoid eye and skin contact. Do not take internally. Ensure all containers are labelled. Keep the containers closed when not in use.

STORAGE CONDITIONS :

Store the containers tightly closed.

8. EXPOSURE CONTROLS/PERSONAL PROTECTION**OCCUPATIONAL EXPOSURE LIMITS :**

This product does not contain any substance that has an established exposure limit.

ENGINEERING MEASURES :

General ventilation is recommended.

RESPIRATORY PROTECTION :

Respiratory protection is not normally needed.

HAND PROTECTION :

Neoprene gloves, Nitrile gloves, Butyl gloves, PVC gloves

SKIN PROTECTION :

Wear standard protective clothing.

EYE PROTECTION :

Wear chemical splash goggles.

HYGIENE RECOMMENDATIONS :

Keep an eye wash fountain available. Keep a safety shower available. If clothing is contaminated, remove clothing and thoroughly wash the affected area. Launder contaminated clothing before reuse.

9. PHYSICAL AND CHEMICAL PROPERTIES**PHYSICAL STATE**

Liquid



MATERIAL SAFETY DATA SHEET

PRODUCT

73551

EMERGENCY TELEPHONE NUMBER(S)

(800) 424-9300 (24 Hours) CHEMTREC

APPEARANCE Clear Colorless
ODOR None
SPECIFIC GRAVITY 0.99 - 1.03 @ 77 °F / 25 °C
SOLUBILITY IN WATER Complete
pH (100 %) 6.6 - 7.0

Note: These physical properties are typical values for this product and are subject to change.

10. STABILITY AND REACTIVITY

STABILITY :
Stable under normal conditions.

HAZARDOUS POLYMERIZATION :
Hazardous polymerization will not occur.

CONDITIONS TO AVOID :
Freezing temperatures.

MATERIALS TO AVOID :
None known

HAZARDOUS DECOMPOSITION PRODUCTS :
Under fire conditions: Oxides of carbon

11. TOXICOLOGICAL INFORMATION

The following results are for the polymer.

ACUTE ORAL TOXICITY :		Test Descriptor
Species	LD50	The following results are for the polymer.
Rat	2,300 - 16,000 mg/kg	
Rating :	Non-Hazardous	

CARCINOGENICITY :
None of the substances in this product are listed as carcinogens by the International Agency for Research on Cancer (IARC), the National Toxicology Program (NTP) or the American Conference of Governmental Industrial Hygienists (ACGIH).

HUMAN HAZARD CHARACTERIZATION :
Based on our hazard characterization, the potential human hazard is: Low

12. ECOLOGICAL INFORMATION

ECOTOXICOLOGICAL EFFECTS :

**MATERIAL SAFETY DATA SHEET****PRODUCT****73551****EMERGENCY TELEPHONE NUMBER(S)****(800) 424-9300 (24 Hours) CHEMTREC**

Proper Shipping Name :

PRODUCT IS NOT REGULATED DURING
TRANSPORTATION

MARINE TRANSPORT (IMDG/IMO) :

Proper Shipping Name :

PRODUCT IS NOT REGULATED DURING
TRANSPORTATION**15. REGULATORY INFORMATION**

NATIONAL REGULATIONS, USA :

OSHA HAZARD COMMUNICATION RULE, 29 CFR 1910.1200 :

Based on our hazard evaluation, none of the substances in this product are hazardous.

CERCLA/SUPERFUND, 40 CFR 117, 302 :

Notification of spills of this product is not required.

SARA/SUPERFUND AMENDMENTS AND REAUTHORIZATION ACT OF 1986 (TITLE III) - SECTIONS 302, 311,
312, AND 313 :

SECTION 302 - EXTREMELY HAZARDOUS SUBSTANCES (40 CFR 355) :

This product does not contain substances listed in Appendix A and B as an Extremely Hazardous Substance.

SECTIONS 311 AND 312 - MATERIAL SAFETY DATA SHEET REQUIREMENTS (40 CFR 370) :

Our hazard evaluation has found that this product is not hazardous under 29 CFR 1910.1200.

Under SARA 311 and 312, the EPA has established threshold quantities for the reporting of hazardous chemicals. The current thresholds are: 500 pounds or the threshold planning quantity (TPQ), whichever is lower, for extremely hazardous substances and 10,000 pounds for all other hazardous chemicals.

SECTION 313 - LIST OF TOXIC CHEMICALS (40 CFR 372) :

This product does not contain substances on the List of Toxic Chemicals.

TOXIC SUBSTANCES CONTROL ACT (TSCA) :

The substances in this preparation are included on or exempted from the TSCA 8(b) Inventory (40 CFR 710)

FOOD AND DRUG ADMINISTRATION (FDA) Federal Food, Drug and Cosmetic Act :

When use situations necessitate compliance with FDA regulations, this product is acceptable under : 21 CFR 173.340 Defoaming Agents, 21 CFR 175.105 - Adhesives, 21 CFR 176.200 Defoaming Agents used in coatings, 21 CFR 176.210 Defoaming agents used in the manufacture of paper and paperboard, 21 CFR 176.300 - Slimicides, 21 CFR 177.1200 - Cellophane, 21 CFR 177.1400 - Hydroxyethyl cellulose film, water-insoluble, 21 CFR 178.1010 - Sanitizing solutions, 21 CFR 178.3120 - Animal glue

This product has been certified as KOSHER/PAREVE for year-round use INCLUDING THE PASSOVER SEASON by the CHICAGO RABBINICAL COUNCIL.



MATERIAL SAFETY DATA SHEET

PRODUCT

73551

EMERGENCY TELEPHONE NUMBER(S)

(800) 424-9300 (24 Hours) CHEMTREC

FEDERAL WATER POLLUTION CONTROL ACT, CLEAN WATER ACT, 40 CFR 401.15 / formerly Sec. 307, 40 CFR 116.4 / formerly Sec. 311 :

None of the substances are specifically listed in the regulation.

CLEAN AIR ACT, Sec. 111 (40 CFR 60, Volatile Organic Compounds), Sec. 112 (40 CFR 61, Hazardous Air Pollutants), Sec. 602 (40 CFR 82, Class I and II Ozone Depleting Substances) :

None of the substances are specifically listed in the regulation.

CALIFORNIA PROPOSITION 65 :

This product does not contain substances which require warning under California Proposition 65.

MICHIGAN CRITICAL MATERIALS :

None of the substances are specifically listed in the regulation.

STATE RIGHT TO KNOW LAWS :

The following substances are disclosed for compliance with State Right to Know Laws:

Water	7732-18-5
Polyalkylene glycol	Proprietary

NATIONAL REGULATIONS, CANADA :

WORKPLACE HAZARDOUS MATERIALS INFORMATION SYSTEM (WHMIS) :

This product has been classified in accordance with the hazard criteria of the Controlled Products Regulations (CPR) and the MSDS contains all the information required by the CPR.

WHMIS CLASSIFICATION :

Not considered a WHMIS controlled product.

CANADIAN ENVIRONMENTAL PROTECTION ACT (CEPA) :

The substances in this preparation are listed on the Domestic Substances List (DSL), are exempt, or have been reported in accordance with the New Substances Notification Regulations.

16. OTHER INFORMATION

Due to our commitment to Product Stewardship, we have evaluated the human and environmental hazards and exposures of this product. Based on our recommended use of this product, we have characterized the product's general risk. This information should provide assistance for your own risk management practices. We have evaluated our product's risk as follows:

* The human risk is: Low

* The environmental risk is: Low

Any use inconsistent with our recommendations may affect the risk characterization. Our sales representative will assist you to determine if your product application is consistent with our recommendations. Together we can implement an appropriate risk management process.