Pacific Gas and Electric Company®

Tom Wilson
Director
Environmental Remediation

3401 Crow Canyon Road
San Ramon, CA 94583

Phone: (925) 415-6340
Mobile: (415) 672-3324
Fax: (925) 415-6352
TCW1@pge.com

October 25, 2011

State Water Resources Control Board
Office of Chief Counsel
Jeannette L. Bashaw, Legal Analyst
P.O. Box 100
Sacramento, CA 95812-0100
jbashaw@waterboards.ca.gov

Chair Hoppin and Members of the State Water Resources Control Board:

Pacific Gas and Electric Company (PG&E) respectfully submits the attached Petition and Request for Stay regarding the Amended Cleanup and Abatement Order No. R6V-2011-0005A1 (WDID NO. 6B369107001) ("CAO") for the Hinkley Compressor Station issued on October 11, 2011 by Executive Officer Harold Singer of the Lahontan Regional Water Quality Control Board ("Lahontan Regional Board").

PG&E has for many years acknowledged with genuine regret its responsibility for chromium contamination in the Hinkley community. PG&E is committed to working cooperatively with the Lahontan Regional Board to expeditiously clean up groundwater contamination resulting from PG&E's historical operations at the Hinkley Compressor Station. We share the mutual goal of ensuring safe, reliable drinking water for the residents of Hinkley to ease their concerns for community health and well-being. To that end, PG&E has taken several steps to address issues raised by the community in Hinkley.

Specifically, PG&E has:

- established a bottled water program that currently delivers water for household use to over 230 households within a mile of the current plume boundary;
- made a public commitment to explore whole household water replacement options with the community and to reporting out frequently on our progress toward this goal;
- established a community resource office in Hinkley to better address the community's questions and concerns;
- established a Community Advisory Committee to provide direct input to PG&E on our remediation activities and other community programs; and
- begun actively working with the Hinkley community to establish an independent panel of technical experts to advise them on technical and other matters of community interest.

While we stand behind these commitments, we nevertheless feel we must submit this petition. This is the first time PG&E has requested action by the State Board on the Hinkley site in more than 20 years of environmental investigation and cleanup conducted under the oversight of the Lahontan Region Board. However, as you will read in the attached petition and request for stay, PG&E has serious concerns that the CAO, as written, sets a troubling precedent for determination of replacement water in cases of groundwater contamination that has implications for water providers and consumers statewide. We have been unable to locate any precedent for the requirements as written, and believe they are not supported by California law. In addition, the CAO sets infeasible requirements for bottled water and permanent whole-house replacement water. Lastly, the CAO timing requirements are such that PG&E believes it will be unable to comply. So, although PG&E is voluntarily responding to the community's concerns and responding to those portions of the CAO that are supported by law and are feasible using current tools and technologies, we must seek relief from the CAO provisions that are either not legally supportable and/or are infeasible. We also believe it is necessary to express concern about several of the CAO provisions which are likely to have wide-ranging and statewide policy implications. Ultimately, we are very concerned that the CAO sends the message to residents throughout California that their tap water and/or the purchased bottled water they currently rely on may not be safe to drink.

Despite this petition and request for stay, we assure members of the State Board that our concerns with this recent Lahontan Regional Board order do not affect our continuing commitment to do what is right for the Hinkley community. PG&E will continue to honor our commitment to provide safe drinking water to the community through our voluntary bottled water program, and we will continue to aggressively test whole-house water technologies and report out on our progress to the Lahontan Regional Board and to the Hinkley community. Our overarching desire is to ensure that the permanent whole house water solution is feasible, implementable and sustainable for the residents of Hinkley. We will also continue to work cooperatively with the Lahontan Regional Board in our ongoing efforts to restore the water quality in Hinkley in accordance with state law and in full compliance with orders previously issued by the Lahontan Regional Board.

Thank you for your attention to this matter,

Sheryl Bilbrey for:
Thomas C. Wilson

TRACY J. EGOSCUE (SB# 190842)
PAUL HASTINGS LLP
515 South Flower Street
Twenty-Fifth Floor
Los Angeles, CA 90071-2228
Telephone: 1(213) 683-6242
Facsimile: 1(213) 627-0705
Email: tracyegoscue@paulhastings.com

J. DREW PAGE (SB# 146437)
LAW OFFICES OF J. DREW PAGE
11622 El Camino Real Ste 100
San Diego, CA 92130
Telephone: 1(858) 433-0122
Facsimile: 1(858)433-0124
Email: drew@jdp-law.com

ROBERT D. INFELISE (SB# 93876)
COX, CASTLE & NICHOLSON LLP
555 California Street 10th Floor
San Francisco, CA 94104
Telephone: 1(415) 262-5125
Facsimile: 1(415) 392-4250
Email: rinfelise@coxcastle.com

Attorneys for Petitioner
Pacific Gas & Electric Company

STATE OF CALIFORNIA

STATE WATER RESOURCES CONTROL BOARD

IN THE MATTER OF LAHONTAN REGIONAL WATER QUALITY CONTROL BOARD AMENDED CLEANUP AND ABATEMENT ORDER NO. R6V-2011-0005A1	REQUEST FOR IMMEDIATE AND EMERGENCY STAY; PETITION FOR REVIEW; MEMORANDUM OF POINTS AND AUTHORITIES IN SUPPORT THEREOF

This *Request for Immediate and Emergency Stay; Petition for Review; and Memorandum of Points and Authorities in Support Thereof* is respectfully submitted to the California State Water Resources Control Board ("State Board") on behalf of Pacific Gas

and Electric Company ("PG&E" or "Petitioner") pursuant to Water Code Sections 13320(a) and 13321, and California Code of Regulations ("CCR") Title 23, Section 2050 *et seq.,* for review of the Amended Cleanup and Abatement Order No. R6V-2011-0005A1 (WDID NO. 6B369107001) (the "CAO") with respect to the Hinkley Compressor Station located at 35863 Fairview Road (APN 048S-112-52) in Hinkley, California (the "Facility"). A copy of the CAO is attached as Attachment I.

Petitioner has for many years acknowledged with genuine regret its responsibility for chromium contamination of local groundwater in the Hinkley, California community.[1] As part of its remediation efforts, Petitioner operates what is believed to be the largest on-site chromium remediation system in the world as well as a large land treatment unit at the Desert View Dairy.[2] Petitioner has also been controlling a portion of the plume with a large fresh water injection system and has recently taken steps to increase plume control pumping by more than 300 percent.[3]

In addition to these and other projects, Petitioner has been actively working to reduce the concerns of Hinkley residents by implementing a number of voluntary programs including, among other things, providing bottled water to all residents within one mile of the plume (currently over 230 residents are participating in this program).[4] As a result, Petitioner was supportive when on January 7, 2011, the Executive Officer of the Regional Water Quality Control Board-Lahontan Region ("Lahontan Regional Board") issued an order requiring bottled water for wells containing chromium levels above natural background. PG&E was already providing bottled water to a much larger number of well owners in the area up to one mile beyond the 3.1 µg/L plume boundary. Independent of PG&E's compliance with any order issued by the Lahontan Regional Board, PG&E also committed in a public meeting on August 24, 2011 to voluntarily work with the community to explore options for providing whole house water replacement to

---

[1] Declaration of Thomas C. Wilson ("Wilson Decl."), dated October 24, 2011, at ¶ 2.
[2] Id. at ¶ 3.
[3] Id.
[4] Id. at ¶ 4(a).

further reduce community concerns. PG&E reported out on progress in this effort at a public meeting on September 28, 2011.

While PG&E stands behind these commitments, the CAO issued on October 11, 2011, by the Executive Officer of the Lahontan Regional Board contains significant and insurmountable challenges. For example, the CAO requires PG&E to provide interim bottled water and permanent whole house water to potentially more than 300 well owners at levels below the recently adopted Public Health Goal ("PHG") level of 0.02 µg/L hexavalent chromium.[5] The bottled water PG&E currently provides comes from Culligan®, a national provider of bottled water. Culligan® tests its water for total chromium and results are non-detect at a detection limit of 0.5 µg/L. Because the CAO requires documentation that replacement water is below 0.02 µg/L hexavalent chromium, the water from Culligan® would not satisfy the requirements of the CAO. As described more fully in PG&E's *Memorandum of Points and Authorities* in its Petition below, it is not feasible to guarantee the water provided to this community will be below the 0.02 µg/L requirement using current technology for either bottled water or permanent whole house replacement water. PG&E will continue to investigate any new technologies or developments in replacement water provisions that could possibly meet the very low levels required by the CAO.

PG&E is concerned also, that use of the 0.02 µg/L hexavalent chromium level to trigger replacement water and as a standard for replacement water are requirements that are not supported by California law. Specifically, the CAO would require that Petitioner:

> (i) provide replacement water that is documented to contain less than 0.02 µg/L hexavalent chromium, (the PHG) and over 100 times lower than the naturally occurring background concentration of 3.1 µg/L hexavalent chromium, when no known bottled water or other source will guarantee that Petitioner can meet this standard;

[5] The CAO notes that for purposes of this standard, replacement water must test below the laboratory reporting limit of 0.06 µg/L due to the limitation of laboratory analysis of low levels of chromium. *See* CAO No. R6V-2011-005 at fn. 1. Even the 0.06 µg/L limit is over 800 times lower than the total chromium MCL and over 50 times lower than the natural background for hexavalent chromium in Hinkley.

(ii) provide replacement water for wells that contain hexavalent chromium concentrations 2,500 times lower than the existing maximum contaminant level ("MCL") for total chromium (which specifically includes hexavalent chromium);

(iii) provide replacement water for wells that contain hexavalent chromium concentrations that are below naturally occurring background levels for the Hinkley area as set by the Lahontan Regional Board; and

(iv) provide whole house replacement water for all purposes even though the CAO acknowledges that the only health risk is from ingestion.

The requirements listed above do not appear to be consistent with California law. Under California law, there are no provisions that grant Regional Boards the authority to require replacement water that is more stringent than natural background and that cannot be supplied using standard commercially-available bottled water that is consumed by residents throughout the State. California law does not allow Regional Boards to require replacement water for wells that already fall significantly below an existing MCL. State Board policy does not outline a process for Regional Boards to require replacement water for wells that are below naturally occurring background levels. There are no provisions of California law that would allow Regional Boards to require water for non-ingestion needs when risks are limited to ingestion.

The legal authority cited by the Lahontan Regional Board for the requirements of the CAO is the 2011 hexavalent chromium PHG produced by the California Office of Environmental Health Hazard Assessment ("OEHHA"). However, OEHHA's fact sheet that accompanies the PHG specifically states that the PHG is not a regulatory standard.[6] In response to written questions from the Lahontan Regional Board Executive Officer pertaining to the CAO in draft form, OEHHA stated: "[t]he PHG is a non-regulatory guideline that does not define an acceptable level of a contaminant in

[6] "Fact Sheet - Final Public Health Goal for Hexavalent Chromium," dated July, 27 2011, *available at* http://oehha.ca.gov/public_info/facts/Cr6facts072711.html.

drinking water."[7] We respectfully disagree that the PHG provides support for the CAO as written.

PG&E is committed to a regulatory process that reflects responsible, fair, and legally enforceable standards for remediation and to meeting as many of the requirements of the CAO as are achievable as quickly as possible. For instance, PG&E voluntarily offers bottled water to all well owners within one mile of the plume,[8] as well as to the Hinkley School and the Hinkley Senior Center.[9] PG&E is also voluntarily investigating potential options to provide whole house replacement water in response to community concerns. Since early 2011, PG&E has undertaken significant efforts to re-engage the community, including establishing and staffing a local customer service office for residents to meet local personnel and to request bottled water; holding community meetings; and forming a project Community Advisory Committee which is engaged in making recommendations to PG&E and responding to concerns regarding household water and PG&E remediation efforts. PG&E is currently working with a subcommittee of the Community Advisory Committee to design an acceptable framework for an Independent Technical Review Panel to assist the community in understanding complex technical issues and addressing other topics of community interest. However, PG&E feels it must challenge the CAO because it is infeasible to comply with its requirements and believes it is unsupported by California law.[10]

California law provides that any person aggrieved by an action of a Regional Board may petition the State Board within a thirty (30) day period in accordance with Water Code section 13320 and CCR tit. 23, section 2050 *et seq*. PG&E has filed this

---

[7] Id. See also, *Memorandum of Points and Authorities* at Section 7.

[8] There is no evidentiary support in the record for the CAO's requirement that PG&E provide bottled water to wells up to one mile away from the chromium plume boundary, or the "Affected Area" in the CAO. However, PG&E voluntarily offers bottled water to this entire area.

[9] Wilson Decl. at ¶ 4(c)-(d).

[10] PG&E raised many of the aforementioned issues in a *Comment Letter on the Draft CAO No. R6V-2011-0005A1*, by Thomas C. Wilson to Harold Singer of the Lahontan Regional Board, dated July 12, 2011 ("Comment Letter"). The Comment Letter is referenced in this document and attached with all supporting exhibits and declarations as Attachment II.

Petition for Review and Request for Immediate and Emergency Stay almost immediately upon the effective date of the CAO. Petitioners did not have the right to the customary thirty (30) day period due to the fact that CAO mandates became legally enforceable within five (5) days of the date of issuance on October 11, 2011. As such, by and through this Petition, Petitioner requests an immediate and emergency stay so that a full review of the issues raised by the CAO may occur.[11]

1. **NAME AND ADDRESS OF PETITIONER**

The contact information for Petitioners is as follows:

Juan Jayo
Pacific Gas & Electric Company
Director of Environmental Remediation and Litigation
One Market Spear Tower, Suite 2400
San Francisco, CA 94105
Phone: 1(415) 973-4377
Fax: 1(415)973-5520
Email: jmj8@pge.com

With a copy to:

Tracy J. Egoscue
Paul Hastings LLP
515 S Flower Street, 25th Floor
Los Angeles, California 9007
Phone: 1(213) 683-6242
Fax: 1(213) 996-63242
Email: tracyegoscue@paulhastings.com

With a copy to:

J. Drew Page
Law Offices of J. Drew Page
11622 El Camino Real Ste 100
San Diego, CA 92130
Phone: 1(858) 433-0122
Fax: 1(858) 433-0124
Email: drew@jdp-law.com

[11] Given the urgency of this response, Petitioner reserves the right to supplement this petition further.

With a copy to:

Robert D. Infelise
Cox, Castle & Nicholson LLP
555 California Street 10th Floor
San Francisco, CA 94104
Phone: 1(415) 262-5125
Fax: 1(415) 392-4250
Email: rinfelise@coxcastle.com

2. **SPECIFIC ACTION OR INACTION FOR WHICH THIS PETITION FOR REVIEW IS SOUGHT**

Petitioner requests review of the actions of the Lahontan Regional Board in connection with the issuance of the CAO, entitled "Amended Cleanup and Abatement Order No. R6V-2011-0005A1 (WDID No. 6B369107001) Requiring Pacific Gas and Electric Company to Clean Up and Abate Waste Discharges of Total and Hexavalent Chromium to the Groundwaters of the Mojave Hydrologic Unit," dated October 11, 2011.

Additionally, Petitioner respectfully requests that the State Board issue an immediate and emergency stay of CAO No. R6V-2011-0005A1 for the reasons delineated in the *Request for Immediate and Emergency Stay* below.

****

REQUEST FOR IMMEDIATE and EMERGENCY STAY

Pursuant to Water Code section 13321 and Title 23, CCR section 2053, Petitioner requests an immediate and emergency stay of the CAO.

Under section 2053 of the State Board's regulations (CCR, tit. 23, § 2053), a stay of the effect of an order shall be granted if petitioner shows: (i) There will be substantial harm to the Petitioner or to the public interest if a stay is not granted; (ii) There will be no substantial harm to other interested persons and to the public interest if a stay is granted; and (iii) There are substantial questions of fact or law regarding the disputed action.

Pursuant to 23 CCR 2053, "a petition for stay shall be supported by a declaration under penalty of perjury of a person or persons having knowledge of the facts alleged." As such, this *Request for Immediate and Emergency Stay* is accompanied by the following declarations which are attached as follows:

- DECLARATION OF THOMAS C. WILSON, PG&E Director of Remediation (Attachment III); and
- DECLARATION OF ANITA BROUGHTON, Haley & Aldrich, Lead Risk Assessor (Attachment IV).

There will be Substantial Harm to the Petitioner or to the Public Interest if a Stay is not Granted

If the CAO is not stayed, Petitioner will suffer substantial harm because compliance with the CAO's mandates is infeasible and inconsistent with state law. Specifically, (1) given that current acceptable, commercial water providers cannot guarantee sufficient supply at the 0.02 µg/L PHG level, this standard and the immediate requirement for replacement water is infeasible and compliance within the timeframe demanded is not possible; (2) the CAO harms the public interest and sets a troubling statewide precedent by suggesting that drinking water supplies across California that currently meet the 50 µg/L MCL safe drinking water standard but are hundreds of times higher than 0.02 µg/L PHG pose an immediate risk to human health; and (3) the order requiring the setting of background levels at each and every potentially-affected well is not feasible, particularly within the outlined timeline.

(1) **Petitioner will be Subject to Inappropriate Liability because an Order for Replacement Water at the 0.02 µg/L PHG level is Infeasible and Inconsistent with State Law.**

The CAO requires that PG&E supply interim replacement water within five (5) days from the date of the issuance of the Order and provide a report within fourteen (14) days that shows that the interim water supply (commercially available bottled water)

meets hexavalent chromium levels of less than 0.02 µg/L.[12] Such replacement water requirements are not feasible "because there is no known technology or combination of technologies that can reliably achieve hexavalent chromium levels at less than 0.02 µg/L" and because such deadlines are impossible to meet.[13]

The CAO's definition of "Affected Area" contains more than 300 wells and, therefore, pending a standard for identifying wells "impacted by the discharger's chromium," could require that Petitioner provide interim replacement water for up to 300 homes in the Hinkley area.[14] Using standard household demand levels, the CAO would require that PG&E provide between an estimated 247,500 to 297,000 gallons of interim replacement water per month within fourteen (14) to forty-five (45) days of the CAO.[15]

Because fourteen (14) to forty-five (45) days does not provide enough time to study or develop long-range or whole house water replacement options, bottled water represents the "best option" for compliance with the CAO.[16] Despite being the "best option," however, the bottled water option would still not guarantee compliance with the CAO.[17]

When PG&E first explored options for providing bottled water to the Hinkley community, it selected Culligan® because it is among the top providers of replacement water and Culligan® routinely tests for total chromium.[18] Even so, PG&E

---

[12] Lahontan Regional Board, CAO No. R6V-2011-0005A1 *at* Order 1.(a)-(b).

[13] Declaration of James DeWolfe ("DeWolfe Decl."), Principal Environmental Engineer, Arcadis U.S., Inc., dated July 8, 2011 *at* ¶ 2(a)-(b) (an original attachment to PG&E's Comment Letter on Draft CAO No. R6V-2011-0005A1, attached as Attachment II.)

[14] DeWolfe Decl. at ¶ 3(a). DeWolfe stated that "because the declared background Cr6 [hexavalent chromium] concentrations in the Hinkley area average 1.2 ppb (µg/L), and have a declared maximum value of 3.1 ppb (µg/L), we assumed that all wells in the 'affected area' . . . will have Cr6 [hexavalent chromium] levels above the Public Health Goal of 0.02 ppb (µg/L)."

[15] Id. at ¶ 3(c) (stating that "the average household was expected to consume 33 gallons per day for drinking and cooking based on estimates for daily per capita faucet use."); *See generally*, CAO No. R6V-2011-0005A1 *at* Order No.(s) 1., 3.

[16] Id. at ¶ 8.

[17] Id. at ¶ 8.

[18] Much bottled water is known to contain concentration levels of chromium hundreds of times higher than the PHG. In a 2008 study of 132 brands of bottled water, tested water had a median concentration of about 8.2 µg/L, which is over 400 times higher than the PHG. *See* DeWolfe Decl. at ¶ 8(a), *citing* Krachler, M. and Shotyk, W. (2008) *Trace and Ultratrace Metals in Bottled Waters: Survey of Sources Worldwide and Comparison with Refillable Metal Bottles*," Science of Total Environment, 407:1089-1096.

has informed the Lahontan Regional Board that Culligan's® current testing demonstrates that its water is non-detect for total chromium at a detection limit of 0.5 µg/L. Although such a detection limit is 100 times lower than the State MCL standard of 50 µg/L and over six times lower than background concentrations of 3.1 µg/L, such testing would not satisfy the requirements of the CAO, because the detection limit exceeds the 0.02 µg/L PHG standard by a factor of 25 (i.e., 0.5/0.02 = 25).

PG&E could attempt to develop its own process for testing the Culligan® bottles of water for hexavalent chromium at this lower detection limit. However, such testing would require an accurate test method, testing would require PG&E to break water bottle seals and expose water to potentially harmful microbes and spoilage,[19] and testing would require systems and infrastructure that could not be established in the mandated time frame.

An additional complication to consider is the difficulty in securing sufficient bottled water at the mandated quality and quantities and to coordinate and achieve delivery of between 247,500 and 297,000 gallons of interim replacement water to up to 300 households within fourteen (14) to forty-five (45) days.

(1) The Public Interest will Suffer in Light of a Troubling Precedent Created by the CAO.

The CAO would also create a troubling statewide public policy precedent with significant ramifications. The CAO's assertion that drinking water above 0.02 µg/L hexavalent chromium could pose an immediate risk to human health could require all entities within the Lahontan Region and throughout California to reconsider the distribution of potable water with greater than 0.02 µg/L hexavalent chromium levels despite the 50 µg/L MCL.[20] According to a California Department of Public Health

[19] *See* DeWolfe Decl. at ¶ 8(c).

[20] The CAO's claim that drinking water above 0.02 µg/L hexavalent chromium poses an immediate health risk is directly contradicted by OEHHA, the author of the PHG. OEHHA specifically stated that the hexavalent chromium PHG does not identify a line between safe and unsafe concentrations of chromium in drinking water. *See* "Fact Sheet - Final Public Health Goal for Hexavalent Chromium," dated July, 27 2011, *available at* http://oehha.ca.gov/public_info/facts/Cr6facts072711.html.

("CDPH") study which looked at hexavalent chromium concentrations in drinking water between 1997-2008, "over 7,000 drinking water sources showed [hexavalent chromium] at or above the 1 µg/L detection limit for the purposes of reporting (DLR) in about one-third of them [over 2,200 sources]."[21]

An overview of CDPH's findings of California wells (including many wells which were screened using a 1 µg/L reporting limit) is presented below in Table 1.

Table 1. Chromium-6 in drinking water sources 1997 through 2008 (Active and Standby Sources)

Peak Level (µg/L)	No. of Sources	% of Detections
>50	6	0.3
41 – 50	5	0.2
31 – 40	14	0.6
21 – 30	61	2.8
11 – 20	231	10.5
6 – 10	456	20.7
1 – 5	1,434	64.9
TOTAL	2,207	100

NOTES:

1. Data are extracted from monitoring results through January 2009. They will change with subsequent updates and should be considered draft.
2. "Sources" are active, standby, and pending sources reporting more than a single detection of chromium-6. Data may include both raw and treated sources, distribution systems, blending reservoirs, and other sampled entities. This table does not include inactive sources, abandoned or destroyed wells, agricultural wells, monitoring wells, or more than one representation of the same source (e.g., a source with both raw and treated entries is counted as a single source).
3. For UCMR sampling, a number of sources may have been screened using a 1 µg/L reporting limit for total chromium. If total chromium was below the screening level, specific analysis for chromium-6 was not required.

Additionally, as part of the same study, the CDPH noted the following findings, which show that the drinking water supplies in many communities across California contain levels of hexavalent chromium well below the 50 µg/L MCL for safe drinking water but hundreds of times higher than the 0.02 µg/L PHG:

- In 1997-98, DHS sampled 10 wells in Merced County. Total

[21] "Chromium-6 in Drinking Water Sources: Sampling Results," California Department of Public Health, last updated July 27, 2011, *available at* http://www.cdph.ca.gov/certlic/drinkingwater/pages/chromium6sampling.aspx.

chromium was 18.8-33.5 µg/L, and chromium-6, 16.8-33.0 µg/L (68-100% of total chromium).

- In 1998, DHS sampled three wells in Los Angeles County. Total chromium was 3.4-10.3 µg/L, and chromium-6, nondetect-5.1 µg/L (<14-54% of total chromium).
- Also in 1998, DHS sampled one well in Contra Costa County. Total chromium was 18.7 µg/L and chromium-6, 1.5 µg/L (8% of total chromium).
- In 1999, a water agency's consultant sampled nine wells in Los Angeles County. Total chromium was 5.3-15 µg/L, and chromium-6, 3.6-11 µg/L (58-100% of total chromium).
- In 2000, DHS sampled eight wells in San Mateo County. Total chromium was 11-28 µg/L and chromium-6 7.4-28 µg/L (64-100% of total chromium).
- In 2000, DHS sampled three wells in Yolo County. Total chromium was 31-54 µg/L, and chromium-6, 24-35 µg/L (44-97% of total chromium).
- In 2000, DHS took four samples from a surface water source in Solano County. Total chromium was 0.5-8.9 µg/L, and chromium-6, was non-detect.[22]

Given the results of these studies, the precedent set by the CAO to set replacement water standards at the PHG would create significant statewide fear and uncertainty of the safety of drinking water, and cast doubt on State-promulgated MCLs for all constituents. It is for this reason (among others) that Petitioner requests action by the State Board in the form of an emergency and immediate stay so that statewide certainty

[22] "Chromium-6 in Drinking Water Sources: Sampling Results," California Department of Public Health, last updated July 27, 2011, *available at* http://www.cdph.ca.gov/certlic/drinkingwater/pages/chromium6sampling.aspx.

about drinking water standards set by CDPH and the Regional Boards' inability to deviate from legally relevant MCLs in replacement water 13304(f) orders can be confirmed.

(2) **PG&E will be Subject to Liability because an Order requiring the Setting of Background Levels at each and every one of the Estimated 320 Potentially Affected Wells is Inconsistent with State Law and Compliance is Infeasible.**

Paragraph No. 3(a) of the Order requires PG&E to "perform an initial and quarterly evaluation of every domestic or community well in the "Affected Area" to determine if detectable levels of hexavalent chromium between the maximum background level and the PHG represent background conditions." Additionally, Finding No. 26 of the CAO states:

> [T]o determine whether hexavalent chromium levels in domestic or community wells are due to naturally occurring background or PG&E's discharge, PG&E must evaluate the hexavalent chromium values in each domestic well in the affected area . . . separately considering a number of factors.

Paragraph No. 3(a) is inconsistent with the regulatory framework established for determining background constituent concentrations. 23 CCR section 2550.7(e) states that when a background study is performed that produces a 95% upper tolerance limit (UTL) (as PG&E has completed at this site), monitoring data are to be compared to the UTL, rather than some other parameter for background. Specifically, 23 CCR section 2550.7(e)(8)(C) states that, "[T]he value for each constituent of concern or monitoring parameter at each monitoring point is compared to the upper tolerance or prediction limit."

The *2007 Background Chromium Study* ("Background Study") determined upper tolerance concentrations of hexavalent chromium and total chromium that could be

present in groundwater throughout the Hinkley Valley Groundwater Basin.[23] The 95 percent UTL concentrations identified in the Background Study provided an upper limit of naturally-occurring chromium that could be detected in any given well sampled throughout the Hinkley Valley.[24] These values are 3.1µg/L for hexavalent chromium and 3.2 µg/L for total chromium.[25] These values were adopted by the Regional Board in Order No. R6V-2008-0002A1.[26]

The CAO's requirement to perform a background assessment on each individual well is not supported by the recognized standard of care for remediation of groundwater contamination nor is it supported by state and federal regulatory guidance.[27] Anita Broughton, a recognized environmental remediation consultant with over 29 years of experience, has been unable to identify any sites where the regulatory entity required individual well background determinations.[28] Additionally, a recent search of available information reveals that no facilities in California require the assessment of individual wells on a site for the determination of multiple background concentrations of a particular contaminant.[29] In fact, Ms. Broughton notes that state and federal regulatory guidance mandate comparing well results to published standards or site-specific maximum background concentrations.[30]

We have been unable to identify any regulatory or statutory basis to require PG&E to determine individual background concentration levels at each and every well, and the CAO disregards the regulatory framework established by State law in addition to generally-accepted remediation practices. Given that there is not an existing precedent to

---

[23] *See* PG&E Background Study Report, Hinkley Compressor Station, Hinkley, CA (2007) ("Background Study"). The Lahontan Regional Board recently asked for a peer review of PG&E's Background Study and received comments from three peer reviewers. The comments provide numerous suggestions on changes or improvements that could be made to the Background Study. PG&E looks forward to working with the Lahontan Regional Board in response to the peer review comments.

[24] *See* "Frequently asked Questions about PG&E's Background Chromium Study in Hinkley," *available at http://www.swrcb.ca.gov/rwqcb6/water_issues/projects/pge/docs/pge_background_study_faq2.pdf.*

[25] Id.

[26] CAO No. R6V-2011-005 at Finding No. 4.

[27] Declaration of Anita Broughton ("Broughton Decl."), dated October 17, 2011, at ¶ 3(a)-(b).

[28] Id. at ¶ 5.

[29] Id. at ¶ 4.

[30] Id. at ¶ 5.

follow, the short timeframe for proposing a method to perform an initial and quarterly evaluation of every single domestic or community well in the "Affected Area" is not feasible. As a result the Petitioner will be subject to liability if a stay is not immediately granted.[31]

Interested Persons and the Public Interest will not be Substantially Harmed if a Stay is Granted

Interested persons and the public interest will not be placed at risk if a stay is granted because existing hexavalent chromium levels in domestic wells all are currently below 6.4µg/L--far below the State-promulgated safe drinking water standard of 50 µg/L.[32] In addition, PG&E has in place a program in which it voluntarily provides bottled water to all well owners within one mile of the plume.[33] This voluntary bottled water program was enacted prior to and satisfied the Lahontan Regional Board's original replacement water order (CAO No. R6V-2011-0005), which required bottled water for residents with drinking water containing hexavalent chromium levels above natural background conditions.[34]

The current CAO acknowledges that "consumptive" use is the only beneficial use at issue with hexavalent chromium in domestic wells.[35] OEHHA confirmed that over 99 percent of the risk from hexavalent chromium in domestic wells in California is due to ingestion.[36] Bottled water eliminates any risk from ingestion. Therefore, PG&E's community bottled water program eliminates the risk of harm to individual persons or the public while a stay is pending.

---

[31] *See* Wilson Decl. at ¶¶ 7-9.

[32] Approximately ten (10) domestic wells currently in use are above natural background and the remaining wells are all below natural background levels. Information on wells can be obtained from Quarterly Groundwater Monitoring Reports and Domestic Well Sampling Results submitted to the Lahontan Regional Board.

[33] PG&E is currently providing water to all well owners within one-half mile of the plume boundary and began voluntarily working to extend that offer to all well owners within one mile of the plume boundary before the CAO was issued.

[34] CAO No. R6V-2011-0005 at Order 1.

[35] CAO No. R6V-2011-0005A1 at Finding No. 32.

[36] Memorandum by OEHHA to Harold Singer, regarding "Proposed PHG for Hexavalent Chromium," dated August 17, 2011, and attached as Attachment V.

Because no domestic wells in Hinkley contain hexavalent chromium levels that approach the State-established MCL for drinking water, and because all persons with wells exceeding the natural background levels presently receive bottled drinking water, there is no concern that the public will be placed at risk if this *Emergency Stay* is granted.

SUBSTANTIAL QUESTIONS OF LAW AND FACT EXIST REGARDING THE DISPUTED ACTION

As explained in the Memorandum of Points and Authorities in Section 7 below and hereby incorporated by reference, there are substantial questions of both law and fact regarding the Lahontan Regional Board's adoption of the CAO.

FOR ALL THE FOREGOING REASONS, Petitioner respectfully requests that the State Board grant an immediate and emergency stay of the effect of Order No. R6V-2011-0005A1 until such time as final action is taken on this Petition.

****

3. **DATE THE REGIONAL BOARD ACTED OR FAILED TO ACT**

The date of the Lahontan Regional Board's action is October 11, 2011, the date the CAO was signed by the Executive Officer of the Lahontan Regional Board.

4. **STATEMENT OF REASONS THE ACTION IS INAPPROPRIATE OR IMPROPER**

The issuance of the CAO was beyond the authority of the Lahontan Regional Board, inappropriate, improper, or not supported by the record, for the following reasons:

(a) The 50 µg/L State Standard for Chromium is the Appropriate Standard for Replacement Drinking Water, Rather than the PHG, and All Wells in the "Affected Area" meet the MCL.

(i) The 50 µg/L MCL Applies to Hexavalent Chromium.

(ii) All Wells in the "Affected Area" Meet the Applicable MCL.

(b) The Lahontan Regional Board Misreads In the Matter of the Petition of Olin Corporation, and Threatens to Set a Troubling Precedent with Statewide Implications.

(c) The Board Lacks the Authority to Require PG&E to Provide Replacement Water that Far Exceeds State and Federal Water Standards.

(d) PG&E Cannot Provide Replacement Drinking Water At The 0.02 µg/L Concentration For Hexavalent Chromium Particularly In The Quantity Or Within The Time Frame Required By The CAO.

(i) It Is Not Feasible to Supply Commercially-Available Bottled Water That Must Be Below 0.02 µg/L Hexavalent Chromium.

(ii) PG&E Cannot Feasibly Supply Permanent Replacement Water with Hexavalent Chromium Concentrations Below 0.02 µg/L with Current Technology.

(e) PG&E Tested Water Samples Using the Test Method Required by the CAO and Could Not Reliably Produce Hexavalent Chromium Concentrations at The CAO Levels.

5. **THE MANNER IN WHICH PETITIONER IS AGGRIEVED**

Petitioner is aggrieved by the Lahontan Regional Board's issuance of a CAO which creates unprecedented mandates that are inconsistent with State law. Such mandates include, but are not limited to, an Order to provide both interim and permanent replacement water at the PHG level rather than the background or the MCL level; and an Order to determine the extent of impacted wells on the basis of a background assessment of each and every well in the "Affected Area" rather than by utilizing a single, established background level.

6. **PETITIONER'S REQUESTED ACTION BY THE STATE BOARD**

Petitioner respectfully requests that the State Board: (1) immediately stay the effect and enforcement of the CAO; and (2) vacate the order to provide interim and replacement water at the PHG level.

Additionally, Petitioner requests that the State Board determine whether or not it is unlawful to require PG&E to provide future replacement water at any level below the MCL and/or the background concentrations for hexavalent chromium.

Finally, Petitioner requests that the State Board consider the lawfulness of the Lahontan Regional Board's order to require PG&E to evaluate the hexavalent

chromium values in each domestic well in the "Affected Area" separately, rather than on the basis of a single background level, as required by regulations.

7. **MEMORANDUM OF POINTS AND AUTHORITIES**

(a) **The 50 µg/L State Standard for Chromium is the Appropriate Standard for Replacement Drinking Water, Rather than the PHG, and All Wells in the "Affected Area" meet the MCL.**

(i) **The 50 µg/L MCL Applies to Hexavalent Chromium.**

California State law mandates that public water systems comply with primary MCLs.[37] The California Department of Public Health ("CDPH"), the entity that sets drinking water standards in California, specifically states that "Chromium-6 is currently regulated under the 50-micrograms per liter (µg/L) MCL for total chromium."[38] In fact, the CDPH notes that, "[t]he total chromium MCL was established to address exposures to chromium-6, which is considered to be the more toxic form of chromium."[39]

In contrast to an MCL, which sets the standard for safe drinking water, a PHG is a non-enforceable, stringent, health-protective goal that CDPH later uses to then develop an enforceable regulatory standard.[40] CDPH can set the MCL above the level of the PHG if it finds that it is not economically or technically feasible to reduce the contaminant to the PHG level.[41] In almost all cases, the PHG associated with a constituent is far lower than the MCL.

The OEHHA announcement of the final PHG for hexavalent chromium, issued on July 27, 2011, reiterates these points in stating:

> An MCL is an enforceable standard. This means that when an MCL is established for a specific contaminant, the level of that contaminant in public drinking water systems must not exceed

---

[37] *See* Cal. Code Regs. tit. 22, § 64431.
[38] "Chromium Six Update," California Department of Public Health, last updated July 27, 2011, *available at* http://www.cdph.ca.gov/certlic/drinkingwater/Pages/Chromium6.aspx.
[39] Id.
[40] "Fact Sheet - Final Public Health Goal for Hexavalent Chromium," dated July, 27 2011, *available at* http://oehha.ca.gov/public_info/facts/Cr6facts072711.html.
[41] Id.

the MCL. The PHG is not an enforceable standard.[42]

An information sheet developed for hexavalent chromium was made available the same day that OEHHA issued its final PHG. To avoid confusion with respect to the PHG, OEHHA stated:

> A PHG is NOT a boundary line between a "safe" and "troubling" level of a contaminant. Drinking water can still be acceptable for public consumption if it contains contaminants at levels higher than the PHG.[43]

As illustrated, California law states that the applicable safe drinking water standard for hexavalent chromium is the MCL of 50 µg/L. The Lahontan Regional Board's order that PG&E supply replacement water for wells in excess of the PHG is contrary to these drinking water regulations. As discussed above, the CAO's reliance on the PHG sets a troubling statewide precedent by asserting that drinking water supplies across California that currently meet the 50 µg/L MCL safe drinking water standard but are many times higher than the 0.02 µg/L PHG pose an immediate health risk.

(ii) **All Wells in the "Affected Area" Meet the Applicable MCL.**

There are no drinking water wells in use in Hinkley with hexavalent chromium levels above the 50 µg/L California drinking water standard. In fact, currently in use there are approximately ten (10) properties in all of Hinkley with domestic wells that have hexavalent chromium levels above the natural background levels of 3.1 µg/L. In addition, the highest current level of hexavalent chromium at any drinking water well is 6.4 µg/L, which falls far below statewide standards for safe drinking water. As such, all wells in the "Affected Area" as defined by the CAO meet the current and applicable MCL.

---

[42] Id.

[43] "Fact Sheet - Final Public Health Goal for Hexavalent Chromium," dated July, 27 2011, *available at* http://oehha.ca.gov/public_info/facts/Cr6facts072711.html.

(b) **The Lahontan Regional Board Misreads In the Matter of the Petition of Olin Corporation aud Standard Fusee, And Threatens To Set A Troubling Precedent With Statewide Implications.**

The Lahontan Regional Board uses a prior State Board Order (Order WQ 2005-0007) to support the use of the hexavalent chromium PHG as the replacement water standard. The State Board decision was rendered in connection with *In the Matter of the Petition of Olin and Standard Fusee* ("Olin"). But Olin does not apply to the CAO now before this Board because:

- Here, there is an MCL covering hexavalent chromium (as a constituent of total chromium), but in Olin there was no MCL for the contaminant in question.

- Here, the 3.1 µg/L background concentration of hexavalent chromium is *above* the PHG, but in Olin the background concentration of the contaminant in question was *below* the PHG.

Olin's importance must be considered in light of the relevant statutes. Water Code section 13304(a) permits a Regional Board to order a discharger to provide replacement water service. Section 13304(f) specifies the characteristics of the replacement water: "Replacement water provided pursuant to subdivision (a) shall meet all applicable federal, State, and local drinking water standards, and shall have comparable quality to that pumped by the public water system or private well owner prior to the discharge of waste."

In Olin, this Board interpreted those provisions in the context of a cleanup and abatement order requiring Olin Corporation and Standard Fusee to provide replacement water service to owners of private domestic wells affected by discharges of potassium perchlorate.[44] At the time Olin commenced the replacement water service in 2002, there was no enforceable State or federal standard for potassium perchlorate, in drinking water. Nonetheless, Olin provided replacement water to owners of domestic wells in which perchlorate concentrations exceeded 4 µg/L, the then-operative notification

[44] State of California, State Water Resources Control Board, Order WQ2005-007.

level set by the Department of Health Services (DHS—now CDPH).[45] There were no known sources of natural background perchlorate, but anthropogenic sources in the area were responsible for background levels between 2 and 5 µg/L.

In April 2004, OEHHA issued a PHG of 6 µg/L for perchlorate.[46] Following the publication of the PHG, Olin sought approval from the Central Coast Regional Water Quality Control Board to raise the level of contamination requiring replacement water service to 6.0 µg/L to match the PHG, and filed a petition with this Board, objecting to the 4.0 µg/L trigger.[47] The State Board concluded, "where no federal, State or local standard yet exists, it is appropriate to use goals developed by agencies with expertise for public health determinations in deciding whether replacement water service is necessary." Thus, in the absence of any standard and with a background level below the PHG, the Regional Board embraced the perchlorate PHG as a replacement water standard.[48]

In contrast to Olin, an MCL exists for total chromium that specifically includes hexavalent chromium. As such, the unique circumstances in Olin that required a Regional Board to literally stand in the shoes of CDPH and improvise a drinking water standard are absent here. Furthermore, Olin involved a typical scenario in which the *background* level of the contaminant in question was *below* an established PHG. In the case of Hinkley, however, the background concentration of hexavalent chromium, 3.1 µg/L, is over 100 times greater than the PHG level of 0.02 µg/L. Thus, unlike Olin, enforcing the hexavalent chromium PHG as a drinking water standard would require PG&E to provide replacement water containing hexavalent chromium far below what naturally exists in groundwater.

The CAO would also go against Board's Resolution No. 92-49, which provides:

---

[45] Id. at 3. Please also note that this document sometimes references the CDPH as "DHS," (Department of Health Services), which was previously tasked with the public health function in California.
[46] Id.
[47] Id. at 4.
[48] Id. at 5-7.

> [U]nder no circumstances shall [policies and procedures for cleanup and abatement of discharges] be interpreted to require cleanup and abatement which achieves water quality conditions that are better than background conditions.[49]

This Board should consider the statewide implications of imposing a PHG as a drinking water standard. At the very least, this Board should not interpret its Olin decision to supplant an enforceable MCL.

(c) The Board Lacks the Authority to Require PG&E to Provide Replacement Water that Far Exceeds State and Federal Water Standards.

The CAO requires PG&E to provide interim and permanent replacement water that "meets primary and secondary drinking water standards and hexavalent chromium levels of less than 0.02 µg/L or the final MCL, once that standard is adopted by CDPH."[50] Therefore, until CDPH adopts a new MCL for hexavalent chromium, the CAO requires that PG&E provide temporary and permanent replacement water with hexavalent chromium at levels below 0.02 µg/L despite a current and applicable MCL of 50 µg/L. This requirement exceeds the Lahontan Regional Board's authority and is inconsistent with applicable law.[51]

According to California law, a PHG is not a drinking water standard: "A PHG is not a regulatory standard. It is only one step in the process of developing an enforceable standard that is set by the California Department of Public Health for drinking water that public water systems must meet.[52] Thus, pursuant to Water Code Section 13304(f), the Lahontan Regional Board cannot require replacement water to meet PHGs.

---

[49] State Water Resources Control Board, Resolution No. 92-49 (as amended on April 21, 1994 and October 2, 1996), at ¶ III(F).

[50] CAO No. R6V-2011-0005A1 at Order No.(s) 1.b, 2.c.

[51] Likewise, the State Board's own enforcement policy is to only require a discharger to abate contamination to background levels. *See* SWRCB Water Quality Enforcement Policy (2010) at 35, *available at* http://www.swrcb.ca.gov/water_issues/programs/enforcement/.

[52] "Fact Sheet - Final Public Health Goal for Hexavalent Chromium," dated July, 27 2011, *available at* http://oehha.ca.gov/public_info/facts/Cr6facts072711.html.

The CAO requires PG&E to provide replacement water with a hexavalent chromium concentration *2,500 times* less than the lowest applicable drinking water standard (and *5,000 times* less than the federal drinking water standard[53]). This is inconsistent with the plain language of Water Code 13304(f) and is improper.

In addition, Water Code section 13304(f) expressly prohibits Regional Boards from requiring that replacement water be more pristine than naturally occurring background water.[54]

Here, the CAO establishes the natural background level for hexavalent chromium as 3.1µg/L. It states: "[h]exavalent and total chromium occur naturally in groundwater at variable concentrations . . . . [t]he 95% upper tolerance limits are 3.09 µg/L for hexavalent chromium and 3.24 µg/L for total chromium."[55] The Lahontan Regional Board has relied on and incorporated these background levels in multiple orders since 2007, including the current CAO.[56] Under Water Code 13304(f) the Lahontan Regional Board cannot require PG&E to provide replacement water at hexavalent chromium levels below the natural background level of 3.1 µg/L.

(d) PG&E Cannot Provide Replacement Drinking Water At The 0.02 µg/L Concentration For Hexavalent Chromium Particularly In The Quantity Or Within The Time Frame Required By The CAO.

(i) It Is Not Feasible to Supply Commercially-Available Bottled Water That Must Be Below 0.02 µg/L Hexavalent Chromium.

Paragraph 1(b) of the CAO requires that interim replacement water meet "State primary and secondary drinking water standards and hexavalent chromium levels

---

[53] United States Environmental Protection Agency, *Basic Information about Chromium in Drinking Water, available at* http://water.epa.gov/drink/contaminants/basicinformation/chromium.cfm (Stating "Ensuring safe drinking water for all Americans is a top priority for EPA. EPA has a drinking water standard of 0.1 mg/L (100 µg/L) for total chromium, which includes hexavalent chromium, and we require water systems to test for chromium. This standard is based on the best available science and is enforceable by law.")

[54] *See* Cal. Water Code § 13304(f) ("Replacement water provided . . . shall meet all applicable federal, state, and local drinking water standards, and shall have comparable quality to that pumped by the public water system or private well owner prior to the discharge of waste.")

[55] CAO No. R6V-2011-0005A1 at Finding No. 6.

[56] Id. at Finding No.(s) 6, 7, 8, 11, 13, 25, 29, 30; Order No. 3.

of less than 0.02 µg/L or the final MCL, once that standard is adopted by CDPH."[57] The CAO notes that replacement water must be lower than an actual reporting limit of 0.06 µg/L.[58] This detection limit is sixteen (16) times lower than the current 1 µg/L hexavalent chromium detection limit used by the State.[59] Compliance with the CAO's replacement water quality requirement is infeasible. As an initial matter, PG&E has been unable to locate any commercial source of bottled water that will guarantee it will meet even the CAO's 0.06 µg/L standard.[60] The lack of supply has also been documented by OEHHA.[61]

Compliance with the CAO is also infeasible because there is no reasonable way to monitor whether bottled water regularly meets the 0.06 µg/L standard, as the bottled water industry does not report hexavalent chromium concentrations.[62] Therefore, a mechanism for identifying a reliable source that meets the criteria in the volumes that would be required does not, as yet, exist. As a result, it is not feasible to meet the terms of the CAO as written, particularly in the timeframe required.

Bulk water delivery – that is, water trucked to water tanks at each location – is not a feasible interim replacement water alternative because it would require at least six months to implement.[63] In the face of this fact, the five (5) to forty-five (45) day deadlines are simply impracticable. Bulk water also raises challenges related to locating a supply with sufficiently low hexavalent chromium concentrations, and concerns about the chemical treatment of the trucks necessary to maintain microbiological quality.[64]

///

///

---

[57] CAO No. R6V-2011-0005A1 at Order No.(s) 1.b., 2.c. For both interim and permanent replacement water, a higher standard is possible if PG&E can prove what the historic levels of chromium six were at a well specific level.
[58] CAO No. R6V-2011-0005A1 at Order No. 1.b., fn. 1.
[59] *See* "Chromium-6 in Drinking Water: MCL Update," California Department of Public Health, last updated July 27, 2011, *available at* http://www.cdph.ca.gov/certlic/drinkingwater/Pages/Chromium6.aspx.
[60] See discussion regarding Culligan in *Request for Immediate and Emergency Stay* above.
[61] A peer reviewer of OEHHA's draft PHG for hexavalent chromium, Professor William Shotyk of Heidelberg University's Institute of Earth Sciences, cited a study that analyzed 132 brands of bottled water. The study found as much as 172 ppb (µg/L) of total chromium, with a median of 8.2 ppb (µg/L) and virtually all total chromium dissolved in water is hexavalent chromium.
[62] DeWolfe Decl. at ¶ 8(b).
[63] *See* Id. at ¶ 15(b).
[64] *See* Id. at ¶ 9(a)-(b).

(ii) PG&E Cannot Feasibly Supply Permanent Replacement Water with Hexavalent Chromium Concentrations Below 0.02 µg/L with Current Technology.

The CAO permanent replacement water quality requirements are the same as those for interim replacement water – permanent replacement water must meet primary and secondary drinking water standards and have less than 0.02 µg/L hexavalent chromium.[65] It is impossible to determine how many wells would qualify for permanent replacement water because the CAO does not define the methodology to determine if a well has been impacted.[66] Assuming that up to three-hundred (300) wells in the "Affected Area" would meet the threshold, PG&E would be required to supply approximately 297,000 gallons per month to serve these wells.[67] This is not feasible.

Advanced technologies would be required to reduce hexavalent chromium below natural background levels in either the local groundwater supply or in other sources. It is not currently known whether PG&E can obtain the necessary permits from CDPH.[68] In addition, each technology analyzed by PG&E raises significant technological and environmental issues that must be explored to determine if they could feasibly meet the CAO requirements.[69] An expert in whole house replacement water, James DeWolfe, explained (in a declaration attached to PG&E's comments on the earlier draft version of the CAO, and attached as part of Attachment II) that no current technology is available to immediately supply whole house replacement water that meets the terms of the CAO.[70]

---

[65] CAO No. R6V-2011-0005A1 at Order No. 2.c.

[66] Id. at Finding No. 26.

[67] Estimates are based on the following two statements from DeWolfe's Declaration: (1) "because the declared background Cr6 [hexavalent chromium] concentrations in the Hinkley area average 1.2 ppb [µg/L], and have a declared maximum value of 3.1 ppb [µg/L], we assumed that all wells in the 'affected area' . . . will have Cr6 [hexavalent chromium] levels above the Public Health Goal of 0.02 ppb [µg/L];" and (2) "the average household was expected to consume 33 gallons per day for drinking and cooking based on estimates for daily per capita faucet use." DeWolfe Decl. at ¶ 3(a)-(c).

[68] *See* Id. at ¶ 11(e).

[69] *See* Id. at ¶¶ 11-13.

[70] DeWolfe Decl. at ¶ 12. An ion exchange system would require additional engineering advances because those currently on the market cannot achieve hexavalent chromium concentrations of 0.02 µg/L. Furthermore, the waste stream created by an ion exchange system, including brines, would likely be considered hazardous waste under federal law, and would require at a minimum an appropriate disposal plan and would not be suitable for disposal in property septic systems. A multi-pass reverse osmosis system faces similar constraints. Up to 75% of the water

PG&E is in the process of procuring equipment to test various filters at the site to see how low they can remove hexavalent chromium. However, the manufacturers of these filters aim to remove chromium at the current state or federal MCL levels (50 or 100 µg/L) and, therefore, it is not yet known whether in the future any current technology might be able to meet the 0.02 µg/L hexavalent chromium level required by the CAO.

(e) PG&E Tested Water Samples Using the Test Method Required by the CAO and Could Not Reliably Produce Hexavalent Chromium Concentrations at The CAO Levels.

PG&E retained Shawn Duffy, an expert in chromium laboratory testing, to determine whether laboratories can reliably test for the ultra-trace hexavalent chromium concentrations required by the CAO using the test method required by the CAO. Mr. Duffy's study concluded that using the test method required by the CAO he could not reliably detect and quantify for hexavalent chromium concentrations at the levels required by the CAO.[71] Requiring a compliance level that is below the level at which laboratories can reliably detect or quantify would set a troubling precedent for cleanups statewide.

Mr. Duffy directed the collection of ten water samples from wells in Hinkley that had previously been reported as non-detect for hexavalent chromium at the Hinkley project detection limit of 0.2 µg/L. Mr. Duffy also directed the manufacture of performance evaluation samples containing known concentrations of total chromium and

---

introduced into such a system will end up as toxic brine, while as little as 25% will be usable as potable water. Like an ion exchange system, reverse osmosis systems on the market are not designed to achieve the hexavalent chromium PHG and would, therefore, require additional engineering advances. Nor is a reduction, clarification and filtration system a viable option; the technology is simply too new and untested. The construction of a central treatment and distribution system would also be infeasible. The design, environmental review, permitting and construction of such a facility would take a significant amount of time. Furthermore, a central treatment system would not itself achieve the 0.02 µg/L standard. The system would have to employ ion exchange and/or reverse osmosis to meet the hexavalent chromium concentrations required by the CAO, thereby raising all the issues described above that plague those technologies. For the same reasons bottled water or bulk water would not be feasible for an interim replacement water supply, a central treatment facility would also not be feasible as a permanent replacement water supply because this treatment would also not achieve the 0.02 µg/L standard. *See generally*, DeWolfe Decl.

[71] *See* Declaration of Shawn Duffy ("Duffy Decl.") at ¶ 3, dated July 9, 2011, and attached as part of the Comment Letter in Attachment II.

hexavalent chromium of 0.01, 0.02, 0.04, 0.06, 0.08 and 0.1 µg/L, as well as the generation of two double-blind field blanks.[72]

The study design samples were split into three identical sample containers and submitted to three California accredited laboratories: Truesdail Laboratory Inc. ("TLI"), Advanced Technology Laboratory ("ATL"), and BCLab Inc. ("BCL").[73] The laboratories were requested to report the hexavalent chromium results to a level of 0.02 µg/L using the test method specified in the CAO (a modified EPA method 218.6-- update 144) and as low as possible for total chromium by EPA Method 200.8.[74]

The test results demonstrate that at the reporting levels required by the CAO, the laboratories frequently failed to accurately detect and/or quantify hexavalent chromium using updated EPA method 218.6.[75] Specifically, the laboratories routinely failed to meet the standard laboratory fortified blank criteria (90% to 110% recovery) for the performance evaluation samples.[76] The laboratories repeatedly produced sampling results that were less than 90% or more than 110% of the known sampling result.[77] Notably, the sample containing the reporting limit of 0.06 µg/L, known as the Limit of Quantitation ("LOQ"), was reported as 0.068 µg/L, 0.071 µg/L, and ND < 0.026 µg/L between the three laboratories.[78] These results are 113%, 118%, and 0% of the known concentration.[79] In other words, not one laboratory met the standard laboratory fortified blank criteria for the very LOQ concentration that the CAO would require PG&E to use.

In addition, the laboratories frequently failed to report similar results for the same Hinkley test samples. For example, for sample H-13-Q2 (a monitoring well) the laboratories reported: 0.091 µg/L (BCL), 0.050 µg/L (TLI), and ND <0.02 µg/L (ATL).[80] Using these varying results for the exact same sample, BCL would require PG&E to

---

[72] Duffy Decl. at ¶ 4(d).
[73] Id. at ¶ 4(g).
[74] Id.
[75] *See* Id. at ¶ 6.
[76] Id. at ¶ 6(a).
[77] Id.
[78] Id. at ¶ 5.
[79] Id. at ¶ 6(b).
[80] Id. at ¶ 5.

supply whole-house replacement water, while TLI and ATL would not. These findings are critically significant as they demonstrate that these three California-accredited laboratories often failed to produce results that were close to one another. The failure to produce at least similar sample results further demonstrates that the laboratories were not able to reliably quantify hexavalent chromium at the ultra-low concentrations required by the CAO.[81]

Finally, Mr. Duffy requested that laboratories analyze the same samples for total chromium, since chemists often use total chromium sample results as a double-check for hexavalent chromium sample results. Unfortunately, the total chromium results frequently failed to correlate, and often failed to approximate the hexavalent chromium results for the same samples. For example, the total chromium levels ranged from non-detect to as high as 0.354 µg/L for the same sample, and rarely correlated or came close to the hexavalent chromium results for the same sample. In addition, none of the laboratories met the standard laboratory fortified blank criteria for total chromium.[82] Thus, the total chromium testing provides further support for Mr. Duffy's conclusion that these laboratories could not reliably detect or quantify hexavalent chromium at the ultra-trace levels required by the CAO.[83] Therefore, based on these results, PG&E could not reliably meet the replacement water hexavalent chromium levels required by the CAO.

8. A COPY OF THIS PETITION HAS BEEN SENT TO THE LAHONTAN REGIONAL BOARD

In accordance with title 23, section 2050(a)(8) of the CCR, the Petitioner mailed a true and correct copy of this petition by First Class mail on October 25, 2011, to the Lahontan Regional Board at the following address:

///

///

[81] *See* Duffy Decl. at ¶ 6.
[82] *See* Id. at ¶¶ 7(d), 8.
[83] Id. at 6.

Harold Singer, Executive Officer
Regional Water Quality Control Board – Lahontan Region
2501 Lake Tahoe Boulevard
South Lake Tahoe, CA 96150-7704

9. ISSUES RAISED IN THE PETITION WERE PRESENTED TO THE LAHONTAN REGIONAL BOARD BEFORE IT ACTED

Petitioner raised many of the issues discussed within this Petition with the Lahontan Regional Board in a comment letter addressed to Harold Singer on July 12, 2011 in response to the Draft Amended CAO No. R6V-2011-0005A1 -- and also in subsequent correspondence.[84]

Although Petitioner raised most of the issues identified within this Petition to the Lahontan Regional Board, to the extent that this Petition includes any arguments not previously raised, Petitioner respectfully requests that the State Board consider that the Petitioner was forced to respond to new issues that were first made public when the Executive Officer signed the final CAO. In fact, the final CAO contains new issues on which Petitioner was not afforded the opportunity to comment prior to its issuance. As described herein, the CAO raises substantial issues of law and fact and Petitioner requests that the State Board review this *Petition* and grant this *Request for Immediate and Emergency Stay*.

[84] Assistant Executive Officer Lauri Kemper submitted a redlined version of the draft CAO to all parties on September 27, 2011. In response, Executive Officer Harold Singer allowed Petitioner to submit a redline version for his consideration on October 5, 2011. Additionally, subsequent to the issuance of the CAO, Executive Officer Harold Singer responded to electronic mail from Petitioner regarding clarifications of the CAO.

DATED: October 25, 2011

TRACY J. EGOSCUE
PAUL HASTINGS, LLP

By: _____
TRACY J. EGOSCUE

Attorney for Petitioner
Pacific Gas & Electric Company

Attachment I:
CAO No. R6V-2011-0005A1

(*Request for Immediate and Emergency Stay; Petition for Review; and Memorandum of Points and Authorities in Support Thereof*)

CALIFORNIA REGIONAL WATER QUALITY CONTROL BOARD
LAHONTAN REGION

AMENDED CLEANUP AND ABATEMENT ORDER NO. R6V-2011-0005A1
WDID NO. 6B369107001
REQUIRING PACIFIC GAS AND ELECTRIC COMPANY TO CLEAN UP AND ABATE WASTE DISCHARGES OF TOTAL AND HEXAVALENT CHROMIUM TO THE GROUNDWATERS OF THE MOJAVE HYDROLOGIC UNIT

San Bernardino County

The Califomia Regional Water Quality Control Board, Lahontan Region (Water Board), finds:

Discharger

1. The Pacific Gas and Electric Company (PG&E) owns and operates the Hinkley Compressor Station (hereafter the "Facility") located southeast of the community of Hinkley in San Bernardino County. For the purposes of this Order, PG&E is referred to as the "Discharger."

Site History and Hydrogeology

2. The Facility is located at 35863 Fairview Road (APN 048S-112-52), one-half mile east of the community of Hinkley in San Bemardino County, in the Harper Valley Subarea of the Mojave Hydrologic Unit. The Facility began operating in 1952 and discharged untreated cooling tower water containing hexavalent chromium to unlined ponds until 1964. Wastewater then percolated through soil to the water table, approximately 80 feet below, creating a chromium plume. In general, the chromium plume extends north from the compressor station to at least Sonoma Road and from east of Summerset Road to west of Mountain View Road. This release of hexavalent chromium is the only known source of anthropogenic or human introduced chromium in the localized area.

3. The hydrogeology in the southern 75 percent and in the northeastern portion of the project area consists of an upper, unconfined aquifer and a lower, confined aquifer separated by a lacustrine clay that forms a regional aquitard. The hydrogeology in the northwestern portion of the project area consists of just the upper, unconfined aquifer, as the lower aquifer and clay aquitard pinch out (terminate against the upward sloping bedrock). In general, groundwater flow is primarily to the north-northwest towards the Harper Dry Lake, with an average gradient of 0.004 feet per foot. The Mojave River contributes more than 80 percent of the natural groundwater recharge to the Hinkley Valley.

4. The soils underlying the Facility are comprised of interbedded sands, gravels, silts, and clays. The depth to bedrock ranges from about 300 feet below ground surface in the southern project area to cropping out (bedrock comes to the ground surface) in the northern portion of the project area. The closest surface water is an unnamed

ephemeral stream, located about 4,000 feet northwest of the plume's northern boundary. In addition, the Mojave River is located less than one mile to the southeast of the Facility.

Chromium Plume

5. The groundwater in the upper aquifer below the Facility contains hexavalent chromium that was discharged from the PG&E compressor station and naturally occurring constituents. The plume is considered to be that portion of the aquifer affected by the discharge. Chromium concentrations in groundwater are highest at the compressor station and become less concentrated towards the north. According to the *Second Quarter 2011 Groundwater Monitoring Report*, the highest level of hexavalent chromium detected in groundwater was 7,800 micrograms per liter (µg/L) at monitoring well SA-MW-05D. A hazardous waste is defined as any waste that contains hexavalent chromium at concentrations that exceed 5,000 µg/L. The plume contains total chromium greater than the state Maximum Contaminant Levels (MCL), or drinking water standard of 50 µg/L in the area from the Facility to Santa Fe Avenue, almost two miles north. Concentrations of hexavalent chromium are present above background levels for at least the next mile north. The chromium plume resides primarily in floodplain sediments originating from the Mojave River and alluvial sediments eroded from local mountains.

6. Hexavalent and total chromium occur naturally in groundwater at variable concentrations, according to the February 27, 2007, document, *Groundwater Background Chromium Study Report, Hinkley Compressor Station*. The mean (or average) background concentrations detected in groundwater are 1.19 µg/L for hexavalent chromium and 1.52 µg/L for total chromium. The work plan for the Study recommended that maximum background concentrations should be expressed as the 95% upper tolerance limits. The 95% upper tolerance limit is the value that is estimated to include 95 percent of the possible detections of natural occurring chromium with a 95 percent confidence level. The 95% upper tolerance limits are 3.09 µg/L for hexavalent chromium and 3.23 µg/L for total chromium.

7. On July 28, 2010, Water Board staff received information from PG&E that hexavalent and total chromium concentrations exceeded 3.1 µg/L at three residential wells and four shallow monitoring wells along Summerset Road, and to the east of Summerset Road, north of Santa Fe Avenue. Three of these wells contained hexavalent chromium ranging from 4 µg/L to 5.5 µg/L.

8. Testing results from the Second Quarter 2011 provided an approximate concentration contour, or outline of hexavalent chromium levels above 3.1 µg/L and total chromium above 3.2 µg/L based on chromium results from the upper aquifer groundwater monitoring wells and short-screen extraction wells. These data indicate that the chromium plume had migrated to locations where the hexavalent chromium levels had previously been detected at levels below 3.1 µg/L.

Regulatory History

9. On August 6, 2008, the Water Board issued Cleanup and Abatement Order (CAO) No. R6V-2008-0002 to the Discharger to clean up and abate the effects of waste discharges and threatened discharges containing hexavalent chromium and total chromium to waters of the State. The CAO, in part, required the Discharger to prevent the chromium plume from migrating to locations where hexavalent chromium is below the background levels.

10. At the November 12-13, 2008 Water Board meeting, the Water Board considered the 2007 *Background Chromium Study*, along with comments and recommendations by interested persons and staff.

11. Following the meeting, the Water Board Executive Officer issued Amended CAO No. R6V-2008-0002A1 (2008 Amended CAO) to establish background concentrations for chromium in Hinkley Valley groundwater as follows:

Maximum background hexavalent chromium = 3.1 µg/L
Maximum background total chromium = 3.2 µg/L
Average background hexavalent chromium = 1.2 µg/L
Average background total chromium = 1.5 µg/L

12. The 2007 *Background Chromium Study* results described in Finding No. 6 have not been subject to an independent third-party review to comment on its accuracy. The study is currently undergoing peer-review through Cal/EPA's scientific peer review program. These background concentrations were set for the purposes of evaluating and eventually setting clean up requirements.

13. On January 7, 2011, the Water Board issued Cleanup and Abatement Order R6V-2011-0005 to PG&E in response to detections of hexavalent chromium above background levels in Hinkley domestic wells. This order required that PG&E provide interim uninterrupted replacement water, such as bottled water, to residences and businesses whose private or community wells were found to contain hexavalent chromium at concentrations exceeding 3.1 µg/L, or total chromium had been detected at 3.2 µg/L. This decision was based on 1) the 2010 testing results that showed concentrations of hexavalent chromium exceeded background levels, and 2) the background levels of chromium memorialized in the 2008 Amended Cleanup and Abatement Order (R6V-20008-0002A1).

Regulation of Hexavalent Chromium

14. On July 27, 2011, the California Office of Environmental Health Hazard Assessment (OEHHA) established a Public Health Goal (PHG) for hexavalent chromium at 0.02 µg/L. This is the first PHG specific to hexavalent chromium. PHGs are based on a risk assessment that identifies a level of exposure at which no known or anticipated adverse effects on health will occur, with an adequate margin of safety (Cal. Health & Safety Code §116365). The PHG is used by the California Department of Public Health (CDPH) to develop the MCL (California Health & Safety Code §116365(a)).

15. Currently, the MCL for total chromium in drinking water is 50 µg/L , which includes all forms of chromium. This MCL was established in 1977. There is no MCL specific to hexavalent chromium.

Authority – Legal Requirements

16. California Water Code section 13304, subdivision (a) states in part:

> *Any person . . . who has caused or permitted, causes or permits, or threatens to cause or permit any waste to be discharged or deposited where it is, or probably will be, discharged to waters of the state and creates, or threatens to create, a condition of pollution or nuisance, shall upon order of the regional board clean up or abate the effects of the waste...*
>
> *...in the case of threatened pollution or nuisance, take other necessary remedial action, including but not limited to, overseeing cleanup and abatement efforts. A cleanup and abatement order issued by the state board or a regional board may require the provision of, or payment for, uninterrupted replacement water service, which may include wellhead treatment, to each owner.*

17. Pursuant to Water Code section 13304, subdivision (f):

> *Replacement water provided pursuant to subdivision (a) shall meet all applicable federal, state, and local drinking water standards, and shall have comparable quality to that pumped by the public water system or private well owner prior to the discharge of waste.*

18. Water Code section 13307.6, subdivisions (a) (4) and (7) state in part:

> *(a) In addition to the requirements of Section 13307.5, the regional board may develop and use any of the following procedures ...if the regional board determines there is expressed community interest in the site...*
>
> *(4) Formation and facilitation of an advisory group.*
>
> *(7) Preparation of a public participation plan.*

19. Water Code section 13267, subdivision (b) states in part:

> *In conducting an investigation [of the quality of any waters of the state within its region] the regional board may require any person who has discharged waste within its region...[to] furnish, under penalty of perjury, technical or monitoring program reports which the regional board requires.*

This Order requires the submittal of workplans, monitoring data, and reports, mainly to document that the replacement water service meets all regulatory requirements. Workplans and technical reports have been required by previous Water Board Orders and are necessary to develop an accurate assessment of the plume of anthropogenic hexavalent chromium in the Hinkley upper aquifer.

20. Section 13304 of the Water Code allows a regional board to hold persons accountable who "cause or permit" any waste discharged in a water of the State. The burden to remediate the impacts of waste falls on the party who is responsible for the discharge, even if their actions alone are not the only source of pollution (*City of Modesto Redevelopment Agency v. Superior Court*, 19 Cal.App.4th 28 (2004)). Likewise, in cases of hazardous waste discharges, the burden to remediate impacts of waste falls on the discharger even if they are not the sole cause of the costs (Browning-Ferris Industries of Illinois, Inc. v. Ter Maat, 195 F.3d 953, 49 Env't. Rep. Cas. (BNA) 1449, 30 Envtl. L. Rep. 20135 (7th Cir. 1999). The Discharger is currently the only known source of anthropogenic chromium in the Hinkley upper aquifer. It is the Discharger's responsibility to remediate the affects of its discharge or to demonstrate that it is not responsible for the contamination or only a legally divisible portion of the contamination.

Replacement Water Service

21. The State Water Resources Control Board (State Water Board) issued precedential Order WQ 2005-0007, In the Matter of the Petition of Olin Corporation and Standard Fusee, Incorporated (referred to as the "Olin Order"). The Olin Order was issued in response to a petition brought by the Olin Corporation and Standard Fusee to provide replacement water service to owners of private domestic wells affected by the discharge of potassium perchlorate from a facility. Because there was no enforceable state or federal standard for perchlorate in drinking water for use in determining when a well is affected such that the user should be entitled to replacement water, the regional board had relied on the notification level for perchlorate of 4 µg/L. After the issuance of a final public health goal issued by the OEHHA of 6 µg/L several years later, The Olin Corporation sought approval to raise the level of contamination requiring replacement water service to 6 µg/L to match the PHG, and the regional board denied the request. The State Water Board determined that *"where no federal, state or local standard yet exists, it is appropriate to use goals developed by agencies with expertise for public health determinations in deciding whether replacement water service is necessary,"* and concluded that the regional board should defer to OEHHA and DHS (now CDPH) in determining the appropriate level of contamination requiring replacement drinking water service. (Olin Order at p. 6-7.) The State Water Board recognized that although the PHG is not a legally enforceable standard, it is appropriate to use the public health goal as the applicable level for determining wells requiring replacement drinking water. (Olin Order at p. 8).

22. The situation facing the Water Board is analogous to that described in the Olin Order in that a drinking water standard specific to hexavalent chromium does not exist but an established PHG exists. Therefore, consistent with the State Water Board's

direction in the Olin Order, it is appropriate for the Water Board to rely on the PHG of 0.02 µg/L for hexavalent chromium as the appropriate level for determining wells requiring replacement water service. This is also consistent with a comment from the CDPH advising the Water Board not to rely on a *draft* PHG. This comment was received during the comment period on this draft Order at which time the OEHHA had not yet established the final PHG for hexavalent chromium. However, since the end of the comment period, the OEHHA has established a final PHG for hexavalent chromium. Once CDPH establishes an MCL for hexavalent chromium, the Water Board may amend this Order to use the MCL as the appropriate level for determining wells requiring replacement water service.

23. In setting the PHG, OEHHA evaluated health risks from hexavalent chromium in domestic water based on a variety of typical household uses of tap water, including drinking, preparing foods and beverages, bathing or showering, flushing toilets, and other household uses resulting in potential dermal and inhalation exposures. Toxicity studies from routes of exposure were categorized according to ingestion, inhalation and dermal contact. Inhalation risks were determined based on studies of the impacts of inhaling hexavalent chromium-contaminated water vaporized in the shower ("shower studies") and were found to be very low.

Many homes in the Hinkley area rely on swamp coolers to provide cooling. These swamp coolers typically use domestic water. The exposure risk associated with the use of water containing hexavalent chromium in swamp coolers was not evaluated as part of the development of the PHG for hexavalent chromium. As such, the Water Board needed independent input on this concern. In a memorandum dated August 17, 2011, the OEHHA advised the Water Board that swamp coolers do not pose any additional exposure risk due to the fact that chromium in water is not converted to the vapor phase in these units.

24. As defined in the Olin Order, wells are "affected" by a discharge of waste when they do not meet federal, state, or local drinking water standards; or where no standards exist, when the discharge does not meet goals developed by agencies with expertise for public health determinations. However, where the naturally occurring background levels of the constituent may exceed the PHG, the Water Board must also consider naturally occurring background levels when considering whether a well is affected. The Water Board can only require replacement water service if the presence and level of the constituent is due to the discharge of waste.

25. The Water Board has established maximum and average background levels of total and hexavalent chromium for the Hinkley area (see Finding Nos. 6, 10, 11 and 12). These levels were established to provide a basis for evaluating cleanup alternatives and were set at levels which had a high probability that any values in excess of these levels were likely caused by the discharge (see Finding No. 6). This criterion, while instructive, is not necessarily appropriate for establishing levels above which replacement water service should be provided. Because these background levels are 50 to 150 times greater than the PHG for hexavalent chromium, it is more appropriate to provide criteria for determining when replacement water service is necessary that is more conservative and protective of public health. Because the 3.1

µg/L hexavalent chromium and 3.2 µg/L total chromium values represent ***maximum*** background levels, hexavalent chromium levels in domestic wells that are below the maximum background levels may have been caused by PG&E's discharge. It is therefore necessary to establish a process to evaluate and determine if hexavalent chromium levels in domestic wells above the PHG, but below the established maximum background level are due to the discharge.

26. Background levels of hexavalent chromium in the Hinkley are variable given the geochemical processes that contribute to the formation of hexavalent chromium in groundwater. Additionally, hexavalent chromium concentrations that are considered background levels in any one well may vary over time. Therefore, because it will be necessary to evaluate each well separately, it is not practicable in this Order to set the hexavalent chromium background values for each domestic well that has been or could be affected by the plume. Rather, to determine whether hexavalent chromium levels in domestic or community wells are due to naturally occurring background or PG&E's discharge, PG&E must evaluate the hexavalent chromium values in each domestic well in the affected area (see Finding No. 30) separately, considering a number of factors, including, but not limited to: changes in hexavalent chromium levels over time, location of well in relationship to the plume and groundwater flow direction, isotopic analysis of hexavalent chromium, and statistical analysis described in Title 27, section 20415(e)(8).

27. The release from the Discharger's facility is the only known source of anthropogenic chromium in the groundwater of the upper Hinkley aquifer. All anthropogenic chromium in this area is considered to be the result of the Discharger's activities.

28. The Discharger is required to abate the effects of its discharge in accordance with Water Code 13304. This includes providing uninterrupted replacement water service to all impacted domestic or community wells. Replacement water service shall have comparable quality to the water pumped prior to the well being affected by the discharge of the waste. There are various methods to provide this replacement water service. Bottled water is not guaranteed to contain hexavalent chromium at levels needed to comply with the Water Code requirement that the replacement water service be comparable to that pumped by the well owner prior to it being affected by the discharge. Similarly, certified treatment systems are also not guaranteed to reduce hexavalent chromium to levels needed to meet the Water Code requirement cited above. Therefore, this Order requires the Discharger to demonstrate that bottled water or the water provided by treatment systems designed to provide replacement water service are of a quality comparable to that which was pumped prior to being affected by the discharge.

29. **Impacted wells** are defined as domestic or community wells in the affected area (see next finding) containing chromium in concentrations (measured at any time) that are above 3.1 µg/L hexavalent chromium or 3.2 µg/L total chromium. Additionally, impacted wells also include those domestic or community wells in the affected area containing hexavalent chromium in concentrations greater than 0.02 µg/L when the analysis performed by the Discharger, in compliance with the approved methods as specified in Paragraph 3.a. of this Order, determines that the

hexavalent chromium is more likely than not, partially or completely, due to the discharge of waste by the Discharger. The Water Board believes this should be a well-by-well comparison and does not intend for any individual hexavalent chromium values to be compared to the average background level.

30. The **affected area** is defined as all domestic wells located laterally within one mile downgradient or cross-gradient from the 3.1 µg/L hexavalent chromium or 3.2 µg/L total chromium plume boundaries based upon monitoring well data drawn in the most current quarterly site-wide groundwater monitoring report submitted by the Discharger. The affected area may change based on new data collected and evaluated each quarter.

Other Findings

31. The Water Board recognizes the significant community interest in this site. It further acknowledges the recent formation of a Community Advisory Group and the challenges that this Group and members of the community may have in evaluating the technical aspects of this site. The Hinkley community is a rural community that includes many different income levels and ethnicities. Therefore, it is important that environmental justice is promoted by ensuring that the cleanup and abatement of the contamination of this area promotes equity and affords fair treatment, accessibility and protection for all members of the community, regardless of their race, age, culture, income or geographic location. In order to effectively participate in these matters, the Water Board believes it is essential that the community have access to independent technical consultants. The cost of this effort should be borne by the Discharger pursuant to Water Code sections 13304 and 13307.6.

32. The Water Board acknowledges that providing bottled water to residences or businesses currently served by affected wells would, on its face, satisfy the requirement for uninterrupted replacement water service, specifically since the beneficial use affected is water for consumptive purpose and bottled water could meet this need. However, environmental justice requires that bottled water not be the permanent solution for this community. In more urban communities, long-term replacement water service would likely consist of replacing the source water, thereby allowing community members total and unrestricted use of all household taps for consumptive use. Relying on long-term use of bottled water for all consumptive uses for residences that previously had the ability to consume water from any household tap interferes with the free use of their property and deprives those persons of prior quality of life expectations. In those situations where the Discharger's actions require replacement water service, it is appropriate to require that not only the quality, but also the long-term replacement water service, be comparable to that which it was prior to the adverse effect to the water supply, even if bottled water must be the source of replacement water service on an interim basis. The fact that replacement water service will likely be in place for many years increases the necessity that there be a requirement in this Order for long-term replacement water service that enables the residents of the community to use their household taps.

33. Pursuant to Water Code section 13304, the Water Board is entitled to, and may seek, reimbursement for all reasonable costs actually incurred by the Water Board to investigate unauthorized discharges of wastes or to oversee cleanup of such waste, abatement of the effect thereof, or other remedial action pursuant to this Order.

34. This Order requires workplans, monitoring, and reports pursuant to Water Code section 13267, subdivision (b). Workplans and technical reports required are essential to design a long-term water replacement plan and implementation schedule to verify compliance with this Order. Monitoring is required to verify that the interim and long-term replacement water service option(s) implemented provides water that meets the quality requirements of the Water Code and this Order.

35. The issuance of this Order is an enforcement action taken by a regulatory agency and is exempt from the provision of the California Environmental Quality Act (CEQA) (Public Resources Code section 21000 et seq.), pursuant to California Code of Regulations (CCR), title 14, section 15321, subdivision (a)(2). In addition, CEQA includes a "common sense exemption" in CCR title 14, section 15061, subdivision (b)(3), which states that where it can be seen with certainty that there is no possibility that the activity in question may have a significant effect on the environment, the activity is not subject to CEQA.

36. In this case, the Discharger may comply with the requirement to provide replacement water service by providing interim bottled water service and developing a permanent replacement water supply by installing wellhead treatment, establishing deeper domestic wells, or installing above-ground tanks (to store hauled water). There is no possibility that these activities would have a significant effect on the environment. Should a community water system be selected as a means of providing long-term replacement water service, the Water Board, if it is the lead agency under CEQA, will address CEQA requirements.

IT IS HEREBY ORDERED, pursuant to Water Code sections 13267 and 13304, the Discharger must:

1. Interim replacement water supply

a. **Within five (5) days from the date of this issuance of this Order, and within five (5) days of the submittal of each quarterly report delineating a revised affected area,** supply interim uninterrupted replacement water service (i.e., bottled water or equivalent), to all those served by domestic and community wells in the affected area where those wells are determined to be "impacted" as defined in Finding No. 30 of this Order and as determined pursuant to Paragraphs 3.a. and 3.b. below. This requirement is suspended once the Discharger provides a permanent replacement water supply or the well meets the conditions specified in Paragraphs 3c or 3.d. below.

b. **Within 14 days from the date of issuance of this Order, and within 14 days of the submittal of each quarterly report delineating a revised affected area** provide a report to the Water Board listing all properties that

have been provided interim uninterrupted water service. The report must include addresses and well numbers. The report must list the bottled water service being used and the water volume being provided. The report must include documentation to show that interim water supply meets state primary and secondary drinking water standards and hexavalent chromium levels of less than 0.02 µg/L[1] or the final MCL, once that standard is adopted by CDPH. The Discharger may propose a higher standard if it can demonstrate that the hexavalent chromium levels in the affected well prior to being impacted by the discharge was higher than 0.02 µg/L. If interim water supply is denied by a property owner or occupant, provide proof or evidence of such refusal.

c. **Within 30 days of the issuance of this Order**, provide a report to the Water Board that is acceptable to the Executive Officer describing how the Discharger intends to provide interim replacement water that achieves the quality limits described in 1.b. above. This report must address the following: source(s) of the replacement water, available information on the variability of the quality of the supply water, supply chain management considerations, proposed testing frequency based on any variability information and supply chain management plans, and a contingency plan. Additionally, the Discharger must provide a report to the Water Board at least 15 days prior to changing any aspect of the method for providing interim replacement water service. However, in the case where the Discharger must change its method due to unplanned or unanticipated quality issues or availability, the Discharger may change its method without first notifying the Water Board if needed to maintain compliance with this Order. In this situation, the Discharger must submit a report to the Water Board within five (5) days of making the change that describes the changes and addresses each of the topics required in the original report.

d. **Quarterly** (as part of its quarterly reports), provide monitoring information on the quality of the replacement water service consistent with the monitoring plan submitted in 1c above or as modified by the Water Board.

2. Permanent replacement water supply

a. **By no later than 30 days from the date of this signed Order**, submit a work plan to prepare the feasibility study required in Paragraph 2.c. below. The Workplan must include a conceptual outline of the analysis of each alternative and a project management schedule for completing each major task in the feasibility study.

b. **By not later than 110 days from the date of this signed Order,** submit a status report on the progress to prepare the feasibility study which should include a

[1] For purposes of this standard, drinking water must test below the reporting limit of 0.06 µg/L due to the limitation of laboratory analysis of low levels of chromium.

summary of results through the first three months and any indications that alternatives may or may not be viable.

c. **By no later than 180 days from the date of this signed Order**, submit to the Water Board a feasibility study on method(s) to provide permanent replacement water supply for all indoor domestic uses for all impacted wells in the affected area. Permanent replacement water must meet all California primary and secondary drinking water standards and hexavalent chromium levels of less than 0.02 µg/L[2] or the final MCL, once that standard is adopted by CDPH. The Discharger may propose a higher standard if it can demonstrate that the hexavalent chromium levels in the affected well prior to being impacted by the discharge was higher than 0.02 µg/L. The feasibility study must include the following:
   1) evaluate various methods to provide replacement water supply including, but not limited to: replacing individual wells with deeper individual wells, storage tanks and hauling water, providing point of entry treatment systems (evaluate at least three systems that use at least two different technologies), and an area wide or community water system by either consolidation with an existing public or private water purveyor, forming a new system (either public or private) or developing a system for two or more residences that may not involve a regulated water purveyor.
   2) Discussion of the feasibility and timing to implement each method including the need and timing for permits, approvals and environmental analysis.
   3) Results of pilot studies of each treatment method that is not certified to reduce hexavalent chromium to levels needed to achieve compliance with this Order.
   4) An evaluation of the quantity of water (gallons per minute) that can be provided by each method and a comparison with typical household supply needs.
   5) An evaluation of the quality of water that can be provided by each method in comparison with California primary and secondary drinking water standards and with levels of hexavalent chromium of less than 0.02 µg/L[3].
   6) An analysis of by-products or wastes that may be generated by each method and disposal options and costs.
   7) An operations, maintenance and, if appropriate, replacement plan.
   8) A water quality monitoring and reporting plan to verify quality and performance of each method.
   9) A complete cost analysis including construction, operations, maintenance and replacement.
   10) A contingency plan to ensure uninterrupted replacement water service.

d. The Discharger must present this feasibility study to the community and determine the acceptability of each method on a community-wide and specifically

---

[2] For purposes of this standard, drinking water must test below the reporting limit of 0.06 µg/L due to the limitation of laboratory analysis of low levels of chromium.

[3] For purposes of this standard, drinking water must test below the reporting limit of 0.06 µg/L due to the limitation of laboratory analysis of low levels of chromium.

from those currently being provided interim replacement water service and, if different, the owners of the impacted wells.

e. **Within 90 days of acceptance of the plan by the Water Board**, the Discharger must implement permanent replacement water service for all impacted wells. This schedule may be extended by the Water Board if it accepts a plan that requires more time to implement as demonstrated by the feasibility study.

f. **Within 120 days from the date the Water Board accepts the plan to provide permanent replacement water service**, provide a report to the Water Board listing all properties that have been provided permanent uninterrupted replacement water service. The report must include addresses and well numbers. State the method used to provide permanent uninterrupted replacement water service and provide evidence to prove that provided water meets state primary and secondary drinking water standards and contains hexavalent chromium in concentrations no greater than 0.02 µg/L[4] or the final MCL, once that standard is adopted by CDPH. The Discharger may propose a higher standard if it can demonstrate that the hexavalent chromium levels in the affected well prior to being impacted by the discharge was higher than 0.02 µg/L. If storage tanks or transportation vehicles are used to store or transport water, provide evidence of state or local government certification. If permanent replacement water supply is denied by a resident or business, provide proof or evidence of such refusal.

g. **Quarterly** (as part of its quarterly reports), provide monitoring information on the quality of the replacement water service consistent with the monitoring plan submitted in Paragraph 2.c.8 above or as modified by the Water Board.

3. Determination of impacted wells

a. **Within 45 days of issuance of this Order**, the Discharger shall propose a method or methods to perform an initial and quarterly evaluation of every domestic or community well in the affected area to determine if detectable levels of hexavalent chromium between the maximum background level and the PHG represent background conditions, or are more likely than not, partially or completely, caused by the discharge of waste by the Discharger. The proposed method or methods should take into consideration the factors listed in Finding No. 26 of this Order.

b. **Within 10 days of acceptance** by the Water Board Executive Officer of the proposal in 3.a. above and as part of all quarterly submittals providing new groundwater and domestic well sampling results, the Discharger shall submit an evaluation of domestic and community wells in the affected area and the results of its determination of whether or not the well is impacted.

---

[4] For purposes of this standard, drinking water must test below the reporting limit of 0.06 µg/L due to the limitation of laboratory analysis of low levels of chromium.

c. The Discharger may remove a well that was determined to be impacted due to total chromium levels above 3.2 µg/L or hexavalent chromium levels above 3.1 µg/L from impacted status if analytical results from four (4) consecutive quarters are below the above levels and the well does not meet the criteria for being designated as impacted by the accepted methods in 3.a. above.

d. The Discharger may remove a well that was determined to be impacted due to an evaluation using the accepted methods in 3.a. above from impacted status if the results of hexavalent chromium from four (4) consecutive quarters demonstrate that the well is no longer impacted based on the approved methods described in 3.a. above.

e. The Discharger may also provide evidence that the concentration of hexavalent chromium that is above 3.1 µg/L in a domestic or community well within the affected area is not due to its discharge and therefore be relieved of the requirement to provide replacement water service.

4. Independent Consultants

a. The Discharger must develop a process to fund an independent consultant(s) that can advise the community on matters subject to regulation by the Water Board. The independent consultant(s) selected by the community must not be involved in any aspect of this site (consulting for PG&E or involved in any litigation) and be acceptable to PG&E and the Water Board.

b. **Within 60 days of issuance of this Order**, the Discharger must develop a formal agreement with the community to implement this requirement. The Community Advisory Committee is the only existing group that may currently be viewed as representing the community. This Committee, a subset of the Committee or a totally different group would be acceptable as representing the community. It is also acknowledged that there are likely many divergent views in the community and that one group may not fully represent the spectrum of these views. The Water Board will monitor the Discharger's progress to implement this requirement and will modify this schedule if it determines that additional time is needed to develop an agreement acceptable to the community and will eliminate this requirement if the community rejects the need for independent consultants.

Order No. R6V-2011-0005

This Order amends Orders 1 and 2 in CAO R6V-2011-0005 for providing replacement water supply and submitting reports to the Water Board. All other Orders in CAO R6V-2011-0005 remain in effect unless later modified by the Water Board, the Water Board's Executive Officer, or his/her designated representative.

Laboratory Analysis

All future analysis of water samples must utilize the most recent testing methods. Testing for Total Chromium analysis must be done using US EPA Methods SW 6010B or 6020A to a reporting limit of 1 ppb. Testing for Hexavalent Chromium must be conducted in accordance with a modified version of EPA Method SW 218.6 with a reporting limit of 0.06 ppb.

The EPA has recently determined that detection limits of 0.02 ppb for hexavalent chromium are possible using a modified version of Method SW 218.6. These modifications allow for improved low concentration measurement and are outlined in Dionex Corp. Application Update 144 "Determination of Hexavalent Chromium in Drinking Water by Ion Chromatography" found at www.dionex.com/en-us/webdocs/4242-AU144_V18.pdf. The EPA determined that these modifications allow laboratories to attain a detection limit as low as 0.02 µg/L and can support a reporting limit of 0.06 µg/L (ppb). Information about the modified version of Method SW 218.6 is available at: http://water.epa.gov/drink/info/chromium/guidance.cfm.

The laboratory used must be certified by the California Environmental Laboratory Accreditation Program (ELAP) for hexavalent chromium analysis in drinking water. A list of certified labs is maintained by ELAP and is available at: (http://www.cdph.ca.gov/certlic/drinkingwater/Pages/Chromium6.aspx)

Liability for Oversight Costs Incurred by Water Board

The Discharger shall be liable, pursuant to Water Code section 13304, to the Water Board for all reasonable costs incurred by the Water Board to investigate unauthorized discharges of waste, or to oversee clean up of such waste, abatement of the effects thereof, or other remedial action, pursuant to this Order. The Discharger shall reimburse the Water Board for all reasonable costs associated with site investigation, oversight, and cleanup. Failure to pay any invoice for the Water Board's investigation and oversight costs within the time stated in the invoice (or within thirty days after the date of invoice, if the invoice does not set forth a due date) shall be considered a violation of this Order. If the Property is enrolled in a State Water Board-managed reimbursement program, reimbursement shall be made pursuant to this Order and according to the procedures established in that program.

Certifications for All Plans and Reports

All technical and monitoring plans and reports required in conjunction with this Order are required pursuant to Water Code section 13267 and shall include a statement by the Discharger, or an authorized representative of the Discharger, certifying (under penalty of perjury in conformance with the laws of the State of California) that the workplan and/or report is true, complete, and accurate. Hydrogeologic reports and plans shall be prepared or directly supervised by, and signed and stamped by a Professional Geologist or Professional Civil Engineer registered in California.

No Limitation of Water Board Authority

This Order in no way limits the authority of this Water Board to institute additional enforcement actions or to require additional investigation and cleanup of the site consistent with the Water Code. This Order may be revised by the Executive Officer as additional information becomes available.

Enforcement Options for Noncompliance with the Order

Failure to comply with the terms or conditions of this Cleanup and Abatement Order may result in additional enforcement action, which may include the imposition of administrative civil liability pursuant to Water Code sections 13350 and 13268 or referral to the Attomey General of the State of Califomia for such legal action as he or she may deem appropriate.

Right to Petition: Any person aggrieved by this action of the Lahontan Water Board may petition the State Water Board to review the action in accordance with Water Code section 13320 and California Code of Regulations, title 23, sections 2050 and following. The State Water Board must *receive* the petition by 5:00 p.m., 30 days after the date of this Order, except that if the thirtieth day following the date of this Order falls on a Saturday, Sunday, of state holiday, the petition must be received by the State Water Board by 5:00 p.m. on the next business day. Copies of the law and regulations applicable to filing petitions may be found on the Internet at: http://www.waterboards.ca.gov/public_notices/petitions/water_quality or will be provided upon request.

Ordered by: Harold J. Singer Dated: Oct 11, 2011

HAROLD J. SINGER
EXECUTIVE OFFICER

Attachment II:

Letter by PG&E to Harold Singer, dated July 12, 2011, regarding Draft CAO R6V-2011-0005A1 "Comment Letter"

(*Request for Immediate and Emergency Stay; Petition for Review; and Memorandum of Points and Authorities in Support Thereof*)

Pacific Gas and Electric Company®

Thomas C. Wilson
Director
Environmental Remediation

3401 Crow Canyon Rd.
San Ramon, CA 94583

925.415.6340
415.672.3324
TCW1@pge.com

VIA HAND DELIVERY

July 12, 2011

Harold Singer
Executive Officer
California Regional Water Quality Control Board
Lahontan Region
2501 Lake Tahoe Blvd.
South Lake Tahoe, CA 96150

Dear Mr. Singer:

Thank you for the opportunity to comment on the Draft Amended Cleanup and Abatement Order No. R6V-2011-0005A1 (the "Draft CAO") which, if adopted, would require PG&E to supply replacement water meeting certain quality standards to Hinkley residents. We respectfully request your careful consideration of the comments contained in this letter:

I. Introduction

PG&E has for many years acknowledged with gennine regret its responsibility for chromium contamination that contributed to the burdens and concerns of members of the Hinkley community. PG&E is committed to continuing to work cooperatively with the Lahontan Water Board, interested agencies and the Hinkley community to address these environmental impacts and community concerns stemming from our past operations at the Compressor Station located in Hinkley.

As part of our responsibility for remediation, PG&E currently operates the largest *in-situ* chromium remediation system in the world, as well as a large land treatment unit at the Desert View Dairy. PG&E has also been controlling the plume with a large fresh water injection system, which PG&E recently expanded. In addition, PG&E recently expanded agricultural pumping to further control plume movement that will result in a more than 350% increase in plume control pumping. PG&E is also actively pursuing additional remedial options as part of what is being called the final remedy.

In addition to these extensive remedial activities, PG&E has been actively working to reduce the ongoing concerns of Hinkley residents. We have undertaken a number of voluntary actions to address and respond to these concerns, including:

- PG&E offered to test chromium levels in any well within one mile of the plume.

- PG&E offered to purchase numerous properties near the plume – and its purchases have all been significantly above the appraised value.

- PG&E has provided bottled water to all well owners with chromium six levels above natural background levels,[1] as well as to the Hinkley School and the Hinkley Senior Center. This existing program satisfied the subsequent requirement of the Lahontan Board's original replacement water Order, which required bottled drinking water for Hinkley residents with chromium six levels above natural background.

- PG&E has also offered to supply bottled drinking water to any resident within one mile of the chromium plume, regardless of whether their well water exceeds background levels.

In light of these actions, which were voluntary and are above-and-beyond prior Lahontan Board direction, there is no reason to believe that any member of the Hinkley community is drinking water that contains an unsafe level of chromium six. Nonetheless, PG&E acknowledges that there is confusion and concern among some members of the Hinkley community regarding the safety of drinking water. We urge the Lahontan Board to partner with public health officials in an effort to educate and inform the Hinkley community and to address their concerns.

PG&E believes that the Board's new Draft CAO represents an unsupported and unreasonable expansion of water replacement requirements, and sends a confusing message to the Hinkley community. It would order interim replacement water for all wells withiu one mile of the plume that have chromium six concentrations greater than approximately 0.6 *percent* of natural background levels. It also mandates the permanent provision of water to any well with increased chromium six concentrations, even if they remain well below the natural background level. The Draft CAO is not supported by state law, science, engineering or public policy. As you observed in a May 26, 2011 e-mail on the subject: "Given that the groundwater in this area is well below the current MCL [Maximum Contaminant Level] for total chromium and that a public health goal is still in draft form, any order would likely have significant technical, legal and policy considerations." This letter outlines several of the significant technical, legal and policy concerns stemming from the Draft CAO.

Draft CAO Not Authorized by State Law

PG&E's position is that for several reasons, any one of which would invalidate the Draft CAO, California law does not authorize the replacement water Draft CAO under these circumstances:

---

[1] A total of twelve domestic wells in Hinkley have been found to contain chromium six above background levels. PG&E has agreements to purchase nine of the affected properties. Two of these wells were recently sampled, even though (i) they are approximately one mile outside the current plume boundary and (ii) there are significant reasons to conclude that the chromium six levels are not related to PG&E's activities. For example, well 34-65 is up-gradient from the plume and a fault acts as a barrier between the plume and that well. Nevertheless, PG&E provided bottled drinking water to both well users within days of the first sample showing chromium six levels above background, and PG&E is working with the Lahontan Board staff to determine whether there is any connection between PG&E's activities and the chromium six found in these wells.

- All domestic wells in Hinkley meet the state drinking water standard and, therefore, there is no legal support for an order requiring replacement water.

- It is improper to rely on a draft Public Health Goal ("PHG") to require replacement water. Doing so would elevate a draft goal to a *de facto* public drinking water standard without having first followed the required regulatory process for a drinking water standard or even for a final public health goal.

- State law does not allow remediation orders, including replacement water, for wells that contain chromium six levels below natural background levels. The Draft CAO's interim requirement to provide replacement water to all users whose wells are above 0.02 ppb chromium six within one mile of the plume, as well as the permanent requirement to provide water to any users whose wells indicate increased chromium six concentrations, are clearly at odds with this law.

- There is no evidence, much less substantial evidence as required by state law, that PG&E has affected the groundwater one mile beyond the plume boundaries. In fact, the Draft CAO would – contrary to state law – require PG&E to demonstrate that it has not impacted each well in the Hinkley area.

- State law does not permit an order, as proposed in the Draft CAO, that requires replacement water that is of higher quality than natural background water.

- The Draft CAO is void for vagueness. For example, it does not define the wells subject to the Draft CAO or the volume of water to be supplied.

Lack Of Scientific Support For The Draft CAO

As explained in more detail in Section II below, the Draft CAO is not supported by science. Based on the draft PHG, the Draft CAO states that chromium six in domestic wells above 0.02 ppb poses an immediate health risk to Hinkley residents. However, the language in the draft PHG, as well as an understanding of the science and purpose behind the PHG, demonstrates that such a conclusion is inappropriate. In addition, the draft PHG does not reflect emerging science indicating that there is a level below which chromium six does not have an adverse effect. The draft PHG's reliance on supposed swamp cooler or similar inhalation risks is not supported by the scientific record. And, even if the draft PHG were a valid regulatory tool, the science behind the draft PHG indicates that the most reasonable way to eliminate the purported risk would be to supply bottled drinking water to well owners with chromium six levels above natural background.

Compliance With The Draft CAO Is Not Feasible

The Draft CAO's requirements are so sweeping and the level of compliance (0.02 ppb chromium six) so low that it is simply not feasible to comply with them given current testing methods and water treatment technologies. A team of experts retained by PG&E concluded that there is no water source or treatment method in existence today that could reliably produce water below 0.02 ppb at the

volumes required by the Draft CAO. Indeed, even many bottled water sources contain chromium above the limits set in the Draft CAO. In addition, it is not even possible to reliably test water at levels below 0.1 ppb chromium six. Finally, as noted below in Section IV(C), the interim and permanent replacement water provisions contradict one another and would result in wasteful and unnecessary activities at best.

Public Policy Concerns

The Draft CAO would create a statewide public policy precedent with significant ramifications. The Draft CAO's claim that drinking water above 0.02 ppb chromium six poses an immediate risk to human health would require all entities within the Lahontan region and, indeed, throughout California to stop distributing water with greater than 0.02 ppb chromium six levels. A large number of California cities, including cities near Hinkley such as Victorville and Apple Valley, have water supplies with chromium six levels hundreds of times higher than the draft PHG. As a result, the precedent set by the Draft CAOcould create significant uncertainty statewide, as could the precedent of relying on a draft regulation.

PG&E asks that you exercise your discretion and not issue the Draft CAO in its current form in light of PG&E's extensive remediation activities and the fact that the Draft CAO is not supported by state law, is contrary to the latest scientific research, is infeasible, and would create far reaching statewide implications.

II. State Law Does Not Provide Authority To Order Replacement Water In Hinkley

A. It Is Improper To Order Replacement Water Where Contaminant Concentrations In Applicable Wells Are Below State Drinking Water Standards

California has a Maximum Contaminant Level ("MCL") that applies to chromium six. As the Department of Public Health – the agency that sets MCLs in California – confirms: "Chromium-6 (hexavalent chromium) is currently regulated under the 50 ppb maximum contaminant level (MCL) for total chromium."[2] The State Board has made clear that it is improper for regional boards to require replacement water for wells that meet established drinking water standards, and it recently rejected a similar attempt by the Central Coast Board to require replacement water in such circumstances.[3]

---

[2] *Chromium-6 MCL Update*, available at http://www.cdph.ca.gov/certlic/drinkingwater/Pages/Chromium6.aspx. *See also* email from Harold Singer to Laurie Kemper, dated May 26, 2011, and attached to this letter as Exhibit A ("I understand that the Water Board Prosecution Team is developing an order to address a directive from the Water Board to evaluate the need for whole-house replacement water for Hinkley residents affected by the historic chromium discharge from the PG&E Compressor Station. *Given that the groundwater in this area is well below the current MCL for total chromium* and that a public health goal is still in draft form, any order would likely have significant technical, legal and policy considerations." (Emphasis added.))

[3] See SWRCB Order 005-0007, p. 6: "Wells 'affected' by a discharge of waste include those wells in which water does not meet the federal, state, and local drinking water standards." (Emphasis added). The State Board further stated: "The logical

Footnotes Continued on Next Page

Therefore, the Draft CAO raises statewide policy questions. All of the domestic wells in the Hinkley area meet the current state MCL for chromium. On this basis alone, the provisions in the Draft CAO requiring PG&E to provide replacement water to occupants whose wells meet existing MCLs should not stand.

B. It Is Improper To Treat A Draft PHG As A Drinking Water Standard

In spite of the existence of the total chromium MCL, the Draft CAO improperly relies on a draft chromium six PHG as a drinking water standard. The Draft CAO attempts to justify its use of the 0.02 ppb draft PHG for chromium six by asserting that "[w]here OEHHA [the Office of Health Hazard Assessment] has established a PHG but DPH has not established an MCL, the State Water Resources Control Board . . . has determined that it is appropriate for a regional water board to require replacement water for wells affected at levels exceeding the PHG." (Draft CAO, Authority – Legal Requirements, § 18, p. 5.) This assertion misrepresents both the current facts and the applicable law.

First, there is no PHG for chromium six. The 0.02 ppb standard created by the Draft CAO is based on a *draft* – and not a *final* – PHG currently under review by OEHHA. Second, even if the draft PHG for chromium six were final, a PHG, unlike an MCL, does not constitute a drinking water standard (Health & Safety Code § 116365(a), (c)) and, thus, cannot form the basis for requiring replacement water. Finally, the Draft CAO misreads *In The Matter Of The Petitions Of Olin Corporation And Standard Fusee* and would actually violate the State Board's direction by setting an enforceable chromium six standard based on the Lahontan Board's independent review of the scientific literature, instead of relying on DPH.

1. The Draft PHG Is A Work In Progress

The Draft CAO elevates the draft PHG for chromium six to a *de facto* drinking water standard. This is improper on multiple fronts. First, the draft PHG is by definition a work in progress, not the final PHG. But it is the final PHG – and not any draft – that DPH will ultimately use to issue or revise any MCL for chromium six in the future. (Health & Safety Code § 116365(b)(1).) Thus, the draft PHG number relied on in the Draft CAO is at least *two* levels (and two public notice and comment processes) removed from being an enforceable drinking water standard. Moreover, OEHHA has already revised the draft PHG for chromium six at least once and is currently revising the draft PHG again in response to public comments, including scientific analysis directly contradicting some of its central

---

result of the Central Coast Water Board's argument that the State Water Board Res. 92-49 requirement for cleanup to background contaminant levels justifies its water replacement levels would routinely require water replacement for groundwater constituent levels that may be many times lower than that determined safe by state and federal agencies. Simply put, while cleaning up to background may be required, that does not mean that replacement water is always necessary until the cleanup is complete, regardless of the amount of contamination." (Id. at 6, fn. 23.)

conclusions.[4] As discussed in detail below, the use of a draft PHG raises serious statewide precedent and public policy issues.

2. A PHG Is Not (And Is Not Intended To Be) A Drinking Water Standard

The Lahontan Board should not rely on even a final PHG to create an enforceable drinking water standard. OEHHA is clear that the PHG for chromium six is not intended to mark the line between safe and unsafe concentrations, and is not intended to create an enforceable drinking water standard: "The PHG is not meant to be the maximum 'safe' level of chromium 6 in drinking water. It represents a stringent health-protective goal that CDPH will use to develop an *enforceable* regulatory standard for chromium 6 in drinking water." (Press Release, *OEHHA Releases Revised Draft Public Health Goal for Hexavalent Chromium*, Dec. 31, 2010; emphasis added.) OEHHA has also explained:

> A PHG is NOT a boundary line between a "safe" and "dangerous" level of a contaminant. Drinking water can still be acceptable for public consumption if it contains contaminants at levels higher than the PHG. A PHG is a health-protective level of a contaminant in drinking water that California's public water systems should strive to achieve *if* it is technically and economically feasible.

(Fact Sheet, *Draft Public Health Goal for Hexavalent Chromium*, Aug, 2009; emphasis in original.)

> The PHG itself recognizes the limitations of a PHG:
>
> PHGs *are not regulatory requirements*, but instead represent non-mandatory goals. Using the criteria described above, PHGs are developed for use by [DPH] in establishing primary drinking water standards (State Maximum Contaminant Levels, or MCLs). Thus, PHGs are not developed as target levels for cleanup of ground or ambient surface water contamination, and may not be applicable for such purposes, given the regulatory mandates of other environmental programs.

(Draft PHG, at iii (emphasis added).)

---

[4] For example, one of the comments OEHHA must address the substantial critique of OEHHA's methodology found in an October 23, 2008 Memorandum by the Department of Toxic Substances Control's Senior Toxicologists. This memorandum, which is attached to this letter as Exhibit B, concluded that "generation of a PHG for hexavalent chromium at this time may be premature as it is not possible to assign a dose-response relationship – other than the default OEHHA assumptions and methods used since 1985. Additional investigations are indicated and should be considered before public release of the PHG value or its documentation." (DTSC October 23, 2008 Memorandum at 9.)

As DPH made clear in its June 27, 2011 letter to you commenting on the Draft CAO:

> The Lahontan Water Board's finding in the draft [CAO] is that the draft PHG level "is an appropriate standard to rely on to protect the public from contaminated drinking water, despite the fact that it has not been formally promulgated." *CDPH thinks it is premature, and has potentially far overreaching implication to domestic water supply wells in the state, to use the draft PHG for this purpose.*

(Emphasis added.)

The Draft CAO would immediately elevate a *draft* PHG to an enforceable drinking water standard and, in so doing, skip over the extensive process California follows to create an enforceable standard. It would eliminate all of the analysis mandated for creating an enforceable standard, including the analysis of technical and economic feasibility of complying with the standard.[5] In so doing, the Draft CAO would create serious public policy concerns statewide.

Tellingly, one of the lead authors of the draft PHG told Hinkley residents that he felt it was safe to use the water in Hinkley. In comments before the Lahontan Board in January 2011, Dr. Robert Howd, one of the lead authors of the chromium six PHG at OEHHA, opined that "people should be safe drinking water at the levels [of chromium six] found in drinking water wells in Hinkley." ("Chromium 6 Plume Boundary Shows Slight Changes," *Desert Dispatch*, 1/27/2011.)

3. The Draft CAO Misreads *In The Matter Of The Petitions Of Olin Corporation And Standard Fusee* And Would Actually Violate Its Direction By Creating A New Drinking Water Standard Without Deferring To The State Agencies Charged With That Task

The Draft CAO relies on a prior State Board decision, but the decision does not support the Draft CAO. SWRCB Order WQ 2005-0007 was rendered in *In The Matter Of The Petitions Of Olin Corporation And Standard Fusee* ("*In re Olin*"). The petitioners had discharged potassium perchlorate

---

[5] As described later in these comments, PG&E's expert team has determined that it is technically and economically infeasible to comply with the Draft CAO. PG&E's team performed the very analysis that is required by state law, but which should be performed by DPH prior to creating an enforceable drinking water standard. (*See* Health & Safety Code § 116365(b)(1) ("The [California Department of Public Health] shall consider all of the following criteria when it adopts a primary drinking water standard: (1) The public health goal for the contaminant adopted by the Office of Environmental Health Hazard Assessment pursuant to subdivision (c); (2) The national primary drinking water standard for the contaminant, if any, adopted by the United States Environmental Protection Agency; (3) The technological and economic feasibility of compliance with the proposed primary drinking water standard. For the purposes of determining economic feasibility pursuant to this paragraph, the department shall consider the costs of compliance to public water systems, customers, and other affected parties with the proposed primary drinking water standard, including the cost per customer and aggregate cost of compliance, using best available technology.")).

into groundwater. Pending the outcome of its investigation, Olin provided replacement water to owners of wells testing at or above 4 ppb. Later, OEHHA issued a *final* perchlorate PHG of 6 ppb; however, DPH had not issued an MCL for perchlorate and no other safe drinking water standard had been issued. Natural background levels for perchlorate were far below the final PHG (unlike in Hinkley, where the natural background chromium levels are far above the draft chromium six PHG).

Olin asked that the Central Coast Regional Board's Order limit required replacement water to owners of wells with perchlorate concentrations of at least 6 ppb, the PHG level. The Central Coast Board refused, and Olin and Fusee appealed to the State Board, contending that the Central Coast Board's order to provide replacement water for owners of wells testing below the "final PHG of 6 µg/L adopted by OEHHA" was an abuse of discretion. (*Id.* at 4.) In ruling on the petition, the State Board directed the Central Coast Board to defer to DPH and to OEHHA in determining safe levels of contaminants in drinking water:

> Any other approach would require regional water boards to make individual, possibly inconsistent public health and toxicological determinations or, in the alternative, to require replacement drinking water whenever there is any detection of a contaminant. This approach ignores the expertise of OEHHA and, in the case of contaminants for which MCLs have been developed, [DPH].

(*Id.*, at 5.) The State Board also noted that a PHG, even one that is final, "is not a legally enforceable standard." (*Id.*)

Here, instead of deferring to DPH, the state agency charged with determining and establishing primary and secondary drinking water standards (Health & Safety Code §§ 116275(c)-(d), 116365), the Draft CAO would set an enforceable chrome six standard based on the Lahontan Board's independent review of the scientific literature.[6] (Draft CAO, Findings, § 25, p. 7 ("The Water Board finds it is appropriate to rely on this [draft] standard based on the vast amount of sound scientific evidence and agency peer review supporting the draft 2010 PHG").)

---

[6] The Draft CAO is unclear as to whether this review was conducted by a toxicologist, an enforcement officer or a person(s) in another staff position. But regardless of the qualifications of the reviewers, "regional water boards should defer to OEHHA and DHS in determining safe drinking water levels." (*In re Olin* at 7.)

The Draft CAO also appears to critique the risk assessment on which the draft chrome six PHG is based, implying that it may not sufficiently protect human health:

> The 2010 draft PHG does not include peer-reviewed scientific studies of the risks associated with the use of hexavalent chromium –contaminated water in other domestic appliances, including swamp coolers . . . Normal household use of these appliances may present additional inhalation risks not accounted for in the 2010 draft PHG.

(Draft CAO, Recent Changes in the Regulation of Chromium 6, § 17, p. 5.) The Draft CAO's independent review of the draft chrome six PHG and determination that it represents "an appropriate drinking water standard" results in the very danger identified by the State Board – the Lahontan Board would be making "individual, possibly inconsistent public health and toxicological determinations" instead of allowing the agencies with expertise in such matters to promulgate regulations addressing the safety of chrome six in drinking water.

DPH is clear and the Draft CAO acknowledges that chromium six is currently regulated under the total chromium MCL. As such, the Lahontan Board's effort to enforce a draft PHG is not the equivalent of the Central Coast's effort to enforce a final perchlorate PHG in the absence of an MCL. *In re Olin* (i) explicitly directs regional boards to defer to DPH and OEHHA in setting safe drinking water standards (*id.*, at 5); (ii) does not even mention, let alone endorse, the use of a draft PHG; (iii) by its own terms is not to be cited as authority for setting replacement water quality standards (*id.*, at 7); and (iv) dealt with perchlorate levels well *above* natural background levels and therefore not inconsistent with Porter-Cologne Water Quality Control Act and the State Board's enforcement policies that a discharger cannot be required to remediate below background concentrations. (*Id.*, at 3). In short, *In re Olin* merely stands for the proposition that where there is no applicable MCL and background levels of a contaminant are below a final PHG, a regional board may refer to that PHG in determining which wells require replacement water.

Until such time as OEHHA finalizes the draft PHG for chromium six and, through the proper processes and with the proper considerations, DPH establishes or revises the MCL for chromium six, the Lahontan Board should calibrate any chromium six concentrations in any proposed replacement water order based on the existing and enforceable MCL for total chromium (which includes chromium six).

C. State Law Does Not Authorize An Order Requiring Cleanup Of, Or Replacement Water For, Wells That Contain Chromium Six Levels Below Natural Background Levels

Any requirement that replacement water be provided to well owners with chromium six levels below background would run afoul of State Board Resolution 92-49, which is explicit that, consistent with the State's non-degradation policy, regional boards may not require cleanup or abatement below background conditions. (Resolution 92-49 § III.F ("under no circumstances shall these provisions be interpreted to require cleanup and abatement which achieves water quality conditions that are better than background conditions").) Replacement water is a component of abatement and mitigation for

discharges and, therefore, governed by this rule. There is no legal basis to require a discharger to respond to natural, pre-existing conditions and the Draft CAO would, therefore, create a dangerous precedent and statewide public policy question.

While the Draft CAO acknowledges the background study that was performed by PG&E in Hinkley, it ignores the results of the study when it requires interim water for all wells above 0.02 ppb and permanent water for all wells with increases in chromium six levels (including wells below natural background levels). In fact, without any basis or analysis, the Draft CAO concludes, "[t]here is no indication that the rising chromium levels are a result of fluctuation in the naturally occurring constituents" and "[t]he rise in chromium levels indicates that the anthropogenic hexavalent chromium plume resulting from the discharge of chromium at the Discharger's compressor station is migrating to new areas in the upper aquifer." (Draft CAO, ¶ 8, p. 3.) The prosecution team ignored the specific instruction to present "any evidence . . . they are relying on to support the draft order." The conclusions asserted in the Draft CAO are not supported by any evidence, and are contrary to the established science and demonstrated record.

As stated by Dr. Brian Schroth, an expert on chromium in groundwater, chromium six is naturally present in Hinkley and throughout the world. "Naturally-occurring hexavalent chromium is ubiquitous in groundwater systems throughout the Mojave Desert and globally, with naturally-occurring concentrations sometimes exceeding 50 µg/L in alluvial aquifers in the western Mojave Desert and elsewhere in central and southern Arizona, and western New Mexico." (Declaration of Brian Schroth, which accompanies this letter, ¶ 3(a).) In the Hinkley area, the Lahontan Board's 2008 cleanup and abatement order established maximum groundwater chromium six background levels at 3.1 ppb. This background level has been used by both the Lahontan Board and PG&E for many purposes since 2008, including plume boundary depictions and final remedy cleanup analysis. Unless and until something changes, this is the background level established by formal research and a regulatory order.[7]

However, the Draft CAO requires interim replacement water for any well within one mile of the plume that is above 0.02 ppb chromium six. This requirement ignores natural background levels, including the established background level of 3.1 ppb. There is no analysis or science to (i) justify ignoring the 3.1 ppb standard that the Lahontan Board has historically accepted and (ii) support using 0.02 ppb as the trigger level for interim water, particularly in areas outside the plume boundary.

---

[7] The Draft CAO states, "[a] final determination of background water quality has not been made." (Draft CAO, Authority – Legal Requirements, § 23, p. 6.). However, PG&E's background study and the 2008 Cleanup and Abatement Order established a maximum chromium six background level of 3.1 ppb. Until and unless it is changed, this remains the operative legal background level.

With regard to wells with increased chromium six levels (even those below natural background), the Draft CAO states without any support that the increase is caused by PG&E. However, Dr. Schroth has collected and reviewed the chromium six data from groundwater supply wells statewide. (*Id.*, ¶ 4.) The vast majority, if not all, of these wells are considered to contain naturally occurring chromium six. (*Id.*, ¶¶ 5-6, 8.) Dr. Schroth's research demonstrates that it is extremely common for wells containing naturally occurring chromium six to increase in concentration (or to decrease in concentration) over time due to natural forces:

> Groundwater quality records collected by the CDPH show that concentrations of Cr(VI) detected in water supply wells vary considerably over time at any given well. As a result, increases or decreases in the concentration of Cr(VI) at a given well do not always signify the arrival or departure of a particular source or plume of Cr(VI). Rather, these changes may be expected as a result of other factors, including sample collection procedures, seasonal changes, changes in well operation, laboratory analysis, variations in annual precipitation, and other factors.

(*Id.*, ¶ 9.) Dr. Schroth's Declaration includes charts showing groundwater data collected by DPH for the Mojave area. (*Id.*, ¶ 10.) The charts demonstrate that it is very common for the concentration of chromium six to vary in a random pattern around a background value. (*Id.*) For example, concentrations of Cr(VI) detected in Hesperia Water District well 15-A have ranged from 2.6 to 7.93 ppb. (*Id.*, ¶ 11.) Similar concentration ranges are reported for Victor Valley Water District well 208 (Cr(VI) ranging from 4.2 and 9.5), Loma Linda University Anderson Well 2 (Cr(VI) ranging from 1.3 to 5.4 ppb), and Anderson Well 3 (Cr(VI) ranging from 2.0 and 4.5). (*Id.*)

If the 0.02 ppb chromium six draft PHG value were adopted as a state drinking water standard, over 50% of the drinking water supply wells in California would likely exceed this concentration. In a SWRCB study of active and standby drinking water supply wells, 3,156 out of 5,943 wells tested between 1997 and 2008 had concentrations exceeding 1 ppb. (*Id.*, ¶ 8(a).) Given that this study used sampling methods with detection limits dramatically higher than 0.02 ppb, it is very likely that significantly more wells would show chromium six detections above 0.02 ppb if the sampling were done using new sampling methods with lower detection limits. (*Id.*, ¶ 7.)

D. State Law Does Not Authorize An Order Requiring PG&E To Provide Replacement Water For Wells That PG&E's Operations Did Not Impact

The Lahontan Board has previously identified the boundaries of the plume impacted by historic releases from PG&E. The existing order defines the plume as those areas where groundwater exceeds 3.1 ppb chromium six and 3.2 ppb total chromium in the upper aquifer. (Draft CAO, Plume Migration, § 7, p. 2.) The Draft CAO, however, would (i) redefine the "affected area" to include all domestic wells located within one mile of the plume (*id.*, Findings, § 29, p 8) and (ii) define "impacted wells" as wells in the "affected area" with chromium concentrations above background concentrations *or* that have statistically significant (by a yet-to-be-determined standard) greater concentrations than past chromium six concentrations in the same well. (*Id.*)

On an interim basis, the Draft CAO would require PG&E to provide replacement water to all residences and businesses with wells in the "affected area" containing chromium six above 0.02 ppb. (*Id.*, Interim Replacement Water Supply, § 1(a), p. 9.) On a permanent basis, the Draft CAO would require PG&E to provide replacement water to "impacted wells" (that is, wells above background or with increased chromium six levels) within the "affected area" (one mile outside the plume boundary). (*Id.*, Permanent Replacement Water Supply, § 2(a), p. 9.) These requirements violate the Porter-Cologne Water Quality Control Act and would create a dangerous precedent and public policy question for cleanup activities statewide.

First, Water Code section 13304(a) permits the State and regional boards to require persons who discharge waste into State waters to provide replacement water to each "affected" public water supplier or private well owner. The Draft CAO exceeds the limits of the law to the extent that it orders PG&E to provide replacement water to well users whose wells are *outside* the identified plume, and to well users whose wells are below natural background levels for chromium six.

Second, section 13304(a) only permits the Lahontan Board to order replacement water to a Hinkley resident if there is "substantial evidence" that PG&E's historic releases are the source of elevated levels of chromium six in a specific well. (*See In the Matter of the Petition of Exxon Company, USA*, 1985 WL 1120860, 6 ("while we can independently review the Regional Board record, in order to uphold a Regional Board action, we must be able to find that finding of ownership [i.e. responsibility] was founded upon substantial evidence."); *In the Matter of the Petitions of Aluminum Company Of America; Alcoa Construction Systems, Inc.; And Challenge Developments, Inc.*, 1993 WL 303166, 3 ("There must be substantial evidence to support a finding of responsibility for each party named").) Substantial evidence to support a finding of responsibility for each party named means credible and reasonable evidence which indicates the named party has responsibility. (*In the Matter of the Petition of Exxon Company, USA* 1985 WL 1120860, 6.) An arbitrary determination of responsibility is an abuse of discretion. (*Id.*; *In the Matter of the Petitions of Aluminum Company Of America; Alcoa Construction Systems, Inc.; and Challenge Developments, Inc., supra*, 1993 WL 303166, 3.)

The Draft CAO defines "affected area" as all domestic wells located within one mile of the plume. (*Id.*, Findings, § 29, p 8). That definition includes wells located *up-gradient* and *cross-gradient* from the source of the historic release which clearly are not related to or impacted by the release. The Draft CAO has no basis to cast so wide a net.

The Draft CAO contains no evidence, science, regulatory guidance or past agency action to support its expansive and revised definition of the "affected area." And the Lahontan Board has provided no other "supporting information" to PG&E. The law does not permit a regional board to define an affected area, or issue related orders, on that basis.

E. The Lahontan Board Would Abuse Its Discretion If It Were To Require Replacement Water That Contained Lower Concentrations Than Natural Background

State law does not authorize orders requiring replacement water to meet standards below natural background. Existing primary and secondary drinking water standards do not prohibit chromium six concentrations above 0.02 ppb. However, according to the Draft CAO, PG&E must provide replacement water meeting "state primary and secondary drinking water standards *and* hexavalent chromium levels of 0.02 µg/L." (Draft CAO, Interim Replacement Water Supply, §§ 1(a), 1(b), p. 9; Permanent Replacement Water Supply, 2(a) and 2(b), pp. 9-10.) The Draft Order highlights that the drinking water standards do not specifically address chromium six separate and apart from total chromium. If the maximum chromium six concentration of 0.02 ppb were subsumed in the drinking water standards, the Draft CAO would not have had to separately define the maximum chromium six concentration. Requiring replacement water with a higher quality than natural background concentrations would create challenges and public policy questions statewide.

Water Code section 13304(f) requires that replacement water "meet all applicable federal, state, and local drinking water standards, and *shall have comparable quality to that pumped by the public water system or private well owner prior to the discharge of waste.*" (Emphasis added.) The State Board's own enforcement policy only requires a discharger to abate contamination to background levels. (*See*, SWRCB Water Quality Enforcement Policy (2010), at p. 35.) Thus, according to Porter-Cologne aud the State Board's own pronouncements, replacement water need only be as clean as the water before the discharge occurred; background levels need not be improved upon.

If the Draft CAO's replacement water standards are a harbinger of future cleanup standards, the Lahontan Board would be creating an expectation within the Hinkley community that cannot be met. As described below, there is no viable means of achieving a cleanup level consistent with the 0.02 ppb chromium six standard. When the time arrives to set goals for the eventual remediation, almost assuredly the Lahontan Board will be forced to backtrack and, in essence, announce that what it considered unsafe in 2011 is, in fact, not a threat to public health.

F. The Draft CAO Is Void For Vagueness

The State Board has long recognized that its orders implicate the "constitutional issues of void-for-vagueness and overbreadth under the due process clause of the Fourteenth Amendment." (*In The Matter Of The Petition Of United States Steel Corporation*, 1976 WL 376714, 6-7.) The State Board further acknowledged the need for its orders to satisfy due process requirements when it explained, "[a]n administrative order such as the present one, which may be enforced by a penal sanction under Water Code Sections 13265 and 13387 may be equally effective as a deterrent to the exercise of constitutional rights as a penal statute." (*Id.*)

The Draft CAO would require PG&E to provide replacement water for an *indeterminate* number of wells insofar as it defines "impacted wells" as those with concentrations "that are statistically significantly greater (at a confidence level to be determined) than past hexavalent chromium

concentrations in that same well." (Draft CAO, Findings, § 29, p. 8.) The Draft CAO does not define the statistical test and specifically states that the confidence level will be determined later. It also orders PG&E to, "*at a minimum*," provide sufficient water for drinking, cooking and swamp cooler needs. (Draft CAO, Interim Replacement Water Supply, § 1(a), p. 9.) But, while the Draft CAO sets a *floor* on the necessary quantum of interim replacement water, it is silent on the *ceiling*. The Draft CAO also fails to adequately identify the universe of Hinkley residents that would receive replacement water.

The Draft CAO is vague on many essential points. How many wells does the Draft CAO encompass, and exactly where are those wells? How much more water than the minimum is required to satisfy the Lahontan Board? The Draft CAO does not provide answers. At best, it says the Lahontan Board will make those decisions at some unspecified time in the future. As a result, PG&E might never know whether it has complied with the Draft CAO. A court reviewing the draft order would surely find it void for vagueness in that it would not provide PG&E with (i) sufficient information to comply, and (ii) notice of when it might be deemed in violation and, therefore, subject to administrative, civil, or criminal penalties. The Draft CAO's ambiguity would set a troubling standard and raise far-reaching public policy questions.

III. Science Does Not Support The Draft CAO

A. The Draft CAO Demonstrates A Misunderstanding Of The Draft PHG And The PHG Process, And Ultimately Misuses The Draft PHG

The Draft CAO wrongly asserts that "hexavalent chromium in domestic wells above 0.02 µg/L *poses an immediate health risk to Hinkley residents* through continued household use of *contaminated water*, including drinking, preparing foods and beverages, bathing or showering, flushing toilets, and other household uses resulting in potential dermal and inhalation exposures."[8] However, the scientific knowledge regarding chromium six impacts comes from studies of workers and laboratory animals exposed to massive concentrations of chromium six thousands and even millions of times higher than the concentrations in Hinkley groundwater. There is no basis for concluding that chromium six levels above 0.02 ppb pose an immediate health risk.

---

[8] The Draft CAO also wrongly asserts that the OEHHA Chronic Inhalation Reference Exposure Level ("REL") demonstrates "*established science* that inhaled hexavalent chromium has *adverse impacts on human health at extremely low levels.*" (Draft CAO, Recent Changes In the Regulation of Chromium 6, § 15, p. 4 (emphasis added).) The correct REL is actually 0.2 µg/m^3 and is based on an animal exposure study of rats exposed to chromium six for 18 hours per day at concentrations ≥ 50 µg/m^3. Regardless, these air exposure levels have no relation to the groundwater chrome six concentrations in Hinkley.

Dr. Joshua Hamilton, an expert in chromium six toxicology, summarized the chromium six toxicology studies of the last eighty years:

> Despite over eighty years of intense study reported in tens of thousands of scientific papers, the only demonstrated adverse health effects of chromium occurred at levels of exposure that are more than a thousand times higher than those that would be encountered in environmental and household settings, including those in Hinkley. Conversely, there are no studies showing any adverse effects of Cr(VI) at levels anywhere near the current MCLs, let alone the background concentrations at Hinkley or the level proposed for the draft PHG.

(Declaration of Joshua W. Hamilton Decl. ("Hamilton Dec."), which accompanies this letter, ¶ 7.)

For regulatory purposes, risk assessors take the high-dose study results and then employ mathematical modeling with extremely conservative assumptions to develop estimates of levels without health impacts from a lifetime of exposure. (*Id.*, ¶ 8.) These estimates produce numbers that are tens of thousands to millions of times lower than the studies upon which they are based. (*Id.*) For example, the lowest concentration of chromium six that caused tumors in animals in the National Toxicology Program ("NTP") study that was used as the basis for the draft PHG was 20,000 ppb. (*Id.*) The proposed PHG level of 0.02 ppb is *one million times lower* than the concentration that caused cancer from a lifetime of drinking water exposure in mice. (*Id.*) There is no way to confirm any of the assumptions that are made in these models or to determine whether there are any measurable health effects at such low levels. (*Id.*) The estimates certainly do not represent "established science" demonstrating "immediate health risks" at the PHG level. (*Id.* ¶¶ 8-10.)

B. The Draft PHG Does Not Incorporate The Latest Science

The draft PHG's reliance only on studies that utilized massive doses on lab animals eliminates any chance to determine whether there is an exposure level below which there are no adverse health impacts. (*Id.*, ¶¶ 7-8.) Many of the comments on the draft PHG recognized this problem. Fortunately, science is moving forward to answer this specific question.

Emerging science is leading to the conclusion that there is a threshold level below which there are no adverse health effects. (*Id.*, ¶ 10(b), (c).) Ninety-day toxicity studies, currently in the final stages of completion, have begun to be published in the peer-reviewed literature. (*Id.*, ¶ 10(c).) As Dr. Hamilton notes:

> Based on the results presented to date, these studies will unequivocally support a threshold mechanism as the Mode of Action ("MOA") for Cr(VI) in vivo via ingestion and inhalation exposure. In fact, these studies were specifically designed to investigate the MOA and to complement the 2008 NTP studies in all respects, including study design. The pending studies are even being conducted by the same scientists that conducted the 2008 NTP studies.

(*Id.*)

In short, the latest science suggests that there is a level below which exposure to chromium six does not cause adverse health effects. Dr. Hamilton recently served on a USEPA expert panel charged with reviewing USEPA's chromium six toxicology. (*Id.*, ¶ 10(b).) The panel strongly urged USEPA to incorporate the new studies in USEPA's toxicology assessments. (*Id.*, ¶ 10(c).) USEPA recognizes that this new science should be considered before any further regulatory changes. (*Id.*, ¶ 11.) Thus, the Draft CAO should not rely on the draft PHG, which does not incorporate the latest science and any such reliance would create a troubling precedent for cleanup orders statewide.

C. Evaporative Swamp Coolers And Similar Devices Do Not Introduce Chromium Six Into The Air

Curiously, the Draft CAO criticizes the draft PHG, despite using it as a benchmark, for not including peer-reviewed scientific studies of the risks associated with the use of chromium six in evaporative swamp coolers and similar household appliances. (Draft CAO, Recent Changes In The Regulation Of Chromium 6, § 17, p. 5.) Not only does this analysis violate that State Board's policy prohibiting regional boards from crafting their own health standards (as discussed above), but it also contradicts the existing science.

In fact, OEHHA *did* consider inhalation risks in the PHG, and determined that they represent an extremely small portion of the calculated risk. (Hamilton Dec., ¶ 12.) OEHHA determined that the principal exposure pathway of concern for household exposure to chromium in drinking water supplies is *ingestion*. (*Id.*) OEHHA also specifically examined the question of inhalation exposure to chromium via showering – which is generally assumed to he the principal inhalation pathway of concern for households with contaminants in drinking water supplies – and included shower inhalation exposure in the draft PHG. (*Id.*) Exposure by inhalation during showering did not contribute significantly to the overall exposure or risk. (*Id.*) And even with conservative assumptions regarding exposure during showering, the contribution to risk from inhalation was 180 times lower than that from drinking water exposure. (*Id.*)

Dr. Hamilton also reviewed the relevant scientific literature and found two peer-reviewed scientific studies that concluded that – because chromium six is not volatile and does not evaporate into the air – evaporative swamp coolers do not introduce chromium six into the air. (*Id.*, ¶ 13(b).) Thus, Dr. Hamilton concludes that exposure to airborne chromium six from swamp coolers is not a pathway of concern for affected households in Hinkley or elsewhere:

> The scientific and regulatory literature confirms that inorganic constituents, including chromium, that may be present in the water used in swamp coolers are not volatile and do not evaporate with the water. Instead, the inorganic constituents remain behind on the filter or, for those units with recirculation versus a drip line and drain, in the sump. Moreover, a 1996 scientific publication by Finley et al. [] examiued Cr(VI)-contaminated water in an evaporative cooler, in a trial experiment in a Hinkley-area house with a typical evaporative cooler. They demonstrated that even using a concentration of Cr(VI) of 20,000 ppb in a unit running for twenty-four hours, there was no increase in the airborne Cr(VI) concentration above the natural outside and indoor backgrounds.

> Thus, there is no basis for any concerns regarding inhalation exposure risk from evaporative coolers, particularly at the concentrations in any impacted Hinkley households, which are more than 4,000 times lower than the levels examined in these experiments.

(*Id.*, ¶ 13(a).)

Dr. Hamilton also concluded that household appliances similar to evaporative swamp coolers would not introduce chromium six into the air:

> Like swamp coolers, other similar appliances (such as humidifiers aud hot water vaporizers) that act by volatilizing heated water or by evaporating water from a filter will not be a potential source of Cr(VI) into indoor air because Cr(VI) will not be volatilized with the water.

(*Id.*, ¶ 14.)

In short, there is no scientific support for attempting to justify the Draft CAO based on alleged risks from swamp coolers aud similar devices.

D. Focusing On The Inhalation Risk Alone, Even If The Draft PHG Were Utilized, The Most Reasonable Solution Would Be Replacement Of Drinking Water Only

It is unnecessary to require replacement water for anything other than drinking and cooking because if inhalation were the sole means of exposure at the levels found in Hinkley, there would be no risk to human health. (*Id.*, ¶ 12, fu 5.) That is true even using the conservative assumptions regarding inhalation in the draft PHG. (*Id.*, ¶ 12.) The PHG associated with chromium six inhalation exposure alone can be calculated from the information in the draft PHG by removing the contribution from oral exposures. (*Id.*, ¶ 12 fn 5.) That value – the PHG associated with inhalation exposure alone – is 3.6 ppb (PHG calculation on page 94 of draft PHG). (*Id.*) Thus, using the extremely conservative assumptions of the draft PHG, the PHG level associated with inhalation risk alone (3.6 ppb) would be higher than the natural background chromium six levels in Hinkley (3.1 ppb). Therefore, even if the draft PHG were somehow a valid regulatory tool, the most reasonable approach would be to require replacement water for drinking and cooking for users of wells testing above natural background levels. Once the ingestion risk calculated by the PHG is removed by using bottled water for drinking and cooking, even the draft PHG would not require replacement water for the remaining risks of chromium six levels below natural backgrouud.

IV. There Are Tremendous Technological And Practical Hurdles To Complying With The Draft CAO That Make It Infeasible

A. It Is Not Feasible To Supply Whole-House Replacement Water That Must Be Below 0.02 ppb Chromium Six And Certainly Not In The Limited Time Allowed By The Draft CAO

PG&E convened a team of experts to assess the available options to meet the Draft CAO's interim and permanent replacement water requirements. James DeWolfe of Arcadis U.S., Inc., headed the team. (Declaration of James DeWolfe ("DeWolfe Dec."), which accompanies this letter, ¶ 1.) The team concluded that there are no feasible means to comply with the replacement water requirements of the Draft CAO. (*Id.*, at ¶ 2.)

The interim replacement water requirements in the Draft CAO would require PG&E to deliver water to between 250 and 300 locations for drinking, cooking and swamp cooler needs.[9] Assuming three occupants per location, drinking and cooking would require 33 gallons per day. (*Id.*, ¶ 3(c).) In addition, swamp cooler needs during warm months would require an additional 40 gallons per day, resulting in a total of 73 gallons per day per location, and between 547,500 and 657,000 gallons per month for the 250 to 300 locations. (*Id.*) Because the Draft CAO would give PG&E only two weeks to provide interim replacement water for drinking, cooking and swamp cooler needs for wells within the "affected area" with chromium six levels above 0.02 ppb, the only conceivable option for interim replacement water is bottled water. (Draft CAO, Interim Replacement Water Supply, § 1(a), p. 9; DeWolfe Dec., ¶ 15(a).) There is no other alternative that could be implemented within two weeks. (*Id.*, ¶¶ 9-16.)

Even if PG&E could surmount the logistical constraints associated with providing such a large volume of bottled water to as many as 300 locations, it would not be feasible to provide bottled water that meets the Draft CAO's quality requirements. (*Id.*, ¶ 8.) Bottled water chromium concentrations are typically significantly greater than 0.02 ppb. (*Id.*, ¶ 8(a).) One study demonstrated that total chromium concentrations in bottled water are significantly higher than what would be permitted under the Draft CAO.[10] PG&E would not be able to monitor at the source whether the bottled water met the 0.02 ppb standard because (i) the bottled water industry does not report chromium six, or even total chromium, concentrations, and (ii) bottled water under one label often comes from different sources and may not be uniformly treated using the same technology. (*Id.*, ¶ 8(b).) Furthermore, the

---

[9] DeWolfe Dec., ¶ 3(a). Because the Draft CAO requires replacement water to be provided to users with wells below chromium six background levels, Mr. DeWolfe's team assumed that all wells within the "affected area," as defined in the Draft CAO, will require interim replacement water. (*Id.*) The Draft CAO does not provide PG&E with sufficient time for testing and analysis to determine the *exact* number of "impacted wells," as defined in the Draft CAO. (*Id.*)

[10] A peer reviewer of OEHHA's draft PHG for chromium six, Professor William Shotyk of Heidelberg University's Institute of Earth Sciences, cited a study that analyzed one hundred and thirty-two brands of bottled water. The study found as much as 1.72 ppb of total chromium, with a median of .082 ppb. Virtually all total chromium dissolved in water is chromium six.

logistical problems associated with testing the bottles after they leave the plant are insurmountable given that the water would come from different sources and would not have been uniformly processed. (*Id.*, ¶ 8(c).)

Bulk water delivery – that is, water trucked to water tanks at each location – is not a feasible option for interim replacement water because it would require at least six months to implement. (*Id.*, ¶15(b).) Bulk water also raises challenges related to locating a supply with sufficiently low chromium six concentrations and concerns with the chemical treatment necessary to maintain microbiological quality. (*Id.*, ¶ 9(a)-(b).)

The permanent water replacement requirements would require whole-house replacement water to one to three homes where the wells tested above background for chromium six[11] and to locations where wells show statistically significant increases in chromium six concentrations, even though the concentrations are still below background. (Draft CAO, Findings, § 29, p. 7-8.) It is impossible to determine how many wells would qualify for permanent replacement water because the Draft CAO does not define what is a "statistically significant" increase. (*Id.*) But assuming that roughly one-third of the wells within one mile of the plume would meet the threshold, PG&E would be required to supply whole-house replacement water to approximately 100 locations.[12] Whole-house replacement would require approximately 660,000 gallons per month. (DeWolfe Dec., ¶ 3(d).) The Draft CAO would require this permanent water supply within 75 days of the Draft CAO being issued. (Draft CAO, Permanent Water Supply, § 2(a), p. 9.) This is not feasible for the following reasons:

First, no permanent replacement water supply source could be implemented within 75 days. (DeWolfe Dec., ¶ 14-16.) Second, due to the advanced technologies that would be required to reduce chromium six below natural background levels in either the local groundwater supply or in Golden State Water Company's sources, PG&E probably would not be able to obtain the necessary permits from DPH. (*Id.*, ¶ 11(e).) Third, each technology analyzed by the Arcadis team raises significant technological and environmental issues.[13]

---

[11] Twelve domestic wells have tested above background for chromium six. However, PG&E has agreements to purchase nine of the affected properties, and two of the wells may be more than one mile from the plume boundary.

[12] DeWolfe Dec., ¶ 3(b). Chromium six concentrations in the Hinkley area wells are known to fluctuate over time in a nearly random pattern. (*Id.*) Thus, Mr. DeWolfe's team assumed that at any given time, well readings are increasing, decreasing, or stable with equal probability. (*Id.*)

[13] *Id.*, ¶¶ 11-13. In a letter addressed to Harold Singer dated June 24, 2011, David Loveday and Pauli Undesser of the Water Quality Association (the "WQA") commented on the Draft CAO. The WQA promotes sales of water treatment devices. According to the letter, the technologies "readily available" to address chromium six reduction include "reverse osmosis (using TFC or CTA membranes), distillation, strong base anion resin, and weak base anion resin." But the assertion that these technologies are "readily available" is entirely undermined by the next sentence of the WQA letter, which states: "However, California requires testing of such technologies to validate performance according to national standards and at this time, none of the best available technologies in a whole house format are [*sic.*] is tested and certified." Thus, none of the technologies is even close to being "readily available." Furthermore, none of the technologies referenced in the WQA letter has been proven to reduce chromium six concentrations to 0.02 ppb. (DeWolfe Dec., ¶ 16(a).) In a similar vein, on July 9,

Footnotes Continued on Next Page

An ion exchange system would require additional engineering advances because those currently on the market cannot achieve chromium six levels of 0.02 ppb. (*Id.*, ¶ 11(d).) Furthermore, the waste stream created by an ion exchange system, including brines, would likely be considered hazardous waste under federal law, requiring an appropriate disposal plan. (*Id.*, ¶ 11(b).)

A multi-pass reverse osmosis system faces similar constraints. Up to 75% of the water introduced into such a system will end up as toxic brine, while as little as 25 % will be usable as potable water. (*Id.*, ¶ 12(b).) Like an ion exchange system, reverse osmosis systems on the market are not designed to achieve the chromium six draft PHG and would, therefore, require additional engineering advances. (*Id.*, ¶ 12(e).) Nor is a reduction, clarification and filtration system a viable option; the techuology is simply too new and untested. (*Id.*, ¶ 13.)

The construction of a central treatment and distribution system would also be infeasible. (*Id.*, ¶ 14.) The design, environmental review, permitting and construction of such a facility would, obviously, take longer than 75 days to complete. (*Id.*) Furthermore, a central treatment system would not itself achieve the 0.02 ppb staudard. (*Id.*) The system would have to employ ion exchange and/or reverse osmosis to meet the chromiun six levels required by the Draft CAO, thereby raising all the issues described above that plague those technologies. (*Id.*)

For the same reasons bottled water or bulk water would not be feasible for an interim replacement water supply, they would also not be feasible as a permanent replacement water supply. (*Id.*, ¶ 15.)

The Draft CAO's failure to follow the state process through which technical and economic feasibility is considered prior to enacting cleanup requirements raises serious public policy questions with statewide impact.

B. It Is Not Even Possible To Reliably Test For Chromium Six In Hinkley At Levels Below 0.1 ppb As Required By The Draft CAO

PG&E retained Shawn Duffy, an expert in chromium laboratory testing, to determine whether laboratories can reliably detect and quantify for chromium six at the ultra-trace levels required by the Draft CAO. Mr. Duffy conducted a study and concluded that it is not possible to reliably detect and quantify for chromium six at the levels required by the Draft CAO. (Declaration of Shawn Duffy

---

2011, Robert Conaway sent an e-mail to Mr. Singer containing fourteen website links to various commercial water treatment device providers. Coincidentally, prior to receiving Mr. Conaway's e-mail a PG&E representative had already contacted six of the fourteen vendors identified by Mr. Conaway. Based on conversations with six of the vendors and a review of all the websites, PG&E has determined that all of these providers use the same technologies that are critiqued in Mr. DeWolfe's Declaration and this letter. Most significantly, none of the product websites claim they can remove chromium six down to 0.02 ppb. In fact, two of the vendors are not certified to provide treatment systems in California. Thus, these technologies do not provide a feasihle means to comply with the Draft CAO.

("Duffy Dec."), which accompanies this letter, ¶ 3.) Requiring a compliance level that is below the level at which laboratories can reliably detect or quantify would set a troubling precedent for cleanups statewide.

Mr. Duffy directed the collection of ten water samples from wells in Hinkley that had previously been reported as non-detect for chromium six at or near the existing project detection limit of 0.2 ppb. Mr. Duffy also directed the manufacture of performance evaluation samples containing known concentrations of total chromium and chromium six of 0.01 ppb, 0.02 ppb, 0.04 ppb, 0.06 ppb, 0.08 ppb and 0.1 ppb, as well as the generation of two double-blind field blanks. (*Id.*, ¶ 4.)

The study design samples were split into three identical sample containers and submitted to three California accredited laboratories: Truesdail Laboratory Inc. ("TLI"), Advanced Technology Laboratory ("ATL") and BCLab Inc. ("BCL"). (*Id.*) The laboratories were requested to report the chromium six results to a level of 0.02 ppb using a modified EPA method 218.6 (as required by the Draft CAO) and as low as possible for total chromium by EPA Method 200.8. (*Id.*, ¶ 4(g).)

The test results from the three California accredited laboratories demonstrated that at the reporting levels required by the Draft CAO, the laboratories frequently failed to accurately detect and/or quantitate chromium six using EPA method 218.6. (*Id.*, ¶ 6.) Specifically, the laboratories routinely failed to meet the standard laboratory fortified blank criteria (90-110% recovery) for the performance evaluation samples. (*Id.*, ¶ 6(a).) The laboratories repeatedly produced sampling results that were less than 90% or more than 110% of the known sampling result. (*Id.*) For example, the sample containing 0.06 ppb (stated to be the Limit of Quantitation ("LOQ")) was reported as 0.068 ppb, 0.071 ppb, and ND < 0.026 ppb by the three laboratories. (*Id.*, ¶ 5.) These results are 113%, 118%, and 0% of the known concentration. (*Id.*, ¶ 6(b).) In other words, not one laboratory met the standard laboratory fortified blank criteria (90%-110% recovery) for the very LOQ concentration that the Draft CAO would require PG&E to use on Hinkley samples.

In addition, the laboratories frequently failed to report the same (or even similar) results for the same Hinkley test samples. For example, for sample H-13-Q2 (a monitoring well) the laboratories reported: 0.091 ppb (BCL), 0.05 ppb (TLI), and ND <0.02 ppb (ATL). (*Id.*, ¶ 5.) Using these varying results for the exact same sample, BCL would require PG&E to supply whole-house replacement water, while TLI and ATL would not in light of the 0.06 ppb testing threshold set forth in Draft CAO. (Draft CAO, Interim Replacement Water Supply, § 1(a), p. 9, fn. 2.) These findings are critically significant since they clearly demonstrate that the three California-accredited laboratories often failed to produce results that were even close to one another. The failure to produce the same (or at least similar) sample results further demonstrates that the laboratories were not able to reliably quantitate for chromium six at the low levels required by the Draft CAO. (Duffy Dec., ¶ 6.)

Finally, Mr. Duffy requested the laboratories to analyze the same samples for total chromium. Chemists often use total chromium sample results as a double-check for chromium six sample results. Unfortunately, the total chromium results frequently failed to correlate, and often failed to approximate the chromium six results for the same samples. For example, the total chromium levels ranged from non-detect to as high as 0.354 ppb for the same sample, and rarely correlated or came close

to the chromium six results for the same sample. (*Id.*, ¶ 8(d).) In addition, none of the laboratories met the standard laboratory fortified blank criteria for total chromium. (*Id.*, ¶¶ 7(d), 8.) Thus, the total chromium testing provides further support for Mr. Duffy's conclusion that laboratories cannot reliably detect or quantify for chromium six at the ultra-trace levels required by the Draft CAO. (*Id.*)

C. The Draft CAO Requires Overly Broad Interim Water Replacement Requirements Followed By Inconsistent Permanent Water Replacement Requirements

The Draft CAO would require interim replacement water for drinking cooking and swamp cooler needs for any well above 0.02 ppb within one mile of the plume. The requirement to supply swamp cooler needs for hundreds of properties would require an enormous water source that could not be practically supplied by bottled water. (DeWolfe Dec., ¶ 8.) As discussed above, this requirement has no basis in science or the data. Nonetheless, interim water would be required from 14 days to 74 days after the Draft CAO is signed.

After 75 days, the Draft CAO would require permanent replacement water for all indoor domestic uses for any well above background or that shows a statistically significant increase (as yet undefined) in chromium six concentrations. It is impossible to tell how many wells would fall within this permanent replacement water requirement because it is undefined. However, it is clear that the number of wells above background or with increased chromium six concentrations is a smaller set of wells than all wells with chromium six concentrations above 0.02 ppb.

The interim and permanent replacement water requirements are not consistent. The interim requirement would require the installation of enormous infrastructure for only 60 days, followed by new and different infrastructure for a different set of wells after 75 days. These requirements are inefficient and unfair, and would not result in any added safety.

V. The Draft CAO Would Create Impossible Statewide Policy And Precedent

A. The Draft CAO Would Create An Impossible Statewide Standard Based On The Unfounded Statement That Chromium Six Above 0.02 ppb Poses An Immediate Health Risk

The Draft CAO would find that "hexavalent chromium in domestic wells above 0.02 ppb poses an immediate health risk to Hinkley residents through continued use of contaminated water including drinking, preparing foods and beverages, bathing or showering, flushing toilets, and other household uses resulting in potential dermal and inhalation exposures." (Draft CAO, Findings, § 26, p. 7.) There is no substantial evidence to support this finding. If adopted, the Lahontan Board would unnecessarily create alarm by proclaiming a danger without justification.

The Lahontan Board should use caution in passing judgment on the Hinkley groundwater given the statewide implications. Does the Lahontan Board intend to order all entities within its jurisdiction to cease providing drinking water with chromium six concentrations above 0.02 ppb? Does the Lahontan Board intend to suggest that the residents of Riverside (drinking water chromium six concentrations, 1.69 ppb), Davis (19 ppb), Victorville (9.5 ppb) and Apple Valley (9.2 ppb) are

encountering an "immediate health risk" every time they use municipal water supplies to drink, shower, and flush toilets? The potential for far-reaching, unintended consequences statewide is precisely why the State Board has ordered regional boards to defer to the agencies with expertise regarding drinking water safety, rather than formulating their own standards. (SWRCB Order 20005-007, at 6.)

B. Reliance On Draft Regulations Is Improper And Would Create Significant Statewide Uncertainty

In our system of governance, *draft* laws, regulations and standards are not enforceable. For example, in *Chaparral Greens v. City of Chula Vista*, 50 Cal.App.4th 1134, 1145 (1996), the petitioner argued that an agency was required to review all "applicable" regional plans, including plans in draft form. The Court of Appeal disagreed, holding that "[a] plan that is in draft form cannot be said to be nonetheless legally applicable, or enforceable, as to a particular project." (*Id.* at 1145 n. 7.) Similarly, in *Laurel Heights Improvement Assn. v. Regents of University of California*, 6 Cal.4th 1112, 1127 (1993), our Supreme Court refused to consider draft amendments to the California Environmental Quality Act Guidelines because the "amendments remain in draft form and have not been adopted." The State Board itself has refused to permit petitioners to rely on draft orders. (SWRCB Order 2010-0016, 2010 WL 2674817, p. 3 n. 2 ("Petitioners also rely on an unpublished draft Board order concerning the American River. Because a draft order has not been adopted by the Board, it does not constitute 'longstanding FAS precedent,' and Petitioners' reliance on it is misplaced").)[14] The Government Code itself forbids any agency from enforcing any regulation until such time as the rulemaking process is complete. (Gov. Code § 11340.5; an agency shall not "issue, utilize, enforce, or attempt to enforce any guideline, criterion, bulletin, manual, instruction, order, standard of general application, or other rule which is a regulation as defined in subdivision (g) of Section 11342,[15] unless [it] has been adopted as a regulation and filed with the Secretary of State pursuant to this chapter.")

The propriety of the Draft CAO, if adopted, may ultimately be tested in a court of law, which would not likely ratify a standard for replacement water based on a *draft* PHG. Consistent with *Chaparral*, *Laurel Heights* and SWRCB Order 2010-0016, a court is unlikely to permit the State and Lahontan Boards to treat a draft PHG as a substitute for "applicable federal, state, and local drinking water standards." (*Id;* Water Code § 13304.)

---

[14] Federal courts have also rejected evidentiary offers based on drafts. *See, e.g., Idaho Rivers United v. F.E.R.C.*, 189 Fed.Appx. 629, 637 (9th Cir. 2006) ("The petitioners' argument is not built on solid ground because it merely relies on the draft [Biological Opinions]" and recognizing the agency can change positions between the draft and final documents); *Idaho Farm Bureau Federation v. Babbitt*, 58 F.3d 1392, 1403-1404 (9th Cir. 1995) (Reliance by agency on draft report improper because it foreclosed consideration of public comment on that draft).

[15] A regulation is defined as a "rule, regulation, order, or standard of general application . . . adopted by any state agency to implement, interpret, or make specific the law enforced or administered by it, or to govern its procedure, except one which relates only to the internal management of the state agency." (Gov. Code, § 11342(b).)

VI. Conclusion

PG&E respectfully requests that you exercise your discretion and not issue the Draft CAO in its current form. If there are any remaining legal, technical or toxicology questions, we would be happy to provide additional information at your request. In addition, we are also informed that DTSC, which has a full staff of toxicologists and is responsible for a number of other sites with potentially contaminated drinking water, and DPH, which also has toxicologists with expertise in drinking water safety, have both expressed views on the proper use of the proposed chromium six PHG in determining drinking water safety. We urge you to confer with the experts in both of these sister agencies and with the San Bernardino County Health Department if you have any further questions regarding toxicology or the potential statewide impacts of the Draft CAO. If after consulting with the experts at the other state agencies you determine that it is appropriate to proceed with the Draft CAO, we urge you to recognize the significant policy questions implicated by the Draft CAO and refer its issuance to the Lahontan Board, where a full public hearing may be had and comments received.

Very truly yours,

Thomas C. Wilson

EXHIBIT A

(to the comment letter)

From: Harold Singer [mailto:HSinger@waterboards.ca.gov]
Sent: Thursday, May 26, 2011 11:48 AM
To: Lauri Kemper
Cc: drew@jdp-law.com; DAG6@pge.com; MCKd@pge.com; Kim Niemeyer; Reed Sato
Subject: Whole-house Replacement Water

Ms. Kemper

Based on recent statements at Water Board meetings and documents in the Water Board files, I understand that the Water Board Prosecution Team is developing an order to address a directive from the Water Board to evaluate the need for whole-house replacement water for Hinkley residents affected by the historic chromium discharge from the PG&E Compressor Station. Given that the groundwater in this area is well below the current MCL for total chromium and that a public health goal is still in draft form, any order would likely have significant technical, legal and policy considerations. Because of this, it is important to provide affected and interested public and PG&E an opportunity to respond to the factual, legal and technical assertions in the order and to develop a record to support any Water Board action relative to this issue. Therefore, I am requesting that if the Prosecution Team proposes an order for whole-house replacement water that this order be transmitted to me in draft form along with a mailing list of potentially affected and interested persons. The transmittal should also include any evidence that the Prosecution Team is relying on to support the draft order. I will transmit the proposed draft order and supporting information to potentially affected and interested persons, along with instructions for providing comments.

thank you
harold singer

7/11/2011

EXHIBIT B

(to the comment letter)

November 2, 2009

Dr. Michael Baes
Pesticide and Environmental Toxicology Branch
Office of Environmental Health Hazard Assessment
California Environmental Protection Agency
1515 Clay St., 16th floor
Oakland, California 94612
Attn: PHG project

Re: Public Health Goal for Hexavalent Chromium in Drinking Water

Dear Dr. Baes:

I am writing to you on behalf of the Hexavalent Chromium Panel (Panel) of the American Chemistry Council regarding the draft Public Health Goal (draft PHG) for Hexavalent Chromium (Cr (VI)) issued on August 20, 2009, by the Office of Environmental Health Hazard Assessment (OEHHA). The Panel consists of members who have an interest in the production or use of hexavalent chromium in North America. After reviewing available toxicity data, OEHHA published the draft PHG of 0.06 parts per billion as being "protective against all identified toxic effects from both oral and inhalation exposure to hexavalent chromium that may be present in drinking water."[1] As discussed below, the Panel urges OEHHA to await additional research findings anticipated in 2010 before finalizing the draft PHG.

In July 2009, Toxicology Excellence for Risk Assessment (TERA) convened a Science Advisory Board (SAB) to provide guidance on research to investigate the potential mode(s) of action (MOA) of hexavalent chromium based on the US EPA Guidelines for Carcinogen Risk Assessment (2005). Based on TERA's scientific recommendations, The Hamner Institute for Health Sciences was commissioned to conduct research on five key areas integral to assessing the MOA(s) for chromium: 1)

[1] The draft PHG is available online: http://www.oehha.ca.gov/water/phg/pdf/Cr6PHGdraft082009.pdf

a 90-day drinking water study; 2) genomic studies on tissues from the 90-day study; 3) pharmacokinetic modeling; 4) in *vivo* mutation analysis and 5) high data content *in vitro* studies. [2] This research is expected to be completed in 2010.

In a memo dated October 23, 2008, from Dr. David Berry, Senior Toxicologist with the Human and Ecological Risk Division of the Department of Toxic Substances Control, to Dr. Jeff Wong, Chief Scientist of the Department of Toxic Substances Control, the Hamner research program was recognized as critical in addressing the mode of action of chromium and the studies should be "prerequisites to any revisions to the OEHHA public health goal for Cr^{+6}." In issuing the draft PHG prematurely, however, OEHHA has failed to consider the anticipated mode of action research. (*See* Appendix A) We agree with DTSC comments regarding how important it is to use current scientific principles and recent advances such as incorporating mode of action are preferable to using outdated default assumptions. We urge OEHHA to await additional research findings anticipated in 2010 before finalizing the draft PHG.

The Panel appreciates your consideration of these comments. If you have questions, or would like to discuss the studies sponsored by the Panel, please contact me at (703) 741-5614 or kristy_morrison@americanchemistry.com.

Sincerely yours,

Kristy L. Morrison

Kristy L. Morrison
Manager, Hexavalent Chromium Panel
Chemical Products & Technology Division
American Chemistry Council

[2] More information on the SAB can be found online here: http://www.tera.org/Peer/Chromium/Chromium.htm

APPENDIX A

Linda S. Adams
Secretary for
Environmental Protection

Department of Toxic Substances Control

Maureen F. Gorsen, Director
8800 Cal Center Drive
Sacramento, California 95826-3200

Arnold Schwarzenegger
Governor

MEMORANDUM

TO: Jeff Wong, Ph.D.
Chief Scientist
Department of Toxic Substances Control
1101 I Street, 25th Floor
Cal/EPA
Sacramento, CA 95814

FROM: David L. Berry, Ph.D. David L. Berry
Senior Toxicologist
Human and Ecological Risk Division
8810 Cal Center Drive
Sacramento, CA 95836-3200

DATE: October 23, 2008

SUBJECT: Hexavalent Chromium Public Health Goal

The HERD was asked to provide review and comment on the "Confidential Pre-Release Draft, Public Heath Goal for Hexavalent Chromium in Drinking Water", prepared by the Pesticide and Environmental Toxicology Branch of the Office of Environmental Health Hazard Assessment (OEHHA) dated September 2008.

General Comments

The toxicity of hexavalent chromium [Cr^{+6}] has been known for at least 180 years and the carcinogenicity in humans of inhaled Cr^{+6} was first reported in the United States in 1948. The inhalation carcinogenicity of hexavalent chromium has been well documented in numerous human epidemiological investigations. The carcinogenicity of hexavalent chromium via the oral route has been a subject of speculation since the late-1960s and a lifetime bioassay in rodents conducted by the National Toxicology Program (2007) with Cr^{+6} in drinking water found an increased incidence of tumors in treated animals.

Human health risk assessments are based on the understanding of two basic components: toxicokinetics and toxicodynamics. Beginning in the late 1980s, the

physiologically-based pharmacokinetic models [PBPK] developed in the pharmaceutical industry for interspecies scaling in drug development began to be applied to other aspects of pharmacology and toxicology including human health risk assessment. These models address the first component (toxicokinetics) and allow consideration of the applied dose and the effective dose at the target organ after taking into account the absorption, distribution, metabolism, and excretion of a compound. These methods depend upon understanding of the route of exposure, partitioning of the compound across biological barriers and compartmentalization in various organs/tissues to scale effective dose across species. The PBPK models mathematically scale dose from a laboratory animal to humans with more precision than the traditional allometric (body surface area) methods promulgated at 22 CCR 12703; PBPK scaling is currently used by the World Health Organization, National Academies of Science, U.S. Environmental Protection Agency [2005], Health Canada, the U.S. Air Force, and the European Union.

The methods used by OEHHA to draft the Public Health Goal for Hexavalent Chromium in Drinking Water are consistent with the methods used in the development of all other Public Health Goals that have been issued by OEHHA. The OEHHA methods are the default protocols that were outlined in the 1985 California Department of Health Services *Guidelines for Carcinogen Risk Assessments and Their Scientific Rationale.* The 1985 methods were to be updated every 5 years, but as of today's date there have been no subsequent revisions or edition of those guidelines. The 1985 default methods ignore recent advances in interspecies scaling and evaluation of the mode of action (MOA or toxicodynamics) of various carcinogens (e.g., formaldehyde) that are utilized routinely by other regulatory agencies in derivation of toxicity factors for a wide range of materials. In the present case, the default methods employed by OEHHA are highly conservative and over-estimate substantially the carcinogenic potency of ingested hexavalent chromium. The reader may appreciate the fact that there are serious consequences associated with overly conservative analyses that fail to account for a carcinogenic MOA.

Most regulatory guidance is based on 'scientific principles' that provide the foundation for that guidance. Situations can occur where strict adherence to default regulatory guidance may violate (or significantly depart from) the basic principle(s) that the guidance was supposed to support. In this regard, it is standard OEHHA practice to assume the animal data can be described by a linear dose-response relationship [LMS], but no data (other than reference to the results of standard short-term tests for genotoxicity) to support that assumption were provided. As written, there is no *a priori* reason to accept the OEHHA assumption that Cr^{+6} –induced tumors of the gastrointestinal tract in rodents can be described most accurately with a statistical model that is linear at low-dose. It is important to remember the difference between the basic principles versus the default assumptions made in the 1985 guidance and to realize that the guidance should be modified in order to be consistent with current scientific principles (and not vice-versa).

Specific Comments

1. The proposed PHG of 0.06 part per billion [ppb] or 60 nanograms/L is well below any method detection limit for Cr^{+6} in drinking water to be found in any commercial or academic analytical laboratory. This has significant implications for warnings required under Proposition 65. Assuming Title 22 is revised at Section 12707(b)(4), and the OEHHA default risk assessment is applied, all potable water supplies with analytically-detectable levels of Cr^{+6} will be required to warn, if not implement mitigation measures. The DPH web site provides the number of domestic water supplies with detectable levels of Cr^{+6}; there are over 2,300 such cases in the State of California and the vast majority of the Cr^{+6} detections in water is associated with naturally occurring sources – including the State Rock (serpentinite) that contains upwards of 1,700 ppm total Cr.

2. The PHG for Cr^{+6} was based on an oral cancer "slope factor" of 0.6 mg/kg-day^{-1}, which OEHHA derived from the data for small intestinal tumors in male mice seen after lifetime ingestion of Cr^{+6} in drinking water [NTP, 2008]. OEHHA then used an occupational study with an inhalation slope factor [510 mg/kg-day^{-1}] derived for industrial conditions [chromium ore refinery] and modeled an exposure assessment for Cr^{+6} exposure during showering. Using a inhalation slope factor based on metal fumes from ore refining with temperatures (1275-1400° C) sufficient to generate chromium fume [Othmer, 2001] extrapolated to a 38°C domestic shower cannot be justified in that the OEHHA-calculated shower Cr^{+6} exposure far exceeds the empirical exposure to Cr^{+6} in shower water droplets [Paustenbach et. al., 2003]. Chromium ore processing conditions and the generation of metal fumes are simply not relevant to domestic showering conditions.

3. The accuracy of the OEHHA discussion of Cr distribution in tissues and organs can be improved by incorporating the PBPK model of chromium in the rat [O'Flaherty, 1996] and its extension to human beings [O'Flaherty et. al., 2001]. Discussion of *in vitro* chromium partitioning in erythrocytes may not be relevant to *in vivo* studies of chromium administered p.o., regardless of the form of chromium. Based on human epidemiological investigations, tumors of the lymphatohematopoietic [blood and lymph] system have not been reported. Use of PBPK modeling for risk assessment is encouraged in EPA's 2005 guidance and is especially important in understanding interspecies extrapolations given the divergent findings in rats and mice and the recognized differences in the human GI including a more acidic stomach.

4. The discussion of Cr kinetics, both trivalent and hexavalent, is incomplete. O'Flaherty [1996] cites relevant papers that are not included in the PHG document that provide an in-depth discussion of the differences in uptake between Cr^{+3} and

Cr^{+6} and that the rapid uptake of chromium in the erythrocyte [as Cr^{+6}] is followed by reduction [to Cr^{+3}]. The kinetics indicate that Cr^{+6} is eliminated differently than Cr^{+3} but that the half-life of Cr^{+6} is greater than a day which is remarkable given the rapid reduction of Cr^{+6} to Cr^{+3}. The loss of Cr^{+6} from the erythrocyte and subsequent uptake into liver and bone marrow suggests that not all Cr^{+6} is reduced to Cr^{+3} as it is distributed into various tissue compartments and eliminated in the urine and feces. The simplistic models proposed in PHG Figures 1 & 2 add nothing to understanding of the toxicokinetics of either Cr^{+6} or Cr^{+3} [see Figure 1, O'Flaherty, 1996].

5. The document places significant weight on the Borneff et al. [1968] study where a single dose level of 500 mg/L of potassium chromate was administered to male and female mice in a three generation study. The fact that only a single dose level was examined precludes any identification of a dose-response relationship, a key piece of evidence required in any assessment of causality. During the course of the investigation, an ectromelia epidemic affected both control and treated groups with significant loss of animals. The reduced numbers of animals severely limits the power of this investigation for both potential adverse reproductive outcomes and potential carcinogenic response. While the Borneff study may be historically interesting, the study is qualitative at best. Only the more recent, audited chronic drinking water study with Cr^{+6} that was conducted by the NTP [2008] can be relied upon for any potential rule making.

6. Inspection of the data generated in the subchronic toxicity study by the NTP [2007a] yields a NOAEL of 15 ppm for mice [1.6 mg/kg-day combined sexes, see pg 27]. In the OEHHA summary, a LOAEL of 1.6 mg/kg-day is reported for the NTP [2007a] study [see pg 76]. The identified LOAEL is actually a NOAEL.

7. The subchronic NTP study [NTP, 2007a] using F344 rats and B6C3F1 mice with sodium dichromate provided the range finding data for the subsequent 2 year chronic bioassay of. Based on these studies, doses of 14.3, 57.3, 172 or 516 mg/L [male and female rats and female mice] and 14.3, 28.6, 85.6, and 257.6 mg/L [male mice] were administered to animals for two years. Non-neoplasic, treatment-related lesions were not observed in male rats. Treatment-related liver toxicity was observed in female rats [fatty involution and chronic inflammation] that increased with increasing dose. Mice [male and female] survived the treatment and the only non-neoplastic lesions observed were diffuse hyperplasia in the duodenum. The NTP study reported no non-neoplastic lesions in the oral cavity of the rat, but no data from the subchronic study were collected for the oral cavity. The NTP reported the results of an additional review of the oral cavity tissues specifically to look for non-neoplastic lesions following observation of the tumors. As the mice failed to develop lesions of the oral cavity and rats are known to be more sensitive to oral cavity tumors than mice (according to NTP's historical data for all chemicals tested),

oral cavity tumors are apparently species-specific and/or a consequence of repeated exposure and associated with the potent chemical oxidizing properties of dichromates and repeat local tissue damage. It is noteworthy that there has not been any increase in oral cavity tumors among workers exposed to Cr^{+6} in any of the numerous epidemiology and clinical studies (e.g., Bloomfield and Blum, 1928; Baetjer, 1950; Gross and Kosch, 1943; Langard and Norseth, 1975; Mancuso and Hueper, 1951). The human data are relevant as chromium workers in historical conditions had ample opportunity for significant oral cavity exposures to inhaled Cr^{+6} in fume, concentrated particulate or aerosol forms [see #14].

8. The NTP two year chronic bioassay of sodium dichromate in F-344 rats and B6C3F1 mice found that rats developed increased incidence of papilloma and carcinoma formation in the oral mucosa and tongue. In mice, the tumors were adenomas and carcinomas found in the ileum, jejunum, and duodenum. These effects were dose-related with the highest dose yielding the greatest tumors per number of animals, only the highest dose yielded increased tumors - except in the case of the male mice. The HERD did not review the actual NTP data and restricted the present review to only the findings presented in the PHG document. The OEHHA combined the respective mouse and rat papillomas, adenomas, and carcinomas to yield a greater tumor response per animal, a statistical method that results in an increased "slope factor" or carcinogenic potency. The high dose tumor effect was also associated with the highest animal mortality and these doses were associated with development of hyperplasia in these tissues in the subchronic studies [NTP, 2007a].

9. The spectrum of tumors indicates that only those tissues with initial Cr^{+6} contact were affected by the treatment. For the rat, the initial tissues contacted by the dichromate in drinking water were the tongue and the oral mucosa. No tumors were observed in the rat forestomach or small intestine. Unlike the rat, the tumors in the mouse were found in the small intestine, an organ with greater residence time and increased opportunity for Cr^{+6} direct tissue contact. Tumors in other organs (including the forestomach) were not detected in the mice, a unique finding for such a chronic study. Although the study was not designed to allow for investigation of the Cr^{+6} MOA, it is clear that tumor development is related to local inflammation and hyperplasia in the target tissue. One candidate MOA concerns the chronic local inflammation induced by the chronic tissue damage inflicted by high-dose chromate and the role of reactive oxygen species. Since the NTP concluded that the lesions in the duodenum in mice were seen in concert with local regenerative hyperplasia, it appears that the highest dose induced overt tissue damage (in addition to the presence of chronic inflammation) and that the tumors arose as a result of that damage. Given that the subchronic investigations revealed hyperplasia in the rat oral mucosa and in the mouse small intestine, the tumor response is very similar to

the promotional response in epithelial cells induced by phorbol diesters. All of these features point to the conclusion that ingested doses of Cr+6 that are insufficient to produce local irritation, tissue damage, inflammation and regenerative hyperplasia are also without additional carcinogenic risk.

10. In the discussion of the results on page 52 of the PHG document the authors mix a human study with the rodent studies. The comparing and contrasting of rodent and human data occurs later in the text.

11. In all of the high dose groups, decreased water consumption and body weight were noted. This observation is consistent with the high dose being unpalatable or due to the effects of systemic poisoning by high-dose sodium dichromate. Thus, only at exposures where either the water would be refused by consumers due to foul taste or at doses sufficiently high to induce gastric or other distress could a practical or measurable increase in carcinogenic risk be measured.

12. The OEHHA weight of evidence discussions are based on human epidemiologic studies of hexavalent chromium considered occupational exposures where the route of administration was primarily via the inhalation pathway. Thirty-one studies were chosen where digestive tract [primarily stomach] tumors were reported. None of the studies cited addressed the oral route contribution to the potential tumor incidence and none of these studies focused on consumption of hexavalent chromium. However, in all of the studies that were cited, tumors of the respiratory tract were observed. In a meta- analysis of chromium exposure and cancer mortality [Cole and Rodu, 2005], at least 84 papers were reviewed relating hexavalent chromium exposure to 10 causes of cancer mortality [lung, stomach, prostate, kidney, central nervous system, leukemia, Hodgkin's disease, lymphatohematopoietic cancers, all cancer and all causes]. Based on the meta-analysis, there is only a weak association between inhaled Cr^{+6} and lung cancer; moreover, there was no significant association of inhalation Cr^{+6} exposure to any of the seven other cancers evaluated [note that the Cole & Rodu (2005) study was excluded by OEHHA].

13. There are limited epidemiological investigations of hexavalent chromium exposure via the ingestion route. Six papers were reviewed that addressed one area in China where a documented exposure to Cr^{+6} in the drinking water occurred. Zhang and Li (1987) evaluated potential relationships between drinking water exposure to hexavalent chromium and the incidence of various cancers and mortality. The OEHHA analysis concluded that the study showed significant increases in stomach and lung cancer and OEHHA reported (Table 8) a summary of epidemiological investigations and concluded there was a relationship between occupational exposure to chrome and increased stomach cancer. OEHHA then calculated rate ratios for the incidence of stomach tumors for these 19 investigations that ranged

from 0.95 to 5.0. However, analyses of these same data by Cole and Rodu [2005] indicated there were no significant increases in stomach or GI tumors associated with Cr^{+6} ingestion and only a very weak association between Cr^{+6} exposure and lung tumors.

14. Based on the tumor data for the F344/N rats and the B6C3F1 mice [NTP, 2008], the mouse appears to be more sensitive to the hexavalent chromium treatment. Hyperplasia was observed in the two year study in the mouse forestomach in a dose-dependent pattern strongly implicating regenerative hyperplasia as a mode of action for the small intestine tumors. Species-specific variability in GI parameters are critical to understanding the relationship between the observations in mice and relevance to low concentration exposures in humans. In contrast, oral cavity tumors are rare in the F344/N rat. Additionally, one cannot expect concordance between the site(s) of tumor development between rodents and humans given the great species-specific variability.

15. OEHHA employed the U.S. EPA BMDS model to fit a dose response curve for tumor incidence in the male B6C3F1 mice and extrapolated from the lower bound to the origin. The combined adenoma and carcinoma data for duodenum or small intestine data were used to generate a mean and lower-bound estimate of the Cr^{+6} exposed mice (ED_{10} and LED_{10}) associated with a ten percent increase in tumors. OEHAA also calculated a dose response curve for female B6C3F1 mice for tumors of the small intestine. Presumably, although not explicitly stated, OEHHA used the data from the male mice for determination of an oral slope factor due to the lower tolerated dose for the male mice.

16. The BMDS generated dose associated with a 10 percent increase in tumor incidence was scaled to a human equivalent dose based on body weight to the 4/3 power [$TD = \alpha \times BW^{4/3}$; allometric scaling]. Subsequently, the data were evaluated using the linearized multistage model [LMS] to develop a slope factor for the oral potency of hexavalent chromium. The OEHHA used the LMS to estimate an oral potency factor for male B6C3F1 mice of 0.6 mg/kg-day^{-1} and calculated an oral slope factor of 0.8 mg/kg-day^{-1} for female mice. The NTP [2008] data clearly illustrate evidence for carcinogenicity in the small intestine of the mouse and oral cavity of the rat. However, the MOA for Cr^{+6} tumorigenicity in the gut is not clear from the NTP data and it has not been addressed by the OEHHA. The tumors of the gastrointestinal tract appear to be related to regenerative hyperplasia [NTP, 2007a] in the target tissue followed by progression to benign tumors and finally carcinoma. This is highly indicative of a promotional mechanism that begs the discussion of a threshold dose-response relationship. The NTP studies cannot provide a basis for the MOA to direct a technical basis for the proper selection of a model to evaluate

the carcinogenic potency of Cr^{+6}. The default application of the LMS model makes the assumption that there is no threshold or dose below which there is no tumor response or increased carcinogenic risk. The LMS model is highly conservative and may greatly over-estimate the potency of Cr^{+6} via the oral route. Without understanding the MOA, it is not possible to assign a rigorous dose-response relationship or develop a justifiable oral slope factor.

17. Evaluation of noncarcinogenic effects associated with dichromate ingestion were based on the classical NOAEL/LOAEL approach based on six selected studies. The NTP [2007a] study was chosen as the study given the most weight for a determination of an RfD for oral Cr^{+6}. The OEHHA chose an uncertainty factor of 1000 [10x for using a LOAEL, 10x for extrapolation between species, and 10x to protect sensitive species]. The default 10x interspecies scaling factor is a practice in regulatory assessments where PBPK is not available or has been rejected. In the present situation, PBPK models are available and if utilized would reduce the uncertainty and increase the accuracy of the Cr^{+6} health risk assessment.

18. The carcinogenic potency discussion of the inhalation route of exposure on pages 79 to 89 would be more appropriate in a separate PHG document for establishing an inhalation toxicity factor. There are published studies (Crump et al. 2004; Gibb et al. 2000; Park et al. 2004; Park & Stayner, 2006) that could be used, or directly provide updated inhalation unit risk factors for Cr^{+6} rather than the current OEHHA slope factor that is based on dated information. The more recent studies were used by OSHA for their 2006 rulemaking.

19. OEHHA Appendix A Carcinogenic Threshold. It is not clear how does this discussion contributes to the understanding of a threshold-based dose-response relationship for ingested dichromate. Clearly, the NTP studies do not indicate the absorption of hexavalent chromium is a consequence of over burdening the ability of the GI tract's capacity to reduce Cr^{+6} to Cr^{+3}. Given the tumor response in the rat and mouse, the most likely threshold effect is the ability of the hexavalent chromium to elicit dose-dependent overt tissue damage, chronic inflammation and local regenerative hyperplasia.

20. OEHHA Appendix B Borneff et al. (1968). As noted above, the Borneff study has many limitations due to confounding factors such as ectromelia and lack of a dose-response relationship. The study is qualitative and the results have not been reproduced and should be viewed as anecdotal. The NTP chronic two-year bioassay is a full GLP investigation with rigorous quality control and assurance and pathology review. The NTP is a much stronger investigation and should be the primary basis for any assessment of carcinogenic risk associated with ingested Cr^{+6}.

21. OEHHA Appendix B Helicobacter Hypothesis. There is no information on the presence of *Helicobacter pylori* in the F344/N or the B6C3F1 animals used in the NTP bioassay. While *H. pylori* affects a significant human population and it may influence the stomach pH, it has not been shown experimentally to affect the ability of the stomach to reduce Cr^{+6} to C^{+3} or to affect absorption of chromium across the gut. Appendix B is speculative, lacks relevance to developing the PHG and it should be eliminated from the document as it is speculation.

Recommendations

The NTP bioassays do not address the MOA of hexavalent chromium via the ingestion pathway. Given the lack of data on the Cr^{+6} MOA in the gut, generation of a PHG for hexavalent chromium at this time may be premature as it is not possible to assign a dose-response relationship – other than the default OEHHA assumptions and methods used since 1985. Additional investigations are indicated and should be considered before public release of the PHG value or its documentation.

Subsequent to the 2007 publication of the National Toxicology Program report on the lifetime carcinogenicity bioassay in rats and mice, the Hamner Institute for Health Sciences (Research Triangle Park, North Carolina) initiated pilot studies to update and revise the rodent: human Cr^{+6} PBPK model and to investigate the Cr^{+6} mode of action (MOA) at the genomic level in order to support rigorous human health risk assessments. At the present time, those pilot studies are only just beginning as well as re-evaluation of the 14 day acute and the 90 day subchronic studies in rats and mice upon which the dose selection for the lifetime bioassay was based. The goal of the preliminary studies are to gain sufficient data to inform the design of protocols designed to define more accurately the risk assessment approach which should be taken with ingested Cr^{+6}. It may well be that at the high doses used in the NTP bioassay, that the properties of chemical oxidation are responsible for the upper gastrointestinal tract tumors, whereas, it may be that a genotoxic MOA may be operational in the small bowel where chronic inflammation may be the initiating event. The hexavalent chromium MOA has simply not been established.

The Hamner Institute is willing to cooperate with Cal/EPA, provided sufficient funding is identified to support collection of the genomic and pharmacokinetic parameters that are necessary to determine the MOA and to scale properly the delivered Cr^{+6} dose to target tissues properly from rodents to humans. Using Magnetic Image Resolution (carried out at the University of North Carolina), the Hamner Institute has been able to measure and quantify the relative contributions of Cr^{+3} and Cr^{+6} in the target tissues. The Hamner Institute already has in hand the original O'Flaherty PBPK model for chromium in rats. As of today's date, there is no PBPK model for mice.

These studies are prerequisites to any revisions to the OEHHA public health goal for Cr^{+6}. In the absence of the empirical data, it is speculative to suggest values other than the default 60 nanogram/L PHG are equally, more or less protective of the public health. Taking the most recent Hamner Institute re-evaluation of the pathogenesis and genomics of formaldehyde-induced nasal carcinomas in rodents as an example, the minimum budget required to measure the genomic changes and to develop and implement the PBPK model for one (1) species was $870,000 (direct and indirect costs combined) over 2 years. Thus, one can anticipate a total cost for collection of the required mode of action data and refinement of the PBPK models for rats and mice would be ~$1.8 M over 2 years.

Relative Source Contribution and Bioavailability

The more common commercially important forms of hexavalent chromium include: the oxide (CrO_2), chromyl chloride, ammonium dichromate, potassium dichromate, sodium dichromate, potassium chromate, sodium chromate, potassium chlorochromate, silver chromate, barium chromate, strontium chromate and lead chromate. Their solubilities in water varies from the completely insoluble lead salt to the very soluble oxide. Chromic oxide (the trivalent Cr_2O_3) predominates in ores (e.g. chromites) from which metallic chromium is produced is completely insoluble in water. Thus, one cannot generalize materials as "hexavalent" chrome; rather, the exact form of the element must be taken into account in human health risk assessments - a situation not unlike that applied to other inorganic elements (e.g., arsenic).

It is common practice to take into account xenobiotic exposures incident to bathing, showering and all other domestic uses of potable water (e.g., toilets) when establishing a maximum contaminant level (MCL) for inorganic (e.g., 22 CCR 64431) and organic (e.g., 22 CCR 64444.5) materials. The contribution to total exposure associated with volatile organics like perchloroethylene, carbon tetrachloride, trichloroethylene and related materials has been quantified and it can be substantial (up to 50%of lifetime average dose in the case of chloroform) (McKone, 1987; McKone and Knezovich, 1991). However, none of the common chromium compounds (either as present naturally in ores or as refined commercially important forms) are volatile.

The fact none of the chromium compounds are volatile begs the question of exposure during use of potable domestic water. Given the lack of volatility and the relative water solubility, the only physical form in which a potassium or sodium chromate can be present in water would be as an aerosol. The OEHHA analysis appears to assume the bioavailability of a dilute chromium aerosol is equivalent to that of chromium fume that can arise during welding, cutting or plating or ore processing. All of the temperature

conditions under which chromium fume or aerosols are generated are substantially greater than those encountered in routine household use of potable water for bathing.

There are no empirical data to substantiate the presence of chromium aerosols (regardless of oxidation state) in drinking water intended for domestic consumption or other incidental use. Therefore, it is not possible to assign a relative source contribution for chromium present during bathing in calculation of potential risk to the public health. No reference to peer-reviewed empirical data concerning bathing and showering contributions to total daily chromium dose was provided in the materials submitted for review. Most important, it is necessary to divide chromium and its inorganic compounds into a number of chemical-specific groupings, each with a specific MCL based on the available exposure, toxicological and epidemiological evidence.

References

Baetjer, A.M. 1950. Pulmonary carcinoma in chromate workers. II. Incidence on the basis of hospital records. Arch. Ind. Hyg. Occup. Med. 2: 505-516.

Bloomfield, J.J. and Blum, W. 1928. Health hazards in chromium plating. Pub.Health Repts. 43: 2330-2347.

Borneff, J., Engelhardt, K., Greim, W., Kunte, H., Reichert, J. 1968. Carcinogens in water and soil. XXII. Mouse drinking water experiments with 3,4-benzopyrene and potassium chromate. Arch. Hyg. Bakteriol. 152(1):45-53.

Clewell, H.J., Andersen, M. E. and Barton H.A. 2002. A consistent approach for the application of pharmacokinetic modeling in cancer and noncancer risk assessment. Environ. Hlth. Perspect. 110(1):85-93.

Cole, P. and Rodu, B. 2005. Epidemiologic studies of chrome and cancer mortality: A series of meta-analyses. Regul. Toxicol. Pharmacol. 43:225-231.

Crump, C., Crump, K., Hack, E., Luippold, R., Mundt, K, Liebig, E., et al., 2003. Dose-reponse and risk assessment of airborne hexavalent chromium and lung cancer mortality. Risk Anal. 23(6):1147-1163.

Fisher, J.W. 2000. Physiologically Based Pharmacokinetic Models for Trichloroethylene and Its Oxidative Metabolites. Environ. Hlth. Perspt. 108(S2): 265-273.

Gibb, H.J., Lees, P.S., Pinsky, P.F. and Rooney, B.C. 2000. Lung cancer among workers in chromium chemical production. Am J. Ind. Med 38(2):115-126.

Gross, E. and Kosch, F. 1943. Lung cancer in the chromate dye industry. Arch. Gewerbepath. Gewerbehyg. 12: 164-170.

Manciso, R.F. and Hueper, W.C. 1951. Occupational cancer and other health hazards in a chromate plant: A medical appraisal. I. Lung cancers in chromate workers. Ind. Med. Surg. 20: 358-363.

McKone, T.E. 1987. Human exposure to volatile organic compounds in household tap water: The indoor inhalation pathway. Environ. Sci. Technol. 21: 1194-1201.

McKone, T.E. and Knezovich, J.P. 1991. The transfer of trichloroethylene (TCE) from a shower to indoor air: Experimental measurements and their implications. J. Air Waste Mgmt. 41(6): 832-837.

National Academy of Sciences. 2006. Assessing the Human Health Risks of Trichloroethylene: Key Scientific Issues. The National Academies Report in Brief, Washington, D.C.

National Toxicology Program, 2007a. NTP Technical Report on the Toxicity Study of Sodium Dichromate Dihydrate Administered in Drinking Water to Male and Female F344/N Rats and B6C3F1 Mice and Male BALB/c and an3-C57B/6 Mice, toxicity Report Series, Number 72, January 2007.

National Toxicology Program, 2008. NTP Technical Report on the Toxicology and Carcinogensis Studies of Sodium Dichromate Dihydrate in F344/N Rats and B6C3F1 Mice (Drinking Water Studies), NTP TR546, NIH Publication No. 08-5887.

O'Flaherty, E.J. 1966. A physiologically based model of chromium kinetics in the rat. Toxicol. Appl. Pharmacol.138: 54-64.

O'Flaherty, E.J., Kerger, B. D., Hays, S.M., and Paustenbach, D.J. 2001. A physiologically based model for the ingestion of chromium (III) and chromium (VI) by humans. Toxicol. Sci. 68:196-213.

Othmer, K. 2001. Chromium and Chromium Alloys, In: Encyclopedia of Chemical Technology 3rd Edition, Vol. 6 pp. 54-60. John Wiley & Son, Inc., New York

Park, R.M., Bena, J.F., Stayner, L.T., Smith, R.J., Gibb, H.J. & Lee, S.J. 2004. Hexavalent Chromium and Lung Cancer in the Chromate Industry: A Quantitative risk Assessment. Risk Anal. 24(5):1099-1108.

Park, R. M. & Stayner, L.T., 2006. A Search for Thresholds and other Nonlinearities in the Relationship between Hexavalent Chromium and Lung Cancer. Risk Anal. 26(1): 79-88.

Paustenback, D.J., Finley, B.L., Mowat, F.S. and Kerger, B.D. 2003. Human health risk and exposure assessment of chromium (VI) in tap water. J. Toxicol. Environ. Health, Part A. 66:1295-1339.

Proctor, D.M., Otani, J.M., Finley, B.L., Paustenbach, D.J., Bland, J.A., Speizer, N., and Sargent, E.V. 2002. Is hexavalent chromium carcinogenic via ingestion? A weight-of-evidence review. J. Toxicol. Environ. Hlth. Part A 65:701-746.

Proctor, D.M., Panko, J., Liebing, E., Paustenbach, D.J. 2004. Estimating historical occupational exposure to airborne hexavalent chromium in a chromate production plant: 1940 – 1972. J. Occup. Environ. Hyg. 1:752-767.

U.S. EPA. 2005. Guidelines for Carcinogen Risk Assessment. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC. EPA/630/P03/001B.

World Health Organization. 2005. Chemical-Specific Adjustment Factors for Interspecies Differences and Human Variability: Guidance Document for Use in Dose/Concentration-Response Assessment. World Heath Organization, Geneva.

World Health Organization. 2008. Principles of Characterizing and Applying Physiologically-based Pharmacokinetic and Toxicokinetic Models in Risk Assessment. Harmonization of Approaches to the Assessment of Risk from Exposure to Chemicals. World Health Organization, Geneva.

Zhang, J. and Li, S. 1987. Cancer mortality in a Chinese population exposed to hexavalent chromium in water. J. Occup. Environ. Med. 39:315-319.

Reviewed by: Calvin C. Willhite, Ph.D.
Staff Toxicologist

Stephen M. DiZio, Ph.D.
Supervising Toxicologist

DECLARATION OF SHAWN DUFFY

DECLARATION OF SHAWN DUFFY

I, Shawn Duffy, declare:

1. I am employed by CH2MHill as a Senior Chemist. My resume is attached to this Declaration as Exhibit A. Pacific Gas and Electric Company ("PG&E") has engaged CH2MHill to assist with issues surrounding the chromium plume in Hinkley, California. I have been working on chromium chemistry and testing related issues for PG&E since 2004. I was asked to lead a team tasked with assessing the reliability and variability associated with the analysis for hexavalent chromium in a typical commercial laboratory setting with a theoretical Method Detection Limit ("MDL") of 0.02 µg/L as required by the draft Cleanup and Abatement Order No. R6V-2011-0005A1 (the "Draft CAO").

2. I have more than seventeen years of experience in the field of analytical chemistry. On behalf of PG&E, I have been responsible for technical oversight of laboratory analytical activities, quality control and quality assurance, and data validation for the Hinkley remediation project. I perform on-site field audits for the evaluation of sampling protocols and sampling team performance, as well as conduct laboratory audits to ensure quality control and compliance with the project's quality assurance program plan. For the last seven years, I have worked with the subcontracted laboratory's analysts and quality control personnel to ensure that the analytical techniques are consistent and correct according to the applicable methods for hexavalent and total chromium analyses to achieve the current reporting levels of 0.2 ug/L for Cr(VI) and 1.0 ug/L for total chromium.

3. My opinion is that:

(a) A reporting level below 0.10 µg/L for Cr(VI) is beyond the capacity of the typical commercial laboratory to detect and accurately quantitate using EPA Method 218.6.

(b) A reporting level at or below 0.10 µg/L for total chromium is beyond the capacity of the typical commercial laboratory to detect and accurately quantitate using EPA Method 200.8.

4. As part of the process of forming my opinions, I conducted the following study:

(a) Ten water samples were collected from Hinkley wells that had previously been reported as non-detect or just above the detection limit; with a 0.2 µg/L detection limit for Cr(VI) and as non-detect with a 1.0 µg/L detection limit for total chromium. Six groundwater monitoring wells and four water supply wells were sampled. In addition, three of the ten wells were randomly selected and field duplicate samples were collected from the three locations. Each sample was filtered through two, certified clean, inline 0.45 micron filters and collected into a single certified clean container using clean hands/dirty hands methodology, a sampling method that prevents contamination of the samples by limiting the contact of the sample and sample containers with any object that could cause contamination. The Cr(VI) sample aliquot for each well was carefully homogenized – that is, shaken after being filtered into a single large container – and an aliquot split into three certified clean bottles pre-preserved with a buffer solution and submitted to three separate California-approved Environmental Laboratory Accreditation Program("ELAP") laboratories that claim to achieve an MDL of 0.02 µg/L or less for Cr(VI) using CFR Part 136, Appendix B procedures. The total chromium (all forms of chromium) portion for each of the samples listed above was also carefully homogenized and split into a certified clean bottle pre-preserved with nitric acid and submitted to the same three California-approved ELAP laboratories that were asked to analyze the total chromium by EPA Method 200.8 and report the results to the lowest possible MDL using CFR Part 136, Appendix B procedures, which has been labeled as a technically-flawed procedure for assessing level of

detection by the U.S. EPA in the *Revised Assessment of Detection and Quantitation Approaches, EPA-821-B-04-005*. (Exhibit B).

(b) All samples were collected under formal chain of custody and placed into coolers, on ice, immediately after sampling. The samples were then transferred to each lab's courier for delivery to the contracted project laboratories. All of the sample containers used for this study were supplied by one laboratory, and the samples were labeled using generic alpha-numeric identifiers that were similar to the double-blind sample identifiers used for the quality control samples included in this study.

(c) The study included positive, negative and duplicate quality control samples, all of which were submitted as double-blinds to each of the contracted project laboratories.

(d) Six of the quality control samples were performance evaluation ("PE") samples containing known concentrations of Cr(VI) from 0.01 to 0.10 ppb. Environmental Resource Associates ("ERA") was contracted to prepare and certify three sets of double-blind PE samples for total chromium and Cr(VI) at concentrations of 0.01 µg/L, 0.02 µg/L, 0.04 µg/L, 0.06 µg/L, 0.08 µg/L and 0.10 µg/L. ERA prepared the PE samples and transferred them into pre-preserved bottles provided by the same laboratory used for the samples, and using deionized water that was demonstrated to be free of total chromium and Cr(VI) to a MDL of <0.02 µg/L (also provided by the same laboratory). ERA was unable to certify the final concentration of the PE samples due to their ultra-trace level concentrations.

(e) The PE samples were delivered to the project team in the field on the afternoon of the field sampling event to insure that all the study samples were shipped together to each of the three contracted project laboratories. Upon receipt of the PE sample by the project

team, fictitious sample labels were affixed to the PE sample containers and the sample identifiers entered on the chain of custody.

(f) Two field-derived blanks were also included. An aliquot of the same chromium-free deionized water provided to ERA was transferred to identical sample containers, labeled and sent as one of the study samples. A second blank was prepared at one of the ten ground water sample locations, an aliquot of the same chromium-free deionized water was put through the same in-line 0.45 micron filters as the samples, split into preserved containers, labeled and sent as one of the study samples.

(g) The samples were sent to three California accredited ELAP laboratories: Truesdail Laboratory, Inc. ("TLI"), Advanced Technology Laboratory ("ATL") and BCLab, Inc. ("BCL"). The three laboratories are listed on the California Department of Public Health web page, subgroup code 103.310, which is specific to the certification for EPA Method 218.6 for Cr(VI). The laboratories were requested to report the Cr(VI) results to a level of 0.02 µg/L using EPA Method 218.6 with a modified buffer, and as low as possible for total chromium by EPA Method 200.8.

5. The three laboratories reported results for total chromium and Cr(VI) as follows:

Trace-Level Chromium Study – June 2011

Sample ID	Sample Type	ATL Cr(VI) µg/L	ATL Cr (initial analysis) µg/L	ATL Cr (Second analysis) µg/L	BC Lab Cr(VI) µg/L	BC Lab Cr µg/L	TLI Cr(VI) µg/L	TLI Cr µg/L
H-01-Q2	0.10 - PE STD	0.096	ND < 0.02	0.117	0.031	ND < 0.85	0.103	0.146
H-02-Q2	0.08 - PE STD	0.082	ND < 0.02	0.096	ND < 0.026	ND < 0.85	0.085	0.396
H-03-Q2	0.06 - PE STD	0.068	ND < 0.02	0.077	ND < 0.026	ND < 0.85	0.071	0.170
H-04-Q2	0.04 - PE STD	0.048	ND < 0.02	0.061	ND < 0.026	ND < 0.85	0.031	0.096
H-05-Q2	0.02 - PE STD	ND < 0.02	ND < 0.02	0.046	ND < 0.026	ND < 0.85	0.046	0.069
H-06-Q2	0.01 - PE STD	ND < 0.02	ND < 0.02	0.035	ND < 0.026	ND < 0.85	0.035	0.140
H-07-Q2	Supply Well	0.41	0.23	0.546	0.34	ND < 0.85	0.429	0.646
H-08-Q2	Supply Well-Dup	0.4	0.21	0.545	0.33	ND < 0.85	0.421	0.635
H-09-Q2	Supply Well	ND < 0.02	ND < 0.02	0.034	ND < 0.026	ND < 0.85	ND < 0.02	0.099
H-10-Q2	Supply Well	ND < 0.02	ND < 0.02	0.029	ND < 0.026	ND < 0.85	ND < 0.02	0.061
H-11-Q2	Supply Well-Dup	ND < 0.02	ND < 0.02	0.042	ND < 0.026	ND < 0.85	ND < 0.02	0.055
H-12-Q2	Supply Well	ND < 0.02	ND < 0.02	0.023	ND < 0.026	ND < 0.85	ND < 0.02	0.510
H-13-Q2	Monitoring Well	ND < 0.02	ND < 0.02	0.190	0.091	ND < 0.85	0.050	0.190
H-14-Q2	Monitoring Well	ND < 0.02	ND < 0.02	0.118	ND < 0.026	ND < 0.85	ND < 0.02	0.140
H-15-Q2	Monitoring Well	ND < 0.02	ND < 0.02	0.130	ND < 0.026	ND < 0.85	ND < 0.02	0.278
H-16-Q2	Monitoring Well	0.23*	ND < 0.02	0.255	0.82*	ND < 0.85	0.422 *	0.304
H-17-Q2	Monitoring Well	ND < 0.02	ND < 0.02	0.103	ND < 0.026	ND < 0.85	ND < 0.02	0.112
H-18-Q2	Monitoring Well	ND < 0.02	ND < 0.02	0.266	ND < 0.026	ND < 0.85	ND < 0.02	0.303
H-19-Q2	Monitoring Well-Dup	ND < 0.02	ND < 0.02	0.218	ND < 0.026	ND < 0.85	ND < 0.02	0.276
H-20-Q2	DI –Blank	ND < 0.02	ND < 0.02	0.207	ND < 0.026	ND < 0.85	0.079	0.288
H-21-Q2	Field-Blank	ND < 0.02	ND < 0.02	0.197	ND < 0.026	ND < 0.85	0.163	0.354

* Cr(VI) samples received in lab at pH < 2

6. The results from the three certified laboratories show the inherent variability of analytical results for samples at these ultra-trace level concentrations, and demonstrate the inability to report quantitatively accurate results for both EPA Method 218.6 Cr(VI) and EPA Method 200.8 total chromium:

(a) The laboratories for the Hinkley project are required to analyze a Cr(VI) low level laboratory fortified blank spiked at the RL (currently 0.2 µg/L) and have routinely met the standard laboratory fortified blank criteria of 90 to 110% recovery (in other words, the labs test a sample with a known concentration and report a result between 90% and 110% of the known concentration). However, only four of the fifteen PE sample results in this study met the same recovery criteria.

(b) Two of the three laboratories, ATL and TLI, reported EPA Method 218.6 Cr(VI) results for the 0.06 µg/L concentration PE standard with recoveries of 113% and 118%, respectfully. The third laboratory reported the results as non-detect at 0.026 µg/L or 0% recovery.

(c) Two of the three laboratories, ATL and TLI, reported EPA Method 218.6 Cr(VI) results for the 0.04 µg/L concentration PE standard with recoveries of 120% and 77%, respectfully. The third laboratory, BCL, reported the results as non-detect at 0.026 µg/L or 0% recovery.

(d) Two of the three laboratories, ATL and BCL, reported EPA Method 218.6 Cr(VI) results for the 0.02 µg/L concentration PE standard as non-detect at 0.02 and 0.026 µg/L, respectfully, or 0% recovery. The third laboratory, TLI, reported the results at 0.046 µg/L, a 240% recovery.

7. At the required level for the MDL of 0.02 µg/L and the required level for the RL of 0.06 µg/L from the Draft CAO, false positive blank detections are a serious problem that can cause erratic non-reproducible trace-level sample results:

(a) ATL analyzed the samples twice for total chromium. During the first round of analysis, the calibration blank showed a high base line and the PE standards were all reported as non-detect. On second analysis, the calibration blank base line was significantly lower and ATL reported all the PE standards with a high bias, 230% at 0.02 µg/L to 117% at 0.10 µg/L.

(b) The 0.01 µg/L PE blank, which was below the MDL for all labs, was reported by one lab at a Cr(VI) concentration of 0.035 µg/L, a clear false positive.

(c) The deionized water blank and the field blank were also reported by one lab at a Cr(VI) concentration of 0.079 and 0.163 µg/L, more clear false positives.

(d) All three blanks were reported by two labs as detects for total chromium at levels between 0.035 µg/L and 0.354 µg/L, an order of magnitude difference.

8. The three labs could not produce inter-laboratory results that would meet the precision (that is, the measurement of reproducibility) or accuracy (the amount of agreement between a measured value and the "true" value) of the methods:

(a) BCL reported the EPA Method 218.6 Cr(VI) results as non-detect for all the PE standards except one. That one standard (0.10 µg/L) was recovered at 31%.

(b) BCL was also unable to detect total chromium, using EPA Method 200.8, below a MDL of 0.85 µg/L, and all samples were reported as non-detect.

(c) One sample (H-13-Q2) had results reported for Cr(VI) ranging from ND <0.02 µg/L to 0.091 µg/L by the three laboratories. While the total chromium was reported ranging from ND < 0.02 µg/L to 0.190 µg/L.

(d) TLI was biased high on all the PE standards for both EPA Method 218.6 Cr(VI) and EPA Method 200.8 total chromium, the result of what I suspect was lab contamination. The high bias included positive detects of 0.140 µg/L, 0.288 µg/L, and 0.354 µg/L for the three blanks and a minimum recovery of 146% for the total chromium.

(e) ATL analyzed the samples for total chromium twice. During the first round of analysis, the calibration blank showed a high base line and the PE standards were all reported as non-detect. During the second round, the calibration blank base line was significantly lower and ATL reported all the PE standards with a high bias, 230% at 0.02 µg/L to 117% at 0.1 µg/L.

I declare under penalty of perjury under the laws of the State of California that the foregoing is true and correct, and that this Declaration was executed on July 9, 2011, at 6:30 PM, Redding, California.

Shawn Duffy

EXHIBIT A

Shawn P. Duffy

Project Chemist

Education

B.A., Biology, Humboldt State University, 1988

Distinguishing Qualifications

- More than 17 years experience in analytical and environmental chemistry
- More than 25 years experience in the field of science
- Experience in a wide range of sampling techniques and situations
- Biological, physical, and chemistry field, office, and lab experience.

Relevant Experience

Project Chemist/CH2M HILL; Mr. Duffy is a chemist with more than 17 years of experience in the field of analytical chemistry. He provides oversight for an analytical laboratory budget of one - two million dollars annually. He provides senior oversight for a CH2M HILL staff of 8 chemists and data managers, providing chemistry, data management, and data validation for the PG&E Program. Mr. Duffy manages subcontracts with laboratories, provides statements of work, manages purchase orders, approves invoices, provides schedules, oversees laboratory corrective actions, and provides a point of contact between the laboratories and CH2M HILL staff. Performs onsite field audit for evaluation of sampling protocols and sampling team performance and conducts laboratory audits to ensure quality control and compliance with the project's quality assurance program plan.

As the lead CH2M HILL chemist on the PG&E Hinkley and Topock Sites; Mr. Duffy has, For the last seven years, worked with the subcontracted laboratory's analysts and quality control personnel to ensure that the analytical techniques are consistent and correct according to the applicable methods for hexavalent and total chromium analyses to achieve the current reporting levels. And with a primary goal of continuing to meeting the forever increasing regulatory requirements and providing cost effect, accurate, low level analyses.

Representative Projects

Pacific Gas and Electric Projects

PG&E Topock Compressor Station - **March 2004 to present.** Provides chemistry support for all aspects of the Topock Compressor Station Project including the Ground Water Monitoring Program, Compliance Monitoring Program, River Monitoring Program, East Ravine Groundwater Investigation, AOC4 Time Critical Removal Action, RCRA facility investigation/remedial investigation, Background Study, and Interim Measures 1, 2, and 3. Oversight of data validation and senior review of data and data quality reports; provides statements of work and coordinates with CH2M HILL contract administrators to supply purchase orders for lab services, and provide information to project managers and field crews to verify that project work is following the work plan and QAPP. Assist with onsite sample collection and sampling coordination.

Shawn P. Duffy

IM3- As the lead chemist on the design and startup of a field laboratory facility at the Topock IM3 site Mr. Duffy has researched, implemented and trained field personnel for the characterization of hexavalent and total chromium at Topock to meet discharge permit requirements. Investigated and evaluated equipment necessary for IM3 onsite lab and provide design requirements. Provides onsite and on-call chemistry support and training of facility operators

PG&E Hinkley Compressor Station - **March 2004** to **present**. Provide chemistry support for all aspects of the Hinkley project, including the Performance Monitoring Program, Nitrate Monitoring Program, Interim Pumping Program, Injection Monitoring Program, and Background Study. Oversees data validation, senior data review, and data quality reports; provides statements of work and coordinates with CH2M HILL Contract Administrators to supply purchase orders for laboratories services; and provides information to project managers and field crews to verify that project work is following the work plan and QAPP.

Other PG&E clean-up and investigative projects include – Antioch investigation, Colusa site investigation and remediation, Humboldt site characterization, Merced MGP, Red Bluff MGP, Selma site remediation, Shell Pond investigation, Wildcat remediation, Woodland MGP, and the Pipeline Hyrotest.

Federal Projects –

Beale Air Force Base, California; Assisted project chemist with data validation and reviewed: evaluated laboratory data, interacted with the laboratories, and provided data quality information to the project chemist and project manager. Drill rig oversight for sampling and monitoring well installation.
Lennar Mare Island; Mare Island, California; Assisted project chemist with data validation. Reviewed and evaluated laboratory data, interacted with the laboratories, and provided data quality information to the project chemist and project manager.
Hill Air Force Base, Utah; Assisted project chemist with data validation. Reviewed and evaluated laboratory data, interacted with the laboratories, and provided data quality information to the project chemist **and** project manager.
Hickam Air Force Base, Utah; Assisted project chemist with data validation. Reviewed and evaluated laboratory data, interacted with the laboratories, and provided data quality information to the project chemist **and** project manager.
Massachusetts Military Reservation (MMR), Massachusetts; Assisted project chemist with data validation.

Experience Prior to Current CH2M HILL

Biologist/Group Lead; Klamath Wildlife Resources, Redding, California; Lead a Spotted Owl Inventory Survey in Plumas National Forest. Identified, photographed, and marked locations of species by map/GPS. Conducted an Aquatic Amphibian survey in the Lassen National Forest, Almanor Ranger District and lead a Northern Goshawk survey in Lassen National Forest, Hat Creek Ranger District. Using a standard compass, GPS, aerial photographs, and topographic maps to run transects thru suitable habitat establishing survey points

On-call Biologist; ENPLAN; Redding, California; Acted as a biological observer during construction of new wetlands for mitigation purposes. Planted seeds gathered from nearby wetlands. Perform monthly hydro- and photo monitoring.

Technical Support Manager/Customer Support Manager; Michrom BioResources, Auburn, California. Provide technical and customer support for an HPLC (High Performance Liquid Chromatography) instrument manufacturer, LC and LCMS (Liquid Chromatography Mass Spectrometry) supplies, and applications technologies company. Performed column packing, column quality assurance/quality check (QA/QC), instrument QA/QC (including HPLC, Auto-sampler, and other OEM instrumentation), and in-house service of instrumentation parts. Provide technical support to in-house and extemal clients for HPLC/LCMS, columns, and chemistry related issues. Travel to customer sites for installation, training, service and equipment demonstrations. Generate company-wide marketing materials including catalogs, calendars, brochures and magazine advertisements. Part of a three-member team responsible for worldwide sales and technical documents. Advise clients on new and improved LC/MS methodologies and applications.

Hydrologic Technician; U.S. Bureau of Reclamation; Shasta Lake City, California. Install, maintain, and calibrate continuous monitoring satellite telemetry sites. Member of a sampling teams collecting various sample types (soil, water, fish, biological/climactic); team members from the U.S. Geological Survey, U.S. Bureau of Reclamation, U.S. Fish and Wildlife Service, California Department of Fish and Game, and CH2M HILL. Organized /summarized data for official documents and publications. Project Manager and investigation team member for the NCAO Shasta Lake Limnology Study. Member of the Interagency Technical Team for the Anadromous Fish Screen Program. Converted historic database information into a Microsoft Access database and advised the Mid Pacific Regional Data Center on the construction of a Microsoft Access regional database.

Biological Science (Fisheries) Technician; U.S. Bureau of Reclamation; Shasta Lake City. Monitor water quality in the regional rivers and reservoirs as part of the Water Quality Branch of the Environmental and Natural Resources Division. Collected and analyzed water samples to monitor levels of Cu, Cd, and Zn from the Spring Creek water shed (Iron Mountain Mine - EPA Hazardous Waste Site). Interfaced with contract laboratory personnel to ensure compliance to project specific holding times, instrumentation methodologies, and reporting limits. Member of the Shasta Lake Limnology Study team (NCAO/Denver).

Team Leader; North State Resources; Redding, California. Perform Forest Inventory for Six-Rivers and Shasta-Trinity National Forests. Located field plots using aerial photos and topographic maps. Identified and recorded data on the size and condition of trees, the species and cover of understory vegetation, ground cover and various site attributes, habitat types, and forest health indicators. Conducted a Vegetation Survey contracted from the U.S. Forrest Service (USFS).

Sample Custody Group Leader/Supervisor/ GC Chemist/Extraction Chemist; CH2M HILL Quality Analytical Laboratory; Redding, California. (Sample Custody) Train and supervise sample custodians. Maintain client relationships and ensure the flow of information between the client's project team and the laboratory personnel. Manage over-the-counter projects and filled in as project manager for Client Service Personnel. Provide all project specific information necessary for sample analysis and subsequent reporting of analytical data. Maintain custody of

the samples within the laboratory and assisted clients with information about analytical procedures and results. Order stock and. Subcontract analyses that required work to be performed at another laboratory. Evaluate and revise laboratory Standard Operating Procedures. Interpret and implement DOT regulations for shipping preserved containers. (GC chemist) Analyze Gas, BTEX, Herbicides, and PNAs. Interpret raw data and providing report information. (Extraction chemist) Train existing and entry-level personnel in environmental sample extraction procedures. Inventory, order, and prepare stock standards and solutions. Perform specialized extraction procedures. Routinely used perform different extraction methods to comply with different program requirements including; SW846, CLP, Calif LUFT, Alaska and Wisconsin Diesel Range Organics (DRO) and TCLP extractions for semivolatile and volatile organic preparations. Member of CH2M Hill Laboratory Health & Safety Committee and confined space entry team.

Assistant Entomologist/Field Inspector; Shasta Mosquito Abatement District; Redding, California. Identify adult and larval forms of the mosquito population for state reports. Conducted inspections on the effectiveness of control measures, lethal dose level (LDL) studies, and used maps and aerial photos to identify and locate mosquito sources.

Lab Assistant; Humboldt State University, Life Sciences Department; Arcata, California.

Lab Assistant; Shasta College, Life Sciences Department; Redding, California.

Professional Development

Certifications

Certified Motorboat Operator,

40 Hr. *HAZWOPER* training,

"Working at a Watershed Level" Training

Instrumentation

HACH Spectrometers, Horiba Multiparameter Unit, Various Ground Water Pumps, Hydro lab, D.O. Meter, Sontek Velocity Meters, HPLC, Gel permeation instrument (ABS model 1000), Continuous liquid/liquid extractors, Zero headspace TCLP extractors (Millipore), CP Thermister, Turbidimeter, Graphite furnace, Sparger units (Tekmar), Sonic Disrupters, Analytical balances, CAM extractors, pH meter, GPS, GCs (Varian models 6000, 3400, 3400 STAR, 3600 and 3700), GC Autosamplers/Detectors - FID, ECD, PID, IR, Fluorometer, and UV Instrumentation (models 8000 and 8010), Nitrogen delivery systems, Kuderna Danish Apparatus,

EXHIBIT B

EPA

Revised Assessment of Detection and Quantitation Approaches

October 2004

Engineering and Analysis Division
Office of Science and Technology
Office of Water (4303T)
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460

EPA-821-B-04-005

October 2004

Disclaimer

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.

Table of Contents

FOREWORD ..... *i*

CHAPTER 1: INTRODUCTION 1-1
- 1.1 Background
- 1.2 Clause 6 Settlement Agreement Requirements
- 1.3 EPA's Approach to Conducting this Assessment
- 1.4 Peer Review of the Agency's Assessment
- 1.5 Proposal and Request for Public Comments
- 1.6 Terminology used in this Document

CHAPTER 2: OVERVIEW AND HISTORY OF DETECTION AND QUANTITATION LIMIT APPROACHES 2-1
- 2.1 Currie's Call for Standardization
- 2.2 Development of the MDL and ML as Practical Embodiments of Currie's Proposal
- 2.3 Other Detection and Quantitation Approaches

CHAPTER 3: ISSUES PERTAINING TO DETECTION AND QUANTITATION 3-1
- 3.1 Analytical Chemistry Approaches to Detection and Quantitation
- 3.2 CWA Regulatory Issues Affecting Detection and Quantitation
- 3.3 Statistical Issues

CHAPTER 4: EVALUATION CRITERIA 4-1
- 4.1 Criterion 1
- 4.2 Criterion 2
- 4.3 Criterion 3
- 4.4 Criterion 4
- 4.5 Criterion 5
- 4.6 Criterion 6
- 4.7 Consensus Principles

CHAPTER 5: ASSESSMENT 5-1
- 5.1 Detection Limit Approaches
- 5.2 Quantitation Limit Approaches

CHAPTER 6 FINDINGS AND NEXT STEPS 6-1

APPENDIX A: LITERATURE SEARCH REGARDING DETECTION AND QUANTITATION LIMIT APPROACHES A-1

APPENDIX B: COMPUTATION OF DETECTION AND QUANTITATION LIMITS B-1

APPENDIX C: EXAMPLE CALCULATIONS C-1

Foreword

EPA has assessed current procedures for determining the sensitivity of test methods and their application to Clean Water Act (CWA) Programs. The assessment was required by a settlement agreement with the Alliance of Automobile Manufacturers, *et al.* We announced the availability of our preliminary assessment for public comment on March 12, 2003. This assessment discussed statistical, chemical, and regulatory issues related to detection and quantitation and different approaches to detection and quantitation. The Agency has revised the preliminary assessment Document to incorporate public comment on that assessment.

In a related action on March 12, 2002, we proposed to revise EPA's method detection limit (MDL) definition and procedure, and codify our minimum level (ML) procedure. The MDL and ML, respectively and in order of increasing magnitude, are the EPA's embodiment of a detection and a quantitation limit.

In this revised assessment, we have:

- Explained why and how we conducted this assessment (Chapter 1),
- Identified relevant concepts to include in the assessment (Chapter 2 of this document),
- Identified issues that may be relevant to the assessment from an analytical chemistry, statistical, or regulatory perspective (Chapter 3),
- Used six criteria to evaluate the ability of each procedure or concept to support activities under the Clean Water Act (Chapter 4),
- Assessed how well each concept meets the evaluation criteria (Chapter 5),
- Summarized our findings and discussed next steps (Chapter 6), and
- With real-world data and several different procedures, calculated and compared detection and quantitation limits, and evaluated the theoretical and practical limitations of each concept (Appendices).

Public comment on the preliminary assessment and the proposed regulatory revisions expressed many divergent views that conflicted with the proposed revisions. Commenters noted that: (1) the MDL does not adequately address analytical variability or systematic error (bias); (2) the MDL does not always achieve a one percent (1%) false positive rate; (3) EPA should provide better guidance on the intended use of the MDL and ML in compliance reporting; and (4) the MDL and ML are not appropriate for all applications in CWA programs. Several commenters expressed support for two alternatives to the MDL and ML that were submitted by a laboratory association and the U.S. Geological Survey, respectively. Although none of the alternative procedures recommended by commenters fully satisfied EPA's needs under the CWA, several procedures contain steps, such as blank correction, that EPA believes warrant further consideration. There was no agreement among commenters as to which of the competing alternatives or revisions to adopt. Commenters suggested that we work together to discuss mutual concerns and possible solutions rather than proceed with the proposed revisions. We agree and recognize that these concerns provide a strong starting point for a continued dialog with stakeholders.

Based on this new information, it is clear that there is a broad interest in improving current procedures and uses, but no consensus for a specific procedure or procedures has emerged among the laboratory, industry, regulatory or regulated communities. In addition, EPA sees merit in alternative procedures suggested by commenters; however, none of these completely satisfy EPA's needs. Thus, we believe that it is appropriate to withdraw the March 2003 proposed revisions, take final action on the 2003 assessment to complete the terms of the settlement agreement, and obtain additional stakeholder

input. In a *Federal Register* notice published on September 15, 2004 [69 FR 55547], we announced that a neutral party is exploring the feasibility of a process by which a broad group of stakeholders would work together to define and address concerns about the way detection and quantitation limits are calculated and used to support CWA programs. This stakeholder process would include stakeholders representing constituencies such as citizens, environmental organizations, permit writers, regulators and regulated industries. We trust that this stakeholder process will address the wide variety of views held by stakeholders and lead to recommendations for possible improvements to current EPA procedures and/or use of alternative procedures.

To facilitate open discussion and consideration of issues, we have made every effort to ensure that this Revised Assessment Document does not prejudge the result of a future stakeholder process. We look forward to further stakeholder participation in this process.

Chapter 1
Introduction

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 Spectrometry* (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 nse of an independent peer review to evaluate the Agency's assessment (Section 1.4), and EPA's March 2003 publication of and request for comment on the February 2003 assessment, and a related proposal concerning potential changes to detection and quantitation limit procedures approved for use under the Clean Water Act (Section 1.5). A brief discussion of the terminology used in this document is provided in Section 1.6.

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 (since amended to November 1, 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 (Note: since amended to November 1, 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 petitioners 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.

In March 2003, EPA published an Assessment Document dated February 2003, and requested comments on the assessment and additional data to support continued evaluation of detection and quantitation limits. Three stakeholders commenting on the assessment also offered to provide EPA with data that would substantiate their views or aid EPA in further evaluating detection and quantitation procedures. These data are further described in Sections 1.3.2.7 - 1.3.2.8 and Section 1.3.3 below.

Although the petitioners 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 1.3.3 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/μL (or μg/mL). Each solution was injected into the GC/MS in triplicate with the mass spectrometer threshold set to zero, 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 Manufacturers 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 (Gihbons *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 μg/L for mercury and selenium on the low end, and from 2.0 and 1000 μg/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 American Council of Independent Laboratories Data

The American Council of Independent Laboratories (ACIL) is a trade association representing independent, commercial scientific and engineering firms. Its members are professional services firms engaged in testing, product certification, consulting, and research and development. On behalf of its membership, ACIL submitted comments on EPA's proposal. To substantiate their comments, ACIL provided EPA with data summary tables consisting of blank analyses used to calculate detection limits. The data provided were performed by a single laboratory using Method 200.7 for five analytes. Because only blank sample analyses were available, not all detection and quantitation limit procedures could be assessed using the data. However, comparisons of the detection limit procedures submitted by ACIL and the US Geological Survey were performed based on these data and are discussed in Appendix C. In addition, because blanks were analyzed approximately two to three times per week, a comparison of long-term to short-term variability was also performed using these blank data. ACIL also submitted an alternative procedure for estimation of a detection limit, which is summarized in sect. 2.3.3 of this document.

1.3.2.8 U.S. Geological Survey Method Detection Limit Data

To assist EPA's assessment of their long-term MDL (LT-MDL) procedure, the US Geological Survey (USGS) provided data from blank sample analyses. These data represented a combination of 78 metals, methods and matrices, and were analyzed approximately twice per month. Unlike the blank data provided by ACIL, these blanks were collected in the field, and, therefore, include more sources of variability. As with the ACIL data, it was not possible to assess all detection and quantitation limit

procedures using the blank data set because some procedures require use of samples spiked at one or more concentrations. The LT-MDL procedure is summarized in sect 2.3.4 of this document.

USGS also submitted spiked sample results along with the blank data. These spikes were limited to a single concentration, and did not sufficiently characterize the region of interest to allow for full evaluation of detection and quantitation levels.

1.3.3 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 available, 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 a brief rationale for using or excluding the data from this assessment, additional discussion is in Appendix B.

After EPA published the February 2003 Assessment Document for comment, ACIL submitted data as described in Section 1.3.2.7 above, and three commenters offered to provide EPA with additional data that would enhance EPA's assessment. EPA requested the data offered by each of these organizations, but received a response from only two of the three (Laucks Testing Laboratories and USGS). After evaluating the data, EPA determined that the data from Laucks Testing Laboratories was not useful because it was incomplete. The Laucks data unfortunately did not include the data from extraction to detection which is needed to compare detection and quantitation approaches. Most of the data sent by USGS was useful and is described in Section 1.3.2.8.

Table 1. Data Sets Suggested by Petitioners and Commenters

Dataset Source and Year	Analytes and technology	EPA's Use of Datasets
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
ACIL 2002-2003	Metals by ICP/AES (200.7)	Used in this assessment and described in Section 1.3.2.7
USGS 2002-2003	Metals by ICP/MS and GFAA	Used in this assessment and described in Section 1.3.2.8
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

Dataset Source and Year	Analytes and technology	EPA's Use of Datasets
EPRI 1994	Al, Be, Tl 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 because the maximum number of replicates (5) in the dataset, is less than the minimum number required (7) to calculate an MDL. Samples spiked with low levels of Aroclors exhibited average recoveries >500%, across 10 laboratories.
Laucks Testing Laboratory 2003	Mercury, 2,4-Dinitrophenol, Hexachlorocyclo-pentadiene, 4-Nitrophenol (OSW 8270)	Not used in this assessment because the dataset was incomplete. Data included only calibration data and not the extraction to detection data needed to compare det/quant procedures.

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 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 Assessment Document. Where appropriate, EPA revised that Assessment Document to reflect specific suggestions and comments offered by the peer reviewers. The revised version of the Assessment Document, reflecting peer reviewer comments, was completed in February 2003, and made available through a public notice on March 12, 2003 (see section 1.5 below). 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 were provided in the public docket supporting the Agency's March 2003 assessment.

1.5 Proposal and Request for Public Comments

On February 28, 2003, the EPA Administrator signed two notices for publication in the Federal Register. These notices fulfilled EPA's obligations under Clause 6(a) of the Settlement Agreement and were published in the Federal Register on March 12, 2003.

The first of these notices announced the availability of EPA's assessment of detection and quantitation procedures that are applied to analytical methods used under the Clean Water Act. It also announced that results of the assessment could be found in the "Technical Support Document for the Assessment of Detection and Quantitation Concepts" (EPA 821-R-03-005, February 2003), requested public review and comment on the assessment. The full text of this notice was published at 68 FR 11791, March 12, 2003.

The second notice requested comment on proposed revisions to the detection and quantitation definitions and procedures at 40 CFR part 136. The proposed changes were based on the assessment and on stakeholder comments received over the years. The full text of this notice was published at 68 FR 11770, March 12, 2003.

1.5.1 Summary of Changes Proposed in March 2003

EPA proposed a number of technical and editorial changes to the definitions specified at 40 CFR 136.2 and to the procedure specified at 40 CFR 136, Appendix B. A detailed description of those changes can be found in the March 12, 2003 public notice (68 FR 11770). Briefly, those proposed changes included:

- A revised definition of the term "detection limit" at 40 CFR 136.2(f) to explicitly equate the term with the "method detection limit" specified in 40 CFR 136, Appendix B; and a revised definition of the term "method detection limit" included in Appendix B to provide technical clarifications and more clearly equate the term with the "critical value" described by Currie (1968, 1995) and the Limit of Detection described by the American Chemical Society (Keith et al., 1980; McDougal et al., 1983). Those concepts are further described in Chapter 2 of this assessment document.
- An expanded Scope and Application discussion in the codified MDL procedure to recognize that there are a variety of purposes and analytical methods for which the MDL procedure may be employed. The proposed revisions provided examples of four 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.) The proposed revisions also clarified that the procedure may not be applicable to certain test methods such as those used to measure pH or temperature.
- Proposed modifications to the considerations for estimating the detection limit in Step 1 of the codified MDL procedure and to the specifications for establishing the test concentration range in Step 3 of the codified procedure.
- Proposed deletion of the optional procedure for calculating a 95% confidence interval estimate for the MDL.

- Proposed changes to the iterative procedure to mandate its use when determining an MDL for a new or revised method or when developing a matrix-specific MDL, but allow it to remain optional when determining an MDL for other purposes, such as verifying lab performance.
- Proposed addition of a new procedural section to address the treatment of suspected outliers.
- Proposed deletion of the discussion of analysis and use of blanks included in Section 4(a) of the codified procedure.
- Proposed changes to the optional pre-test described in Section 4(b) of the procedure to improve the utility of results from this test.
- Editorial changes to the codified version of the MDL. Examples of these editorial changes include addition of a summary section, clarifications, reorganization of steps, simplified presentation of calculations, and deletion of the reporting section.
- Proposed addition of a definition of the ML at 40 CFR 136.2
- Proposed addition of a procedure (including a definition) of the ML to 40 CFR 136, Appendix B
- Explicit allowance of alternative detection and quantitation procedures, provided that the resulting detection and quantitation limits meet the sensitivity needs for the specific application. The objective of this proposed allowance was to provide greater flexibility in establishing or improving the sensitivity of methods for use under CWA and facilitate approval of analytical methods from other agencies or organizations that utilize alternate detection and quantitation concepts.

In addition to requesting comment on the assessment and the proposed revisions, EPA also specifically requested comment on several aspects of the proposal, including alternative actions that could have been taken. With respect to the ML, for example, EPA explicitly sought comment on the proposed addition of the ML definition to 40 CFR 136.2 and procedure to 40 CFR 136, Appendix B vs an alternative option of not incorporating the definition at 40 CFR 136.2, but instead continuing to specify the ML on a method-by-method basis. EPA encouraged commenters to support their views with data or information that would assist the Agency in making a final decision.

1.5.2 Impact of Comments on the Assessment

EPA provided a 120-day period following publication of the notices for submission of comments (from the date of publication of the notices to July 10, 2004). In response to requests from stakeholders, EPA re-opened this comment period on July 16, for an additional 30 days (68 FR 41988).

During the comment periods, EPA received comments from 126 individuals or organizations representing the diversity of the stakeholder community on this issue. They included 23 laboratories, 31 water treatment plants, 3 federal agencies, 11 state and county agencies, 23 industrial firms, 3 instrument manufacturers, 19 trade organizations, 4 consultants, 8 individuals, and the law firm representing the petitioners. Comments offered by these groups addressed more than 25 different issues. A complete summary of the comments and EPA's responses to those comments can be found in Appendix B to this Assessment Document. These comments are discussed at various locations throughout this document, and include discussion of:

- Additional detection and quantitation limit procedures suggested by commenters. (Chapters 2, 3, and 5)
- Public comments received on chemical, regulatory, and statistical issues, along with EPA's consideration of these issues in light of the comments received. (Chapter 3)
- Comments received on each of the evaluation criteria used in EPA's assessment and EPA's response to those comments. (Chapter 4)

- Potential process for additional stakeholder involvement on the evaluation of detection and quantitation limit procedures (Chapter 6)

Appendix C contains a detailed analysis of the detection and quantitation limit procedures evaluated through computation of limits using the data described in Section 1.3.2. This analysis has been revised to reflect new data and comments on the original version of the assessment, which was published as Appendix C to the February 2003 version of EPA's Assessment Document.

1.6 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.

Although the Settlement Agreement refers to the word "sensitivity" to describe detection and quantitation limits, we have avoided such use of the term "sensitivity" in this document because the term is widely used by analytical chemists to describe something other than detection and quantitation capabilities. Traditionally, analytical chemists have referred to the term "sensitivity" as meaning instrument signal units per concentration units, such as is given for a calibration slope or a response factor. For example, in ion selective potentiometry, the sensitivity is 59 millivolts per decade change in concentration for monovalent species and half that for divalent species. Sensitivity is a performance characteristic, but it differs from detection limits. For example, one might compare the sensitivity of instruments. Obtaining a sensitivity of 10,000 counts per ppb indicates a properly functioning Sciex 250, while a Perkin-Elmer 6000's sensitivity would be 100,000 counts per ppb. Another performance characteristic of sensitivity is that it may vary in an expected pattern as with mass to charge ratio in mass spectrometry or atomic number for x-ray fluorescence spectrometry.

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. In March 2003, EPA published two FR notices that met the terms of Settlement Agreement Clause 6a. Comments were received on those notices over a 4 month period ending in August 2003. EPA evaluated all comments received, and revised its assessment as appropriate to reflect these comments. This document details this revised assessment.

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 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 bave 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%

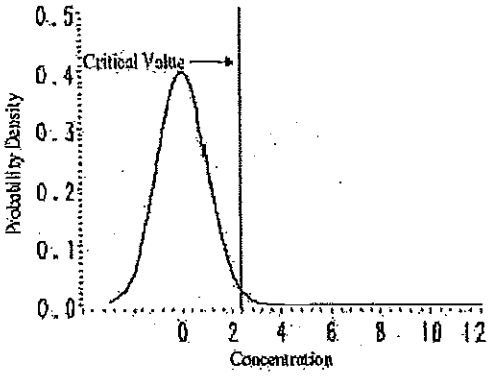

Figure 2-1

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 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%.

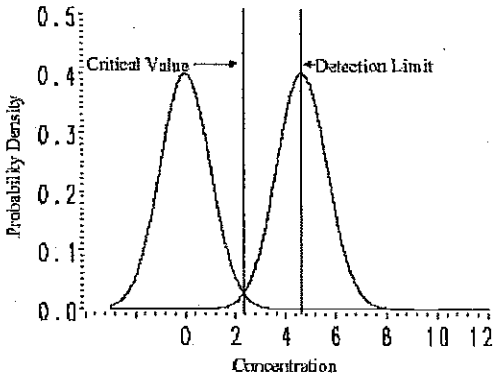

Figure 2-2

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.

Some members of the regulated industry and others have criticized the MDL because:

- There are some inconsistencies between the definition and the procedure
- It does not account explicitly for false negatives
- It does not always yield a 1% false positive rate
- It does not sufficiently account for blank bias
- A prediction or tolerance limit adjustment is not provided
- It does not account for interlaboratory and temporal intralaboratory variability, and
- It allows discretion in the use of the optional iterative procedures

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 footnote s 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 MDL[1]. 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, 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) x 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) x 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.

Some members of the regulated industry and others have criticized the ML because it:

- Does not account for interlaboratory and temporal intralaboratory variability, and
- Is based on a multiple of the estimated standard deviation which is assumed to be constant in the region of detection and quantitation, rather than a fitted model as suggested by the regulated industry.

These concerns are discussed later in this document.

---

[1]The 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 for newer analytical methods.

2.3 Other Detection and Quantitation Approaches

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 in 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 situations. 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., ±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 aud 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.

Lowest Concentration Minimum Reporting Level (LCMRL) and Minimum Reporting Level (MRL): Recognizing the potential for improvements over the PQL approach, and mindful that confidence in quantitation depends on measurement precision as well as accuracy, EPA's Office of Ground Water and Drinking Water has recently developed a standardized procedure for the determination of the "Lowest Concentration Minimum Reporting Level (LCMRL)" and a companion procedure for laboratories to establish their ability to quantify analytes at a "Minimum Reporting Level" (MRL).

The Lowest Concentration Minimum Reporting Level (LCMRL) is defined as the lowest true concentration for which the future recovery is predicted to fall, with high confidence (99%), between 50 and 150% recovery. A result below the LCMRL is an estimated value that does not satisfy these data quality objectives. However, it may be appropriate to report "estimated" data (i.e., below the LCMRL), depending upon the objectives of the study being conducted. The proposed LCMRL procedure is an iterative process that uses results from three or more different concentrations, of at least five to seven replicate reagent water samples at each concentration. The average recovery, standard deviation, number of replicates, and Student's *t* value are used to calculate a prediction interval of results that takes into account accuracy and precision at the level tested. For a concentration level to pass criteria, the prediction interval of results must be contained within the boundaries of a predefined quality control interval.

The Agency also has developed a procedure for use in the drinking water program which permits laboratories to confirm that they are capable of meeting a required MRL during their initial demonstration of capability. The MRL validation procedure will involve the analysis of one set of at least seven replicate reagent water samples spiked at the required MRL. To be validated at the MRL, the calculated prediction interval of results must be contained within the predefined quality control interval.

The Agency anticipates using standardized LCMRL/MRL procedures to support the monitoring required under the Safe Drinking Water Act for unregulated contaminants. Requirements for this monitoring are expected to be proposed in the *Federal Register* late in 2004. This proposal will include a full description of the LCMRL/MRL procedures.

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 ?Why? 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 in some cases may yield conservative (high) estimates,
- 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 complexity and expense the statistical procedures involved in calculating an IDE and IQE could be a barrier to innovation and method development.

These concerns are discussed in detail later in this document.

In December 2002, the Inter-Industry Analytical Group (IIAG) submitted a proposal to EPA that recommends (1) a sensitivity test intended to "replace the MDL as a test of whether an individual laboratory is performing adequately," and (2) an interlaboratory validation study design intended to characterize precision and accuracy of methods used for regulatory compliance.

IIAG's proposed sensitivity test includes the provision that EPA first determine the lowest calibration point of a method, prescribe a dilution of that calibration point as the spike level (e.g., at one-half or two-thirds the lowest calibration point), specify a required number of replicates, and set a quality

control acceptance criterion. IIAG asserts that this test would provide all laboratories with a single spike level and an "unambiguous pass or no-pass test." EPA solicited comment on approaches that might be considered appropriate for such determinations (i.e., the lowest calibration point of a method, an appropriate dilution, a number of replicates, and an acceptance criterion for standard deviation between measurements of the replicates). IIAG's proposed "full range" validation study is intended to determine precision and bias across the entire working range of an analytical method (i.e., from a blank to the upper end of the working range) and would account for variability between laboratories. IIAG recommends that results of such a study be used to establish an interlaboratory method detection level.

At the time IIAG's submitted the sensitivity test and full range validation study, EPA did not have the opportunity to evaluate IIAG's proposal against the criteria discussed in Chapter 4, but included the complete text of the recommendations in the regulatory record supporting the February 2003 Assessment Document. EPA is including an assessment of this proposal in Chapter 5 of 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.

During the comment period on the February 2003 assessment document, the American Council of Independent Laboratories (ACIL) submitted a procedure that was developed to address bias that may arise in the estimation of detection limits under certain conditions. The ACIL procedure separates estimation of a detection limit into two cases: analyses that always produce a numeric result, even so-called "blank" samples (i.e., zero analyte added), and analyses that do not always produce a numeric result, i.e. blank samples appear to produce no signal. We will call these Case I and Case II. Analysis of samples for metals with inductively coupled plasma optical emission spectroscopy (ICP-OES) is an example of ACIL Case I; analysis of samples for PCBs with gas chromatography and mass spectrometry (GC/MS) is an example of ACIL Case II.

For Case I analyses, ACIL suggests making use of the numeric results obtained from the analysis of blank samples which laboratories routinely run as a quality control measure. ACIL provided a detailed set of instructions for conducting the analyses and doing the MDL calculation. Differences between the ACIL Case I calculation and the EPA MDL calculation include: (1) use of blanks rather than low-level spikes to estimate standard deviation, (2) the calculation of both a critical level and a long-term MDL, where the MDL is based on adding the mean of the blank results to 2 times the product of the standard deviation and t-statistic, (3) a bias offset correction that adds the mean of the blank results to the calculated critical level and MDL, (4) recommends the use of results from a minimum of 20 analyses, and (5) analysis over the course of a year from routine daily operations (rather than on one day). ACIL's Case I procedure is similar, but not identical, to the USGS procedure that is described in Section 3.3.4 of this document. The ACIL Case I procedure has no explicit limits on the amount of contamination allowed in the blanks before a laboratory is considered to be "out of spec."

For Case II, blanks cannot be used to estimate the standard deviation because they do not provide a response. Thus, Case 2 recommends an iteration of multiple low level spikes somewhat similar to the requirements in the EPA MDL procedure. However, the calculation of an MDL from the results of these spiking experiments differs significantly from the EPA MDL procedure. The procedure also specifies a sensitivity check for which some of the details are not as explicit compared to the Case I part of the ACIL detection limit procedure.

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

NWQL has since developed a variant of the MDL called the long-term MDL (LT-MDL) that has been in routine use since 1999. The LT-MDL determination ideally employs at least 24 spiked samples prepared and analyzed by multiple analysts on multiple instruments over a 6- to 12-month period at a frequency of about two samples per month.

Unlike EPA programs that rely on hundreds of commercial, Federal, State, and local laboratories for sample analysis, the samples collected for USGS water programs are analyzed by the USGS National Water Quality Laboratory in Denver, Colorado. 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 from the spike concentration down to zero, 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 three primary differences between the EPA MDL and the USGS LT-MDL procedures are the (1) larger minimum number (24) of spike samples, (2) longer time period, and (3) mixing of instruments and analysts in the determination of the LT-MDL. Because the MDL and LT-MDL approaches otherwise are so similar, EPA did not evaluate the long-term MDL approach in the February 2003 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).

In the LT-MDL procedure the low-level spike used for each analyte and instrument is recalibrated at least once a year or when an anomaly occurs. USGS has enhanced the LT-MDL procedure by using their large volume of uncensored blind laboratory blank data, which also is collected yearly, as a reality-check on the spike-based LT-MDL. In cases where the standard deviation used to calculate an LT-MDL based on blind blank data is significantly different (especially when greater) than the standard deviation used to calculate the spike-based LT-MDL, the blank data are used to calculate the LT-MDL. Blind blank data also are used to evaluate whether the calculated LT-MDL requires an off-set correction for blank bias, i.e, LT-MDL = (s * Student t) + median or mean blank concentration. This offset is similar, but not identical, to the ACIL Case I procedure described in Sect. 2.3.3 of this document. The LT-MDL offset correction compensates for a blank distribution that is not centered on zero (as assumed by the EPA MDL formula).

The NWQL has found that this blank bias correction to the LT-MDL is especially important for blank-limited analytes, including some metals, total organic carbon, phenol, and nutrients. The NWQL also uses a data reporting convention that incorporates a higher reporting level (called the laboratory reporting level; LRL) that is set at two or more times the LT-MDL. However, this convention also includes reporting of data between the LT-MDL and LRL.

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, EPA considered 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 the plan for conducting this assessment, EPA identified other issues that should be considered. With the exception of the first issue, these issues are discussed in this chapter and include:

- Concepts of the lower limit of measurement (discussed in chapter 2),
- 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.

These issues are organized 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 the issues discussed in Sections 3.1 - 3.3.

3.1 Analytical Chemistry Approaches to Detection and Quantitation

This section explains the key analytical chemistry issues involved in the development of detection and quantitation limits. These include: (1) nonzero sample blanks, (2) instrument censoring, (3) matrix effects, (4) analyte recovery, and (5) temporal variability of the measurement system.

3.1.1 Nonzero Sample Blanks

Analytical chemists rarely state 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 that some small residual amount of the substance may be present and contribute to the measurement result. The inability 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 this potential problem was published by 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" (or "positive bias" from blanks) 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 several EPA analytical methods, including EPA Method 1631 for measurement of trace-level mercury.

The most common analytes for which contamination problems are encountered in environmental measurements are metals, primarily zinc because of its ubiquity in the environment. With the exception of some volatile organic compounds, such as methylene chloride and acetone, that are used as solvents in laboratories, contamination in the measurement of organic compounds is less of a problem than contamination of samples for metals analyses. 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 necessarily mean zero concentration, and through careful control and evaluation, it is possible to make measurements for which the blank contribution is sufficiently small to be considered negligible.

In the February 2003 version of this document, EPA noted that 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. Stakeholders commenting on EPA's 2003 assessment of these approaches added that, for many blank analyses, there is a measurable response (blank bias) that can be attributed to reagents, sample vessels, and other contamination sources, and that the MDL procedure failed to take these blank responses into account. Several commenters suggested that the mean of the blank results should be added to the formula used to calculate the MDL. The American Council of Independent Laboratories (ACIL) submitted a procedure to use blanks rather than spikes to estimate a detection limit. The USGS submitted a long term MDL procedure that uses either the mean or median of blank results as a lower bound reality check on the MDL whenever an MDL computed from the low level spiking experiments is sufficiently less than the blank results.

Following a careful evaluation of these comments and further consideration of this issue, EPA recognizes that, under certain conditions, it may be appropriate to account for blanks in establishing detection and quantitation limits with certain limitations. For example, a procedure to handle blanks should account for negative results, and should limit and control sample and laboratory contamination. Negative blanks are possible and can be caused by a blank-subtracted calibration in which the result for the calibration blank is greater than the results for the blanks used to establish the MDL. If such negative blanks were to result in a negative mean blank, adding the mean blank result to the formula could result in an unattainably low MDL. Conversely to eliminate unnecessarily high MDLs, laboratories also would need to ensure that the results of blank samples are not excessive. The laboratory would need to use "clean" and other techniques to control contamination to the lowest possible levels and/or use a second or higher order calibration function to preclude high results for a calibration blank from exerting undue influence on the sample results. In addition to working out some of the details of necessary bounds on blank correction and contamination, differences between the procedures submitted by USGS and ACIL need to be evaluated.

3.1.2 Analytical Instrument Thresholds: Data Censoring

Certain analytical instruments typically employ "thresholds" to eliminate spurious or background signals so that analysts can be relieved of the burden of removing or compensating these small signals. As a result, the instrument itself may return a response of zero to a blank (a "non-response"). 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 filter out the static 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 a 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 application of an instrument threshold (sometimes known as the "Episode 6184 study"); 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 eliminated) and with the threshold off (i.e., there is no suppression of signals). Samples were analyzed at decreasing concentrations, including a blank concentration level, with triplicate determinations at each concentration. With the threshold turned on, all of the measurements made on the blank were reported as zero. This is not surprising, given the purpose of the instrument threshold. Without the threshold off, only 27 of 230 measurements on the blank (11%) were reported as zero, and no negative results were reported.

Instrument thresholds have a direct and indirect impact on estimating detection and quantitation limits. The main direct impact is that it is not possible to estimate the standard deviation of measurements at zero. However, by definition, the standard deviation at zero is required to calculate the Currie critical value (CRV; L_c). The EPA MDL procedure was constructed to deal with this problem by providing a way to estimate a standard deviation at a low concentration, and including instructions for determination of a concentration as close to zero as is possible that will generate a measurement.

To calculate an MDL using the 40 CFR 136, Appendix B, MDL procedure, it is necessary to find the lowest concentration at which the analytical system will return results. Many laboratories have run repeated measurements in order to find this concentration. The challenge of finding this lowest concentration manifested itself in EPA's variability versus concentration (Episode 6000) studies. Technologies for determination of organic, conventional pollutants, and metal analytes were evaluated in the Episode 6000 studies. The MDL procedure suggests iteration until the calculated MDL is within a factor of five of the spike level. For the Episode 6000 studies, EPA instructed laboratories to use a factor of three instead of five in an attempt to more narrowly define the lowest spike level at which measurements could be made. This change to a factor of three also was suggested by one of the peer reviewers charged with evaluating EPA's 2003 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 peer 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 then the critical level, including possibly blanks"* (Rocke, 2002).

The reviewer's calculations suggest that the MDL may be strongly inflated for a spike level of five times the MDL, but only moderately inflated at a spike level of three times the MDL. However, during the Episode 6000 studies, several laboratories asked for relief from the factor of three requirement because a large number of iterations were required to meet it. In response, EPA reinstated the factor of five for these laboratories. If the reviewer's example calculations are correct and a practical procedure for determining the MDL using the factor of three were implemented, it could exacerbate the concern from the regulated community that MDL values are too low.

Several stakeholders commenting on EPA's 2003 assessment suggested that approaches to detection and quantitation should address methods that do not always produce an instrument response, e.g. so-called blanks never produce a response because of electronic censoring by the instrument, and that EPA's approaches do not do so. These stakeholders prefer that the MDL not be applied to methods for which an identifiable analyte signal cannot be established using method blanks, where pattern recognition is required (e.g., Method 608 for PCBs as Aroclors), or where the method requires more than one signal for an analyte to be positively identified (e.g., the use of multiple ions in GC/MS methods). EPA recognizes that additional guidance needs to be developed for these methods. One commenter, the American Council of Independent Laboratories, submitted a draft set of procedures designed that partially addressed methods that do not produce an instrument response at zero concentration. EPA evaluated the ACIL procedure, which involves a complex set of iterative spiking experiments, and found that it needs further refinement. EPA agrees that this issue warrants further examination.

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.

"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 from the matrix 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

Some stakeholders have suggested 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 medium (water, soil, biosolids, tissue). Some commenters on EPA's 2003 assessment believe that any form of a detection limit study should require demonstration of the lowest level of detection achievable in an interference-free matrix and should relate this to what is actually detectable in a highly complex matrix. One commenter stated that EPA should provide an objective set of procedures that a permittee can follow to avoid liability when faced with an MDL or ML it legitimately cannot achieve because of the unique nature of its effluent. EPA notes that permittee liability was not a goal or purpose of our assessment of detection and quantitation approaches and issues. Although EPA recognizes stakeholder concerns about the matrix effects, there are several problems associated with the approach suggested by some commenters. These problems include:

- Many "real-world" matrices contain the target pollutant at levels well above the detection or quantitation levels, 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.

- 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.

- Few environmental analyses are conducted on actual samples of reagent water or other reference matrices and there may be matrix-specific limitations to the sensitivity of any given analytical method. From a practical standpoint, it would be very impractical to evaluate method sensitivity in every possible matrix to which a method might be applied, or to establish a subset of all possible matrices that would satisfy the concerns of every regulated discharger.

- The cost of determining detection and quantitation limits in every possible matrix would be prohibitive.

Because of these difficulties a reference matrix (or reference matrices) is an appropriate and practical first choice to establish detection and quantitation limits. And the procedures for defining these limits should allow for evaluation of data collected in the specific matrices of concern. Laboratories or data users are most able to determine which matrices might be considered to be "highly complex" based on the matrices that are typically analyzed in a given laboratory. EPA's detection and quantitation procedures do not preclude laboratories from determining MDLs in matrices other than reagent or "blank" matrices, and the Agency encourages laboratories and others to determine matrix-specific MDLs when all efforts to resolve matrix interferences have been exhausted. The existing procedure at 40 CFR 136, Appendix B, includes a discussion regarding determination of matrix-specific MDLs for this reason. Laboratories usually are very capable of eliminating or compensating matrix interferences if tasked to do so. However, given the degree of concern about this issue it is appropriate for all parties to continue to search for additional solutions to the "real world" matrices issue.

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 an 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, it is appropriate to leave the development and maintenance of standard reference materials (SRMs) and certified reference materials (CRMs) to NIST and the commercial marketplace. These reference materials are a useful means of challenging a test method and EPA 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 SRMs and CRMs could be useful in establishing detection and quantitation limits, 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 of Analytes from the Sample Matrix

In the preceding two sections, we discussed bias (errors) from blank contamination and matrix effects. Errors from recovery effects ("recovery bias") are discussed in this section. Recoveries are a measure of the amount (usually expressed as a percentage) of analyte that is recovered from the sample matrix and measured by the analytical system. Chemists sometimes use the phrase "accuracy of the method" when listing the percent recovery of an analyte. A goal of analytical chemistry is to achieve recoveries as close as possible to 100%. When this is not achieved, recovery correction may be used. The purpose of recovery correction is to adjust the measured concentration for the amount by which the measured concentration differs from the true concentration (if known). Recovery "factors" are initially determined by analysis of a sample containing a known (spiked) amount of the analyte. These factors are applied to measurements of samples with an unknown amount of the analyte in the same or a similar matrix. To illustrate the potential need for recovery correction, consider analytes, such as organic bases

(e.g., benzidine) and acids (e.g., phenols) in a water sample, that are either not totally (100%) recovered in the extraction process, or are adsorbed on the surface of a GC column at very low (nanogram) levels. As a result, the measured concentration of these analytes 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, and the amounts measured in unknown samples also should be doubled to allow for recovery correction.

Several stakeholders have stated that understanding this "recovery bias" is particularly important when reporting results near the limit of detection, and is critical when reporting quantifiable results. These stakeholders believe that even if recovery bias is not controlled at the detection level, the approach to determining detection and quantitation limits should compensate for it.

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, however, has been included in several of the most recently developed approaches for detection and quantitation. For example, the minimum detectable value (MDV) 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 ISO minimum detectable value approach, recovery is treated as a linear function versus concentration, and an extrapolation is used to estimate the recovery at zero concentration. 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 minimum detectable value, an minimum detectable value very close to zero, or a negative minimum detectable value. 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 minimum detectable value. The difference is that an unweighted model is used in the minimum detectable value, 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 minimum detectable value.) 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), "bias" means the amount by which the estimated sample standard deviation differs from the true population standard deviation. This use should not be confused with a common use of "bias" in analytical chemistry measurements to denote the deviation of a result from the true value (usually expressed as percent.)

Recovery correction may be appropriate if (1) when developing method detection and quantitation limits, 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 conditions are rarely observed. 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%.

Although some stakeholders disagree, EPA believes that the normal condition in environmental analytical measurements is that variability (as standard deviation) between sample results 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. For examples of the effect of poor recovery on precision, see the quality control (QC) acceptance criteria for the semivolatile organic compounds in Table 8 of EPA Method 1625 (see 40 CFR part 136, Appendix A). This increase in relative variability is not the result of measurements being made at lower levels, as is the normal case, but is a result of variability in the extraction (partitioning) process. One stakeholder commenting on EPA's 2003 assessment stated that EPA's statement that variability remains approximately constant as recovery decreases may not hold true in all cases, and recommends that, if recovery falls within a specified level (e.g., less than 70%), detection limits should be adjusted accordingly. EPA acknowledges that there may be instances in which this general condition does not hold true.

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. Measurements of dioxins/furans, PCBs, and pesticides can be made to very low (femtograms per liter) concentrations, with no decrease in recovery compared to recoveries observed at much, much greater (microgram per liter) concentrations. (One microgram is equivalent to one million femtograms). The ability to measure dioxins/furans, PCBs or pesticides down to these low concentrations demonstrates that recoveries for these compounds do not decrease with decreasing concentration. There also are chemicals, such as the nitrophenols and benzidine, that are not recovered reliably at sub microgram per liter levels. But these instances are known and recognized in the instructions for conducting these measurements.

MDLs are established and listed in methods based on the determined (measured) concentration (not the spike concentration), and laboratories and others that are required to verify MDLs, verify based on the determined concentration. If EPA estimated and listed MDLs in methods based on the spike (true) concentration, logic would require that the true (recovery corrected) concentration be used for regulatory compliance with the result that all results, not just the MDL, would be greater than nonrecovery corrected results.

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 an additional procedure for recovery-correction in determining detection or quantitation limits for such methods could result in double-counting of analytical bias.

Another procedure for dealing with bias (errors) from blank contamination, matrix effects, or errors from recovery effects ("recovery bias"), is to assure that the detection or quantitation limits which is determined meet the data users data quality needs for both precision and accuracy, without any correction. As described previously in chapter 2, EPA's drinking water program is developiug an approach to setting quantitation levels called the minimum reporting level or MRL. The MRL addresses these issues by setting a data quality objective for minimum and maximum permitted inaccuracy arising from these effects.

The issue of bias in determination of detection or quantitation limits be it from blanks, matrices or other than 100% recovery of an analyte is a longstanding issue. All parties should continue to collaboratively work to develop other solutions or approaches to mitigate bias effects.

3.1.5 Temporal Variability of Analytical Measurements

As with most other areas of technology, 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 of 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, commercially available high resolution mass spectrometer systems were able to achieve a detection limit of approximately 4 pg/L and a 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 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.

Although there is no such thing as a "perfect" measurement, the idea that "practice makes perfect" (i.e., analytical results get better with practice) applies to the quality of measurements made with a given method over time. We can demonstrate this 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 followed by a decrease back to normal level after analysts become sufficiently experienced with the analytical method. .

The use of quality control (QC) charts as a means of tracking method and laboratory performance 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 or stability, 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 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.

Stakeholders commenting on EPA's 2003 assessment of procedures for characterizing the detection and quantitation limits of analytical methods expressed concern that EPA's MDL and ML are determined using a single batch of samples representing a "snapshot" in time, and do not account for the temporal variability that can occur in a laboratory from day to day (e.g., due to use of multiple analysts and instrumentation, changing laboratory conditions). Although the codified version of the MDL does not preclude laboratories from incorporating temporal variability into the procedure (e.g., it allows the use of more than 7 replicates and does not require that the replicates be analyzed in a single batch), many users understand the MDL to be a single batch procedure. EPA encourages, where appropriate, use of data gathered over an extended period of time to calculate an MDL because measurement capabilities tend to improve and laboratory conditions tend to vary. Detection and quantitation limit calculations can be supported by procedures that allow laboratories to affordably characterize such changes and improvements.

3.2 CWA Regulatory Issues Affecting Detection and Quantitation

Section 3.2.1 provides a brief overview of Clean Water Act activities that involve chemical measurements and are, therefore, directly impacted by detection and quantitation limit approaches. Specific issues to 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, both directly and indirectly, by detection and quantitation limit approaches. Stakeholders commenting on EPA's assessment of detection and quantitation procedures stated that, because of the differing technical demands and regulatory and laboratory uses of detection and quantitation levels, the procedures for determining these values should be based on sound science. These stakeholders urged EPA to consider the implications of each technical decision it makes regarding determination of detection and quantitation values on the practical implementation of its regulations.

Several commenters believe that EPA, permit holders, and laboratories would be better served if different approaches to detection and quantitation were taken for each use. Commenters specifically cite uses as a start-up test in a single laboratory, as a value characterizing a given analytical method, as a test for approving a method modification or alternate test procedure (ATP), compliance monitoring, and as a permit compliance level. The Inter Industry Analytical Group, in particular, has recommended the following 3-part approach:

- A sensitivity test (as a test of start-up proficiency),
- A long-term MDL approach (for laboratory reporting),
- Full-range validation study (such as the ASTM IDE/IQE) for validation of new methods and for setting quantitation levels that will be used as permit compliance levels.

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 also considers alternate testing procedures (ATPs). If EPA deems these ATPs to be acceptable for use, they 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) encourages 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 led organizations that develop methods to use different approaches, and many organizations have changed approaches over the years. Some stakeholders, who commented on the 2003 assessment, believe that method-specific detection and quantitation limits should account for interlaboratory variability, and therefore should be based on interlaboratory data. Other stakeholders believe that such a requirement would be overly restrictive and burdensome, resulting in fewer approved methods and technologies. 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.

A solution to this lack of consensus would be to require that all methods promulgated at 40 CFR part 136 contain uniform approaches for detection and quantitation. However, taking such action would be disingenuous and confound methods promulgation because:

- To date, no single detection and quantitation limit approach has emerged to meet the needs of all organizations for all applications.
- If EPA's were to select an approach that differs from those of other organizations, those organizations would be required to conform their method to accommodate the EPA approach. Doing so would mean that these organizations would have to invest additional laboratory resources to develop detection and quantitation limits that conform to OW definitions.
- If outside organizations decided against conforming their approaches to that of EPA, 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 encourage the development of improved measurement techniques, and provide the stakeholder community with a variety of measurement options whenever possible, it would be counter-productive to allow method developers the choice of only one detection or quantitation limit approach, or to only promulgate those methods that contain this single approach. However, there are real benefits to standardization, all new methods developed by EPA for promulgation at 40 CFR part 136 should reflect such standardization, and EPA should strongly encourage outside organizations to include these standardized approaches in their methods. However; there was no clear consensus as to what this standardized approach should be. Industry advanced IDE/IQE procedures but others did not necessarily support them.

3.2.1.2 Verification of Laboratory Performance

Just as sensitivity is important for evaluating 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 as a measure of method sensitivity. 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 laboratory performance evaluation samples such as those used in laboratory accreditation studies. Several commenters on EPA's 2003 assessment recommended that a simple "sensitivity" test (e.g., determination of analyte recovery in a sample containing a low-spike concentration of the analyte) be used to evaluate or establish laboratory performance. Although at least two commenters submitted some ideas for conducting such a test, none were sufficiently or clearly detailed. However; EPA is open to consideration of approaches to verify lab performance.

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. These approaches were developed to establish detection and quantitation limits that could be met by any laboratory that participated in the study. However; the IDE/IQE cannot be used to verify individual laboratory performance.

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 are intended to capture a wide range of sources of variability such as temporal variability, and 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 a voluntary consensus standards body, such as ASTM International. EPA will explore approaches for lab performance verification through the stakeholder process.

3.2.1.3 National Pollutant Discharge Elimination System

The National Pollutant Discharge Elimination System (NPDES) serves as a means by which EPA, States, and Tribes control point source releases of pollutants into the nation's waters. Under this system, individual facilities are issued NPDES permits that provide effluent limitations that restrict the quantities, discharge rates, and concentrations of pollutants that may be legally discharged. Typically, these limitations are based on technology-based standards. If, however, these technology-based limits 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 must 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 and affordable treatment technologies. Functionally, these industry-specific gnidelines 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. A discharge that causes, has reasonable potential to cause, or contribute to an excursion of an applicable water quality standard requires a water-quality based limit. Such a water-quality based limit shall be established at levels that derive from and comply with applicable water-quality standards and must be consistent with the assumptions and requirements of any available waste load allocation for the discharge, approved by EPA pursuant to 40 CFR 130.7.

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 (e.g., 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. NPDES permits generally specify the use of measurement methods promulgated at 40 CFR part 136 under the Clean Water Act Section 304(h) for purposes of compliance monitoring and other reports submitted under NPDES permits.

Detection plays a role in compliance monitoring because of concerns with measurement results at the low end of any analytical 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, 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."

Stakeholders are particularly concerned with the use of the detection and quantitation limits for compliance purposes (e.g., judging whether a discharger is in compliance or whether a waterbody complies with its water-quality standards), for which a high level of reporting consistency and confidence in the data is required. Several commenters on EPA's assessment stated that procedures used to determine these limits should provide the certainty required to make regulatory decisions.

Several commenters suggested that there should be a single compliance benchmark for detection of each analyte that is independent of laboratory or method capabilities; laboratories used for compliance reporting would be required to demonstrate that they can detect at or below this level. These commenters note that such an approach would be particularly useful and appropriate for analytes with water quality (or other) standards set below current technological capabilities.

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 rather than specifying specific performance benchmarks that a laboratory must meet to demonstrate and maintain proficiency. 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 by 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 initially may not be able to achieve these standards. However, most laboratories facing this problem would try to improve and achieve these standards by investing in staff training, improved equipment, a stronger quality assurance program, or clean room practices and higher quality maintenance and operations.

There is 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 contamination and variability at these low levels, thus raising detection and quantitation limits to a level that is higher than desired for protection of public health and the environment.

3.2.3 Compliance Evaluation Thresholds

When technology-based effluent limitations are developed, the limits typically are 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 that permit limits may be set below the detection or quantitation limit of the most sensitive, approved analytical method. This is particularly true for pollutants that are toxic in extremely low concentrations or that bioaccumulate. A recommended approach for these situations is to include in the permit, the appropriate permit limit derived from the water quality model and the waste load allocation for the parameter of concern, regardless of the proximity of the limit to the analytical detection level, along with an indication of the specific analytical method that should be used for monitoring (See Technical Support Document for Water Quality-based Toxics Control, EPA/505/2-90-001, March 1991). Both the MDL and ML have been used as reporting limits or compliance evaluation thresholds in NPDES permits. EPA promulgated regulations for NPDES permits for dischargers to the Great Lakes basin that require the use of the ML for compliance assessment purposes (See Appendix F, Procedure 8, Part B of 40 CFR 132). EPA has recommended for most permitting situations that the compliance level be defined in the permit as the ML (See Technical Support Document for Water Quality-based Toxics Control, EPA/505/2-90-001, March 1991). Outside of the Great Lakes basin, it is important to note, however, that states that implement the NPDES permits program have not always followed EPA's guidance. The inconsistent use of the MDL and ML as reporting limits or compliance evaluation thresholds in NPDES limits suggest that EPA should develop additional implementation guidance.

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 compliance evaluation threshold for the situation in which the WQBEL is less than a 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 has been an exceedance of the ambient criterion. Use of the ML as a compliance evaluation threshold is appropriate because it is the point at which the measurement could be considered reliable.

3.2.4 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. EPA recognizes 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 that State and local governments use different approaches, a change by EPA with regard to this assessment of detection and quantitation procedures may have an unanticipated impact on those States and local governments.

3.2.5 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. Users of methods 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. Personnel working for the method developer generally include a project manager, measurement analysts, who are experienced in several measurement technologies or very experienced in a specific, complex technology, and at least one statistician. Operating laboratories typically will not have a statistician, and the breadth and dept of the analyst experience may be less than in a method development laboratory, because an operating laboratory is focused on obtaining reliable results in the analysis of a given sample using a well-tested measurement technology.

3.2.5.1 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 the developers's 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 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 and the required skill of the personnel who design, conduct and interpret the results of 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 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.

3.2.5.2 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 to participate in a round robin study for a VCSB or 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 most environmental testing laboratories.

Highly complex procedures are usually more costly to implement than simple procedures. As noted in Section 3.1.5, method performance generally improves over time. This means 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. Mandating interlaboratory studies using complex detection and quantitation procedures means that laboratories lacking statistical support staff, and seeking to 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. This limitation has the effect of hindering method development and improvement.

3.2.6 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.

In the 2003 assessment, EPA argued that 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 stakeholders, such as regulators, permittees, and the laboratory community. To minimize this confusion, EPA suggested that a single pair of related detection and quantitation procedures could meet the needs of all CWA applications. Some commenters disagreed with this approach and recommended that at least two distinct procedures should be used, one for method development and one for verifying laboratory performance.

3.2.7 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

except 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). If EPA already has standards in place for detection and quantitation approaches, EPA is not obligated by NTTAA to replace these with VCSB standards. Although some stakeholders commenting on EPA's 2003 assessment encouraged EPA to allow use of alternative procedures for determining detection and quantitation levels, commenters in general did not support eliminating continued use of the MDL or ML.

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 detection and quantitation approaches evaluated in this document, including the approaches submitted by commenters on the 2003 assessment, also are prescriptive, not performance, standards.

To effect a performance-based approach to estimating detection and quantitation limits, an option that EPA may consider is to allow method developers, laboratories, and others the choice of 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. That these different approaches (prescriptive standards) yield different answers would be immaterial if EPA evaluates the answers relative to a specific decision, i.e. the benchmark becomes a performance rather than a prescriptive standard. That evaluation should not be divorced from knowledge of the decision to be made (e.g., the regulatory limit for a given pollutant).

3.2.8 Alternative Procedures

One of the peer reviewers who evaluated a draft version of the February 2003 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)

EPA currently assesses and has approved at 40 CFR Part 136 methods that employ an alternative procedure for establishing method sensitivity. This approval process includes an overall evaluation of the suitability of the method in entirety and thus includes the detection or quantitation approach used to establish the performance specifications listed in the method.

The peer reviewer is referring to the system of reference methods used under the Clean Air Act. This system 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 (i.e. is a method appropriate for CWA applications), whereas there remain fundamental theoretical issues surrounding the relative merits of the various detection and quantitation approaches that are the subject of this document.

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 (e.g., a compliance limit) or environmental study. 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.

A stakeholder commenting on EPA's 2003 assessment recommended that EPA adopt alternative procedures in Appendix B of 40 CFR 136 as site-specific alternatives to the MDL and ML when such an alternative is determined to be necessary by a discharger and/or regulatory agency (e.g., in special cases when more scientifically rigorous procedures are needed). As noted previously, EPA has reviewed and

approved at 40 CFR Part 136 methods that employ an alternative procedure for establishing method sensitivity as part of an overall evaluation of the suitability of the method. EPA has done so without need of any revisions to appendix B at 40 CFR part 136.

3.3 Statistical Issues

This section provides 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 several ways in which some of these sources of 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. Many of these methods also provide detailed steps for performing all sample handling and analysis activities, and detailed requirements for analysis of specific quality control samples with corresponding QC acceptance criteria.

Even when prescribed EPA method requirements and guidance are used, however, it is not possible to completely eliminate all variability that can occur within or between laboratories. Even with procedures in place to control quality and reduce variability, it should be recognized that some laboratories, analysts, and instruments 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.

Many of these sources of variability are considered in establishing detection and quantitation limits for analytical methods used under EPA's Clean Water Act programs because these detection and quantitation limits are first established in single-laboratory studies, then evaluated or verified in multiple laboratories, and, where necessary, further evaluated in an interlaboratory study. These studies include evaluation of method performance characteristics, including detection and quantitation capabilities, in multiple laboratories using multiple matrices, analysts, instrumentation, reporting activities, standards, and reagents. Although detection and quantitation are not evaluated in the various matrices used in these studies, EPA's MDL procedure includes instructions for determination matrix-specific MDLs.

Some stakeholders commenting on EPA's assessment of approaches to detection and quantitation believe that accounting for these sources of variability when determining detection and quantitation limits is necessary because relative variability increases as the lower sensitivity limits of a method are approached. Some stakeholders believe, for example, that a methodology for detection and quantitation has to address the variability that occurs across laboratories (interlaboratory variability) using the same

analytical method. Other stakeholders believe, however, that interlaboratory variability is not an issue because detection and quantitation decisions are made in a single laboratory. Some stakeholders believe that procedures should address the long-term variability that can occur within a single laboratory over time. As discussed in section 3.1.5 of this chapter, EPA encourages, where appropriate, gathering data to address temporal variability. EPA acknowledges that interlaboratory variability is very important during the methods development process and should be incorporated, as appropriate, during the process. EPA also recognizes that within lab variability should be considered when establishing laboratory performance.

Over the years, stakeholders have noted that the variability that can result from application of analytical methods to different matrices also should be addressed by procedures for determining method-specific detection and quantitation limits. However, it has been EPA's experience that matrix effects typically can be overcome using various sample processing procedures. In EPA's interlaboratory validation studies of the 600-series wastewater methods, the recoveries of some organic analytes from real-world matrices were closer to 100% than were recoveries 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. EPA does not believe it is appropriate or feasible to aggressively pursue matrix effects in establishing detection and quantitation limits (i.e., EPA has not attempted to find worst-case matrices in order to maximally exacerbate matrix effects). Instead, EPA considers the type of matrices that would be regulated under the Clean Water Act (e.g., the effluents that are discharged from properly designed and operated secondary treatment plants). Further discussion of matrix effects can be found in Section 3.1.3.

Because detection and quantitation limits focus exclusively on the capabilities of the measurement process, a source of variability that is ***not*** considered in any of the detection and quantitation limits is the variability that is associated with sample collection. If the sample is not 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 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 this result may not be typical, and EPA does not mean to diminish the importance of understanding measurement error in the region of detection and quantitation, EPA believes it is important to understand it in the context of the overall sampling and analysis error.

3.3.2 False Positives and False Negatives

3.3.2.1 False Positives and False Negatives in Making Detection Decisions

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 (Lc), 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 statistically different from blanks ("not detected"). Measured values that are no less than

the critical value are judged to be statistically different from the blanks ("detected"). Denoting measured values that are less than the critical value as non-detects constitutes censoring and is discussed in more detail in Section 3.3.5.

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 in statistics as Type I error ("false positive"). 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 at the hypothesized concentration but the measurement procedure fails to indicate its presence. The occurrence of (iv) is defined as the Type II error ("false negative").

The following table summarizes possible situations:

	State of the Sample	
Decision	*Concentration=C, where C>0* *Analyte Present*	*Concentration=0* *Analyte Not Present*
Concentration=C, where C>0	Correct (i)	Type I error (ii)
Concentration =0	Type II error (iv)	Correct (iii)

Calculating the probability of a Type I error only requires assumptions regarding the distribution of observations under the hypothesis that the concentration is equal to zero. In the terminology of statistical decision theory, Concentration = C, where C>0 corresponds to a true value is referred to as the "Alternative Hypothesis" (see, e.g. *Introduction to Mathematical Statistics*, by Hogg and Craig, 5th edition, [1995]). When C is hypothesized, assumptions need to be made about the distribution of observations at Concentration=C for the probability of Type II error to be evaluated.

In analytical chemistry, the probability of Type I error is often called the "false positive" rate and the probability of Type II error is often called the "false negative" rate. The statistical alternative hypothesis should be specified before introducing the false negative rate. An error common to some published discussions of false negative rates and detection and quantitation concepts is to state that use of Currie's detection limit as a reporting limit or action level will somehow "control" the rate of false negatives. This is both incorrect and counter-productive, because a single level cannot control false negative rates.

Currie introduced the idea of a Detection Limit, Ld, in place of a statistical alternative. The Detection Limit is *not* a part of the detection decision process (i.e., is the concentration in the sample statistically different from the blank?). The Detection Limit is defined such that when the true concentration of an analyte is equal to the Detection Limit, there is a small probability that a measured value will be less than the Critical Value (detection decision-making level in this case), and thereby result in the false negative decision.

One of the peer reviewers of EPA's 2003 Technical Support Document (the TSD) stated:

> *"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...)* " (Piegorsch, 2002)

To illustrate the intent of Currie's detection limit, consider a case where the detection decision-making level is set equal to Currie's critical value, and a sample is spiked at 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 detection decision-making level, and 1% of the measurement results will be reported as being measured below this level. Knowledge of the lowest true concentration that will routinely produce acceptable results (e.g., 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.

3.3.2.2: Effect of Bias on Rates of False Positives

The presence of bias in a method can have a strong effect on the rate of false positives associated with detection limit estimates. For example, in defining the critical level, Currie assumed that blank results follow a Normal distribution centered about zero (0). However, for some methods and analytes, this assumption may not hold due to factors that can and should be controlled, such as calibration errors and high background contamination. In many cases, bias can lead to either under- or over-estimation of detection limits. In cases such as these, not taking bias into account when determining detection and quantitation limits (using the mean or median of the results, for example) may influence false positive rates.

3.3.3 Use of Multiple Replicates

Existing detection/quantitation procedures are based on estimating the standard deviation of blank or spiked replicates. Statistical estimates tend to be less variable when the number of replicates increases. Some commenters on EPA's 2003 assessment believed that use of only seven replicates over a short period of time results in a substantial underestimation of the MDL. However EPA's MDL procedure does not limit the maximum number of samples that the laboratory may use to estimate the MDL; the procedure

simply sets a minimum number of seven replicates. Laboratories may choose to improve their estimates of the standard deviation that is used to calculate the MDL by analyzing more than seven replicates.

3.3.4 Statistical Prediction and Tolerance

To define a critical value, a detection limit, or a quantitation limit, different descriptive terminology is used to distinguish differences in the numeric value of the limit. The following example uses a critical value, but the questions motivating detection and quantitation limit decisions may be phrased in a similar fashion.

In setting a critical value, do we want a critical value that tells us how likely it is that:

- A measurement result was produced by measuring a blank sample,
- The next measurement result will be produced by measuring a blank sample, or
- 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:

- Percentiles;
- Prediction intervals; and
- 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*, 1991, and Pratt and Gibbons, *Concepts of Non-parametric Theory*, 1981. Hahn and Meeker describe at length the different statistical intervals including their properties, applications, and 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 2003 assessment stated:

> *"Tolerance intervals are inappropriate for environmental monitoring. The main issues here are 1) is the true concentration greater than some specified safe 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 this assessment stated:

"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):

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

"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. It 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.

While the theoretical construct underlying Currie's critical level is clear and straightforward, EPA recognizes that estimating this level from limited data is less straightforward and the choice of an appropriate statistical methodology involves policy judgements that might legitimately differ for different uses of the MDL.

3.3.4.1 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-α)100% confidence. Random variables that are the lower and upper ends of the interval, L_1 and L_2, respectively, are referred to as tolerance bounds. A tolerance bound is therefore the endpoint of an interval of random length that is determined on the basis of having a specified probability of 1-α that its coverage of the population is at least equal to a specified value P. The quantity 1-α is referred to as the confidence level for the interval and P is the minimum proportion of the population contained in the interval. Tolerance bounds are not estimators of values such as a mean or a percentile but rather bounds that are always guaranteed to contain 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 their interpretation and application.

In effect, the determination of what, if any, interval to use is a policy decision. The choice 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 levels make tolerance intervals useless.

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

Statistical intervals can be, and have been by a number of authors, 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.

The error rates in ASTM's IDE Standard Practice are based on statistical tolerance intervals (i.e., the nominal Type 1 error rate is 5% (5%=100%-95%), and the nominal Type 2 error rate is 10% (10% = 100%-90%)). Several stakeholders have commented that the use of a tolerance interval approach can protect, at a 99% level of confidence, against false positives and false negatives, and that tolerance intervals become increasingly important with a decreasing sample size. For example, if the sample standard deviation is determined with 7 measurements and all sources of variance are properly represented in the 7 measurements, then there is approximately a 5% chance that the true population standard deviation will be more than two times the sample standard deviation. For a typical ICP determination of 20 or more elements this means that at least one is likely to have a calculated MDL two times lower than it should be. Obviously the false positive rate for this element will be large.

The use of prediction and/or tolerance limits in setting detection and quantitation limits should be evaluated in the context of the specific application 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.

3.3.5 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 indicates a coneentration 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 suppression of results that are 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 be useable for certain purposes).

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 loses information about low-level measurements and can introduce bias into the data set. If all low values are eliminated, then the average (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, because they convey useful information about the distribution of results.

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

Although the Snperfund approach provides the greatest degree of flexibility for data users, it should be nsed 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 2003 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 an incentive to the analytical and regulated community to pursue measurements at the lowest levels which analytical methods are capable of achieving. However, EPA also cautions that this approach effectively censors measurements made below the MDL and could yield an overestimate of the concentration of the analyte of concern.

Several stakeholders have requested that EPA provide specific guidance and procedures regarding data censoring and reporting, particularly when data are reported for compliance evaluation. EPA notes that the decision to censor data is a data reporting and data use policy issue, not a laboratory issue. This holds without regard to 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, but also establishes this limit as the reporting level for non-detects to better encourage development of lower MDLs. However, EPA also recognizes that laboratory methodologies and data reporting and use policies are interrelated.

3.3.6 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 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 when analyzing some samples, the associated measurement results may stand out as outliers. A graphic example is provided in Figure 3-1, which shows measured concentrations of aluminum versus spike concentrations for analytieal results obtained using EPA Method 1620. At a spike concentration of 250 µg/L, one of the measured values is approximately 750 µg/L. This result visually stands out from the rest of the values, and may be an outlier.

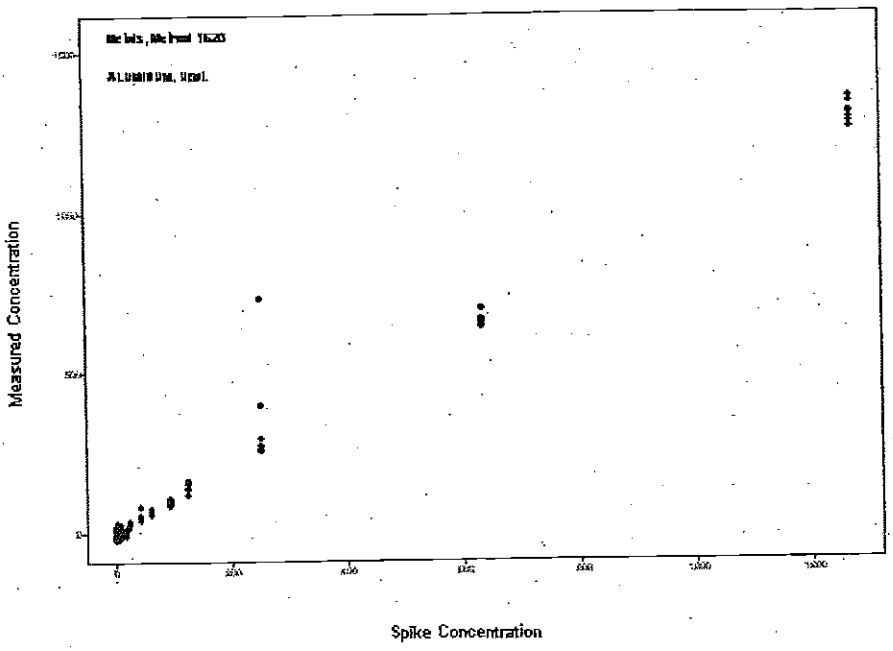

Figure 3-1

Stakeholders commenting on EPA's assessment of detection and quantitation procedures generally believed that outliers should be identified and removed from data used to determine detection and quantitation limits. Commenters added that, although it would be helpful to have specific instructions for identifying outliers, application of the instructions should be optional (i.e., to the discretion of the data user).

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 described in ASTM Practice D-2777.

Because extreme values can be expected to occur on occasion, it may not be appropriate to exclude them from the results used to develop detection or quantitation values. As recommended in the ASTM procedure, the first step is to contact the laboratory to try to determine and resolve the cause. 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, or there is a known or suspected analytical error, 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. EPA believes that results that are associated with spurious errors that cannot be corrected will invalidate the measurement and should not be incorporated into the MDL

determination.

3.3.7 Detection and Quantitation Studies

3.3.7.1 Study Design

The issues associated with the design of detection and quantitation studies include:

- how effectively a selection of spike concentrations can be used to correctly determine which model type should be used to model variability,
- the extent to which 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.

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 would provide 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 this paper does not apply to all models, it indicates principles that do apply. Increasing the number of replicate measurements 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.

It is also important to note that modeling of variability introduces modeling error, and direct measurements of the variance in the region of interest may provide a more appropriate estimate of variability, especially where the change in variance over this region is small.

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 to 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 (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 replicate samples analyzed at each concentration. 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 1 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 or practical. 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). Although this was done for the Episode 6000 study, this approach would not be practical in most cases. 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 2003 assessment 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.

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.

3.3.7.2 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.

3.3.7.2.1 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 3-2 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 of detection and quantitation.

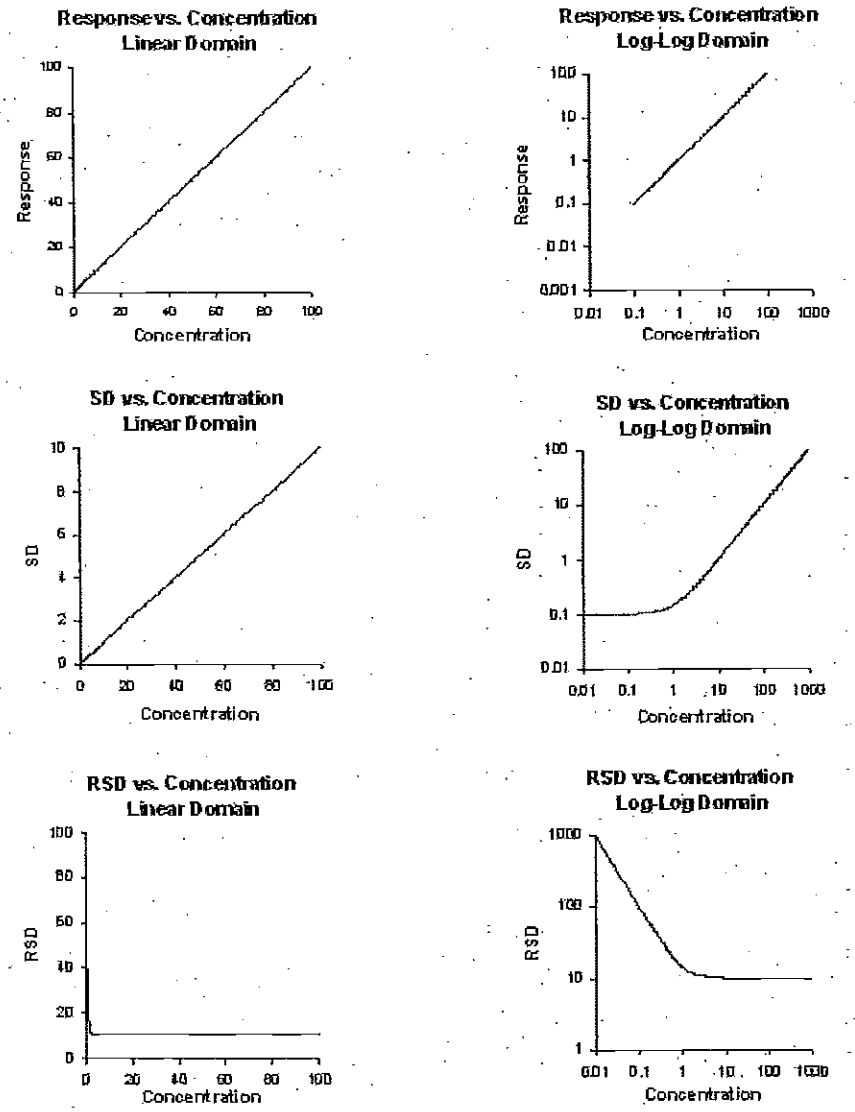

Figure 3-2

3.3.7.2.2 Detection Limits Using Variability at Low Concentrations

The middle two graphs in Figure 3-2 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 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.

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

Both 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 at any concentration level 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 measured by Method 625. The RSD at this quantitation limit could be 30, 50, or 70 percent. Limiting the RSD associated with the quantitation limit to some arbitrary 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 a 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 enough to a compliance limit that measurement error could influence the compliance determination. On the other hand, having widely varying RSDs for different analytes within the same method may be confusing to permitting and enforcement authorities who may not appreciate the subtleties of reporting violations in light of the underlying RSDs.

Another means of arriving at a limiting RSD is to graph RSD versus concentration, as shown in the bottom two graphs of Figure 3-2. 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.

3.3.7.2.4 Criteria for Selecting Models

Both statistical and graphical procedures have been proposed for selecting between models for predicting measurement results based on 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.

Both graphical and statistical criteria will be strongly affected by the number and choice of spike concentrations used to fit the different models. Too few spike concentrations will lessen the statistical power of significance tests for slope and curvature from which decisions on the type of model will be made. In addition, the amount of subjectivity with which decisions are made using graphs increases when fewer concentration levels are used. For example, the judgement of whether a residual plot depicts "random scatter" is essentially impossible when only five concentration levels are used (i.e., the residual plot will include only five points). The number of results from which standard deviations are calculated will also have an effect on how models are selected. This set of results may include analysis of multiple replicates at a single laboratory or analysis of one or more replicates from multiple laboratories. If data are obtained from too few laboratories or replicates, the standard deviation estimates will be less reliable, which could lead to incorrect model selection based on statistical or graphical criteria.

3.3.7.2.5 Assessment of Current Models

EPA plotted variability versus concentration data to evaluate the extent to which 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*, of the February 2003 Technical Support Document (EPA-821-R-03-005, February 2003). Four sets of composite scatter plots for all combinations of analytical technique, analyte, and 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. 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 straight line can be visualized easily. All the graphs are contained in attachments to

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 will increase proportionally with concentrations in the higher concentration range.

Because the large number of resulting plots makes it difficult to draw general conclusions, for the most part, conclusions must be considered on a case-by-case basis.

3.3.7.3 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 resulting in 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 estimation and 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 especially where variability does not change much over the region of interest.

Chapter 4
Evaluation Criteria

This chapter presents the criteria developed by EPA as a means for evaluating and 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, including EPA's needs under CWA programs. A total of six criteria were established, and are discussed in Sections 4.1 - 4.6. The six evaluation criteria are:

Criterion 1: The detection and quantitation limit approaches should be scientifically valid.
Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.
Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.
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 well-operated laboratories.
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.
Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act (CWA), and should support state and local obligations to implement measurement requirements that area at least as stringent as those set by the Federal government.

Section 4.7 presents additional principles recommended by stakeholders commenting on EPA's assessment.

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 pertaining to assessments of scientific validity when considering the admissibility of expert scientific testimony[2]. These conditions also are directly relevant to EPA's needs.

Some stakeholders supported the use of objective criteria for determining scientific validity, but questioned the appropriateness of using criteria that were designed for courts and juries to support scientific decisions made by scientific experts. EPA carefully reviewed the Court's conditions for demonstrating the scientific validity of an expert's reasoning or methodology, and believes that these conditions are appropriate for demonstrating the scientific validity of any scientific approach or methodology, including those that might be used to establish detection and quantitation limits under CWA. EPA further believes these criteria are consistent with the EPA Science Policy Council's assessment factors for evaluating the quality of scientific and technical information (EPA 100/B-03/001, June 2003), including the extent to which technical information and data are peer reviewed and appropriately tested. However, EPA is willing to consider alternative or supplemental criteria for evaluating scientific validity as it moves forward with the stakeholder process.

Stakeholders agree that detection and quantitation levels should be based on sound scientific principles, and note that low-cost and/or simple approaches should not he selected if inaccurate or unmeasurable limits may result. Stakeholders also noted that some of the conditions listed above (e.g., the condition that an approach or methodology should have attracted widespread acceptance within a relevant scientific community) have the potential for favoring concepts already adopted and required by regulatory agencies. EPA agrees that this is a valid concern, and therefore, will consider the overall validity and practicality of new approaches.

4.2 Criterion 2

Criterion 2: *The approach should address realistic 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.

Early 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 an EPA research laboratory in Cincinnati, Ohio with specialized experience in the analytical chemistry of drinking water. This laboratory also established method detection and quantitation limits which, in many instances, initially could not be achieved in other laboratories. Over time, however, the difficulty in acbieving these limits was overcome as analysts gained experience with the use of these new methods.

Stakeholders have suggested that detection and quantitation limits be developed using data from multiple laboratories in order to account for the routine inter- and intra-laboratory variability that can occur over time. Although compliance measurements are made in single laboratories, EPA agrees that detection and quantitation limits in methods that will be widely used by many laboratories should consider

---

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

these sources of variability. For this reason, after the development in a single laboratory of a new or modified analytical method with an initial estimate of detection and quantitation limits, EPA's Office of Science and Technology evaluates and verifies these limits in multi-laboratory studies.

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. Single-laboratory studies at a specialized research laboratory may produce detection and quantitation limits that are lower than those produced by studies that gather data from many laboratories that may or may not be experienced with the method. EPA believes that a realistic expectation of method and laboratory performance likely lies somewhere in between that provided by a specialized single-laboratory study and that provided by an interlaboratory study with no pre-qualification requirements. Estimates of detection and quantitation limits should consider the inherent variability of the measurement process, but not be based on the lowest common denominator, e.g., data from inexperieneed or unqualified analysts and laboratories.

EPA expects that laboratories must meet some minimum standards of performance and experience with a method, and sets performance criteria in methods. Examples of such criteria include measures to demonstrate that a laboratory is producing accurate results at a eoneentration 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 detect pollutants at low concentrations (i.e., at the method 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 of laboratory performance is the means for rejection of outliers to more accurately estimate routine variability. True outliers can occur in laboratory data, and some means of resolving outlier issues should 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 QA/QC 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 tbe exclusion on technical grounds.

In examining each approach against this criterion, EPA will evaluate whether the approach can be used to provide 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 for determining detection and quantitation limits at a single laboratory should be simple, with detailed instructions, and cost-effective to implement (i.e., it should be reliable and "laboratory-friendly"). Laboratories that use detection or quantitation procedures range from large laboratories and laboratory chains with a wide range of technical capabilities, to small laboratories operated by one or a few people with limited statistical skills. While this range of laboratory capability

places a premium on simplicity and ease, EPA agrees with stakeholders that data reliability and quality are also important. A suitable approach or procedure for detection and quantitation incorporates the right balance between the need for valid data and the need for the procedure to be simple and inexpensive to perform. EPA also believes that if a procedure is complicated, it will be prone to error in 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 EPA wishes to encourage development and use of innovative techniques that improve measurement performance or lower measurement costs, the Agency should 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:

- demonstrate that a method works in a particular matrix at the levels of concern (i.e., demonstrate the absence of matrix effects),
- characterize improvements in measurement capabilities in terms of detection and quantitation capabilities, and
- characterize the detection and quantitation capabilities 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 detection and quantitation capability of an analytical method in every possible matrix in which it may be used. At a minimum, it is unlikely that EPA or any other organization or laboratory 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 detection and qnantitation 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 real-world or reference matrices (e.g., reagent water, sand, diatomaceous earth) to establish QC acceptance criteria that verify performance of the method at mid-to-high concentrations. Each analytical method includes QC acceptance criteria for such real-world and reference 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 would prefer to utilize detection/quantitation concepts that allow for similar characterization of detection/quantitation capabilities in representative matrices and that are supported by simple, cost-effective procedures that would allow individual laboratories to evaluate the effects of specific matrices on these capabilities on an as needed basis. Because methods approved at 40 CFR part 136 already contain a suite of quality control procedures and QC acceptance criteria that control laboratory performance, EPA believes that it is not necessary to verify detection and quantitation limits in eacb and every batch of each and every matrix analyzed. Rather, such testing can be done on an as-needed basis when it is suspected that matrix interferences may preclude reliable measurements at low levels.

Another consideration influencing the need for simplicity and practicality is that measurement capabilities generally improve over time. As is discussed in Section 3.1 of this document, and as has been noted by stakeholders, this is attributable to a variety of factors, including:

- increased staff experience with a given technique,
- technological upgrades or improvements in the instrumentation used for analysis, and
- development of new instrumentation or techniques that improves detection/quantitation, 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 with these analyses than a laboratory that rarely performs these measurements. 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 and measure previously undetected pollutants. Such improvements 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 by providing practical and affordable procedures for evaluating method performance.

In evaluating this criterion, EPA will favor affordable and easy-to-use approaches and procedures that allow analysts to 1) determine matrix-specific variations when necessary, based on realistic data, and 2) demonstrate lower detection and quantitation limits associated with improvements in measurement capabilities. Procedures for establishing the detection capabilities of new methods or associated with improved measurement capabilities should be practical enough to encourage such development. However, EPA recognizes that some uses for detection and quantitation limits may require a more comprehensive approach involving multiple laboratories. These procedures should specify the nature, minimum number, and concentration levels of the samples to be used, 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 estimate the theoretical 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 by a well-operated laboratory as being present really is present. Historically, approaches to making detection decisions have set the criterion for detection at 99 percent confidence (i.e., with 99% confidence that the analyte concentration is greater than zero). 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 (this is a Type I error), of one percent. The procedure also should be capable of generating a detection level when the substance of interest is not present in a blank and/or when instrument thresholds are used

in routine operation. A well-operated laboratory is a laboratory that routinely monitors performance through QC analyses, control charts, and other measures to rapidly identify and correct deteriorating or poor performance, and with analysts experienced with method sample preparation, analysis, and detection procedures.

In evaluating this criterion, EPA will favor approaches and procedures that reflect routine analytical conditions in a well-operated laboratory.

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. In evaluating different approaches to estimating quantitation limits, 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.

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 tolerance 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 (i.e., a measurable or recognizable signal), 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 3.3.5. Given the difficulty in achieving consensus on an appropriate means of establishing a quantitation 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 perceive that they understand a quantitation limit to mean the lowest concentration at which an analyte can be identified and quantified with some degree of certainty. This understanding necessarily involves use of the sound judgment of a qualified analytical chemist.

Therefore, EPA will continue to monitor developments on this subject, and if appropriate, re-evaluate this issue if and when it becomes practical and 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 laboratories develop and report individual estimates of uncertainty. Criterion 5 reflects this belief.

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 evaluate detection limit or quantitation capabilities for methods approved at 40 CFR part 136 for the following applications:

- Ambient and effluent permitting and compliance monitoring under NPDES and the pretreatment program and under State and local programs,
- Quality control in analytical laboratories, and
- Method development, promulgation, and modification.

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, as recommended by some stakeholders commenting on EPA's assessment. In the 2003 assessment, EPA stated that this would increase confusion, record keeping burdens, and laboratory testing burdens. EPA also stated that data generated under a single procedure can be used for development of detection and quantitation limits that are applicable to more than a single use. For example, the data used to determine the capabilities of multiple laboratories using a given method may also be used to develop method-specific detection and quantitation limits. For these reasons, EPA recommended the adoption of a single pair of related detection and quantitation procedures used to address all or most Clean Water Act applications. Some stakeholders recommend the use of different approaches for different CWA applications. For example, these stakeholders would prefer a more rigorous approach to determining detection and quantitation limits for method development than for verifying laboratory performance. They would like to include a procedure that is based on a multilaboratory approach rather than a single laboratory approach to define detection and quantitation capabilities of analytical methods. EPA recognizes that the complexity and statistical rigor appropriate for a detection and quantitation approach for method development and validation would be greater than that needed for demonstrating laboratory proficiency. EPA plans to seek additional stakeholder input on whether different approaches are needed for different CWA purposes (see Chapter 6).

Although EPA prefers to identify a manageable set of detection and quantitation limit approaches to meet CWA needs, EPA believes that any reasonable approach advanced by other organizations should be acceptable for use provided it meets the needs of the specific application for which it would be used. Allowing use of detection and quantitation approaches developed by other organizations provides the stakeholder community with increased measurement options that may help reduce measurement costs or improve measurement performance for specific situations. This approach also is consistent with the intent of the National Technology Transfer and Advancement Act.

The Clean Water Act authorizes State or local governments to implement specific aspects of the Act, with the provision 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 to ensure consistency with approaches adopted 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. 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, and in most cases, 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 (e.g., analytical methods are not available that can reliably measure at levels necessary to protect receiving water), permitting authorities must decide how to evaluate and report pollutant concentrations at these levels. 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.

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. As discussed in Chapter 6, EPA believes that additional discussion about this criterion is appropriate based on negative comments from stakeholders regarding the use of a single pair of detection and quantitation limit approaches to meet all CWA needs.

4.7 Consensus Principles

Some stakeholders commenting on EPA's assessment of approaches to detection and quantitation expressed their support of a set of "consensus principles" submitted by 36 signatories representing industry and laboratory communities. EPA agrees with certain consensus principles such as the principle that detection and quantitation levels should be based on sound scientific principles, and that low-cost and/or simple approaches should not be used if invalid data will result (see Criterion 1 above). As another example, EPA incorporated routine variability, the rate of false positives, precision, and matrix effects in several criteria, and considered these aspects in its assessment of detection and quantitation concepts. Some of these consensus principles are included in the criteria discussed in this chapter. Other consensus principles have clarified or highlighted existing aspects of approaches to detection and quantitation and provide a framework for additional consideration.

For ease of consideration, the consensus principles recommended by commenters have been separated by EPA into technical and policy considerations and include:

Technical Considerations

- Detection and quantitation levels must be based on sound scientific principles. Low-cost and/or simple approaches must not be selected if inaccurate compliance determinations or unmeasurable permit limits may result.
- The definition of "quantitation" must account for both precision and bias.
- Detection limit procedures must take into account the variability and bias of method blank results.
- False positives (Type I errors), false negatives (Type II errors), and precision must all be addressed by detection concepts and reporting of analytical results for regulatory purposes.

- Precision, bias, and qualitative identification (where appropriate) must all be addressed by the definition and concepts of quantitation and by the reporting of analytical results for regulatory purposes.
- Detection limit procedures must include procedures for ongoing demonstration of sensitivity, preferably incorporated into the routine analytical quality control as a check against false negatives.
- Detection and quantitation levels must take into account routine inter- and intra-laboratory variability within a laboratory over time.
- In its procedures for establishing detection and quantitation levels, EPA must develop guidance on how to account for the effects of various matrices.

Policy Considerations

- The L_C, L_D, and L_Q are three distinct points, each of which has unique criteria that must be satisfied. For consistency with international standards, EPA must adopt the definitions of L_C (critical value), L_D (detection limit), and L_Q (quantification limit) of IUPAC (International Union of Pure and Applied Chemistry) that are being adopted by international standards organizations (e.g., the International Organization of Standardization (ISO)).
- The definitions of and procedures for determining detection and quantitation levels must take into account that quantitation levels are used as regulatory compliance levels in NPDES permits.
- EPA should specify consensus standard procedures for establishing significant figures and for rounding data.
- EPA must strive for consistency across all EPA offices (the Office of Water, Office of Research and Development, Office of Ground Water and Drinking Water, and Office of Solid Waste and Emergency Response) in defining and applying detection and quantitation levels.

Chapter 5
Assessment

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),
- ISO/IUPAC minimum detectable value (MDV; Section 5.1.5),
- American Council of Independent Laboratories (ACIL) Critical Value (ACIL L_C; Section 5.1.6),
- United States Geological Survey (USGS) Long-term Detection Limit (USGS LT-MDL; Section 5.1.7), and
- Inter-industry Analytical Group (IIAG) Sensitivity Test and Full-Range Validation Study (Section 5.1.8).

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's 2003 assessment of detection and quantitation limit approaches focused on approaches developed or published by ASTM International, the American Chemical Society (ACS), ISO/IUPAC, and EPA. Stakeholder commenting on the initial assessment suggested that EPA should include additional approaches in the next assessment. In addition to the initial four approaches, EPA has included three additional approaches in this Revised Assessment document. These approaches are: the long-term MDL developed by USGS, a new detection limit procedure developed by the American Council of Independent Laboratories (ACIL), and a paired approach involving a sensitivity test and full-range validation study submitted by the Petitioners (the Inter-industry Analytical Group). Several commenters advocated these as approaches that more realistically reflect measurement variability. These additional approaches are discussed and assessed in Sections 5.1.6 - 5.1.8 of this chapter.

Some stakeholders commenting on EPA's 2003 assessment believed that the evaluation criteria used by EPA were written to favor the MDL and ML over other approaches to detection and quantitation. EPA disagrees. The criteria were written to reflect EPA's needs for detection and quantitation approaches under the CWA, and it is not necessary that an acceptable approach meet all of these criteria under all conditions. Because the MDL and ML were developed to address EPA's needs, it should not be surprising that the MDL and ML procedures generally meet the criteria EPA set out to assess detection and quantitation procedures. EPA has frankly assessed the MDL and ML against these criteria and notes

that the MDL and ML procedures do not meet all of these criteria under all operating conditions (see Sections 5.1.1 and 5.2.1 below). Due to the variability and unpredictability inherent in measurement science, it is unlikely that any procedure would meet all of EPA's criteria under all conditions. However, EPA is open to further discussions with stakeholders about the appropriateness of the evaluation criteria described in Chapter 4, in particular, the issue of whether EPA should adopt different approaches for different applications, as discussed in Chapter 6.

5.1 Detection Limit Approaches

Sections 5.1.1 through 5.1.8 describe EPA's assessment of eight 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: Of the six assessment criteria in Chapter 4 four (Nos. 1, 2,3 and 6) pertain to both detection and quantitation limit approaches. One criterion (No. 4) pertains only to detection limit approaches, and one criterion (No. 5) pertains only to quantitation limit approaches. Therefore, the following discussion of each detection and quantitation limit approach applies only the five applicable criteria.

5.1.1 Evaluation of the MDL

Section 5.1.1.1 is 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 apply to detection limit approaches.(i.e., Criteria 1-4, and Criterion 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 = 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, and
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 (S^2_A is the larger of the two variances). 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 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.

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. Over the years, as stakeholders have sought to improve upon or identify alternative procedures, the MDL has been the subject of a number of studies and comparisons, including this assessment. As a result, the MDL is one of the most widely tested detection limit procedure in the history of detection approaches. (See Appendix A for a list of literature references concerning the MDL and other detection limits.)

Critics of the MDL have noted that the detection limit produced with the MDL procedure can vary depending on the spike level used. It is true that an initial MDL may be calculated using any spike level, regardless of how high. Although a high initial spike level will result in an initially high MDL, the self-correction cheek in the MDL procedure requires the final spike level to be within a certain range of the reported (i.e. final) MDL. Specifically, Step 1 of the MDL procedure focuses the spiking level on the lowest concentration at which measurements can be made, and the factor of 5 requirement in Steps 3 and 4 assure that the determined MDL will be at or near this concentration. Therefore, the requirements included in Steps 1, 3 and 4 guard against an artificially high MDL being produced due to the choice of a high initial spike level. EPA also recognizes the concern that the iterative procedure in step 7, which provides a reality check on the results obtained in steps 1 - 6 is optional. EPA will consider whether additional guidance on this aspect of the procedure is needed.

In preparation for the assessment of detection and quantitation approaches, EPA 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 MDL procedure is capable of reasonably estimating the lowest level at which measurements can be made when the factor of 5 requirement is met.

One of the stakeholders commenting on EPA's 2003 assessment suggested that the MDL failed to meet this condition because EPA should have tested it in "real world" matrices. EPA does not agree with this suggestion for several reasons. First, it is not practical or possible to test detection limits in every real world matrix, and there is no consensus as to which real world matrix would represent an appropriate real world matrix for testing. Second, 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. In theory, the sample could be diluted to dilute the target pollutant, but in practice sample dilution would also likely dilute any interferences that might be present, thereby defeating the purpose of using a real world matrix. The current EPA approach, which exhaustively tests the MDL procedure in a reference matrix using multiple techniques and ten different concentrations that span the entire region of interest, is more than adequate to constitute "testing" of the MDL procedure. On the other hand, where data suggests that matrix interferences may significantly affect achievable quantitation and detection limits, this should be considered by a permit writer on a case by case basis.

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). The MDL procedure has been included at 40 CFR part 136, appendix B since 1984. Values resulting from this procedure have been included, published, and tested in many analytical methods since promulgation, including methods publisbed by EPA and other Federal agencies, and by consensus standards organizations and trade associations such as ASTM International, and APHA, AWWA, and WEF.

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 0.01(1%). Therefore, the MDL meets this condition. In addition, the Step 7 of the MDL procedure suggests calculating a 95% confidence interval for the determined MDL, providing additional estimation about the uncertainty (i.e., error) of the MDL determined using the procedure.

The US Geological Survey (USGS) provided a dataset of spiked and blank sample data that EPA used to evaluate the error rate associated with the MDL. (Error rates associated with the ACIL and USGS detection limit procedures also were evaluated and are discussed in Sections 5.1.6 and 5.1.7.) Although the sample size was insufficient to conclusively demonstrate the error rate of the MDL, the results suggest the actual error rate is close to the intended 1%. In this case, the observed mean error rate was 2.9%. Readers are referred to Appendix B for a discussion of two factors affecting this estimate - relatively small sample size and some added long-term variability.

In the 2003 assessment, EPA suggested deleting the procedure for calculating the 95% confidence interval because it appeared to be rarely, if ever, used. No commenters specifically agreed with this suggestion, but several commenters responded that it should be retained. One commenter, arguing in favor of the procedure, stated that "It has long been recognized that a 95% confidence level is appropriate to establish standards and other regulatory requirements." Considering these comments, EPA now believes there is no compelling reason to remove this procedure.

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. 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.

Notwithstanding the preceding, the MDL procedure would be improved with additional guidance, particularly with respect to initial spike levels, handling outliers, the optional reasonableness step (Step 7), and multi-analyte test methods. The MDL procedure does not contain a 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 simply select the best seven results of a dataset. 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 guidance on how and when to address each of these issues.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The MDL 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 detection approaches. 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, it is the most widely used approach of detection within the environmental chemistry community.

Stakeholders commenting on EPA's 2003 assessment of detection and quantitation procedures noted that the extent to which the MDL has been used is a result of EPA's approval and inclusion of the procedure in 40 CFR part 136, and does not necessarily demonstrate that the MDL procedure produces an accurate assessment of detection. EPA agrees that the extent of use could be attributed, in part, to promulgation of the procedure at 40 CFR part 136. For this reason, EPA has not relied on widespread use of the MDL as a sole or over-riding argument for its continued use. Rather, EPA views widespread use of the MDL as one of many factors to be considered when evaluating which concept or concepts best meet the Agency's needs under the Clean Water Act. For example, EPA agrees that the ability of a procedure to produce an accurate assessment of detection capabilities is an important consideration, and addresses this issue repeatedly throughout the assessment. In this chapter, for example, the ability of a procedure to produce an accurate assessment of detection capabilities is addressed in

- Criterion 1, condition 3, which concerns error rate,
- Criterion 1, condition 4, which concerns use of standards to control operation of the procedure,
- Criterion 2, which addresses the ability of the procedure to realistically reflect laboratory and method performance, and

- Criterion 4, which addresses the ability of the approach to identify the concentration at which users can be confident a substance reported as present is really present.

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

The MDL procedure is designed to demonstrate laboratory performance with a given analytical method, and can be applied to a broad variety of physical and chemical methods. 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.

When the MDL procedure is followed as intended (i.e., all sample processing and analysis steps of the method that are applied to routine analyses are included in determination of an MDL), the demonstrated MDL will include some of the routine variability associated with the laboratory and the method.

Stakeholders commenting on EPA's assessment stated that, because the MDL procedure is performed in a single laboratory, on the same day, by the same analyst, in a single matrix, using a minimum of 7 replicates, the procedure does not account for all sources of variability. These commenters believe that the procedure does not address inter- or intralaboratory, long-term, concentration range, analyte/method, or matrix variability. EPA notes that the MDL procedure does not include the restrictions noted by these stakeholders (e.g., users are not restricted to use of only seven replicates; to analysis of all replicates on the same day; or to determination of MDLs only in reagent water). The MDL procedure includes, for example, instructions for determining a matrix-specific MDL and specifies that the procedure requires a complete, specific, and well-defined analytical method. However, EPA also recognizes that in practice the MDL procedure may be performed in the manner described by these comments and that doing so will limit the amount of routine variability reflected in the results.

The MDL procedure provides users with the flexibility needed for multiple applications. For example, if a laboratory desires to evaluate its performance using a single method to analyze a particularly difficult matrix over a period of time (e.g., one year), the MDL procedure allows such an evaluation. However in some cases, the MDL procedure might benefit with specific provisions for including sources of variability that may not be addressed when following the minimum requirements of the MDL procedure.

Stakeholders commenting on EPA's assessment directed most of their concern at the lack of long-term variability in the MDL procedure. These commenters pointed to the American Council of Independent Laboratories (ACIL) procedures for calculating the critical level and long term-MDL (LT-MDL) and to the US Geological Survey's (USGS) procedures for generating their LT-MDL. These procedures include the collection of blanks over a long period of time to include this source of variability. The commenters stated that the lack of long-term variability leads to underestimates of Currie's critical value (L_c), and one commenter included sets of blanks collected over 3 months to demonstrate this effect.

EPA assessed the effect of long-term variability on calculated limits by simulating multiple 7-replicate subsets from the full dataset offered by the commenter, and compared these short-term critical levels to the critical level calculated using the full data set. Although the range of days from which the sets of 7 replicates were simulated varied from between one week to greater than 3 weeks, a graphical

analysis of the data did not reveal any effect of time on the resulting L_c. The total number of blanks also did not seem to have an effect on the percentage of short-term L_c results that exceeded the overall L_c. Details of this assessment are provided in Appendix C, along with possible reasons why expected differences were not observed.

As noted in Section 3.3.3 of this RAD, a larger number of replicates will yield better estimates for standard deviations, and therefore, better estimates of Currie's L_c and EPA's analogous MDL. However the analysis performed in Appendix C demonstrates that MDLs estimating L_c based on 7 replicates are not biased low. These values are merely less precise than those based on a larger number of replicates. As noted previously, the current MDL procedure does not restrict laboratories to using 7 replicates (to the contrary, the procedure specifies a minimum of 7 replicates), nor does it restrict laboratories to performing the replicates on a single day. Laboratories that wish to perform more tests or to conduct their tests over a longer period of time should be encouraged to do so.

Due to the variability inherent in measurement science, instrumentation, and the humans conducting analyses, laboratories may routinely obtain detection limits that are lower or higher than those obtained in another laboratory. Thus, when an MDL is determined during method development, it is important to determine that MDL in more than one laboratory to ensure the MDL published in the method reflects demonstrated expectations of method performance in a community of laboratories. It is not necessary for this community to include the entire universe of all possible laboratories that might desire to practice the method. Rather, during the stages of method development and validation, this community only should include well-operated laboratories with analysts who are experienced with the techniques used in the method, and have some familiarity conducting all of the steps in the new method before generating MDLs that will be published with the new 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 MDLs 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 (laboratory MDLs ranged from 0.04 to 0.18 ng/L), the value published in the promulgated version of Method 1631.

The MDL procedure addresses demonstrated expectations of laboratory and method performance, including routine variability, and users should not be restricted to the minimum requirements of the MDL procedure. If the MDL procedure is 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 procedure is among the most practical and affordable procedures that have been suggested for determining detection limits because of the reasonable number of minimum replicates (seven) and the relative ease with which the spiking experiments can be designed and the resulting data

analyzed. The MDL is designed for use by a single laboratory, and can be performed by a single analyst using a single instrument. And the MDL procedure also allows MDLs from several analysts or instruments within a laboratory, or between laboratories to be pooled and provide an estimate of the range of MDLs that might be routinely expected.

Use of the optional iterative procedure would increase the number of analyses by at least seven each time the procedure is implemented. If the procedure is implemented two times in reagent water, a minimum 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).

5.1.1.2.4 Criterion 4: The detection level approach should estimate the theoretical 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 MDL meets this condition as described under Section 5.1.1.2.1, Condition 3 of this document in many cases. However, EPA recognizes that there are cases where this does not hold, and that users of the MDL procedure see this as a significant problem. EPA sees merit in blank correction procedures developed by ACIL and USGS to address these cases. In future stakeholder consultations, EPA plans to discuss these and other alternative solutions to this problem.

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 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 published in 1997 by ASTM International as standard D6091. The IDE was developed with support from members of the regulated industry to provide a comprehensive detection limit procedure that addressed the concerns of the regulated industry, statisticians, and analysts involved in ASTM Committee D19 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 tbis 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 % eonfidence upper statistical tolerance limit.

5. Nontrivial Amount of Censored Data - this section addresses the effect of "non-detects" or "less-than." 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-than 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. The Electric Power Research Institute provided input into the design of EPA Method 1631 and 1638 Validation Studies for the purpose of calculating IDEs and IQEs. EPRI also calculated IDEs and IQEs based on these data. These two datasets include a total of ten metal analytes and therefore do not cover a wide range of analytical techniques and methods. Other than these two datasets, 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 [IDE_0] and multilaboratory analyses of multiple concentrations of each matrix of interest surrounding IDE_0). 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, 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, it meets 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 does not meet 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)

One stakeholder, however, stated that concerns about the complexity and subjectivity in the IDE (and IQE) procedures were unimportant, in part, because IDEs calculated using different models were generally very close, and in part because "user-friendly software is available that will automatically perform the IDE and IQE calculations." To consider the merit of this comment, EPA calculated single-laboratory variants of the IDE using each of the four major model types using the Episode 6000 data set, and true interlaboratory IDEs for each model type using the Method 1631 and 1638 interlaboratory study data sets. Results of these calculations, along with the RSDs between the different IDE values obtained for each analyte, are presented in Appendix B. Based on the calculated RSDs, there is a large amount of variability between the single-laboratory variants of the IDEs calculated using the different models. Generally, the IDEs calculated using the constant model were much greater than those calculated using the

other models. The hybrid model generally yielded the lowest IDEs, and the IDEs calculated using the hybrid and exponential models were quite similar for some analytes, but quite different for others. While one might hope that the variability between models would decrease if interlaboratory variability were included in the calculations (as designed), EPA found this was not the case. To the contrary, RSDs between the IDEs calculated from the interlaboratory datasets suggest that variability between model estimates appears to increase when the additional variability between laboratories is included.

To evaluate the commenters' statement that the complexity and subjectivity of the procedures was not important because the calculations can be automatically performed using "user-friendly software," EPA evaluated the two software packages offered hy the commenter. One package was a DOS-based program called "QCalc" and the other was an Excel spreadsheet that calculates IDEs based on Excel functions, macros, and the Solver add-in function. EPA calculated single laboratory variants of the IDE for a random subset of 20 analytes from the Episode 6000 study using 1) the QCalc package, 2) the Excel spreadsheet, and 3) the suite of SAS programs EPA has been using to calculate IDEs as part of this assessment. To ensure that differences between results were due to the programs themselves, the same data were used for each program. Results of this comparison are provided in Appendix B to this Revised Assessment Document.

One immediate problem was that comparisons could not be made between IDEs calculated using QCalc and the other software packages for all of the models because the QCalc package only performs the IDE calculation using two of the models (exponential and hybrid). The ASTM IDE procedure suggests that one of three models be used (constant, linear, and exponential). No explanation was provided as to why the software was limited to two models instead of three, or why one of the two models (i.e., the hybrid model) used in the software was not one of the three models recommended by ASTM. (The hybrid model used in QCalc is recommended by ASTM for calculation of an IQE but not for an IDE.)

Although similarities were generally observed among the various software packages when the same model type was applied to the same set of data, EPA did observe strong differences in the values calculated using the hybrid model across the various software programs. The Excel values generated using the hybrid model were slightly higher than those determined using EPA's programs and approximately twice as high as those determined using QCalc. Possible explanations for these differences are given in Appendix C.

Perhaps the most significant problem with the assumption that use of the automated software packages alleviates the complexity and subjectivity in the IDE procedure is that the various packages do not always select the same model for the same set of data. ASTM's IDE procedure (D 6091) specifies that the fitting to the constant model should be attempted first. If this fitting fails, a straight-line model should be attempted, and if that fails, the exponential model should be fitted and evaluated for reasonableness and lack of fit. EPA's SAS programs were coded to preferentially select the constant, linear, and exponential models for the IDE, according to this scheme. However, QCalc and Excel packages each follow a different scheme. As a result, the EPA and QCalc programs selected the same model type to calculate the IDE for only 1 of the 20 analytes, the Excel and QCalc programs selected the same model type for only 6 of the 20 analytes, and the Excel and EPA programs selected the same model type for only 1 of the 20 analytes. Details and possible explanations for these underlying differences can be found in Appendix C.

Based on these differences in selecting and fitting models, it does not appear that the two available software programs remove all complexity and subjectivity from the IDE calculation. Instead, they appear to introduce new issues by using steps not included in the ASTM procedures. The results support EPA's conclusion that such conditions make it impossible to estimate the actual error associated with the IDE, and that the IDE, as currently constructed, does not meet this condition 3.

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.

There are several "gray areas" in the published procedure. The most significant of which 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 problems noted in preceding Condition 3 concerning the use of different model selection strategies among three different programs (the QCalc and the Excel software packages provided by a commenter and EPA's SAS programs) is a direct reflection of the subjective nature of model selection likely to result from the lack of guidance in the procedure. 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.

The IDE standard gives procedures that are inconsistent with procedures 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.

Although the IDE is supported by a published procedure, EPA found that the procedure will not adequately control its operation because of the degree of subjectivity involved in implementing the procedure and inconsistencies with its IQE counterpart. Therefore, the IDE does not meet this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The IDE was published by ASTM, International in 1997. ASTM, International is a voluntary consensus standards organization that constitutes part of the relevant scientific community, however, seven years after publication no new or revised ASTM standard has included detection limits using the IDE approach. EPA is not aware of an IDE that has been published in the open literature or in an analytical method. Thus, the IDE partially meets this criterion.

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

The IDE procedure, D6091, 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 Section 6.3.3.2 of D6091) 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. Information about single-laboratory (or within-laboratory) variability is very important because assessments of laboratory performance is based on the variability (uncertainty) of the dat produced at that laboratory. Compliance measurements are made in a single laboratory and the results are reported with the uncertainty (variability) associated with that dataset.

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 as those used in the Excel spreadsheets discussed in Section 5.1.2.2.1. 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 IDE_0, 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 estimate the theoretical 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 .01 by several orders of magnitude. This protection against false positive results is excessive and would yield numerical values of little practical value for making the detection decision.

Although there is an estimate of Currie's Lc included in the IDE procedure, it is unclear where the detection decision is made (it really should be an ICE/IDE procedure). If one focuses on the IDE and not the Lc estimate, this criterion not met. Therefore, it is not clear whether the IDE would meet this criterion (No. 4).

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 the CWA, it may not be appropriate for all uses. The IDE is an implementation of Currie *detection level* or *minimum detectable value*, and may in practice yield results 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 the lack of a supporting procedure for establishing an LOD, and a conceptual dependence on the variability associated with measuring blanks. For example, there is no detailed instructions, similar to those in the IDE and the MDL procedures, 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 meets this condition because the LOD definition was published in the peer-reviewed journal *Analytical Chemistry* in 1980 and 1983.

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 does not meet this condition, because it 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.

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, the LOD meets this condition.

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

The LOD approach is designed to address realistic 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 does not meet this criterion, because it is not supported by a clearly defined procedure for establishing the LOD.

5.1.3.2.4 Criterion 4: The detection level approach should estimate the theoretical 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 ($S_t - S_b$) > 3σ." Therefore, the LOD meets this criterion.

5.1.3.2.5 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 LOD values, the ACS LOD does not meet this criterion because it cannot be used in a regulatory context unless it is assumed to be functionally equivalent to the MDL (i.e., use the MDL procedure 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 prodnce a response, it is impossible to calculate a CRV because the standard deviation of multiple zero results is zero. One solution for testing the approach is to assume that the CRV is about 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 does not meet 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 counted the number of papers in published journals that include CRV values, but EPA's literature search for detection and quantitation approaches did not produce many 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 realistic 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 realistic expectations of laboratory and method performance in theory, but in practice, provides no direct means for demonstrating this performance. Therefore, 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 does not meet this criterion.

5.1.4.2.4 Criterion 4: The detection level approach should estimate the theoretical 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.

CRV suggests $\alpha = 0.05$, resulting in $1-\alpha$ of 0.95 or 95 % probability of detection . Therefore, the CRV does not meet 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 does not meet this criterion because it cannot be used in a regulatory context.

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:

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

where:
z is the score variable
1-β is the false negative error rate, recommended at 5 % (β = 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, the ISO/IUPAC detection limit in this assessment will be referred to 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 approacb 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, the MDV probably 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 does not meet 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 realistic 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 does not meet 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 does not meet this criterion.

5.1.5.2.4 Criterion 4: The detection level approach should estimate the theoretical 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 reduces the probability of false positives to a value smaller than 1% by several orders of magnitude. . 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 does not meet 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 does not meet to meet this criterion because it cannot be used in a regulatory context.

5.1.6 Evaluation of the American Council of Independent Laboratories (ACIL) Critical Value

During the comment period on the February 2003 assessment document, the American Council of Independent Laboratories (ACIL) submitted a procedure that was developed to address errors, which are referred to as "bias", that may arise under certain conditions when estimating detection limits. The ACIL

procedure separates estimation of the detection limit into two cases; cases where analyses always produce a numeric result (i.e., even so-called "blank" samples produce a signal), and cases where tests do not always produce a numeric result (i.e., blank samples appear to produce no signal). Blanks that do not produce a signal may do so either because they really are blanks, or the instrument is suppressing the signal. For convenience, EPA refers to these as Case I and Case II, respectively. Analysis of metals with inductively coupled plasma optical emission spectroscopy (ICP-OES) is an example of ACIL Case I, and analysis of organic pollutants with gas chromatography/ mass spectrometry is an example of ACIL Case II. Although the ACIL procedure appears to be a work-in-progress, it has some interesting approaches for the use of blanks, and is similar in some respects to the USGS LT-MDL procedure.

5.1.6.1 Description of the ACIL Approach and Procedure

For Case I analyses, ACIL offers procedures for calculating a limit that approximates Currie's critical value (L_C) and procedures for calculating a limit that approximates Currie's detection limit (L_D). As discussed in Chapter 2 and noted again in Section 5.1.5 above, Currie's L_D was 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 L_D 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 (L_C), which is determined from measured values. The approach of a true concentration L_D and its associated allowance for false negatives is of little practical value in making the actual detection decision. For this reason, EPA focused its assessment of ACIL's procedure on the ACIL version of Currie's critical value rather than the ACIL version of L_D.

For Case II analyses, ACIL suggests a procedure that does not rely on the Currie L_C and L_D framework. Instead, the procedures involve picking an initial spike value, adjusting that level up or down based on whether the analyte was detected, and spiking seven replicates at the new level.

A brief description of each procedure is provided below.

ACIL's Case I Critical Value (ACIL L_C)

As with EPA's MDL, the ACIL L_C is an attempt to approximate Currie's critical value. Whereas EPA's MDL is based on the standard deviation of blank samples spiked with low levels of the target analyte, ACIL's Case I detection limit is based on the standard deviation of the blank samples run as part of the laboratories ongoing QC program. (Because some methods will not yield a result when blanks are analyzed, ACIL's L_C procedure is accompanied by a spiked sample approach that can be used with those methods.)

Although ACIL does not formally define ACIL L_C, a footnote 2 to the procedure describes it as

> "very similar to Currie's critical level, L_C (Anal. Chem. Vol. 40 No 3, March 1968, p586). It is the level at which there is a given confidenee that a result can be distinguished from the blank."

Key features of the ACIL Case I detection limit are as follows:

- The procedure relies on the use of blanks (instead of low-level spikes) to estimate standard deviation.
- When a sufficient number of blanks are used in the calculation, the mean blank result is added

into the calculation to account for high bias exhibited in the blanks.

- ACIL states that at least 7 blanks should be used, but recommends more (as many as 100). If the number of replicates is small, ACIL recommends using a tolerance interval calculation for estimating ACIL L_c. Instead of defining exactly what constitutes a "small" number of replicates, ACIL loosely defines it as fewer than 20 or 30. The confidence level for the tolerance interval also is not specified. If the tolerance level approach is used, the mean blank result is not included in the calculation (unlike the calculation used when there are more than 20 to 30 results).
- If multiple instruments are to be used for the same test and will have the same reporting limit, a minimum of 7 blank results from each instrument should be used, and the results should be combined to generate the standard deviation.
- It is acceptable (and expected) that some results will have negative values, and these negative values should not be censored. Outlier removal is allowed, using a statistically accepted test, if appropriate cautions are taken to guard against excessive or inappropriate rejection of data.
- ACIL provides a verification procedure that is based on comparing the variance of the blank results to results from a new set of blanks.
- ACIL suggests reporting all results that meet or exceed the ACIL L_c.

The formula for ACIL L_c is:

$$LC = \overline{X} + (t_{0.99,n-1} * s)$$

Where $\overline{X}$ is the mean of blank results
s is the standard deviation of blank results, and
n is the number of blank results

ACIL's Case II Detection Limit

For Case II analyses, ACIL's procedures involve picking an initial spike value, adjusting that level up or down based on whether the analyte was detected, and spiking seven replicates at the new level. Details of this procedure are as follows:

- Unlike the procedures used for methods that yield numeric results, ACIL Case II procedures would use spiked samples to determine the detection limit for methods that do not always yield numeric results.
- An initial spike value is chosen based on prior experience. (Detailed guidelines are not provided.)
- One replicate at this level is analyzed; if the analyte is detected, a new sample should be prepared at ½ the initial spike value. If the analyte is not detected at the original level, a new sample should be prepared at 2x the initial spike value. This process is repeated to find the lowest level that can be detected
- Once that level is identified, a minimum of 7 replicates spiked at the lowest level at which that analyte was detected are analyzed, and the replicates must be analyzed in three different batches. If the analyte is detected in all replicates, the Case II MDL is set to this spike value. If the analyte is not detected in all 7 replicates, at least 7 additional replicates are prepared and analyzed at twice this value. If the analyte is detected in all 7 replicates spiked at this higher concentration, the Case II MDL is set to this higher spike value. This process is repeated until the analyte is detected in all 7 replicates.

- The ACIL procedure includes a verification step that consists of spiking the reference matrix at 1 to 3 times the Case II MDL (or 1 to 4 times for multi-analyte methods) to verify that the analyte(s) can be detected. If not, the test is repeated at increasing spike levels until detection, and setting the Case II MDL to the level where the analyte(s) were first detected.
- ACIL suggests reporting all results that meet or exceed the Case II MDL.

5.1.6.2 Assessment of the ACIL L_C against the Evaluation Criteria

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

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

Condition 1: It can be (and has been) tested. Although ACIL had not conducted an exhaustive study to test the ACIL L_C, ACIL did apply data generated from member laboratories to the procedure in order to calculate ACIL L_C values. ACIL also compared those values with values produced by EPA's MDL using the same procedure. The results of these tests are included in the public docket supporting this assessment. As part of its own assessment, EPA also tested the procedure using data obtained from the U.S. Geological Survey. In this testing, EPA generated ACIL L_C values, compared those values with values produced by other procedures, and calculated error rates associated with each of the values. Given these studies, the ACIL L_C meets this condition.

Condition 2: It has been subjected to peer review and publication. The ACIL procedure was developed to support ACIL's comments on EPA's 2003, assessment, and it has been subjected to limited peer review within ACIL's member community. Although ACIL references publication of the procedure on the ACIL website, EPA made repeated attempts to locate the procedure on the website over a period of several months, and was unable to locate it. Given the limited peer review beyond the member community, and the lack of publication in a publicly accessible medium, the ACIL procedure does not meet this criterion.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The ACIL procedure meets this condition. According to the formula used for estimating ACIL L_C, the error rate, is specified by α, with a suggested value of 0.01(1%). EPA was able to evaluate this error rate using a small set of data provided by the US Geological Survey. The data included spiked and blank sample results for 18 pollutants, most of which were analyzed by multiple methods, yielding 75 unique analyte/method combinations. For each combination, 25 - 52 blanks were provided. EPA used these blanks to calculate the ACIL L_C, and compared the results of individual blanks with the calculated ACIL L_C. (Details of this assessment are provided in Appendix C.) In theory, no more than 1% of the blanks should have produced a result that exceeded the ACIL L_C. Although the sample size was insufficient to conclusively demonstrate the error rate of the ACIL L_C, the results suggest the actual error rate is close to the estimate of 1%. In this case, the observed mean error rate was 1.9%, and the highest error observed for any method/analyte combination was only 3.8%. Given the small sample size, failure of a single blank could (and did) result in a 3.8% failure rate, suggesting that this study may yield an error rate that is larger than that which would be observed in a larger study. Regardless, it is clear that the ACIL L_C meets this condition because the estimated error rate is given as part of the procedure, and the actual error rate can be calculated through studies such as the one described above.

Condition 4: Standards exist and can be maintained to control its operation. The ACIL L_C is supported by a written procedure (standard) to control its operation. However, the procedure appears to be in draft form, is somewhat difficult to follow and interpret, and contains inconsistencies and ambiguities that are typical of a draft document. In particular, the instructions for Case II are not as clear or detailed as those for Case I.

As an example of the inconsistencies, a footnote to the ACIL L_C states that a tolerance interval will be a more reliable estimate of the ACIL L_C if the number of blanks is small (i.e., fewer than 20 or 30). This implies that the tolerance interval calculation and preferred ACIL L_C will converge as the number of blank results increases. However, this is not the case. The tolerance interval calculation will almost always yield a higher result than the preferred ACIL L_C calculation. The only way that the tolerance interval calculation will result in an ACIL L_C that is either lower or equal to the original ACIL L_C is when blank contamination is high (unlike the preferred ACIL L_C calculation, the tolerance interval calculation does not include the mean of the blanks). It is unclear why the reliability of one calculation compared to the other depends on the number of blank results.

An example of the ambiguities in the procedure is that the alternative calculations, such as the tolerance interval calculation, are presented as suggestions instead of requirements. This could lead to confusion, as now written, if, as ACIL recommends that, the ACIL L_C be used as a reporting limit.

A different type of ambiguity in the procedure concerns the lack of sufficient detail to ensure consistent application. For example, it is not clear exactly when the tolerance interval calculation is to be used because the procedure defines small as 20 - 30 samples. When would 20 samples be sufficient and when would 30 samples be sufficient? Moreover, the tolerance interval calculation does not specify the confidence level used. In an example, both 99% and 95% are given as possibilities. In comparison, the critical value calculated in ASTM's IDE sets the confidence level at 90%. Setting the confidence level at 99% will yield an ACIL L_C value between 11% and 37% higher than one calculated at 95%, based on the numbers of blank results for which the tolerance interval approach is suggested.

Given these problems, the current ACIL procedure does not meet this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The ACIL L_C was supported by a large number of commenters, most of whom came from the ACIL member community or the environmental laboratory community. Of note, however, is that supporters included instrument vendors, consultants, and several members of the industrial community, including the Inter-industry Analytical Group which offered its own approach to detection and quantitation and which has been highly supportive of the ASTM IDE and IQE approaches. Therefore, EPA believes that the ACIL L_C meets this condition.

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

The ACIL L_C is designed to address realistic expectations of laboratory and method performance, including temporal variability, instrument variability, analyst variability, and high bias observed in blank results. Based on EPA's analysis of the ACIL L_C presented in Appendix C, EPA believes that the approach meets this criterion provided it is interpreted and applied consistently. (Concerns about the need for clarification of the procedure are described in Section 5.1.6.1, Condition 4).

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

The ACIL L_C meets this criterion. It is similar to the EPA MDL procedure, but it relies on the use of QC data generated during routine laboratory operations, thereby making it even more cost effective than the MDL.

5.1.6.2.4 Criterion 4: The detection level approach should estimate the theoretical 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

Footnote 2 to the ACIL procedure describes the ACIL L_C as "very similar to Currie's critical level, L_C (Anal. Chem. Vol. 40 No 3, March 1968, p586). It is the level at which there is a given confidence that a result can be distinguished from the blank." According to the formula used for estimating ACIL L_C, the error rate is specified by α, with a suggested value of 0.01(1%). This alpha value means that, if the analyte is not present in the sample, it will be reported as present (i.e., a false positive) no more than 1% of the time. In lay terms, this suggests 99% confidence that, if a substance is reported as present, it really is present. Therefore, the ACIL L_C meets this criterion.

5.1.6.2.5 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

If EPA's interpretation of the ACIL procedure is correct, the ACIL L_C appears to meet this criterion.

5.1.7 Evaluation of the USGS Long-term Detection Limit (USGS LT-MDL)

The USGS National Water Quality Laboratory (NWQL) began using the EPA MDL procedure in 1992. USGS NWQL has since developed a variant of the MDL called the long-term MDL (LT-MDL) that has been in routine use by the NWQL since 1999. The procedure for calculating the LT-MDL is described in Section 5.1.7.1 below. Section 5.1.7.2 describes EPA's assessment of the LT-MDL against the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1 - 4, and Criterion 6).

5.1.7.1 Description of the USGS Approach and Procedure

As described in the USGS Open-File Report 99-193, the LT-MDL is a modification of the EPA MDL designed to "capture greater method variability," thereby leading to higher detection limits than those obtained using the EPA MDL procedure. As described by USGS, and noted in Chapter 2, the LT-MDL is based on many of the same fundamental assumptions as the MDL, namely:

1. Normal data distribution,
2. Constant standard deviation from the spike concentration down to zero, 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 LT-MDL is determined using low-level spikes of reagent water. The three primary differences between the EPA MDL and the USGS LT-MDL procedures are:

1. Larger minimum number (24) of spike samples,
2. Longer time period, and
3. Combining results from different instruments and analysts in the determination of the LT-MDL.

The USGS Open File Report does not provide an example of the exact calculation used for the LT-MDL. EPA originally presumed that the standard deviation of the results from the 24 spiked sample analyses is multiplied by the Student's *t*-value appropriate for 23 degrees of freedom (t=2.499).

However, USGS comments submitted in response to EPA's assessment of detection and quantitation approaches included a copy of a presentation from the USEPA Region 6 12th Annual Quality Assurance Conference, in Dallas, Texas in August 2002. That presentation provided significant additional information on the calculation of the LT-MDL. Specifically, the LT-MDL uses "F-pseudosigma" (F_σ) in place of S, the sample standard deviation, used in the EPA MDL calculation. F-pseudosigma is a non-parametric measure of variability that is based on the interquartile range of the data. The LT-MDL may be calculated using either the mean or median of a set of long-term blanks, or from long-term spiked sample results, such that:

$$LT\text{-}MDL = M + (t_{0.99,n-1} \times F_\sigma)$$

where:

M = mean or median of blank results
n = number of spiked sample results, and
F_σ = F-pseudosigma, a nonparametric estimate of variability calculated as:

$$F_\sigma = \frac{Q_3 - Q_1}{1.349}$$

where:

Q_3 and Q_1 = the 75th percentile and 25th percentile of spiked sample results, respectively.

USGS believes that the use of F_σ provides an estimate that is more robust and not influenced by outliers.

Like the EPA MDL, the LT-MDL is designed to limit the chance of a false positive result to ≤ 1%. However, the LT-MDL is designed to be used in conjunction with a "laboratory reporting level" (LRL) as part of an overall reporting scheme for the NWQL. As described by USGS, the LRL is set as a multiple of the LT-MDL. The multiplier varies, depending on the mean/median recovery of the analyte in the spiked samples used for the LT-MDL. If the mean or median recovery is 100%, then the multiplier is 2. At 75% mean or median recovery, the multiplier increases to 2.7, and at 50% recovery, the LRL multiplier increases to 4. In each of these cases, the multiplier is essentially equivalent to dividing twice the LT-MDL by the mean recovery (i.e., 2.7 LT-MDL ~ 2 LT-MDL/75%).

The LRL is designed to achieve a risk of ≤ 1% for both false negatives and false positives. The reporting scheme used at the NWQL with the LT-MDL and LRL does not censor results at the LRL, and the laboratory reports all results between the LT-MDL and the LRL with a lab-specific flag.

The USGS presentation from the 2002 meeting describes how USGS enhanced the LT-MDL procedure by using their large volume of uncensored blind laboratory blank data as a reality-check on the LT-MDL derived from spiked reagent water samples. In cases where the standard deviation used to calculate an LT-MDL based on blind blank data is significantly different (especially when greater) from the standard deviation used to calculate the spike-based LT-MDL, the blank data are used to calculate the LT-MDL. Blind blank data also are used to evaluate whether the calculated LT-MDL requires an off-set correction for blank bias, i.e, [LT-MDL = (S x Student's *t*) + median or mean blank concentration]. This offset is similar, but not identical, to the ACIL Case I procedure described in Section 2.3.3 of this document. The LT-MDL offset correction compensates for a blank distribution that is not centered at zero (an assumption in the EPA MDL procedure).

The NWQL has found that this blank bias correction to the LT-MDL is especially important for blank-limited analytes, including some metals, total organic carbon, phenol, and nutrients. In practice, the NWQL recalculates the LT-MDL annually, and compares the results between years using Levene's test of equal variance, which they have found to be less influenced by departures from normality than the F-test – an important consideration given that the LT-MDL is based on a non-parametric estimate of variability.

5.1.7.2 Assessment of the USGS LT-MDL against the Evaluation Criteria

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

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

Condition 1: It can be (and has been) tested. The LT-MDL meets this condition.

USGS has tested and used the LT-MDL since October 1998. Evaluation and use of the LT-MDL began with four methods in use by the NWQL for low-level volatiles by GC/Ms, trace metals by ICP/AES, Kjeldahl nitrogen, and phosphorus. According to the Open File Report, the LT-MDL was scheduled for testing in 17 additional methods, including semivolatile organics, organochlorine pesticides, organophosphorus pesticides, pesticides analyzed by HPLC, metals by ICP/MS, metals by GFAA, and ion chromatography.

EPA used a combination of blank and spiked data submitted by USGS to compute the USGS LT-MDL and compare it to the EPA MDL The blanks were analyzed by USGS over a period of one year and represented a combination of 78 analytes, methods, and matrices, while the spiked sample results represented 39 analytes, methods, and matrices. The analytes were all metals or wet chemistry parameters such as phosphorus and nitrate/nitrite.

Condition 2: It has been subjected to peer review and publication. The LT-MDL does not appear to meet this condition.

Information on the LT-MDL is relatively limited and EPA is not aware of additional USGS publications beyond Open File Report 99-193 and the August 2002 presentation. EPA did not identify any additional publications regarding the LT-MDL in its earlier literature search. The Open File Report itself does not provide any indication that it was subject to a peer-review process.

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 value of 0.01(1%). Therefore, the LT-MDL may meet this condition.

In its evaluation of USGS data submitted as comments (see Appendix C) EPA found that the mean percentage of blanks results that exceeded the detection limit estimate (LT-MDL) ranged from 3.7% to 4.4%, depending on whether the mean or median blank result was used to estimate LT-MDL. These rates exceeded that of the EPA MDL Therefore, although EPA's evaluation found that the error rate for the LT-MDL exceeded the theoretical error rate designed into the procedure, the error rate can be estimated from actual data.

Condition 4: Standards exist and can be maintained to control its operation. The LT-MDL may partially meet this condition, in that the NWQL may have formal procedures in place that more fully describe the LT-MDL. However, as noted above, the information in the Open File Report does not include an explicit formula for calculation of the LT-MDL, nor are other details of the overall procedure, such as the choice of spiking levels, provided in a clear and consistent fashion. The August 2002 presentation provides critical information about the use of Fσ that is not present the Open File Report. The LT-MDL could meet this criterion, if the procedure were clearly documented by USGS.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The LT-MDL does not meet this condition

EPA believes that the LT-MDL is only used at the NWQL. Several commenters, including ACIL, suggested that EPA examine the USGS LT-MDL more closely, specifically in regards to its inclusion of long-term variability. There is, however, no evidence in the comments that the concept has achieved a large following among laboratories or other agencies.

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

EPA believes that the LT-MDL meets this criterion because it incorporates the variability of responses over a long time period, and where a laboratory has multiple instruments and analysts running the same analysis, it incorporates variability across instruments and analysts.

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

The LT-MDL partially meets this criterion. However, the LT-MDL is not a detailed readily available "procedure". Also, the LT-MDL requires data collected over a 12-month period. Given that many State regulatory programs require that laboratories provide an annual demonstration of capabilities, including demonstrating their detection limits, the use of the LT-MDL would have to he limited to those laboratories that already have a year's worth of data available. Some other single-lab approach would have to he used for an initial demonstration of method performance.

5.1.6.2.4 Criterion 4: *The detection level approach should estimate the theoretical 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*

According to the formula used for calculating the USGS LT-MDL, the error rate is specified by α, and the LT-MDL is designed with a value of 0.01(1%). Because the method uses a nonparametric estimate of S, it may not always yield a 1% false positive rate. EPA empirical analysis indicates false positive rates in the range of 3.7% to 4.4%. This compares favorably with the performance of other methods. Thus, the LT-MDL adequately meets this criterion at least in practice.

5.1.7.2.5 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 LT-MDL may meet this criterion. The LT-MDL is designed as part of a broader reporting scheme and it is unclear that EPA, States, and local authorities would be willing or able to use results reported according to that scheme in enforcement scenarios (e.g., "flagged" data).

5.1.8 Evaluation of the Inter-industry Analytical Group (IIAG) Full-Range Validation and Sensitivity Test

In December 2002, the Inter-Industry Analytical Group (IIAG) submitted a proposal to EPA that recommends (1) a sensitivity test intended to "replace the MDL as a test of whether an individual laboratory is performing adequately," and (2) an interlaboratory validation study design intended to characterize precision and accuracy of methods used for regulatory compliance. Although their approach was received too late for consideration prior to publication in the 2003 Assessment Document, EPA provided notice of the approach, requested public comment on it, and agreed to evaluate the IIAG approach in updating the 2003 assessment. Section 5.1.8.1 describes the IIAG approach, and Section 5.1.8.2 describes EPA's assessment of the IIAG approach against the applicable evaluation criteria.

5.1.8.1 Description of the IIAG Approach and Procedure

Full Range Validation

IIAG has proposed that, EPA commit to performing interlaboratory method validation studies designed to produce a "full range" of data, including precision and accuracy, from the point of instrument detection to the upper end of the working range. IIAG has indicated that "such a full range validation will enable EPA to consider DL/QL options in light of data quality objectives without being constrained by a limited database." IIAG suggests that, at a minimum, EPA should commit to performing such full range validation studies for all new methods that it develops and that all organizations submitting new methods for EPA approval should be required to provide the full range data as well.

Sensitivity Test

IIAG also has proposed that EPA consider the use of a "sensitivity test" instead of the MDL to demonstrate that a laboratory is capable of performing according to EPA expectations at the lower range of a test method. IIAG's suggested process for developing this test is as follows:

- EPA would first identify the lowest concentration at which the entire analytical system gives a recognizable signal and acceptable calibration point.
- EPA would then select a simple dilution of that concentration, and develop QC criteria based on the test results from several laboratories performing the test at that dilution (in the same way that QC criteria are developed by EPA for initial precision and recovery demonstrations in methods

such as Method 1631).

- Laboratories could then perform an "Initial Performance Demonstration" (IPD) of their capability to achieve the desired sensitivity by (1) analyzing several replicates of the same sample dilution (using the full method), (2) using the results to compute the standard deviation, and (3) confirming that the results fall within the QC criteria range. IIAG emphasizes that the dilution level would not be considered the detection level, but rather a performance level.

IIAG further suggests that this multi-replicate IPD test would be verified on an ongoing basis. To minimize complexity, IIAG suggests that the ongoing test be conducted at the same spike level as their "Initial Performance Demonstration." IIAG did not suggest a specific frequency for conducting these ongoing tests.

Finally, IIAG suggests that EPA commit to using this IPD sensitivity test in lieu of the MDL, and that EPA express a willingness, subject to funding availability or a third party commitment, to perform testing as necessary to develop "sensitivity" QC criteria, and to modify the few existing Part 136 methods that require the MDL for IPD.

Section 5.1.8.2 below discusses EPA's evaluation of the scientific elements proposed IIAG approach.

5.1.8.2 Assessment of the IIAG Approach against the Evaluation Criteria

The following six subsections discuss the IIAG approach and procedure in the context of the six evaluation criteria. The first three criteria apply to both detection and quantitation limits, Criterion 4 applies to detection limits only, Criterion 5 applies to quantitation limits only, and Criterion 6 applies to both. Because the IIAG full range validation and sensitivity test approach applies to both types of limits, all 6 criteria are discussed below.

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

Condition 1: It can be (and has been) tested. To EPA's knowledge, the IIAG sensitivity test approach has not been tested by any organization, including IIAG. The IIAG approach is still a rough framework, and basic details, such as the number of replicates required and the actual spiking levels to be used, still need to be specified. Testing of the approach in its current framework is possible but would be very expensive, one might have to conduct tests with multiple spiking levels and with varying numbers of replicates, for example, to be sure that the tests will reflect the final sensitivity test procedure. If the procedure were refined to describe the exact steps and requirements, it could be tested more efficiently.

IIAG's full validation study approach can be and has been tested. For example, EPA conducted a full interlaboratory validation study of Method 1631 prior to promulgating the method at 40 CFR 136. That study, which involved 12 participating laboratories, yielded an overall mean percent recovery of 93 and an overall relative standard deviation of 13 percent across all samples.

IIAG has stated that "Although the full-range interlaboratory is aimed at characterizing a method's ability to quantify rather than to detect a pollutant concentration, the study could be used to establish an interlaboratory detection level as well" and "The best solution for performing a full-range validation to establish detection and quantitation levels and precision and bias for promulgating nationwide standards and compliance levels is the ASTM IDE/IQE approach."

The ASTM IDE and IQE are constructed by fitting a model to variability versus concentration data, rather than being derived from the standard deviation of replicate measurements of a single concentration level. As discussed in Section 5.1.2 and detailed in Appendix C, EPA used data from the Episode 6000 study to compare IDEs calculated using data from all 16 concentration levels reported to IDEs calculated using data from only 5 of the concentrations (i.e., at 5, 10, 20, 40, and 80 times the standard deviation of replicate measurements of a blank sample or the lowest level at which measurements could be made). Results of the comparison are summarized in Table 9 of Appendix B to the draft TSD. The results show that the median 16-point IDE is approximately 1.3 times greater than the median 5-point IDE, indicating that data resulting from measurements of concentration levels in the region of detection and quantitation in some cases may yield lower IDE's than data from a wider range of concentration data.

EPA refers readers of this document to Sections 5.1.2.2.1, 5.2.2.2.1, 5.2.2.2.2, and Appendix B for a discussion of additional reasons why EPA believes the ambiguities and inconsistencies in IDE/IQE procedures preclude these procedures from being the best solution for performing a full range study to estimate detection and quantitation limits.

Condition 2: It has been subjected to peer review and publication. The IIAG procedure does not meet this criterion. EPA is not aware of any peer review or publication of the document in a peer reviewed journal. The IIAG document was submitted directly to EPA by the Petitioners, and EPA made the document available to the public for comment.

Condition 3: The error rate associated with the procedure is either known or can be estimated. At present the IIAG's approach consists of a proposed framework rather than a detailed procedure. It lacks key specifics, such as how many replicates would be used in the IPD phase of the test, and what spiking levels would be used. IIAG suggests that EPA would select these levels, and suggests "probably 4 - 7" for the number of replicates.

While IIAG suggests the dilution would be a simple dilution of the lowest calibration standard, offering "1/3 or ½, for example", they also state that "It is not absolntely necessary to reduce the spike level below the lowest calibration point, however, and the sensitivity test could be performed with a spike at the lowest calibrations standard instead of at a dilution of it." No guidelines are offered for which of these levels (or other levels) shonld be chosen, nor are guidelines offered for the number of replicates needed.

Given the lack of detail, the current framework would be subject to different interpretations by different readers or users, and the error rate associated with the procedure would vary depending on how the procedure was implemented. Because the error rate is neither known, nor can it be estimated, the IIAG approach does not meet this condition.

The IIAG procedure is a framework with interesting aspects for further consideration by the full scientific and regulatory community. EPA would be willing to work with IIAG and other stakeholders to identify the details needed to augment this framework to where it would meet this condition.

Condition 4: Standards exist and can be maintained to control its operation. As previously noted, the IIAG approach consists of a proposed framework rather than a detailed procedure framework, and lacks key details that are needed to control its operation. Given the lack of detail, the current framework does not meet this condition.

Again this procedure is a framework with interesting aspects for further consideration by the full scientific and regulatory community. EPA would be willing to work with IIAG and other stakeholders to identify the details needed to augment this framework to where it would meet this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The IIAG procedure does not meet this condition. It was suggested by a limited group of the relevant scientific community (industry firms that comprise the "Inter-industry Analytical Group" and whose wastewater discharges are regulated under the Clean Water Act), and comments on the their approach were mixed. Excluding comments submitted by IIAG itself, EPA received comments from:

- Three electric power producers whose discharges also are regulated under CWA,
- Two publicly owned wastewater treatment systems which regulates industrial discharges to their system under CWA and whose own discharges are subject to regulation under CWA,
- Two commercial environmental laboratories that utilize the methods and detection limit procedures approved at 40 CFR 136 to serve their client's needs,
- One trade council, and
- One private citizen.

All three electric power firms supported the IIAG approach. The two publicly owned treatment systems offered mixed reviews. One supported the sensitivity test and offered suggestions for further consideration; the other opposed the sensitivity test but offered limited support of the interlaboratory validation studies, suggesting that they be limited to the relatively small group of priority pollutants whose water quality based effluent limits are below the method reporting levels. Both of the environmental laboratories were opposed to the IIAG approach, and the trade council suggested that it should be used "as an alternative procedure for dischargers to implement... on a site-specific basis, at their discretion", noting that "As an alternate method, facilities would be able to deal with this on a case-by-case basis and would not need to utilize numerous laboratories to develop the more elaborate detection limits and quantitation limits that the IIAG proposes".

Given these comments, it would appear that acceptance may be widespread within the industrial discharger community, but it is not widespread among the *entire* relevant scientific community.

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

In principle, the IIAG sensitivity test meets this criterion because it is intended to provide realistic information about laboratory and method performance, both with an initial demonstration and with follow-up demonstrations that provide information concerning routine variability. However, and as previously noted, the procedure is not sufficiently detailed to allow laboratories to meet this criterion. To clearly meet this criterion, detailed specifications to allow for consistent implementation of the procedure throughout the laboratory community need to be developed.

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

If the IIAG framework was developed into a detailed procedure, this sensitivity approach would meet this single laboratory criterion. This could complement the IIAG full range validation study, which does not meet this criterion because it is an interlaboratory procedure. The sensitivity test, once detailed, could be performed by a single laboratory and used to evaluate method performance.

5.1.8.2.4 Criterion 4: *The detection level approach should estimate the theoretical 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.*

Because the spiking level to be used in IIAG's sensitivity test is not defined it is not possible to evaluate whether that test meets or does not meet this criterion. IIAG also suggests that a full-range validation study should be used to establish an interlaboratory detection limit, and recommends use of the ASTM IDE procedure as the best means of doing so. If this is the case, the full range validation study would fail this criterion for the reasons given in Section 5.1.2.2.4 regarding the IDE.

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

The IIAG's proposed sensitivity test requirement is likely to meet this criterion once details regarding the procedure are specified. Depending on the spiking levels that are specified in the final procedure, however, it is very likely that the IIAG sensitivity test 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.

5.1.8.2.6 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*

IIAG's suggested use of a full range validation study meets this criterion because such validation studies provide useful information about the performance of the method. As noted previously, EPA typically conducts interlaboratory validation studies at multiple concentrations ranges before promulgating a method for nationwide use at 40 CFR part 136. However, for the reasons discussed elsewhere in this document, EPA does not agree that data collected across the full range of the method should be used to establish detection or quantitation levels.

In the absence of a detailed procedure that could be use to fully evaluate IIAG's, it is difficult to determine if the IIAG sensitivity test meets this criterion.

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. These criteria are Nos. 1 -3, 5 and 6; No. 4 only is applicable to detection limits.

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 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, or 5) x 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 3.18 multiplier is derived by dividing 10 by the value of the t-statistic for seven replicates. Laboratories that choose to perform MDL studies with more than the required minimum of seven replicates follow the instructions in appendix B of 40 CFR part 136 to select the t-statistic value for the number of replicates used. Therefore, the 3.18 multiplier for the ML calculation should be proportionally adjusted. Similarly, the Student's *t*-value is 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 tbe 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.

EPA also has tested the MDL and ML procedure with ten 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. EPA recognizes that additional guidance may be necessary on the selection of the initial spiking level and uses of the iterative procedure.

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 can be verified (tested) by a laboratory when it calibrates the instrument used to analyze samples by the method.

One of the stakeholders commenting on EPA's 2003 assessment suggested that the ML failed to meet this condition because EPA should have tested it in "real world" matrices. EPA does not agree with this suggestion for several reasons. First, it is not practical or possible to test detection limits in every real world matrix, and there is no consensus as to which real world matrix would represent an appropriate real world matrix for testing. Second, 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. In theory, the sample could be diluted to dilute the target pollutant, but in practice sample dilution would also likely dilute any interferences that might be present, thereby defeating the purpose of using a real world matrix. The current EPA approach, which exhaustively tests the ML procedure in a reference matrix using multiple techniques and ten different concentrations that span the entire region of interest, is more than adequate to constitute "testing" of the ML procedure. On the other hand, where data suggests that matrix interferences may significantly affect achievable quantitation and detection limits, this should be considered by a permit writer on a case by case basis.

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 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, 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. If rounding is not considered, the error can be easily estimated. The calculation is still straightforward, but tedious, when the ML rounding procedures are employed. Given these caveats, the ML partially meets this condition.

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 embodied in the definition.

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

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

The ML procedure meets this criterion. It 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).

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.

Due to the variability inherent in measurement science, instrumentation, and the humans conducting analyses, laboratories may routinely obtain limits that are lower or higher than those obtained in another laboratory. Thus, when an ML is determined during method development, it is important to determine that ML in more than one laboratory to ensure the ML published in the method reflects demonstrated expectations of method performance in a community of laboratories. It is not necessary for this community to include the entire universe of all possible laboratories that might desire to practice the method. Rather, during the stages of method development and validation, this community only should include well-operated laboratories with analysts who are experienced with the techniques used in the method, and have some familiarity conducting all of the steps in the new method before generating MDLs that will be published with the new 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 meets this criterion. It 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), by a simple multiplication of the MDL and a application of a rounding procedure.

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 depends on identification of 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.

Notwithstanding the preceding, analysis of Episode 6000 data (see appendices) produced anomalous results from two methods (EPA 502.2 and 524.2) that employ instrument thresholds. For 17% of EPA 502.2 and 49 % of EPA 524.2 analytes the calculated ML was below the concentration at which all seven spiked replicates were detected, i.e. less than the lowest MDL spike. The Episode 6000 dataset is not reflective of a typical compliance measurement or method development study because the range of concentrations studied encompassed several orders of magnitude and included concentrations well below the MDL. This atypical range was employed to push the limits of the instrumentation and the theory underlying determination of the variability of measurements.

In a qualified operating laboratory, or during a method development study, if MLs were calculated to be less than the concentration at which all seven spiked MDL replicates were detected, the laboratory would take corrective measures. When a method is developed for EPA's CWA program, each laboratory in a multi-laboratory study would consult with EPA and take corrective measures, such as calibration adjustments so that reported MDLs are above the signal threshold. In these cases, the calculation of ML = 3.18 * MDL always yields a value greater then the MDL and meets the criterion of "recognizable signal".

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 in Clean Water Act programs since 1984.

5.2.2 Assessment of the IQE

The Interlaboratory Quantitation Estimate (IQE) was published by ASTM, International in 2000 as standard D 6512. The IDE was developed with support from members of the regulated industry in an attempt to provide a comprehensive quantitation limit procedure that addresses the concerns of the regulated industry, statisticians, and analysts. 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-than." 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-than 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. The Electric Power Research Institute provided input into the design of EPA Method 1631 and 1638 Validation Studies for the purpose of calculating IDEs and IQEs. EPRI also calculated IDEs and IQEs based on these data. These two datasets include a total of ten metal analytes and therefore do not cover a wide range of analytical techniques and methods. Other than these two datasets, 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, the IQE procedure can be tested if suffieient resources are invested.

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 review and ballot by members of ASTM Committee D 19, many of whom are qualified peer reviewers. Thus, the IQE meets the intent of this condition (i.e., submission to scrutiny of the scientific community). In addition, the IQE 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, 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 subjective, particularly with respect to selection of an appropriate statistical 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 B to this Assessment Document.) This subjective variability eliminates the ability to estimate the actual error associated with the IQE. Therefore, the IQE does not meet this condition.

As discussed in Section 5.2.2.1, Condition 3, regarding the IDE, one stakeholder stated that concerns about the complexity and subjectivity of the IQE (and IDE) procedures were unimportant, in part, because IQEs calculated using different models were very close, and in part, because "user friendly-software that will automatically perform the IDE and IQE calculations. EPA obtained copies of such software from the commenter and used that software to evaluate the validity of this comment. As described at length in Section 5.2.2.1, EPA concluded that 1) the subset of models used varies among the software packages, 2) the software packages do not always apply the same model to the same data sets, and 3) even if the same model is used, there is a large amount of variability between the results produced when applying the different software packages to the same set of data. Based on these differences, EPA concluded that the available software programs do not remove all complexity and subjectivity from the IQE calculations. Instead, the software programs appear to introduce new issues by using steps not included in the ASTM procedures.

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.

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. 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 is never explicitly stated. As discussed under Condition 4 of Section 5.1.2.2.1 (scientific validity of the IDE procedure), there appear to be inconsistencies between the IDE and IQE that suggest conceptual conflicts between these two standards.

Based on these findings (along with those discussed under Criterion 2 below), 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 does not meet this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The IQE was published by ASTM four years ago (2000). 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 realistic 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, it is not clear that IQE procedure would 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) does not consistently decrease with increasing concentration (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, such as the Excel files discussed in Condition 3.

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 does not meet 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 he achievable by experienced staff in a well-operated laboratory. Therefore, the IQE meets this criterion.

However, similar to the discussion of criterion 5 for the ML (section 5.1.2.4) anomalous results occur. Analysis of episode 6000, analysis of Episode 6000 data (see appendices) produced anomalous results from two methods (EPA 502.2 and 524.2) that employ instrument thresholds.. For 9% of EPA 502.2 and 59 % of EPA 524.2 analytes the calculated single-lab IQE was below the concentration at which all seven spiked replicates were detected. These results indicate that an IQE study coordinator, after calculating IQE from multi-labs results, would have calculated IQEs below the instrument threshold. The IQE procedure is silent on what happens in this case. As previously noted, the Episode 6000 dataset is not reflective of a typical compliance measurement or method development study because the range of concentrations studied encompassed several orders of magnitude and included concentrations well below the detection limit. And this dataset was not developed according to the procedures in D 6512 (the IQE).

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

There is no database of IQE values for CWA analytes that were calculated according to D 6512. These are the data with which one would compare existing CWA limits and thereby assess the effect of using IQEs as reporting and compliance limits in CWA programs.

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 ±30% (10σ ± 3σ). This uncertainty statement is based on σ equal to 10% of the LOQ.

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 the lack of a supporting procedure for establishing an LOQ, and a 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 multiple zero-value results is zero. One solution for testing the approach is to assume that the LOQ is approximately equivalent to the ML as the blank signal approaches zero. If this is a reasonable assumption, the ML procedure is a viable means for testing the LOQ approach, and the LOQ would meet 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 LOQ meets this condition. The definition of the LOQ specifically estimates the uncertainty associated with a concentration at the LOQ as ±30% based on 10% RSD (Kd = 10). Other choices may be made based on study requirements, policy judgments and/or specific results.

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 does not meet 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 realistic expectations of laboratory and method performance, including routine variability

The LOQ approach is designed to address realistic expectations of laboratory and method performance, including routine variability, and therefore, it appears to meet this criterion. Because the ACS has not published a procedure to implement the approach, in practice the LOQ provides no direct means for demonstrating this performance. The ACS LOQ, the 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.

Because the ACS LOQ approach is not supported by a clearly defined procedure for establishing the LOQ, it does not meet 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 does not meet this criterion because it cannot be used in a regulatory context. The LOQ passes this criterion only if it is assumed to be approximately 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 the lack of a supporting procedure for establishing an LOQ, and a 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 approximately equivalent to the ML as the blank signal approaches zero. If this is a reasonable assumption, the ML procedure is a viable means for testing the LOQ approach, and 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 ±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, the IUPAC/ISO LOQ does not meet this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Acceptance of this approach by the scientific community is currently not known. Acceptance would be indicated by use of the LOD in ISO methods. EPA's 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.

5.4.2.2.2 Criterion 2: The approach should address realistic 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, the IUPAC/ISO LOQ passes this criterion for at least some measurement techniques.

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

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

5.4.2.2.4 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.

Assuming a relationship of the IUPAC/ISO LOQ to the ML, 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.

5.4.2.2.5 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

In the absence of a procedure for determining LOQ values, the ISO/IUPAC LOQ does not meet 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.

Chapter 6
Findings and Next Steps

What are EPA's findings in this revised assessment?

In this revised assessment of detection and quantitation approaches, the Agency has evaluated the codified MDL procedure, the ML procedure that EPA proposed to codify in 2003, and several alternative procedures. Some of these alternative procedures were submitted to EPA during the comment period on EPA's 2003 assessment, which was detailed in the February 2003 Technical Support Document (EPA-821-R-03-005). In today's assessment, we have:

- Identified relevant procedures to include in the assessment (Chapter 2);
- Identified issues that may be relevant to the assessment from an analytical chemistry, statistical, or regulatory perspective (Chapter 3);
- Used six criteria to evaluate the ability of each procedure to support activities under the Clean Water Act (Chapter 4);
- Assessed how well each procedure meets the evaluation criteria (Chapter 5);
- With real-world data and several different procedures, calculated and compared detection and quantitation limits using, and evaluated the theoretical and practical limitations of, each procedure (Appendices B and C).

The assessment of the theoretical and practical applications of each procedure (Appendices B and C) suggests that different procedures produce different detection and quantitation limits. Observed differences are largely due to different sources of variability accounted for among the procedures. The overall assessment of each procedure against each of the evaluation criteria suggests that no single pair of detection and quantitation limit procedures perfectly meets EPA's six evaluation criteria. Although the MDL and ML procedures are closest to meeting these criteria, as discussed under EPA's next steps, we recognize that this is not the end of our consideration of future improvements to EPA procedures and/or adoption of specific alternative procedures.

In response to stakeholders who suggested that EPA clarify or revise some steps in these procedures, we proposed modest revisions to the MDL procedure and proposed to codify an ML definition and procedure in conjunction with the 2003 assessment. We also proposed to codify an existing option that allows use of other detection and quantitation procedures to develop detection and quantitation limits. Public comment on both the 2003 assessment and the proposed revisions expressed many divergent views that conflicted with the proposed modifications to the procedures. Commenters suggested that we work with stakeholders to discuss mutual concerns and possible solutions rather than proceed with the proposed revisions. Some commenters submitted detailed, alternative procedures or regulatory revisions. However, there was no agreement among these commenters as to which of the competing alternatives or revisions to adopt, and none of them fully satisfied EPA's needs under the CWA. We have therefore decided to withdraw the proposed revisions.

What are EPA's next steps?

We believe that it is appropriate to withdraw the proposed revisions, take final action on the 2003 assessment, and explore the feasibility of using a stakeholder process to facilitate a resolution of the technical and policy issues raised during the public comment period. It is in the best interest of all parties to solicit additional stakeholder input through consultations. In a *Federal Register* notice published on September 15, 2004 [69 FR 55547], we announced that a neutral party is studying the feasibility of a process by which a broad group of stakeholders would work together to define and address concerns about the way detection and quantitation limits are calculated and used to support CWA programs. This potential stakeholder process will expand the list of interested stakeholders to include state, tribal and local governments, environmental groups and other interested parties. We trust that this potential stakeholder process will address the wide variety of views held by stakeholders and may lead to recommendations for possible improvements to current EPA procedures and/or use of alternative procedures.

To facilitate open, frank and inclusive discussions, we have made every effort to ensure that this Revised Assessment Document does not prejudge the result of the potential stakeholder process. In particular, we recognize that the following stakeholder issues or suggestions provide a strong starting point for a continued dialogue with stakeholders.

Assessment Evaluation Criteria Issues

The February 2003 assessment identified and discussed six criteria the Agency used to evaluate several different approaches to detection and quantitation. The six evaluation criteria are:

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

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

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

Criterion 4: The detection limit approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present (i.e., a one percent false positive rate) when the analytical method is performed by experienced staff in a well-operated laboratory.

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 a well-operated laboratory.

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.

Stakeholders commented that these six criteria favored the MDL and ML procedures. Some stakeholders noted instances where criterion four fails for the MDL, i.e., does not represent the limit at which there is a 99% confidence that the observed signal is not a false positive. Stakeholders also disagreed with EPA's reliance on only one detection and one quantitation procedure, the MDL and ML (see criterion six discussion at 4.6 in this document) Stakeholders suggested that different detection and

quantitation procedures with different levels of rigor be developed and applied to the disparate uses of these limits in CWA programs. Uses of these limits include verification of laboratory performance, method validation, and as a guide for reasonable bounds on values to consider for permit limits. EPA recognizes that the complexity and statistical rigor appropriate for a detection and quantitation approach for method development and validation would be greater than that needed for demonstrating laboratory proficiency. Although EPA believes that the six evaluation criteria are suitable for purposes of this assessment, they need not be the only starting point for future stakeholder evaluations of revised or alternative detection and quantitation procedures.

Technical and Policy Issues

Some of the major comments on the MDL and ML procedures that influenced our decision to withdraw the proposed rule, and to seek additional stakeholder input, include: (1) the MDL does not adequately address analytical variability or systematic error (bias); (2) a need for better guidance on the intended use of these limits in CWA programs; (3) the need for different procedures for different CWA applications, such as method development, laboratory performance checks, and permit limits. Commenters also asked for clearer guidance on specific steps in the MDL procedure, such as selection of initial spike concentrations, and use of iterative and outlier procedures.

The technical issues of analytical variability and bias attributable to blanks encompass a range of concerns. Stakeholders have suggested that detection and quantitation procedures should:

- vary in the nature and extent of statistical rigor and performance verification checks depending on the end;
- use of a calculated limit;
- account for more sources of variability, such as the variability between and within laboratories;
- require more then seven samples and collect samples over a long period of time; and
- use routine blank samples collected over long periods of time to account for background signals and temporal variability (e.g., ACIL and USGS procedures).

EPA believes these suggestions merit serious consideration, and plans to use the stakeholder process to consider ways to address them.

Conclusion

This Revised Assessment Document addresses comments and concerns received from stakeholders and peer reviewers. Based on this new information, EPA believes that discussion of alternatives or improvements to current detection and quantitation concepts or procedures and uses should continue. It is clear that there is a broad interest in improving current procedures and uses, but no consensus for a specific procedure or procedures has emerged among the laboratory, industry, regulatory or regulated communities. We look forward to further stakeholder participation in this process.

Appendix A
Literature Search Regarding
Detection and Quantitation Limit Approaches

Introduction

Beginning in 2001, DynCorp conducted a search of published literature to identify articles that discuss detection and quantitation limit approaches. This literature search effort was conducted under EPA Contract No. 68-C-01-091 to support an evaluation of detection and quantitation limit approaches by the EPA's Office of Water.

The principal goal of this literature search was to determine if any new detection or qnantitation limit approaches had been published since an earlier search conducted for EPA by Science Applications International Corporation (SAIC) in 1997 and 1998. That search resulted in an annotated bibliography developed by SAIC and delivered to EPA in 1998.

In August 2002, EPA included the literature search results in a draft Technical Support Document (TSD) that was submitted for formal peer review. As part of the charge to the peer reviewers, EPA asked them to identify any additional references. Following EPA's review of the suggested additional references, references relevant to the TSD were added.

How the search was conducted

This search was conducted using two major techniques:

- a search of an on-line citation index (an index of articles cited by other authors), and
- a general on-line search of literature.

On-line citation index search

Because the search was intended to identify detection and qnantitation limit approaches and not specific numeric limits associated with a particular analytical method, DynCorp began by searching for references to the major approaches known to EPA. These included the Agency's method detection limit (MDL) and any other terms that have been suggested to the Agency as alternative detection or quantitation limit approaches. In addition to searching for these approaches, DynCorp also searched the citation index to identify references to the original authors of these approaches and for any other authors who either cited the original approaches, the original papers underlying those approaches, or the authors of those approaches. DynCorp used a similar approach to find any papers that cited the references identified in the earlier literature search by SAIC.

DynCorp staff evaluated the full title of each identified citation to determine its relevance to EPA's objective. Where available electronically and at no additional cost, DynCorp staff also reviewed the abstract and/or full paper to further characterize relevance. All papers that were determined to be relevant, or even possibly relevant, were obtained in hardcopy or electronic format for evaluation by EPA.

After reviewing all papers determined to be relevant to EPA's objective, DynCorp examined all of the references cited in those papers to identify additional papers of interest. These, too, were obtained in hardcopy or electronic format for evaluation by EPA, except where noted below.

General on-line literature search

DynCorp performed an on-line direct search of published literature (e.g., a literature database of published articles, not a citation index) using general terms such as "detection limit," "quantitation limit," or "calibration." As expected, this approach returned a very large numbers of papers that mention these terms, even if the focus of the paper was on something far removed from the development or assessment of approaches about detection and quantitation, and proved to be of limited value in serving EPA's objectives for the search. Therefore, DynCorp discontinued this effort and narrowed the on-line literature search to a search for additional, uncited works by authors of the approaches known to EPA or identified through the citation index approach.

Papers determined to be relevant to EPA's objective were obtained in electronic or hardcopy format for evaluation by EPA, except where noted below.

How the results are presented

DynCorp identified a total of 161 relevant publications using the approach described above. Thirty-three (33) of these publications were also identified in the earlier search by SAIC. Of the 128 remaining publications, 35 were published since the SAIC search was completed.

The peer reviewers suggested additional publications covering a variety of topics, including: quality control, analysis of mercury, and approaches to dealing with censored data. EPA reviewed the citations from the peer reviewers and determined that 20 directly addressed detection or quantitation approaches. In particular, EPA noted that the issue of censored data applies regardless of the specific detection or quantitation limit associated with the data, so those citations dealing with censored data were not included.

Each of the 181 publications identified in the search is listed in Attachment 1, which provides the title, year of publication, authors, and source citation. The citations for the 33 papers identified in the earlier search by SAIC are included in the attachment, and can be identified by the phrase "annotated only" in parentheses after the title of the paper.

The final column of the attached spreadsheet is labeled "Category." All of the citations identified in the SAIC literature search and the current search conducted by DynCorp were placed in one of the six following categories, based on the principal characteristic of the article:

- Background - The citation discusses background information (including early works by Currie, Kaiser, and others).
- Calibration concept - The citation primarily deals with calibration of analytical instrumentation
- Critique - The major thrust of the citation is to critique one or more approaches, as opposed to introducing a new approach
- Multi-laboratory approach - The citation describes an approach to developing detection and/or quantitation limits that relies on multi-laboratory measurements
- Single-laboratory approach - The citation describes an approach to developing detection and/or quantitation limits that relies on single-laboratory measurements
- Single-laboratory, multi-level approach - The citation describes an approach to developing detection and/or quantitation limits that relies on single-laboratory measurements but explicitly includes multiple concentrations.

Although there is some degree of overlap between categories, and some papers could probably be classified in more than one category, each citation was classified into only one category for the purposes of this search.

A seventh category called "Not found" was used for three papers that were identified in the literature search, but for which copies could not readily be obtained. One paper is from a German journal that was not available via interlibrary loan. A second article also was not available via interlibrary loan. The third citation is an abstract by Currie, from 1983. Given that the work of Currie is well-represented in the other citations and the fact that this citation appears to be only an abstract, additional efforts were not expended to obtain a copy.

The 20 publications suggested by the peer reviewers were all included at the end of the list, under an eighth category called "Suggested by a peer reviewer."

The references presented in the table were sorted by category and year of publication and are displayed with the most recent citations in each category first.

Summary

The principal goal of this literature search effort was to determine if any new detection or quantitation limit approaches had been published in the literature since the search by SAIC in 1997 - 1998. As anticipated, citations were identified that relate to the recent efforts of the International Organization for Standardization (ISO), the International Union of Pure and Applied Chemists (IUPAC), and the ASTM International. Additional articles critiquing various approaches were identified as well. However, no previously unknown detection or quantitation limit approaches were uncovered as a result of this effort. Likewise, the references suggested by the peer reviewers provided additional details and applications of existing detection and quantitation approaches, but did not suggest any approaches that had not already been identified.

Title	Year	Author	Source	Category
Classical and Inverse Regression Methods of Calibration	1967	R.G. Krutchkoff	Technometrics 9: 425-439	Calibration
The Interpretation of Certain Regression Methods and their Use in Biological and Industrial Research	1939	C. Eisenhart	The Annals of Mathematical Statistics 10: 162-186	Calibration
The Three "Rs" for Relevant Detection, Reliable Quantitation and Respectable Reporting Limits	2000	Ann Rosecrance	Env. Testing & Anal. 9: 13,50	Critique
Detection and Quantification Capabilities and the Evaluation of Low-Level Data: Some International Perspectives and Continuing Challenges	2000	L.A. Currie	Journal of Radioanalytical and Nuclear Chemistry 245: 145-156	Critique
Realistic Detection Limits from Confidence Bands	1999	J.R. Burdge, D.L. McTaggart, S.O. Farwell	Journal of Chemical Education 76: 434-439	Critique
Response to Comment on "An Alternative Minimum Level Definition for Analytical Quantification"	1999	Henry Kahn, William Tellard, Chuck White	Env. Sci. & Tech. 33: 1315	Critique
Comment on "An Alternative Minimum Level Definition for Analytical Quantification"	1999	H.G. Rigo	Env. Sci & Tech. 33: 1311-1312	Critique
Response to Comment on "An Alternative Minimum Level Definition for Analytical Quantification"	1999	Reben Gibbons, David Coleman, Ray Maddalone	Env. Sci. & Tech. 33: 1313-1314	Critique
Comment on "An Alternative Minimum Level Definition for Analytical Quantification"	1998	Henry Kahn, William Tellard, Chuck White	Envir. Sci & Tech 32: 2346-2348	Critique
Response to Comment on "An Alternative Minimum Level Definition for Analytical Quantification"	1998	Robert Gibbons, David Coleman, Ray Maddalone	Envir. Sci & Tech 32: 2349-2353	Critique
A Discussion of Issues Raised by Lloyd Currie and a Cross Disciplinary View of Detection Limits and Estimating Parameters that are Often At or Near Zero	1997	C.H. Spiegelman	Chemometrics and Intelligent Laboratory Systems 37: 183-188	Critique
A Mock Trial for Critical Values (Detection Limits)	1997	C.H. Spiegelman, P. Tarlow	STATS: The Magazine for Students of Statistics 20: 13-16	Critique
Comment on "An Alternative Minimum Level Definition for Analytical Quantification"	1997	David Kimbrough	Envir. Sci. & Tech. 31: 3727-3728	Critique
The Smallest Concentration	1997	R.F. Moran, E.N. Brown	Clinical Chemistry 43: 856-857	Critique
A Statistical Overview of Standard (IUPAC and ACS) and New Procedures for Determining the Limits of Detection and Quantification: Application to Voltammetric and Stripping Techniques (Technical Report)	1997	J. Mocak, A.M. Bond, S. Mitchell, G. Scollary	Pure and Applied Chemistry 69: 297-328	Critique
Response to Comment on "An Alternative Minimum Level Definition for Analytical Quantification"	1997	R.D. Gibbons, D.E. Coleman, R.F. Maddalone	Envir. Sci. & Tech 31: 3729-3731	Critique
Some Conceptual and Statistical Issues in Analysis of Groundwater Monitoring Data	1996	R.D. Gibbons	Environmetrics 7: 185-199	Critique
Some Statistical and Conceptual Issues in the Detection of Low Level Environmental Pollutants	1995	Reben Gibbons	Environ. & Ecol. Statistics 2: 125-167	Critique

Title	Year	Author	Source	Category
Comment on "Method Detection Limits in Solid Waste Analysis"	1995	D.E. Coleman	Environmental Science & Technology 29: 279-280	Critique
Comment on "Method Detection Limits in Solid Waste Analysis"	1995	Janice Wakakuwa, David Kimbrough	Envir. Sci. & Tech. 29: 281-282	Critique
"You Can't Compute With Less-Thans"	1994	Ken Osborn, Ann Rosecrance	East Bay Municipal Utility District, Core Laboratories	Critique
Limits of Detection	1994	N. Cressie	Chemometrics Intelligent Laboratory Systems 22: 161-163	Critique
Conflicting Perspectives About Detection Limits and About the Censoring of Environmental Data	1994	M.J.R. Clark, P.H. Whitfield	Water Resources Bulletin 30: 1063-1079	Critique
Limit of Discrimination, Limit of Detection and Sensitivity in Analytical Systems	1994	R. Ferrus, M.R. Egea	Analytica Chimica Acta 287: 119-145	Critique
Discussion of: A Study of the Precision of Lead Measurements at Concentrations Near the Method Limit of Detection	1994	B.R. Nott, R.R. Maddalone	Water Environment Research 66: 853-854	Critique
Limits of Detection Methodologies	1993	J. Lindstedt	Plating and Surface Finishing 80: 81-86	Critique
Method Detection Limits in Solid Waste Analysis	1993	David Kimbrough, Janice Wakakuwa	Enviro. Sci. & Tech 27: 2692-2699	Critique
Defining the Limits	1993	G. Stanko, W. Krechta, A. Stanley, T. Dawson, K. Hittig, R. Javick, R. Obrycki, B. Hughes, F. Sakse	Environmental Lab 1: 16-20	Critique
A Study of the Precision of Lead Measurements at Concentrations Near the Method Limit of Detection	1993	P.M. Berthouex	Water Environment Research 65: 620-629	Critique
Detection Limit Concepts: Foundations, Myths, and Utilization	1992	D.A. Chambers, S.S. Dubose, E.L. Sensintaffar	Health Phys. 63: 338-340	Critique
Difficulties Related to Using Extreme Percentiles for Water Quality Regulations	1991	P. M. Berthouex, Ian Hau	Research Journal WPCF 63: 873-879	Critique
A Simple Rule for Judging Compliance Using highly Censored Samples	1991	P. M. Berthouex, Ian Hau	Research Journal WPCF 63: 880-886	Critique
Current Method for Setting Dioxin Limits in Water Requires Reexamination	1990	J. LaKind, E. Rifkin	Env. Sci. & Tech 24: 963-965	Critique
Kaiser 3-Sigma Criterion - A Review of the Limit of Detection	1990	L.S. Oresic, V. Grdinic	Acta Pharmaceutica Jugoslavica 40: 21-61	Critique
MCL Noncompliance: Is the Laboratory at Fault?	1990	Steven Koorse	AWWA p53-58	Critique
Qualitative or Quantitative Characterization of Spectrographic Methods? The Detection and Determination Limits	1990	Karol Florian	Chemia Analityczna 35:129-139	Critique
False Positives, Detection Limits, and Other Laboratory Imperfections: The Regulatory Implications	1989	Steven Koorse	Environmental Law Reporter 19: 10211-10222	Critique
Evaluation of Detection Limit Estimators	1988	F.C. Garner, G.L. Robertson	Chemometrics and Intelligent Laboratory Systems 3: 53-59	Critique
Chemometrics - Measurement Reliability	1988	K. Castaneda-Mendez	Clinical Chemistry 34: 2494-2498	Critique
The Detection Limit	1988	P.S. Porter, R.C. Ward, H.F. Bell	Environmental Science & Technology 22: 856-861	Critique

Title	Year	Author	Source	Category
Estimation of Detection Limits for Environmental Analytical Procedures - A Tutorial	1988	Curr Kirchmer	ACS Symposium Series 361: 78-93	Critique
Limits of Detection	1984	J.K. Taylor	Analytical Chemistry 56: 130A	Critique
Clarification of the Limit of Detection in Chromatography	1984	J.P. Foley, J.G. Dorsey	Chromatographia 18: 503-511	Critique
Limit of Detection: A Closer Look at the IUPAC Definition	1983	Gary Long, J.D. Winefordner	Analytical Chem. 55: 712-724	Critique
Trace Analyses for Wastewaters	1982	C.J. Kirchmer	Envir. Sci. & Tech. 16: 430A	Critique
A comparison of statistical and empirical detection limits	1998	G.C.C. Su	Journal of AOAC International 81: 105-110	Multilab
Challenges in Regulatory Environmetrics	1997	C.B. Davis	Chemometrics Intelligent Laboratory Systems 37: 43-53	Multilab
Determining Quantitation Levels for Regulatory Purposes	1996	P.F. Sanders, R.L. Lippincott, A. Eaton	Journal American Water Works Association 88: 104-114	Multilab
Defining Detection and Quantitation Levels	1993	Raymond Maddalone, James Rice, Ben Edmondson, Babu Nott, Judith Scott	Water Envir. & Tech. Jan,93: 41-44	Multilab
Concept 2000-A Statistical Approach for Analytical Practice - Part I: Limits of Detection, Identification, and Determination	1999	Hadrich J et al.	Deutsche Lebensmittel-Rundschau 1999, 95(10), 428-436	not found
Statistics and Environmental Policy: Case Studies from Long-Term Environmental Monitoring Data	1999	Goudey R et al.	Novart FDN Sym 220; 144-157	not found
The Many Dimensions of Detection in Chemical Analysis	1983	Currie LA	Abstracts of Papers of the American Chemical Society, 185 (Mar), 63-PEST	not found
A Practical Strategy for Determining and Verifying Detection Limits	2001	T. Georgian, K.E. Osborn	Env. Testing & Analysis 10: 13-14	Single lab
Review of the Methods of the US Environmental Protection Agency for Bromate Determination and Validation of Method 317.0 for Disinfection By-Product Anions and Low-Level Bromate	2001	D.P. Hautman, D.J. Munch, C. Frebis, H.P. Wagner, B.V. Pepich	Journal of Chromatography A 920: 221-229	Single lab
Comparison of Detection Limits in Environmental Analysis - Is it Possible? An Approach on Quality Assurance in the Lower Working Range by Verification	2001	S. Gelb, J.W. Einax	Fresenius Journal of Analytical Chemistry 370: 673-678	Single lab
On the Assessment of Compliance with Legal Limits, Part I: Signal and Concentration Domains	2001	E. Desimoni, S. Mannino, B. Brunetti	Accreditation Quality Assurance 6: 452-458	Single lab
Capability of Detection - Part 2	2000	ISO	ISO 11843-2	Single lab
Nomenclature in Evaluation of Analytical Methods Including Detection and Quantitation Capabilities (IUPAC Recommendations 1995)	1999	L.A. Currie	Analytica Chimica Acta 391: 105-126	Single lab
New Reporting Procedures Based on Long-Term Method Detection Limits and Some Considerations for Interpretations of Water-Quality Data Provided by the U.S. Geological Survey National Water Quality Laboratory	1999	C. J. Oblinger Childress, W. T. Foreman, B. F. Connor, and T. J. Maloney	USGS Open-File Report 99-193, 19 pages.	Single lab

Title	Year	Author	Source	Category
Analyses of Polychlorinated Biphenyls and Chlorinated Pesticides in Biota: Method and Quality Assurance	1999	P. Cleemann, G.B. Paulson	Journal of AOAC International 82: 1175-1184	Single lab
Detection Limits of Organic Contaminants in Drinking Water	1998	W.M. Draper, J.S. Dhoot, J.S. Dhaliwal, J.W. Ramey, S.K. Perera, F.J. Baumann	Journal of the American Water Works Association 90: 82-90	Single lab
Detection: International Update, and Some Emerging Dilemmas Involving Calibration, the Blank, and Multiple Detection Decisions	1997	L.A. Currie	Chemometrics and Intelligent Laboratory Systems 37: 151-181	Single lab
Regulations - From an Industry Perspective or Relationships Between Detection Limits, Quantitation Limits, and Significant Digits	1997	D. Coleman, J. Auses, N. Grams	Chemometrics and Intelligent Laboratory Systems 37: 71-80	Single lab
Capability of Detection - Part 1	1997	ISO	ISO 11843-1	Single lab
Determination of Site-Specific Effluent Detection Limits	1996	George Nesterka, Harold Taylor	Water Env. Res. 66: 115-119	Single lab
Multivariate Detection Limits Estimators	1996	R. Boque, F.X. Rius	Chemometrics and Intelligent Laboratory Systems 32: 11-23	Single lab
Nomenclature in Evaluation of Analytical Methods including Detection and Quantification Capabilities	1995	Lloyd Currie	Pure & Appl. Chem. 67: 1699-1723	Single lab
Reporting Low-Level Analytical Data Third Draft (1995-11-08) -- New Project of Commission V.I., International Union of Pure and Applied Chemistry	1995	William Horwitz	IUPAC	Single lab
IUPAC Recommendations for Defining and Measuring Detection and Quantification Limits	1994	LA Currie, W. Horwitz	Analuses Magazine 22: 24-26	Single lab
Recommendations for the Presentation of Results of Chemical Analysis	1994	L.A. Currie, G. Svehla	Pure & Applied Chemistry 66: 595-608	Single lab
Detarshi-A Program for Detection Limits with Specified Assurance Probabilities and Characteristic Curves of Detection	1994	L. Sarabia, M.C. Ortiz	TRAC-Trends in Analytical Chemistry 13: 1-6	Single lab
Quality Control Level: An Alternative to Detection Levels	1994	D.E. Kimbrough, J. Wakakuwa	Environmental Science & Technology 28: 338-345	Single lab
Multivariate Decision and Detection Limits	1993	A. Singh	Analytica Chimica Acta 277: 205-214	Single lab
A Model of Measurement Precision at Low Concentrations	1993	P.M. Berthouex, D.R. Gan	Water Environment Research 65: 759-763	Single lab
Robust Procedure for Calibration and Calculation of the Detection Limit of Trimipramine by Adsorptive Stripping Voltametry at a Carbon Paste Electrode	1993	M.C. Ortiz, J. Arcos, J.V. Jurerros, J. Lopez-Palacios, L.A. Sarabia	Analytical Chemistry 65: 678-682	Single lab
Nondetects, Detection Limits, and the Probability of Detection	1991	D. Lambert, B. Peterson, I. Terpenning	JASA 86: 266-277	Single lab
Detection Limits: For Linear Calibration Curves with Increasing Variance and Multiple Future Detection Decisions	1991	R.D. Gibbons, F.H. Jarke, K.P. Stoub	Waste Testing and Quality Assurance: ASTM STP 1075, D. Friedman, Ed., American Society for Testing and Materials, Philadelphia 3: 337-390	Single lab
Limits of Detection in Multivariate Calibration	1991	G. Bauer, W. Wegscheider, H.M. Ortner	Fresenius Journal of Analytical Chemistry 340: 135-139	Single lab
Estimating Detection Limits in Ultratrace Analysis. Part I: The Variability of Estimated Detection Limits	1991	C.L. Stevenson, J.D. Winefordner	Applied Spectroscopy 45:1217-1224	Single lab

Title	Year	Author	Source	Category
Reporting Low-Level Data for Computerized Data Bases	1988	M. Brossman, G. McKenna, H. Kahn, D. King, R. Kleopfer, J. Taylor	ACS Symposium Series 361:317-327	Single lab
Detection Limits with Specified Assurance Probabilities	1987	C.A. Clayton, J.W. Hines, and P.D. Elkins	Analytical Chemistry 59: 2506-2514	Single lab
Limit of Detection and Limit of Determination - Application of Different Statistical Approaches to an Illustrative Example of Residue Analysis	1987	J. Vogelgesang	Fresenius Zeitschrift Fur Analytische Chemie 328: 213-220	Single lab
Determining the Lowest Limit of Reliable Assay Measurement	1983	L. Oppenheimer, T.P. Capizzi, R.M. Weppelman, H. Mehta	Analytical Chemistry 55: 638-643	Single lab
Principles of Environmental Analysis	1983	L.H. Keith, W. Crummett, J. Deegan Jr, R.A. Libby, J.K. Taylor, G. Wentler	Analytical Chemistry 55: 2210-2218	Single lab
Trace Analyses for Wastewaters	1981	John Glaser, Denis Foerst, Gerald McKee, Stephen Quave, William Budde	Env. Sci. & Tech 15: 1426-1435	Single lab
Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry	1980	MacDougall, Daniel, *et al.*	Anal. Chem. 52: 2242-2249	Single lab
Sensitivity and Limit of Detection in Quantitative Spectrometric Methods	1974	J.D. Ingle Jr.	Journal of Chemical Education, 51, 100-105.	Single lab
Decision and Detection Limits for Linear Calibration Curves	1970	A. Hubaux, G. Vos	Analytical Chemistry 42: 849-855	Single lab
Limits for Quantitative Detection and Quantitative Determination	1968	L.A. Currie	Analytical Chemistry 40: 586-593	Single lab
A Statistical Method for Evaluation of Limiting Detectable Sample Concentrations	1967	P.A. St. John, W.J. McCarthy, J.D. Winefordner	Analytical Chem. 39: 1495-1597	Single lab
Initial Evaluation of Quantitative Performance of Chromatographic Methods Using Replicates at Multiple Concentrations	2001	M.A. Castillo, R.C. Castells	Journal of Chromatography A 921: 121-133	Single lab - multilevel
Multivariate Detection Limits with Fixed Probabilities of Error	1999	R. Boque, M.S. Larrechi, F.X. Rius	Chemometrics and Intelligent Laboratory Systems 45: 397-408	Single lab - multilevel
Evaluation of Approximate Methods for Calculating the Limit of Detection and Limit of Quantification	1999	M.E. Zorn, R.D. Gibbons, W.C. Sonzogni	Environmental Science & Technology 33: 2291-2295	Single lab - multilevel
Limits of Detection, Identification and Determination: A Statistical Approach for Practitioners	1998	J. Vogelgesang, J. Hadrich	Accreditation Quality Assurance 3: 242-255	Single lab - multilevel
Weighted Least-Squares Approach to Calculating Limits of Detection and Quantification by Modeling Variability as a Function of Concentration	1997	M.E. Zorn, R.D. Gibbons, W.C. Sonzogni	Analytical Chemistry 69: 3069-3075	Single lab - multilevel
Detection Limits in GC-MS Multivariate Analysis	1997	Boque R et al.	Quimica Analytica, 16(2), 81-86	Single lab - multilevel
An Alternative Minimum Level Definition for Analytical Quantification	1997	Robert Gibbons, David Coleman, Raymond Maddalone	Environmental Science & Technology 31: 2071-2077	Single lab - multilevel

Title	Year	Author	Source	Category
A Two-Component Model for Measurement Error in Analytical Chemistry	1995	David Rocke, Stefan Lorenzato	Technometrics 37: 176-184	Single lab - multilevel
Practical Quantitation Limits	1992	R.D. Gibbons, N.E. Grams, F.H. Jarke, K.P. Staub	Chemometrics and Intelligent Laboratory Systems 12: 225-235.	Single lab - multilevel
Experimental Comparison of EPA and USATHAMA Detection and Quantitation Capability Estimators	1991	C.L. Grant, A.D. Hewitt, T.F. Jenkins	American Laboratory 23: 15-33	Single lab - multilevel
High Pressure Liquid Chromatography Determination of the Intermediates Side Reaction Products in FD&C Red No. 2 and FD&C Yellow No. 5: Statistical Analysis of Instrument Response	1978	C.J. Bayley, E.A. Cox, J.A. Springer	J. Assoc. Off. Anal. Chem 61: 1404-1414.	Single lab - multilevel
Method Detection Limits: Application to Organic Environmental Chemistry Data	2002	Wade, T. L., J. L. Sericano, Y. Qian, G. Wolff, and G. Denoux	Presentation at ACS Symposium, Boston, MA, August 2002	Suggested by a peer reviewer
Detection limits	2002	van der Voet, H.	Encyclopedia of Environmetrics, 1, El-Shaarawi, A. H. and Piegorsch, W. W. (eds.), 504-515. Chichester: John Wiley & Sons	Suggested by a peer reviewer
Statistical Approaches to Estimating Mean Water Quality Concentration with Detection Limits	2002	Shumway, R., R. Azari, and M. Kayhanian	Environmental Science and Technology 36: 3345-3353	Suggested by a peer reviewer
Detection Limit of Isotope Dilution Mass Spectrometry	2002	Yu, L. L., J. D. Fassett, and W. F. Guthrie	Analytical Chemistry, 74: 3887-3891	Suggested by a peer reviewer
ISO 17025 Requirements: How to Evaluate Uncertainty for Dioxin Analysis in Food and Feed from Validation Data?	2002	G. Eppe, and E. De Pauw	Proceedings of the 22nd International Symposium on Halogenated Environmental Organic Pollutants and POPs, Barcelona, Spain, August 12-18, 2002, Vol. 59, pp. 403-406, 2002	Suggested by a peer reviewer
Are Target Dioxin Levels in Animal Feedingstuffs Achievable for Laboratories in Terms of Analytical Requirements? Results of an Interlaboratory study	2002	G. Eppe, and E. De Pauw	Proceedings of the 22nd International Symposium on Halogenated Environmental Organic Pollutants and POPs, Barcelona, Spain, August 12-18, 2002, Vol. 59, pp. 407-410, 2002	Suggested by a peer reviewer
Environmental Statistics with S-PLUS	2001	Millard, S. P., and Neerchal, N. K.	Chapman & Hall/CRC, Boca Raton, FL	Suggested by a peer reviewer
Quantifying Uncertainty in Analytical Chemistry	2000	Ellison, S. L. R., M. Rosslein, and A. Williams, eds.	EURACHEM/CITAC Guide 2000	Suggested by a peer reviewer
Development and Harmonisation of Measurement Uncertainty Principles, Protocol for Uncertainty Evaluation from Validation Data	2000	Barwick V.J., Ellison S.L.R.	VAM Project 3.2.1	Suggested by a peer reviewer
Statistical Procedures for Analysis of Environmental Monitoring Data & Risk Assessment	1998	McBean, E. A., and Rovers, F. A.	Prentice Hall PTR, Upper Saddle River, NJ	Suggested by a peer reviewer

Title	Year	Author	Source	Category
Foundations and future of detection and quantification limits	1996	Currie, L. A.	Proceedings of the American Statistical Association, Section on Statistics and the Environment, 1-8	Suggested by a peer reviewer
Statistical Methods for Groundwater Monitoring	1994	Gibbons, R. D.	John Wiley & Sons, New York	Suggested by a peer reviewer
Spatial chemostatistics	1994	Cressie, N.	Environmental Statistics, Assessment, and Forecasting, Cothern, C. R. and Ross, N. P. (eds.), 131-146. Boca Raton, FL: Lewis Publishers	Suggested by a peer reviewer
Hypothesis testing with values below detection limit in environmental studies	1994	Slyman, D. J., de Peyster, A., and Donohoe, R. R	Environmental Science & Technology 28: 898-90	Suggested by a peer reviewer
A new approach for accommodation of below detection limit data in trend analysis of water quality	1994	Nagaraj, N. K., and Brunenmeister, S. L	Environmental Statistics, Assessment, and Forecasting, Cothern, C. R. and Ross, N. P. (eds.), 113-127. Boca Raton, FL: Lewis Publishers	Suggested by a peer reviewer
Errors and detection limits	1992	Adams, M. J.	In Methods of Environmental Data Analysis, Hewitt, C. N. (ed.), 181-212. Amsterdam: Elsevier Applied Science	Suggested by a peer reviewer
Statistical inference from multiply censored environmental data	1991	El-Shaarawi, A. H., and Naderi, A	Environmental Monitoring and Assessment 17: 339-347	Suggested by a peer reviewer
Less than obvious: Statistical treatment of data below the detection limit	1990	Helsel, D. R	Environmental Science & Technology 24: 1766-1774	Suggested by a peer reviewer
Environmental tests: Are they valid?	1990	Maynard, A. W.	Chemical Technology 20: 151-155.	Suggested by a peer reviewer
Detection in Analytical Chemistry; Importance, Theory, and Practice	1988	Currie, L. A.	American Chemical Society, New York	Suggested by a peer reviewer

Appendix B
Computation of Detection and Quantitation Limits

INTRODUCTION

This appendix supports the Revised Assessment Document (RAD) for EPA's assessment of detection and quantitation approaches. It presumes that the reader has read chapters three through five of the RAD.

We have compared detection and quantitation limits computed from data gathered by EPA or submitted to EPA by stakeholders commenting on EPA's February 2003 (EPA-821-R-03-005) assessment. The comparison shows that, in general, detection limits derived from a single concentration level such as EPA's MDL are, on average, approximately the same as detection limits derived from similar approaches such as the ACS LOD and LOQ and ISO/IUPAC CRV and MDV, and are approximately three times lower than a single-laboratory variant of ASTM's IDE; and that all quantitation limit approaches, such as EPA's ML, the ACS and ISO/IUPAC LOQ, and a single-laboratory variant of ASTM's IQE, produce quantitation limits that are generally only slightly different.

EPA's Approach to Establishing Detection and Quantitation Limits in Analytical Methods

The Engineering and Analysis Division (EAD) within EPA's Office of Science and Technology develops analytical methods for use in EPA's Clean Water Act (CWA) programs. In developing these methods, EAD first conducts a single-laboratory study in which an MDL and ML are determined followed by multiple single-laboratory studies in which the MDL and ML are either verified or if necessary, revised. If resources, time, and applications of the method warrant, an interlaboratory study is conducted in which the MDL and ML are further verified or, if necessary, revised.

To set an MDL, which is both conservative and achievable by qualified laboratories, we generally select the highest MDL from among the MDLs determined or verified by laboratories in the various studies. For example, EPA determined the MDL in Method 1631 (mercury by cold-vapor atomic fluorescence) as 0.05 ng/L in a single laboratory and revised this MDL to 0.2 ng/L based on multiple single-laboratory studies. All laboratories verified the MDL of 0.2 ng/L in an interlaboratory study. Unlike a single-lab MDL and ML computed in a laboratory quality-control setting, the interlaboratory MDL established during method development is set as a high-biased estimate of Currie's Lc. Thus, the single-lab MDL and resulting ML, when scaled up with the interlaboratory MDL data, are very conservative. This interlaboratory scaling up protects against unrealistically low values, and responds to concerns that the MDL is a single-laboratory approach that produces unrealistically low MDLs.

DETECTION AND QUANTITATION LIMITS ASSESSED

EPA used several datasets to evaluate various approaches to determining detection and quantitation values. These data are described in the Data section of this Appendix.

In the original Assessment Document (EPA, February 2003), four different detection and three different quantitation limits were evaluated and compared. The detection limits were the EPA method

detection limit (MDL), the International Standards Organization/International Union of Pure and Applied Chemistry (ISO/IUPAC) critical value (CRV) and minimum detectable value (MDV), and a single-laboratory variant of the ASTM interlaboratory detection estimate (IDE). The quantitation limits were the EPA minimum level of quantitation (ML), the ISO limit of quantitation (LOQ), and a single-laboratory variant of the ASTM interlaboratory quantitation estimate (IQE).

Several stakeholders commenting on EPA's assessment of data expressed difficulty in replicating EPA's calculations supporting these evaluations. Based on these comments, EPA reviewed the computer programs used to calculate the various limits, and compared results obtained using these programs to calculation results and software packages submitted by commenters. EPA concluded that many of the discrepancies between EPA and commenter calculations were due to differences in the datasets and software used (see Software Comparison, later in this appendix). As a result of this review, EPA did, however, find some discrepancies which have been resolved in this document. Revisions are listed below:

- In calculating the single-laboratory IDE (SL-IDE) and single-laboratory IQE (SL-IQE) based on the Exponential model using the Episode 6000 and Method 1631 and 1638 validation study data, incorrect weights were used when modeling recovery. Because the majority of the SL-IDEs were calculated using this model, most of the SL-IDEs presented in Tables 2, 6, 7 and 8 have changed. Because the SL-IQEs were not calculated based on the exponential models, these values did not change.

- When calculating MLs based on the Episode 6000 data, the resulting ML was incorrectly rounded up for many analytes. This has been corrected, and many of the calculated MLs in Tables 4 and 5 have changed.

- In the 2003 assessment, blank results were included in the calculations of the ISO/IUPAC CRV, MDV and LOQ. Upon further review, it was decided that it was invalid to use blank results included in the Episode 6000 study, because these blanks were used to assess carry-over, and would not be representative of routine blank analyses. Therefore, the ISO/IUPAC limits were re-calculated using the lowest spike concentration in place of blank results.

- For two analytes in the Episode 6000 data (uranium and thallium by Method 200.8), incorrect formatting caused multiple spiking levels to be combined improperly. This affected the calculation of all limits for these analytes. This calculation has been fixed, and the limits have changed slightly for these two analytes.

- After completion of the Original Assessment Document, a new version of the IDE procedure (D6091-03) was published by ASTM. This procedure included the use of a standard deviation bias correction factor which was not included in the prior version (D6091-97). Therefore, all IDEs calculated using the Episode 6000 and Methods 1631 and 1638 validation study data were re-calculated using this correction factor. For the majority of analytes, the resulting IDEs increased slightly (by approximately 4%).

The effect of these changes on the analyses are discussed in the Results of Computations section of this Appendix. To better explain how calculations were run, Appendix C gives example calculations of the SL-IDE, SL-IQE, MDL and ML for one analyte.

Along with comments on EPA's assessment, both the American Council of Independent Laboratories (ACIL) and USGS submitted data and procedures. ACIL submitted a procedure for calculating a critical level (CRV) and Long-Term MDL (LTMDL). USGS submitted its procedure for calculating a long-term MDL (USGS LT-MDL). Both the ACIL critical level and USGS LT-MDL are estimates of Currie's Lc, and are therefore comparable to the EPA MDL. Both the ACIL and USGS procedures, however, are based on results collected over a long period of time. The ACIL critical level is based on blank results, and the USGS LT-MDL is based on spiked results. The formula for the ACIL critical level is identical to that of EPA's MDL, except that the mean of the blanks is added to the product of the standard deviation and t-statistic. The USGS procedure does not use a sample standard deviation, but instead uses a non-parametric estimate of variability that is based on the interquartile range. The USGS LT-MDL procedure also allows addition of the mean or median of blank results to the LT-MDL.

ACIL also suggested a separate CRV procedure (ACIL "Case 2") for calculating estimates for those methods for which analysis of blank samples does not produce a signal. For these methods, ACIL suggested an iterative procedure that first determines the lowest level at which all 7 replicates are detected, and then estimates the CRV as the lowest of the observed results of 7 spikes. The analogue of Currie's Ld is estimated as this lowest spike level. EPA finds merit in the idea of dividing the methods into two groups (depending on the presence or absence of a signal from analysis of blank samples) and in the idea of estimating the detection level of the instrument, and plans to further investigate the ACIL approach. However, the particular implementation of the ACIL Case 2 procedure has some conceptual problems that precluded it from evaluation at this time. These problems are described later in this Appendix (see "Episode 6000 Data").

EPA provides further discussion of these approaches and the Agency's reasons for selecting them in Chapters 1 and 2 of the RAD.

Commonality of Approaches

The EPA, ACS, and ISO/IUPAC approaches are all multiples of the standard deviation of either replicate measurements of a blank or of the lowest spike concentration that produces positive (non-zero) results for all 7 replicates. Similarly, the ACIL and USGS approaches are based on multiples of a parametric or nonparametric estimate of variability of replicate measurements, with the difference that the given estimate includes greater sources of variability than those of the other single-concentration approaches.

Other subtle distinctions are that (1) ISO/IUPAC suggest a false positive rate of 5 % ($\alpha = 0.05$) for the CRV and MDV, whereas EPA specifies a false positive rate of 1 % ($\alpha = 0.01$) for the MDL and (2) the EPA MDL was calculated by pooling data from two concentration levels after determining that the variabilities of the two concentration levels are not significantly different (as provided as an option in step 7 of the MDL procedure), thereby increasing the degrees of freedom to 12 from the 6 nsed in computation of the ISO/IUPAC CRV and ACS LOD. The consequence of distinction (1) is that an approach with a higher allowed false positive rate ($\alpha = 0.05$) will produce a lower detection limit than an approach with a lower false positive rate ($\alpha = 0.01$). The consequence of distinction (2) is that a detection limit resulting from pooling at two levels will be more stable and likely somewhat lower than a detection limit at a single level (given the same variability at each level) because the degrees of freedom are increased in the t statistic.

The ACS and ISO/IUPAC approaches specify replicate measurements of blank samples. In computing detection and quantitation limits from the Episode 6000 data, blank results were not used, as blanks analyzed in this study included carry-over effects, and were therefore not representative of routine blank results. Therefore, the lowest spike concentration (or, in the case of the MDL, two lowest spike concentrations) that produced a non-zero result was used for computation of all approaches. This simplification condensed the EPA MDL and the ACS LOD to a single approach subsequently termed the EPA/ACS DL. Similarly, the EPA ML and ACS LOQ were condensed to a single approach, termed the EPA/ACS QL.

The remaining single-concentration approaches are the ISO/IUPAC CRV, MDV, and LOQ, the ACIL critical level and the USGS LT-MDL. The ISO/IUPAC CRV differs from the EPA/ACS DL because of its suggested use of a false positive rate of 5% ($\alpha = 0.05$) versus use of a false positive rate of 1% ($\alpha = .01$) in the EPA/ACS DL. The ISO/IUPAC MDV also differs from the EPA/ACS DL because of (1) its suggested use of a false positive rate of 5% ($\alpha = 0.05$), (2) its stated false negative rate of 5 % (β=0.05), and (3) recovery correction (estimated using a linear regression). Therefore, the ISO/IUPAC CRV and MDV were each treated separately (were not combined with the EPA or ACS approaches) from the other detection limit approaches in the data analysis. The ISO/IUPAC LOQ is also different from the other quantitation limit approaches and was treated separately from these approaches. The ACIL critical level differs from the EPA/ACS DL in its inclusion of long-term variability and the addition of the mean blank result to the limit. The USGS LT-MDL differs from the EPA/ACS DL in its inclusion of long-term variability, the addition of the median or mean blank result to the limit, and the use of a nonparametric estimate of variability in place of the sample standard deviation. Because of the lack of long-term variability and representative blank results in the Episode 6000 data, the ACIL critical level and USGS LT-MDL could not be calculated using the Episode 6000 data. Assessments of these approaches in comparison to the EPA/ACS DL were done using blank and spiked sample data that were submitted to the Agency by ACIL and USGS.

The ASTM IDE and IQE were treated separately because they are constructed by fitting a model to variability versus concentration data, rather than being derived from the standard deviation of replicate measurements of a single concentration, (as are the EPA, ACS, ISO/IUPAC and ACIL approaches). Similar to some of the ISO/IUPAC approaches, the ASTM IDE and IQE include "protection" against false negatives and recovery correction (see section 3.3 of the Revised Assessment Document for a discussion on EPA's concerns about false negative protection). The IQE, but not IDE, also includes an added correction for the bias associated with an estimate of the true standard deviation at each concentration. In the context of the IQE, the word "bias" means the amount by which the estimated sample standard deviation is low compared to the true population standard deviation, and should not be confused with common use of the word "bias" in an analytical measurement.

Single-laboratory Variants of Interlaboratory Approaches

Because the EPA, ACS, and ISO/IUPAC approaches are single-laboratory approaches, and the ASTM IDE and IQE are interlaboratory approaches, the ASTM approaches could not be computed using the single-laboratory data in the Episode 6000 studies. To solve this problem, single-laboratory variants of the IDE and IQE were used. These single-laboratory variants were termed the SL-IDE and SL-IQE for "single-laboratory IDE" and "single-laboratory IQE," respectively. The SL-IDEs and SL-IQEs were constructed using the overall standard deviation within a single laboratory at each concentration rather than the overall standard deviation across all laboratories at each concentration.

Attempted Application to Interlaboratory Data

EPA attempted to apply the various approaches to interlaboratory study data in response to a request by the Petitioners to the Settlement Agreement and so that detection and quantitation limits could be compared. However, because the EPA, ACS, and ISO/IUPAC approaches are single-laboratory approaches, whereas, the ASTM approaches are interlaboratory approaches, it was not possible to compute directly comparable detection and quantitation limits from the same data.

It was possible, however, to compare the detection and quantitation limits produced by EPA and the Electric Power Research Institute (EPRI) from the EPA Method 1631 and EPA Method 1638 interlaboratory study data. Although the resulting detection and quantitation limits are either single-laboratory (EPA) or interlaboratory (ASTM), as appropriate to the particular approach, a comparison of the resulting limits can be informative. The EPRI detection and quantitation limits are presented in EPRI reports of the results of the Method 1631 and Method 1638 studies.

DATA

Datasets Evaluated

The datasets used to evaluate the detection and quantitation approaches discussed above are described in this section. EPA computed EPA/ACS DLs and QLs; ISO/IUPAC CRVs, MDVs and LOQs; and single-laboratory variants of ASTM IDEs (SL-IDEs) and IQEs (SL-IQEs) using the Episode 6000 data. EPA also computed IDEs and IQEs for the Method 1631 and 1638 interlaboratory study data. EPA computed ACIL's critical level, USGS's LT-MDL and EPA's MDL based on a combination of blank and spiked data submitted by USGS, and performed an assessment of the effect of long-term variability based on blank data submitted by ACIL.

EPA's Variability versus Concentration Studies ("Episode 6000")

In 1997 and 1998, EPA conducted a study of variability vs. concentration for a number of analytical methods. Six laboratories were employed for the analyses; each analyte and method combination was tested by one of these laboratories. For nearly all of the technologies, the studies were conducted by spiking reagent (i.e., blank, presumably "clean") water at 16 different concentrations per analyte, ranging from 100 times an initial estimate of the MDL to 0.1 times the initial estimate. A total of 198 analytes were measured, generally with seven replicates analyzed at each concentration. 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), and in Appendix C of the February 2003 Assessment document. 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)

EPA's 2003 assessment of detection and quantitation examined a dataset populated with the results of this study, the object of which was to characterize analytical variability as a function of concentration over a wide range of concentrations, analytes, and analytical methods. Data from this study, including many tables and plots, were provided in the record supporting EPA's original assessment and discussed in EPA's "Technical Support Document for the Assessment of Detection and Quantitation Approaches," EPA 821-R-03-005, February 2003. The database developed contains a total of approximately 22,000 data points. This study was conducted in contract laboratories. EPA performed a contract compliance review of these studies at the time the studies were conducted, but not a point-by-point review of each of the tens of thousands of data points.

In the study, an initial (range finding) MDL was determined for each combination of analyte and analytical technique using a revised draft of the MDL procedure. The revised draft had three significant changes: (1) the definition was more closely conformed to the MDL procedure; (2) optional iterative step 7 of the MDL procedure was made mandatory; and (3) the spike concentration to MDL was reduced from 5 to 3 in an attempt to narrow the resulting MDL. During data gathering, two laboratories complained that the reduction in spike to determined MDL ratio (from 5 to 3) caused a large number of iterations and stated that 5 was more reasonable. Subsequently, EPA returned to the spike to MDL ratio of 5 published in the 40 CFR 136, Appendix B procedure.

After determining the initial MDL, each laboratory analyzed 7 replicates of samples spiked at concentrations of 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. Results associated with the replicate analyses at each concentration level were obtained, as often as possible, using the same calibration that was used in determining the initial MDL. Where laboratory reports indicated that multiple calibrations were conducted, the association between each result and its calibration was used 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.

A variant of the iterative MDL procedure 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 7 replicates at decreasing concentrations until signal extinction, then select the

concentration(s) appropriate for 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 monotonically decreasing concentrations described above and a few selected concentrations to achieve the desired spiking levels.

Some commenters on the 2003 assessment noted possible errors. EPA reviewed these comments and examined the individual data values and other aspects of the assessment that commenters thought were in error. Commenters commented most frequently on measurements of organic compounds by EPA Methods 502.2 (halogenated and aromatic volatiles by GC with photoionization and electrolytic conductivity detectors in series) and 524.2 (volatiles by GC/MS) that were included in the Episode 6000 dataset. EPA performed a more comprehensive review of these data points, and found that the calculated recoveries of some of the compounds are higher or lower than would be expected for the analytical technologies employed. There also appear to be low background concentrations of some compounds in the reagent (blank) into which the analytes were spiked. Backgrounds are commonly observed in determinations of metals, radionuclides, and some volatiles.

Without the raw data for the analyses in question, it is not possible to unequivocally determine the root cause(s) of the high or low recoveries and possible backgrounds. However, atypical recoveries may have been the result of (1) laboratories making measurements at levels as much as 50 times below the lowest level to which they would normally calibrate to establish MDLs and MLs at as low a level as could be measured, and (2) EPA's request that the laboratories use a single calibration (rather than multiple) to prevent discontinuities in the variability vs concentration trends that were the object of these studies.

Another possible explanation for the low apparent recoveries is the setting of thresholds in the GC and GC/MS analyses. If a small constant area of a GC response peak is removed by thresholding, the relative amount of area that is removed will increase as the concentration is reduced, resulting in lower apparent recoveries at the lower concentrations. This effect would be consistent with observations in some of the data.

As for possible backgrounds for volatiles or metals, these backgrounds likely were either present in the reagent (blank) water used by the laboratories for the MDL determinations, or by carry-over from one sample to another. To test for carry-over, some laboratories analyzed one or more blank sample between spike levels after verification of calibration. Instances in which multiple blanks were analyzed often show decreasing small concentrations for some of the analytes. However, these backgrounds resulting from carry-over mean that blank results should not be used to assess false positive rates of the different limits calculated using the Episode 6000 data.

Interlaboratory Study Data

EPA used data from two interlaboratory method validation studies to calculate IDEs and IQEs for a total of 10 metal analytes. These studies were conducted by EPA to evaluate performance of EPA Methods 1631 and 1638, and to gather data to evaluate existing performance specifications, including detection and quantitation limits. To expand the scope of the study, the Electric Power Research Institute (EPRI) funded the distribution of additional samples to study participants. Each study included multiple participant laboratories: twelve for Method 1631 and eight for Method 1638.

The two studies were designed so that each participating laboratory would analyze sample pairs of each matrix at concentrations that would span the analytical range of the method. Each laboratory was provided with multiple sample pairs, including samples measured in filtered effluent, unfiltered effluent, marine water, filtered freshwater, and spiked reagent water. Each laboratory analyzed reagent water sample pairs for each analyte at five different concentration levels. The results of the reagent water analyses were used to fit variability functions and calculate IDEs and IQEs.

Data from these studies also are discussed in Chapter 1 of this document.

Data Submitted by Stakeholders

EPA also used datasets containing results from analysis of blank samples provided by two stakeholders. Blanks analyzed over a period of three months for five analytes using Method 200.7 were provided by the American Council of Independent Laboratories (ACIL), while blanks analyzed over a period of one year representing 78 analytes were provided by the US Geological Survey (USGS). In addition to these blank results, USGS sent results of the analyse of spiked samples representing 39 analytes. Because spiked samples were analyzed only at a single concentration level, many of the different detection and quantitation limits, such as the SL-IDE and SL-IQE, cannot be calculated using these data. However, a comparison of the critical level suggested by the ACIL, the LT-MDL suggested by USGS, and the EPA MDL was performed using the blank and spiked results.

The data submitted by ACIL and USGS also are discussed in Chapter 1 of this document.

Datasets Not Evaluated

The Petitioners and Intervenor to the Settlement Agreement provided the list of datasets shown in Table 1 and suggested that EPA evaluate detection/quantitation limit approaches using the datasets on the list. However, in reviewing the datasets suggested, EPA determined that many were developed for characterizing the behavior of an analyte or analytes across the analytical range of a method, rather than in the region of detection and quantitation. For example, any dataset developed prior to the advent of the IDE and IQE would be inappropriate because there could not have been an estimate of IDE_0 or IQE_0 (i.e., an initial estimate of the given limit; see Section 6.2.2.1 of D6091 and D6512). This eliminates all datasets in Table 1 except the EPA/EPRI Method 1631, the EPA/EPRI Method 1638 dataset, and the MMA 2001-2 dataset. It is possible that some spike level in one or more of the datasets developed prior to the advent of the IDE and IQE would fortuitously meet the IDE/IQE criteria. But the IDE and IQE can be circular; i.e., once developed from a given dataset, there may be a spike level in the dataset that can be construed to meet the criteria. Datasets developed without following the IDE and IQE procedures, particularly without making an *a priori* estimate of IDE_0 or IQE_0, do not meet the requirements of the IDE and IQE procedures, regardless of whether the data in them can be construed to have met those requirements after the fact.

In addition, these datasets do not lend themselves to the comparisons used in this report because the developers of these datasets did not apply the measurements needed to establish an MDL and ML. Therefore, MDLs and MLs could not be determined for comparisons (see the section titled "EPA's Approach to Establishing Detection and Quantitation Limits in Analytical Methods").

The EPA 6000 dataset is comprehensive in coverage of analytes, analytical techniques, and a concentration range from 0.1 to 100 times the MDL, whereas the datasets suggested by Petitioners focus on metals, two Aroclors, and concentrations across the analytical range of the method. The range of data used for construction of an IDE or IQE is particularly important. As detailed in the discussion of the "Effect of number and spacing of concentrations for determination of the SL-IDE and SL-IQE" below, including data across the analytical range in calculation of an SL-IDE significantly raises the SL-IDE.

After EPA published the February 2003 Assessment Document for comment, three commenters offered to provide EPA with additional data that would enhance EPA's assessment. EPA requested the data offered by each of these organizations, but received a response from only two of the three (an analytical laboratory and USGS). After evaluating these data, EPA determined that the data from the analytical laboratory were not useful because they were limited to calibration data and did not include the data from extraction that is needed to compare detection/quantitation approaches.

Michigan Manufacturers Association (MMA) Dataset

In March of 2002, John Phillips of Ford Motor Company submitted a report of results from a study of two Aroclors (PCBs) by the Michigan Manufacturers Association (MMA) for EPA's consideration in evaluating detection and quantitation limit approaches. EPA did not use this dataset because of problems, such as the dataset was limited to a maximum number of five analytical results per spike level, which is inconsistent with the minimum number of seven analytical results per spike level required for determining an MDL, and other values that are determined using non-ASTM approaches. In comments on EPA's evaluation, Hunton and Williams (a law firm representing the Inter-industry Analytical Group), stated that EPA should not have excluded the MMA dataset from its assessment of detection and quantitation approaches. EPA notes, however, that because of the insufficient number of analytical results, comparison of various detection and quantitation approaches is not possible with this dataset, and has not included the dataset in this evaluation. In addition, MMA samples spiked with low levels of PCBs as Aroclors produced an average recovery on the order of 500% at the lowest spike concentration whereas PCBs are recovered at approximately 80% from water in this concentration range (see the recovery data in EPA Methods 608 and 1668A). A logical explanation for the 500% recoveries in the MMA study is that the samples were contaminated by the sample preparation laboratory, by many of the participant laboratories, or both. A single and simple test, which was not conducted in the MMA study, of an aliquot of the prepared water samples using a method, such as EPA Method 1668A, would have demonstrated that the samples were free from contamination and contained the stated spike concentrations at the time that the samples were prepared.

COMPUTATIONS

All computations were carried out using Statistical Analysis System (SAS) version 8.01. The equations for all approaches were programmed into the SAS software by a senior statistician, with assistance from senior analysts. There is some ambiguity in the IUPAC/ISO and ASTM detection and quantitation limit approaches and in interpretation of results from the ASTM approaches. Several formulas are given in the IUPAC/ISO documentation, but none are defined to be the official ISO/IUPAC detection and quantitation limit approaches. Therefore, calculations for the CRV, MDV, and LOQ were chosen because they were most representative of Lloyd Currie's definitions of a critical valne, detection

limit and quantitation limit. Ambiguity in results from the ASTM approaches is attributable to the subjective nature of interpreting residual plots for each analyte. To resolve this issue, IDE and IQE models were chosen using significance tests for slope and curvature.

References used for the IUPAC/ISO approaches were those published by Currie in *Pure and Applied Chemistry* 67:10, 1699-1723 (1995) as updated by *Analytica Chimica Acta* 391 105-126 (1999). Where needed, the ASTM approaches were programmed as single-laboratory variants of the Practices D 6091 (IDE) and D 6512 (IQE). EPA has included the SAS program code on the CD-ROM that supports this document.

To assess stakeholder comments about calculations of the IDE and IQE that were performed and summarized in the original assessment document, EPA requested additional software packages offered by commenters who use the software to determine these limits. On April 20, 2004, EPA received copies of two software packages written for the purpose of determining the IDE and IQE from a representative of Ford Motor Company. The first of these is Qcalc (version 1.0), a DOS-based program. The second of these is an Excel spreadsheet which utilizes Excel functions, macros and an add-in function to determine IDEs and IQEs. These two programs were compared to the SAS programs used by EPA by calculating IDEs and IQEs based on a subset of the Episode 6000 dataset. The results of this comparison are described later in this Appendix (see section titled "Comparison of IDE/IQEs Calculated Using Different Software Packages").

Calculation of the ACIL CRV, USGS LTMDL, and EPA MDL was done using analytical results of blank and spiked samples submitted by USGS. Specific details of these calculations are described in the section titled "USGS Blank and Spiked Metals and Nutrient Data" later in this Appendix.

RESULTS OF COMPUTATIONS

Detection and quantitation limits are presented in a set of tables corresponding to the Episode 6000 study, a single table corresponding to the Method 1631 and Method 1638 studies, and a single table summarizing limits calculated using data submitted by USGS. Within the Episode 6000 dataset, results for detection limits are compared followed by results for quantitation limits. Within the comparison of limits (detection or quantitation), the first table compares the actual limits followed by a table of percent differences between limits.

Episode 6000 data

Table 2 compares detection limits produced by four approaches (EPA/ACS DL; ISO/IUPAC CRV; ISO/IUPAC MDV; and ASTM SL-IDE) and Table 3 presents the percent difference between these approaches, using the formula given below:

$$\%\ difference = \frac{(Lim - DL)}{(Lim + DL)/2} * 100\%$$

where: DL is the EPA/ACS DL for the given analyte, and
Lim is the corresponding limit (CRV, MDV, or SL-IDE) being compared to the DL.

The median percent difference between the EPA/ACS DL and each of the other three limits was compared to 0% using two significance tests: the sign test and Wilcoxon rank-sum test. The sign test evaluates whether the given limit exceeds the EPA/ACS DL 50% of the time. The Wilcoxon rank-sum test is a more powerful test which, unlike the sign test, takes into account the magnitude of the difference between the two limits by ranking the percentage differences presented in Table 3.

The ISO/IUPAC CRV was less than the corresponding EPA/ACS DL for 97% of the analytes and methods, with a median percent difference of -35.7%. The median percent difference of ISO/IUPAC CRV to EPA/ACS DL was significantly less than 0% based on both the sign and Wilcoxon tests with α = 0.05 ($p<0.0001$ for both tests). The major reason for this difference is most likely the different Type I error rate for the two approaches ($\alpha = 0.01$ for the EPA/ACS DL and $\alpha = 0.05$ for the ISO/IUPAC CRV).

The median percent difference between the ISO/IUPAC MDV and the EPA/ACS DL is 8.8% with the MDV exceeding the DL for 53% of the analytes. The median percent difference between the ISO/IUPAC MDV and EPA/ACS DL did not differ significantly from 0% based on the sign test ($p=0.523$) or the Wilcoxon rank-sum test ($p=0.164$) with $\alpha = 0.05$. The likely reason that the two approaches do not yield significantly different results is that the correction for false negatives and recovery correction in the MDV ($\beta = 0.05$) are counteracted by the smaller Type I error rate for the EPA/ACS DL.

The median percent difference between the ASTM SL-IDE and the EPA/ACS DL is 108.7%; i.e., the single-laboratory variant of the ASTM IDE is, on average, three times as large as that of the EPA and ACS approaches. The SL-IDE was greater than the corresponding EPA/ACS DL for 92% of the analytes and methods. The median ratio differed significantly from 1, based on both the sign and Wilcoxon tests with $\alpha = 0.05$ ($p<0.0001$ for both tests). The median ratio and percent of SL-IDEs exceeding the corresponding EPA/ACS DL both increased slightly compared to the calculations presented in the original assessment document, due to the correction of the exponential model calculations for the SL-IDE and the use of the standard deviation bias correction. It is not surprising that the SL-IDE results were generally greater than the EPA/ACS DL, as the SL-IDE is an estimate of Currie's L_D, whereas the EPA/ACS DL is an estimate of Currie's L_C. In addition, the use of two tolerance interval limits in the IDE calculation likely also led to the large difference between the SL-IDE and EPA/ACS DLs.

Table 4 compares quantitation limits produced by the three approaches (EPA/ACS QL; ISO LOQ; and ASTM SL-IQE) and Table 5 compares the percent difference between these approaches taking the EPA/ACS QL as reference. Similarly to the detection limit approaches, the median percent difference was compared to 0% using the sign and Wilcoxon tests. The median percent difference between the ISO/IUPAC LOQ and the EPA/ACS QL is -4.2%, and the median percent difference between the ASTM SL-IQE and the EPA/ACS QL is 19.6%. The ISO LOQ and ASTM SL-IQE are greater than the corresponding EPA/ACS QL for 47% and 62% of the analytes and methods, respectively. The median ratio between the LOQ and QL did not differ significantly from 0% based on the sign test ($p=0.390$), but did based on the Wilcoxon test ($p=0.043$) at $\alpha=0.05$. The median ratio between the SL-IQE and QL differed significantly from 0% based on both the sign test ($p=0.001$) and the Wilcoxon test ($p<0.0001$).

For the SL-IQE comparisons, this result is different from those presented in the original assessment document, due to the fixed rounding issue in the ML calculations (see discussion under Detection and Quantitation Limits Assessed). Because the EPA/ACS QL and the SL-IQE are both estimates of Currie's L_Q, the reason for this difference is not clear. One possible reason for this significant difference is that the SL-IQE does not assume that variability at the quantitation limit is equal to variability of the blank, whereas the EPA/ACS QL does. However, it is worth noting that the difference seems to be strongly affected by which model was used to calculate the SL-IQE. The median percent difference between the QL and SL-IQE is -7.7% when the hybrid model is used to calculate the SL-IQE compared to 67.9% and 179.6% for the linear and constant models, respectively. While use of the constant model assumes that the variability is constant between the blank and quantitation limit, this model type is generally chosen only when there are unnsually high results at one or more of the lower spike levels for a given analyte. Therefore, the SL-IQEs calculated for these analytes are likely somewhat biased high.

Although the Episode 6000 dataset is not ideal for evaluating the ACIL Case 2 iterative approach for those methods/instruments for which analysis of blank samples does not produce a signal, EPA estimated the ACIL Case 2 CRV using the lowest concentration level at which all 7 replicates were observed to test if the conceptual problem with ACIL's implementation of Case 2 occurs in practice. EPA noticed that, because the estimate of Currie's L_C is based on measured values and the estimate of Currie L_D is based on spike level, the estimate of L_D could theoretically fall below L_C for methods with recovery that systematically exceeds 100% or for data with some contamination. Looking at Episode 6000 data, EPA confirmed that this problem may occur in practice. In fact, it occurred for 35 of the 146 analytes (24%) measured using methods that do not always result in signals from analysis of blank samples.

EPA/EPRI Method 1631 and 1638 Interlaboratory Method Validation Study Data

Table 6 compares detection and quantitation limits computed from data generated in the Method 1631 and Method 1638 interlaboratory studies. MDLs and MLs are those listed in EPA Methods 1631 and 1638. EPA computed IDEs and IQEs for the purpose of preparing this assessment. IDEs and IQEs computed by EPRI are from the EPRI reports on the Method 1631 and Method 1638 interlaboratory studies.

In reviewing these data, it must be recognized that the EPA MDLs and MLs are the result of selecting the highest MDL in EPA's single-laboratory studies or interlaboratory study, whereas the IDEs and IQEs are the result of a statistical process that includes recovery correction, correction for bias in the sample standard deviation (IQE only), allowance for prediction and tolerance intervals, interlaboratory variability, and model selection. The most significant reason for the instances of a large disparity between the EPA-determined IDEs/IQEs and the EPRI-determined IDEs/IQEs is model selection. EPA selected the model based on a strict application of the IDE and IQE procedures by a senior statistician. For those instances in which EPA and EPRI selected the same model, the IDEs and IQEs are nearly the same.

Table 7 compares IDEs and IQEs resulting from the four main model types described in the ASTM IDE and IQE procedures. IDEs and IQEs resulting from the constant model were the highest for all analytes. IDEs and IQEs resulting from the other three models were almost equal for some analytes (lead, for example), and differed by more than an order of magnitude for others (mercury, for example). For two analytes, the IDE and IQE estimated using the linear model were negative. This was due to a negative intercept estimate in the precision model. The ASTM IDE and IQE procedures dictate that the linear model should not be used in this situation.

Table 7 also includes RSDs between the IDEs and IQEs calculated using the different model types. This was done based on commenter statements that the choice of model had only a minimal effect on the resulting IDE or IQE. This analysis is discussed later in this Appendix (see "Comparison of IDE and IQEs calculated using Different Models").

USGS Blank and Spiked Metals and Nutrient Data

USGS supplied EPA with blank data collected over a period of one year for 78 metals and nutrient analytes and spiked data collected over a period of one year for 39 metals and nutrient analytes. These results were used to calculate both the USGS LT-MDL and ACIL critical level. The ACIL critical level was calculated using the blank results for the given analyte and method. The USGS LT-MDL was calculated based on the spike results for the given analyte and method. In addition, the LT-MDL was calculated in two ways: by adding the mean of the blank results for the given analyte and method, and by adding the median of the blank results for the given analyte and method.

The EPA MDL also was calculated for each analyte/method using the spiked sample results provided by USGS. Because MDLs are typically calculated using fewer replicates than the 15 to 24 analyzed by USGS, EPA calculated the MDL by simulating different subsets of 7 replicates. Subsets were created by taking each set of 7 consecutive spiked results, i.e., the first 7 samples analyzed would be one subset, the 2nd through 8th samples analyzed would be another subset, etc. This yielded a total of n-6 subsets, where n is the number of total samples for that analyte. The MDL was then determined by randomly choosing one of the n-6 subset MDLs. While the use of only seven replicates run consecutively in each subset minimized the effect of long-term variability, it is worth noting that the amount of temporal variability in each subset is still greater than that typically included in the EPA MDL (i.e., MDL datasets typically are generated in a single day); the time interval between the first and last replicate analyzed within a subset ranged from 30 to 48 days. Therefore, the calculated MDLs are likely somewhat higher than those that would be calculated using results generated over a single day.

After calculation of these limits, the percentage of blank results included in the dataset that exceed each limit for each analyte was calculated. Because all limits were calculated at the 99% confidence level, it would be expected that the average percent of blanks exceeding each limit would be approximately 1% when the blank results follow a Normal distribution centered at 0. Limits based on each of the calculations are presented in Table 10.

Generally, the percentage of blanks exceeding the ACIL critical level was lower than the percentage exceeding the other limits (see summary table following Table 10). The percentages of blanks exceeding the EPA MDL were slightly higher compared to the percentages exceeding the ACIL critical level, due to a small subset of analytes with notable blank bias. The USGS LT-MDL had higher rates of blank exceedance than either the ACIL or EPA limits, regardless of whether the mean or median was added to the limit. This suggests that the effect of blank bias was smaller than the effect of the method of estimating variability. USGS uses the nonparametric calculation to lessen the effect of outliers on the estimate of variability. Because those blanks that exceed a given limit are likely to be outliers themselves, this can lead to inflated exceedance rates. However, it is worth noting that, for the majority of analytes where blanks exceeded the calculated USGS limits, multiple blank results were greater than the associated limit. This suggests that some non-outlying blank results also are exceeding the USGS limits for some analytes.

DISCUSSION

Negative detection limits for the ISO/IUPAC MDV

The calculated ISO/IUPAC MDV was negative for 29 analytes in the Episode 6000 data. Negative MDVs are attributable to the use of a regression model to estimate recovery at each concentration. The standard errors and correlation of the regression parameters are included in the calculation of the MDV. Analytes for which the MDV was negative seemed to coincide with an unusually large standard error of the regression intercept, which generally occurred when the estimated intercept was strongly negative. The large standard error of the intercept was likely due to extrapolating the recovery model to zero concentration; the error around a regression line is greatest for concentrations furthest away from the mean spike level. The effect of this extrapolation also may be seen in the Episode 6000 data. No negative results were used in the MDV and LOQ calculations, yet the median recovery intercept for the analytes analyzed in the Episode 6000 dataset was equal to -0.11. The standard errors of the intercept and slope estimates were generally high (intercept median= 0.27, slope median=0.011), and therefore the estimated intercept and slope terms were frequently not significantly different from 0 and 1, respectively (intercept: not different from zero for 167 analytes/methods; slope not significantly different from 1 for 106 analytes; both intercept and slope not significantly different for 79 analytes). Because the recovery model parameters are not significantly different from 0 or 1 for the majority of analytes, and both the estimated slope and the standard errors of the slope and intercept are included in the calculation of the MDV and LOQ, the inclusion of the recovery model estimates may bias the calculated limits, to the point that the resulting MDV can be negative.

Effect of number and spacing of concentrations for determination of the SL-IDE and SL-IQE

Tests in the Episode 6000 studies were conducted at 16 concentration levels. The IDE procedure suggests using at least 5 concentration levels. Based on statistical theory we would expect the number and spacing of concentration levels to affect the outcome, with a larger number of concentrations producing a more reliable estimate. EPA used the Episode 6000 dataset to test this hypothesis.

The IDE procedure suggests spike concentrations at 0.5, 1.0, 2, 4, and 8 times an initial estimate of the IDE (IDE_0). IDE_0 is estimated at 10 times the standard deviation of analytical results of blanks or replicates of the lowest level that can be measured. EPA's Episode 6000 database contain results of analysis of at least 7 replicates at each of at least 16 concentration levels from 0.1 to 100 times the initial estimate of the MDL (a factor of 1000). Between 0.1 and 10 times the MDL, the data are spaced a factor of approximately 1.5 apart. Above 10 times the MDL, the data are spaced at 10, 20, 50 and 100 times the MDL. The reason for the narrow spacing between 0.1 to 10 times the MDL was to attempt to allow more precise characterization of variability in the region of the MDL.

The SL-IDEs and SL-IQEs in Tables 2 and 4, respectively, were computed and reported using all 16 concentration levels because data were available at all of these levels. However, to determine the effect of the IDE procedure, a separate data analysis was performed. In this separate analysis, concentration levels were limited to a total of 5, and the 5 levels were selected to be as consistent as possible with the levels specified in the IDE procedure; i.e., at 5, 10, 20, 40, and 80 times the standard deviation of replicate measurements of a blank or the lowest level at which measurements could be made. The statement "lowest level at which measurements can be made" was interpreted to mean inclusion or

exclusion of results containing zeros and/or negative numbers. For purposes of this evaluation, concentrations that produced results containing a zero or negative number were excluded; i.e., the lowest concentration that contained no zeros or negative numbers was chosen as the concentration at which the standard deviation would be calculated for the purpose of estimating IDE_0 and IQE_0. Zeros and negative numbers were used in all of the other steps in calculating SL-IDEs and SL-IQEs.

The SL-IDE was calculated after selecting the levels based on IDE_0, and the results were compared to results produced when all 16 levels were included in calculating the SL-IDE. Results are summarized in Table 8. This table shows that the median percent difference between the 6-point IDE and the 16-point IDE is approximately -24.9% (where negative percent differences indicate that the 5-point IDE is less than the 16-point IDE). For those instances in which the same model was chosen (108 out of 198), the median percent difference was -35.6%, which was significantly different from 0% based on both the Wilcoxon rank-sum test and the sign test ($p < 0.0001$ for both tests). For those instances in which a different model was chosen (90 out of 198), the median percent difference was 1.3%, which was not significantly different from 0% based on either test (Wilcoxon: $p=0.85$; sign test: $p>0.99$). Because the choice of model can have a confounding effect on any differences between 16-point and 5-point SL-IDEs, the focus should be on the instances in which the same model was chosen. For these instances, the results indicate that only data in the region of detection and quantitation should be used to establish a detection or quantitation limit.

A similar comparison was performed between SL-IQEs (10%) calculated using all concentration levels to SL-IQEs (10%) calculated using only 5 concentration levels. Results of this comparisons are summarized in Table 9. While differences between the two calculations were not significant based on either the sign test ($p=0.567$) or the Wilcoxon test ($p=0.345$), the differences were larger than those between SL-IDEs, as seen by the larger median percent difference of -194.6%. Unlike the IDE comparison, a different model was used to calculate the 5-point SL-IQE than was used to calculate the 16-point SL-IQE for most analytes. For these 145 analytes, the percent differences were large (median percent difference = 613.9%) but not systematically positive or negative (sign test: $p=0.507$, Wilcoxon: $p=0.606$). For the 50 analytes for which the same model was used to calculate the 5-point and 16-point SL-IQEs, the percent differences were strongly negative (median percent difference = -2,442.7%) and significantly less than 0 (sign test: $p=0.015$, Wilcoxon: $p=0.0007$).

The reason for the use of 5 versus 16 concentration levels yielded significantly different results for the SL-IDE, but not for the SL-IQE, was likely due to the different model types that are recommended in the ASTM IDE and IQE procedures. Systematic differences in the calculated limit appear to occur when the same model type is applied to the 5-point and 16-point datasets. Because the exponential model is chosen based on the significance tests for most analytes in the IDE procedure, the model type used rarely differs between the two sets. There was less consistency in selecting models in the IQE procedure, and the choice of model differed between the 5-point and 16-point SL-IQE for approximately 75% of the analytes. Some of these differences, such as using the constant model instead of the hybrid model for the 5-point SL-IQE calculation, appeared to result in higher SL-IQEs, while others, such as using the linear of hybrid model in place of the constant model for the 5-point calculation, appeared to yield lower SL-IQEs. Therefore, while differences in the selected model resulted in large percent differences, these differences were not consistently positive or negative.

Relative Standard Deviation at the ML and SL-IQE in the Episode 6000 Study

The minimum level of quantitation (ML) is directed at the level at which 10% relative standard deviation (RSD) is attained. However, because the ML is not established at exactly 10% RSD, but is determined by multiplying the standard deviation that is obtained in determination of an MDL by 10 (as recommended by both ACS and Currie for ACS and ISO/IUPAC LOQs), the resulting RSD may not be 10%. The Episode 6000 data provided the opportunity to determine the actual value of the RSD at the ML. For analytes that did not have a spike concentration at the ML, the RSD was determined by linear interpolation between spike levels. Results of the determination showed that the overall median RSD at the ML across all analytes in the Episode 6000 study was 9%, and the median RSD for the 10 analytical techniques ranged between 4 and 16 percent. For 29 analytes, no RSD could be calculated because signals were not generated for samples spiked at the ML. This was likely due to limitations with this dataset that are discussed earlier in this Appendix (see "EPA's Variability versus Concentration Studies"). For 114 of the 169 remaining analytes, the RSD fell between 5% and 15%. Among the analytes that fell outside this range, 28 had RSDs below 5% and 27 had RSDs greater than 15%.

Because IQEs target a specified RSD, RSDs were also calculated based on the SL-IQEs calculated for the Episode 6000 data. Unlike the ML, the SL-IQE procedure does not contain a rounding step and, therefore, the calculated value never corresponded to one of the spike levels used in the study. For this reason, interpolation was required to calculate RSDs at the given SL-IQE value. The overall median RSD based on the SL-IQEs was 7%, with method-specific median RSDs ranging from 6% to 11%. No RSD could be calculated for 9 analytes because signals were not generated for samples spiked immediately above or below the SL-IQE. Similarly to the ML, this was likely due to issues with this dataset that are discussed earlier in this Appendix.

Effect of Outliers on Detection/Quantitation Calculations

The detection and quantitation limits based on the Episode 6000 dataset presented in Tables 2 through 5 were calculated without removing any outlying results. This decision was made based on several reasons. There were generally only 7 results per spike level for each analyte, which is a very small dataset for which to apply outlier tests and removal. In addition, MDL and ML procedures do not include outlier removal steps and, therefore, removing outliers for any of the other procedures would hinder comparisons of the calculated limits themselves. However, based on stakeholder comments, an assessment of the effect of outlier removal procedures on the different detection and quantitation limits was added to this Appendix.

Table 11 shows MDLs and SL-IDEs calculated after Grubbs outlier test (Grubbs F.E. "Procedures for Detecting Outlying Observations in Samples," *Technometrics*, vol. 11 No. 1 1969) was applied to the data. Grubbs test was run at the 5% significance level, and a maximum of one result per spike level was removed based on the results of the test. The choice of outlier test and the associated significance level follows instructions in ASTM-D2777. However, a significance level of 1% is more appropriate for outlier removal tests, as a small sample size coupled with the significance level of 5% can lead to inappropriate removal of outliers. This is true especially for studies evaluating multiple concentrations. For example, in the Episode 6000 study, there were 16 concentrations and 149 of the 198 analytes considered had an outlier present at one or more concentrations based on application of Grubbs test with 5% significance level.

For each analyte, the percent difference of the SL-IDE or MDL calculated using all data compared to the SL-IDE or MDL (calculated using the data after outlier removal) was determined. Summary statistics of these ratios are presented in Table 11. Analytes without outliers are not included in the table or the analyses discussed in this section.

Generally, SL-IDEs decreased slightly when outliers were removed. This is not surprising, as the removal of an outlying result decreases the variability at that spike level. The decrease in the SL-IDEs was not large, however, as the median percent difference comparing SL-IDEs calculated with and without outlier removal was 14.3%, where a positive percent difference indicates that the SL-IDE calculated without outlier removal was greater than the SL-IDE calculated after outlier removal. For a few analytes, removing outliers led to a change in the choice of model used to calculate the SL-IDE. In these cases, the presence of the outliers generally forced the constant model to be used; when outliers were removed, the exponential model was used. Therefore, the change in the calculated SL-IDE for these analytes was greater (median percent difference = 114.7%).

Removal of outliers only changed the MDL results if outlier removal changed the choice of spike levels used to calculate the MDL, or occurred at one of the spike levels from which the original MDL was calculated. This occurred for 60 of the 149 analytes for which any outliers were removed. In these cases, the decrease in the MDL was slightly larger than the change in the SL-IDEs (median percent difference = 30.2%).

For a small subset of analytes, either the SL-IDE or MDL increased after outlier removal. Generally, these increases were very small, and were likely due to increased tolerance factors or decreased mean recoveries for the SL-IDE, or to increased t-statistics for the MDL.

SL-IQEs and MLs calculated with and without outlier removal are presented in Table 12. The effect of outlier removal on calculated SL-IQEs and MLs was generally similar to that on the SL-IDEs and MDLs. For the SL-IQE, the choice of model changed more frequently than for the SL-IDE (31 analytes compared to 8 for the SL-IDE). However, the median percent difference was almost equal to that for the SL-IDE (16.3%). The calculated ML changed based on outlier removal for only 31 analytes, compared to 60 for the MDL. This number was smaller than for the MDL because the ML rounding frequently overshadowed the effect of outliers. However, for the ML, the changes that did occur were greater (median percent difference = 66.7%).

Evaluation of IDE/IQE Procedures

Comparison of IDE and IQEs calculated using Different Models

In the February 2003 Assessment Document, EPA expressed concern about the large amount of variability between calculated IDEs and IQEs resulting from the four different model types, and the subjectivity involved in selecting the most appropriate model. One stakeholder commented that this concern was not valid, and that IDEs calculated using different models were generally very close. To test this statement, EPA calculated SL-IDEs and SL-IQEs using each of the four major model types, and calculated RSDs between the different values for each analyte ("cross-model RSDs"). The resulting SL-IDEs are presented in Table 13. Median RSDs calculated for all analytes are presented at the bottom of the table. For several analytes, the calculated SL-IDE based on the linear model was negative due to the negative intercept of the fitted model. Because the ASTM procedure for calculating the IDE states

that the linear model should not be used in these instances, the SL-IDE based on the linear model was not included in these RSD calculations.

There is a large amount of variability between RSDs calculated with these data using the different models. Generally, SL-IDEs calculated using the constant model were much greater than those calculated using the other models. The hybrid model yielded the lowest SL-IDEs, excluding cases where the linear model SL-IDE was negative. The SL-IDEs calculated using the hybrid and exponential models were quite similar for some analytes, but quite different for others. When examined separately by method, the variability between models was generally smaller for metals methods than organics methods. However, there was a large difference in cross-model RSDs between the two metals methods, (i.e., IDEs across models in Method 1620 had a median RSD of 27%, whereas IDEs across models in Method 200.8 had a median RSD of 88%).

RSDs between SL-IQEs calculated using the different models are included in Table 14. The variability between the different model estimates was similar to that of the SL-IDEs, with a median RSD of 136% between SL-IQEs (10%). Method-specific median cross-model RSDs among SL-IQEs (10%) ranged from 24% for Method 1620 to 166% for Method 524.2.

To assess the effect of interlaboratory variability on the differences between estimates calculated from different models, cross-model RSDs were determined between the different IDEs and IQEs calculated based on the interlaboratory validation studies of Methods 1631 and 1638. These RSDs are presented in Table 7. Based on these data, the variability between model estimates appears to increase when the variability between laboratories is included. Cross-model RSDs between the IDEs calculated from the different model types ranged between 61% and 162%, with a median of 123%. These RSDs are greater than those calculated using the single-laboratory metals data in Episode 6000. Variability between IQEs was smaller than the variability between IDEs. Cross-model RSDs between IQEs ranged between 50% and 190%, with a median of 99%.

Comparison of IDE/IQEs Calculated Using Different Software Packages

A stakeholder commenting on EPA's February 2003 data assessment stated that the Agency's concerns about the complexity and subjectivity in the IDE and IQE procedures were unimportant due in part to the availability of software that will automatically perform the IDE and IQE calculations. EPA obtained two software packages from this stakeholder (see the section titled "Computations") to aid in responding to this and other comments regarding the calculation of IDEs and IQEs in the February 2003 TSD.

EPA compared these two software programs using a random subset of 20 analytes from the Episode 6000 dataset. To ensure that differences between results were due to the programs themselves, the same data were used for each program. Table 15 presents a comparison of the IDE and IQE_{10} (IQE at 10% RSD) results based on the two software packages, along with limits calculated using SAS programs (the latter limits match those presented in Tables 2 and 4). In addition, summary statistics of this comparison are presented in Table 16. Comparisons between IDEs and IQEs calculated using QCalc and the Excel software could not be done for all models, because QCalc only performs each calculation using two of the four models (exponential and hybrid for the IDE calculation, and linear and hybrid for the IQE calculation).

Generally, IDEs and IQEs calculated using SAS programs were very close to those determined using QCalc based on the same model type. The median ratio of the IDE or IQE calculated using SAS compared to the IDE or IQE calculated using QCalc equaled 0.99 or 1.00 for all model types. For two analytes (1,1,-dichloroethene and selenium by Method 1620) the hybrid IDEs and IQEs differed greatly between QCalc and the SAS programs. This appeared to be because the intercept term estimated by QCalc was negative for these analytes (resulting in negative IDEs and IQEs), whereas the intercept term estimated by SAS was approximately the positive absolute value of this estimate (resulting in positive IDEs and IQEs).

IDEs and IQEs calculated using the Excel file were generally comparable to those calculated using the SAS programs and QCalc for the constant, linear, and exponential models. The differences between the values calculated using the Excel file and other packages, however, were much greater for the hybrid model. As seen by the median ratios, the estimated IDEs and IQEs determined based on the hybrid model using Excel were slightly higher than those determined using SAS, and approximately twice those determined using QCalc. Part of this difference is due to the negative values calculated by QCalc for two analytes. However, the calculated values differed greatly, as the resulting IQEs calculated by Excel using the hybrid model ranged from less than 0 to more than 6 times greater than that calculated using the SAS programs. These differences seem to be due to how the hybrid model is fit using Excel. The Solver add-in function used by Excel does not seem to follow the same Newton's Non-Linear Least Squares algorithm described in the ASTM procedures and followed by EPA's SAS programs and QCalc.

In addition to differences in calculated limits based on the same model type, the different programs may yield different IDEs or IQEs based on which model type is indicated as most appropriate by a particular software package. QCalc and the Excel file both automatically suggest the same model type for the IDE and IQE. However, EPA often used a different model type for calculating the IDE and IQE. This was done because the ASTM IDE procedure lists constant, linear, and exponential as the three major model types to be considered, whereas the ASTM IQE procedure lists the constant, linear, and hybrid as the three major model types. Therefore, while the exponential model was used by EPA to calculate most IDEs, it was not used to calculate any of the IQEs. Because of this, while EPA and QCalc selected the same model type to calculate the IDE for only one analyte, the same model type was selected to calculate the IQE for 17 of the 20 analytes.

The Excel file frequently chose a different model type than QCalc and the EPA SAS programs to calculate the IDE and IQE. The Excel file selected a different model type than QCalc for 14 of the 20 analytes, and selected a different model than EPA's SAS program to calculate the IDE and IQE for 19 and 17 analytes, respectively. The reason for this appears to be that the Excel file suggests that the appropriate decision be based on which model has the smallest sum of squared residuals. This is different from the statistical tests of slope and curvature used by QCalc and the SAS programs and also described in the ASTM procedures. While both QCalc and the Excel file also include graphs to aid in model selection, and could potentially yield more consistent model selection through these graphs, it is likely that many users would prefer the clearer answer provided by statistical tests or comparisons of sums of squared residuals.

Based on these differences in selecting and fitting models, it does not appear that the two available software programs remove all complexity and subjectivity from the calculation of IDEs and IQEs. Instead, they appear to introduce new issues by using steps not included in the ASTM procedures. While QCalc appears to follow the ASTM procedures more closely than the Excel file, it does not perform

calculations for all model types and, therefore, may introduce greater subjectivity by only providing calculated limits based on inappropriate models.

Effect Of Long-Term Variability

Several stakeholders commenting on EPA's assessment expressed concern about the lack of long-term variability included in the MDL procedure. Commenters state that the lack of long-term variability leads to underestimates of Currie's critical value (L_C). In addition, ACIL included datasets containiug results of blank samples analyzed over three months for 5 analytes to show this effect. These commenters pointed to the ACIL procedures for calculating the critical level (CRV) and long-term MDL (ACIL LT-MDL) and the USGS procedure for calculating the long-term MDL (LT-MDL), which include the collection of blanks over a long period of time.

EPA assessed the effect of long-term variability on calculated limits by simulating multiple 7-replicate subsets from the full dataset, and comparing these short-term CRVs to the CRV calculated using the full dataset. Subsets were created by taking each set of 7 consecutive blanks, (i.e., the first 7 blanks analyzed would be one subset, the 2nd through 8th blanks analyzed would be another subset, etc.). This yielded a total of n-6 subsets, where n is the number of total blanks for that analyte. Because a blank will be used in as many as 7 subsets, the variability of the short-term CRVs was lower than what would be expected; however, the approach was used to yield the greatest number of simulated subsets. The CRV was then calculated for each subset:

$$CRV_i = \overline{X}_i + s_i * t_{(0.99,6)}$$

where CRV_i, $\overline{X}_i$, and s_i are the critical value, the mean, and the standard deviation for the i^{th} simulated subset of blank results, respectively. The overall CRV was calculated using the same formula, using the mean and standard deviation based on all blank results and a lower t-statistic based on the greater number of blank replicates. Table 17 shows the results of the comparison of calculated short-term and long-term CRVs for the five analytes.

While the range of days from which sets of 7 replicates were simulated varied from between one week to greater than 3 weeks, graphical analyses did not show any effect of the number of days on the resulting CRV. The total number of blanks also did not seem to have an effect on the percentage of short-term CRVs that exceeded the overall CRV. The mean short-term CRV was generally very close to the overall CRV for each analyte. However, for three of the five analytes, the majority of the short-term CRVs exceeded were lower than the overall CRV, indicating that long-term variability did have an effect on the resulting limit. For the other analytes, the effect of any added variability was counteracted by the smaller t-statistic used in the calculation. These t-statistics ranged between 2.4 and 2.5 between analytes, well below the 3.14 used when only 7 replicates are available.

One possible reason for the number of short-term CRVs falling below the overall CRV was the presence of outliers. The ACIL procedure permits the use of an outlier procedure to remove outlying high or low blanks. EPA used Grubbs test and identified 3 blank results for silver, and 1 blank result each for barium and chromium, as outliers. After removal of these results, the overall and short-term CRVs were re-calculated for these 3 analytes. The results of these calculations are given in Table 18.

Because an outlying result is used in the calculation of the overall CRV (but only for a maximum of 7 of the short-term CRVs), the effect of outlier removal was greater for the overall CRV than on the short-term CRVs. For all 3 analytes, the majority of the short-term CRVs were above the overall CRV, and the mean short-term CRV was slightly higher than the overall CRV. This was consistent with the results of cadmium and copper shown in Table 17, for which no outliers were detected. Because no information was available about why these results could have been outlying, it is not known if they were the result of a known error, or were in fact the result of the long-term variability included in the study. However, it appears that the effect of long-term variability is generally not large when compared to the effect of using more replicates on the t-statistic multiplier.

As stated in Section 3.3.3, a greater number of replicates will yield improved estimates of standard deviation and, therefore, better estimates of Currie's L_c. Based on this, although, EPA does not feel estimations of L_c based on 7 replicates are biased low, these estimates may be less precise than those based on greater replicates. The large variability of the 7-replicate CRVs can be seen in the large ranges of short-term CRVs calculated with and without outlier removal. The use of the higher t-statistic also seems to counteract the added long-term variability. The ACIL procedure suggests 7-replicate CRVs are underestimates, and should therefore be multiplied by a factor of 2. The short-term CRV calculated in the ACIL procedure is based on blanks analyzed in a single batch and, therefore, are not comparable to the short-term CRVs simulated by EPA. However, such a multiplier is not necessary in calculating the MDL, even if long-term variability is not included in the analyses.

SUMMARY

Public comment on the February 2003 Assessment Document and the proposed regulatory revisions expressed many divergent views about the merits and usefulness of EPA's 2003 assessment and proposed regulatory revisions. We recognize that there is a broad interest in improving current procedures and uses, but no consensus for a specific procedure or procedures has emerged among the laboratory, industry, regulatory or regulated communities. Thus, we have withdrawn the March 2003 proposed revisions and, to meet the terms of the settlement agreement that is described in chapter 1, are taking final action on the 2003 Assessment Document in this Revised Assessment Document. This is not the end of our efforts to work together, as stakeholders have suggested, to discuss mutual concerns and possible solutions. We look forward to continued stakeholder participation in an ongoing dialog about the development and use of detection and quantitation limits in CWA programs.

In this appendix, we have compared detection and quantitation limits computed from data gathered by EPA or submitted to EPA. This comparison shows that, in general, detection limits derived from a single concentration level such as EPA's MDL are, on average, approximately the same as detection limits derived from similar approaches such as the ACS LOD and LOQ and ISO/IUPAC CRV and MDV, and are approximately three times lower than a single-laboratory variant of ASTM's IDE; and that all quantitation limit approaches, such as EPA's ML, the ACS and ISO/IUPAC LOQ, and a single-laboratory variant of ASTM's IQE, produce quantitation limits that are generally only slightly different. In addition, the following are general statements about the datasets and/or analyses described in this appendix.

1. *Variability of Results*

Comparisons of detection and quantitation limits show high variability among the limits calculated using the different approaches, even with data containing 7 replicates at 16 concentration levels (see

the summary statistics at the end of Tables 3, 5, and 7). The net effect is that the systematic differences among detection and quantitation limits produced by the various approaches are overwhelmed by variability; i.e., there is a small systematic difference among the approaches but great variability in the detection and quantitation limits for a given analyte. This result is not surprising given the variability of data in the region of detection and quantitation. However, it is difficult to postulate a solution to the problem. Gathering more data in the region of detection and quantitation would appear to be a solution, but 91 data points were gathered for each analyte in the region between 0.1 and 10 times the MDL in the Episode 6000 studies, and it is unlikely that any organization could afford to gather even this amount of data for determination of a detection limit. Given the high degree of variability of the data, EPA's approach of conducting a single-laboratory study to gain a first estimate, followed by multiple single-laboratory studies to verify or revise the estimate, and an interlaboratory study, where warranted, to further verify and revise the estimate, is a reasonable means of establishing detection and quantitation limits because of the checks and balances that occur at each step.

2. *Regression Analysis*

Using a regression line to estimate a recovery correction at zero concentration causes great swings in the resulting detection and quantitation limits such as the ISO/IUPAC MDV and LOQ. The estimated regression parameters for the recovery models were often not significant, and the inclusion of the estimated slope and the standard errors of the slope and intercept will, therefore, unnecessarily bias the calculated MDV and LOQ, such that the calculated MDVs may be negative (see Discussion section "Negative detection limits for the ISO/IUPAC MDV, and Table 2 for instances of negative detection limits"). The estimated recovery model used in calculating the IDE and IQE is also strongly affected by the chosen model of variability vs. concentration (see Tables 13 and 14). Even though a linear regression is used to model recovery in each case, the weights used in the model are calculated based on the variability model, and can vary greatly when the number of concentrations used is low. For the Episode 6000 data, the median RSD of the recovery slopes from the four different models used in the IDE calculations for a given analyte and method was 5%. In addition, for 77 of the analytes and methods (39%), at least one estimated recovery slope was greater than 1, and at least one was less than 1. This suggests that the method could be considered to be high biased (and the final IDE and IQE would be decreased by the recovery correction) and low biased (and the final IDE and IQE increased) for the analyte, depending on the chosen precision model. For many analytes the slopes were not significantly different from 1, suggesting that a recovery correction may not be appropriate at all. This is in addition to the philosophical issue as to whether recovery correction is warranted. If there is to be a correction for recovery, it may be better to use some average or median value than a regression, or use a measured value near the region of interest.

3. *IDE and IQE*

Additional development of the ASTM IDE and IQE is needed before they can be used routinely, not only because of the complexity of the procedures, but also because of the ambiguity in determining that the correct model has been selected. While different software packages are available that perform most of the calculations, there are many inconsistencies between these programs, and between the programs and the ASTM procedures, that add another area of subjectivity to the determination of IDE and IQEs. (For the consequences of model selection, compare the IDEs and IQEs determined by EPA and EPRI in Table 6, and the IDEs and IQEs calculated from the different

model types in Table 7. Some differ considerably as a result of model selection in application of the IDE and IQE procedures by different statisticians. In addition, the use of different software may lead to the selection of different models, as seen in Table 15.)

4. *Quantitation Limit Approaches*
Quantitation limit approaches such as EPA's ML and the ACS and ISO/IUPAC LOQ that are directed 10% RSD actually produce RSDs that are in the range of the 10% intended (see the discussion in the Section titled "RSD at the ML in the Episode 6000 Study"). The median RSDs for each method in the Episode 6000 dataset ranged from 6% to 16%, and 58% of the individual analyte RSDs fell between 5% and 15%.

Commenters on our February 2003 Assessment Document suggested that procedures submitted by a laboratory association (ACIL) and the U.S. Geological Survey as alternatives to the MDL and ML should be considered. We agree, have evaluated these procedures in this Revised Assessment Document, and believe they provide a starting point for continued stakeholder discussions.

Regarding these two procedures, we note the ACIL CRV generally yielded lower false positive rates than the USGS LT-MDL. This likely was due to the nonparametric estimate of variability used in the USGS procedure. False positive rates for the EPA MDL, which uses a parametric variability but does not include the mean blank result, were lower than the USGS LT-MDL, which does include the mean blank result. The ACIL procedure states that calculated CRVs are based on fewer replicates and/or short-term variability are biased low, and includes optional alternate calculations to use in these situations. However, comparison of CRVs calculated with full set of long-term blanks to those calculated with subsets of 7 blanks suggest that the absence of long-term variability is counteracted by the larger t-statistic used when the number of blank results is smaller.

ACIL also included a separate procedure for methods for which analysis of blank samples does not always produce a signal. The idea of dividing methods into two groups has merit. However, the current ACIL procedure for these methods often generates estimates of Currie's L_C that are above the estimate of Currie's L_D when contamination is present.

TABLES

Table 1. Datasets Suggested by Petitioners

Dataset and year	Analyte and technology
AAMA 1996-7	Metals by ICP/AES (200.7)
AAMA 1996-7	Mercury by CVAA (245.2)
AAMA 1996-7	PCBs by GC/ECD (608.2)
MMA 2000-1	PCB 1216 and 1260 by GC/ECD
EPA/EPRI 1997-8	Mercury by CVAF (1631)
EPA/EPRI 1997-8	Metals by ICP/MS (1638)
EPRI 1987	Metals by GFAA (EPA 200)
EPRI 1990	Metals by ICP/AES (EPA 200.7)
EPRI 1994	Ai, Be, Ti by GFAA (EPA 200)
EPRI 1996	Cd, As, Cr by GFAA (EPA 200)

Table 2. Comparison of Detection Limits (µg/L except where footnoted) for the Episode 6000 Dataset

Analyte	Method	Procedure	EPA/ ACS DL	ISO CRV	ISO MDV	ASTM SL-IDE
1,1,1,2-tetrachloroethane	502.2	ELCD	0.041	0.005	0.009	0.034
1,1,1,2-tetrachloroethane	524.2		0.052	0.039	-0.047	0.244
1,1,1-trichloroethane	502.2	ELCD	0.012	0.009	0.017	0.041
1,1,1-trichloroethane	524.2		0.055	0.021	0.003	0.308
1,1,2,2-tce+1,2,3-tcp	502.2	ELCD	0.064	0.047	0.086	0.179
1,1,2,2-tetrachloroethane	524.2		0.132	0.131	0.128	0.436
1,1,2-trichloroethane	502.2	ELCD	0.024	0.004	0.006	0.032
1,1,2-trichloroethane	524.2		0.075	0.043	0.040	0.319
1,1-dichloroethane	502.2	ELCD	0.010	0.007	0.014	0.083
1,1-dichloroethane	524.2		0.033	0.020	0.016	0.229
1,1-dichloroethene	502.2	ELCD	0.038	0.030	0.073	0.234
1,1-dichloroethene	524.2		0.054	0.035	-0.037	0.335
1,1-dichloropropanone	524.2		5.184	3.146	5.635	6.372
1,1-dichloropropene	524.2		0.045	0.012	-0.030	0.287
1,2,3-trichlorobenzene	502.2	ELCD	0.048	0.034	0.065	0.134
1,2,3-trichlorobenzene	502.2	PID	0.057	0.042	0.088	0.115
1,2,3-trichlorobenzene	524.2		0.070	0.040	0.031	0.275
1,2,3-trichloropropane	524.2		7.328	0.046	0.033	1.263
1,2,4-trichlorobenzene	502.2	ELCD	0.022	0.014	0.030	0.088
1,2,4-trichlorobenzene	502.2	PID	0.070	0.038	0.080	0.124
1,2,4-trichlorobenzene	524.2		0.053	0.050	0.052	0.224
1,2,4-trimethylbenzene	502.2	PID	0.095	0.053	0.119	0.125
1,2,4-trimethylbenzene	524.2		0.012	0.009	0.017	0.144
1,2-dibromo-3-chloropropane	524.2		1.457	0.391	0.701	1.749
1,2-dibromoethane	502.2	ELCD	0.096	0.007	0.013	0.164
1,2-dibromoethane	524.2		0.127	0.117	0.170	0.326
1,2-dichlorobenzene	502.2	ELCD	0.035	0.031	0.061	0.065
1,2-dichlorobenzene	502.2	PID	0.033	0.024	0.054	0.148
1,2-dichlorobenzene	524.2		0.030	0.023	-0.016	0.130
1,2-dichloroethane	502.2	ELCD	0.017	0.003	0.005	0.042
1,2-dichloroethane	524.2		0.039	0.024	0.013	0.258

Table 2. Comparison of Detection Limits (µg/L except where footnoted) for the Episode 6000 Dataset

Analyte	Method	Procedure	EPA/ ACS DL	ISO CRV	ISO MDV	ASTM SL-IDE
1,2-dichloropropane	502.2	ELCD	0.023	0.014	0.029	0.043
1,2-dichloropropane	524.2		0.056	0.030	0.026	0.247
1,3,5-tmb+4-chlorotoluene	502.2	PID	0.067	0.045	0.100	0.114
1,3,5-trimethylbenzene	524.2		0.011	0.008	0.008	0.135
1,3-dichlorobenzene	502.2	ELCD	0.035	0.005	0.010	0.118
1,3-dichlorobenzene	502.2	PID	0.093	0.077	0.170	0.126
1,3-dichlorobenzene	524.2		0.023	0.016	-0.014	0.143
1,3-dichloropropane	502.2	ELCD	0.016	0.008	0.015	0.047
1,3-dichloropropane	524.2		0.038	0.024	-0.015	0.202
1,4-dichlorobenzene	502.2	ELCD	0.026	0.005	0.009	0.061
1,4-dichlorobenzene	524.2		0.023	0.017	-0.044	0.140
1-chlorobutane	524.2		0.020	0.016	0.018	0.220
2,2-dichloropropane	524.2		2.376	0.103	-0.159	0.691
2-butanone	524.2		0.417	0.297	0.511	0.833
2-chlorotoluene	502.2	ELCD	0.108	0.029	0.056	0.175
2-chlorotoluene	502.2	PID	0.238	0.135	0.302	0.230
2-chlorotoluene	524.2		0.016	0.009	0.002	0.136
2-hexanone	524.2		1.316	0.148	0.231	0.902
2-nitropropane	524.2		0.901	0.275	0.452	1.082
4-chlorotoluene	502.2	ELCD	0.110	0.027	0.050	0.149
4-chlorotoluene	524.2		0.010	0.008	0.007	0.123
4-isopropyltoluene	524.2		0.010	0.008	0.003	0.117
4-methyl-2-pentanone	524.2		0.812	0.480	0.733	1.195
Acetone	524.2		0.859	0.440	0.804	2.120
Acrylonitrile	524.2		0.863	0.444	0.653	1.333
Allyl Chloride	524.2		0.032	0.026	0.005	0.229
Aluminum	1620		29.555	15.043	28.666	206.975
Aluminum	200.8		19.145	1.690	3.547	12.747
Ammonia as Nitrogen[1]	350.3		0.010	0.007	0.014	0.014
Antimony	1620		1.552	0.801	1.754	4.260
Antimony	200.8		0.178	0.003	0.007	0.019
Arsenic	1620		1.065	0.917	1.375	1.410

Table 2. Comparison of Detection Limits (µg/L except where footnoted) for the Episode 6000 Dataset

Analyte	Method	Procedure	EPA/ ACS DL	ISO CRV	ISO MDV	ASTM SL-IDE
Arsenic	200.8		0.226	0.137	0.272	0.366
Barium	1620		1.702	1.337	1.831	1.837
Barium	200.8		0.033	0.029	0.061	0.084
Benzene	502.2	PID	0.030	0.029	0.067	0.079
Benzene	524.2		0.014	0.014	0.026	0.125
Beryllium	1620		0.528	0.339	0.408	0.448
Beryllium	200.8		0.007	0.004	0.006	0.024
Boron	1620		15.387	10.356	17.792	21.161
Bromobenzene	502.2	ELCD	0.131	0.093	0.186	0.765
Bromobenzene	502.2	PID	0.012	0.009	0.019	0.050
Bromobenzene	524.2		0.044	0.036	-0.060	0.211
Bromochloromethane	502.2	ELCD	0.013	0.012	0.024	0.482
Bromochloromethane	524.2		0.125	0.113	0.159	0.345
Bromodichloromethane	502.2	ELCD	0.004	0.003	0.005	0.075
Bromodichloromethane	524.2		0.043	0.026	0.019	0.205
Bromoform	502.2	ELCD	0.006	0.003	0.001	1.513
Bromoform	524.2		0.123	0.065	0.031	0.400
Bromomethane	502.2	ELCD	0.267	0.219	0.358	7.293
Bromomethane	524.2		0.068	0.055	0.056	0.280
Cadmium	1620		0.127	0.079	0.134	0.191
Cadmium	200.8		0.004	0.007	0.012	0.022
Calcium	1620		36.726	35.822	72.397	41.358
Carbon Disulfide	524.2		0.027	0.015	-0.040	0.239
Carbon Tetrachloride	524.2		0.038	0.027	-0.040	0.314
Carbontet+1,1-dcp	502.2	ELCD	0.029	0.008	0.016	0.072
Chloroacetonitrile	524.2		0.919	0.773	1.527	1.569
Chlorobenzene	502.2	ELCD	0.011	0.010	0.022	0.460
Chlorobenzene	502.2	PID	0.030	0.025	0.055	0.064
Chlorobenzene	524.2		0.025	0.022	0.012	0.133
Chloroethane	502.2	ELCD	0.108	0.008	0.009	2.598
Chloroethane	524.2		0.066	0.041	0.038	0.395
Chloroform	502.2	ELCD	0.043	0.006	0.009	0.032

Table 2. Comparison of Detection Limits (µg/L except where footnoted) for the Episode 6000 Dataset

Analyte	Method	Procedure	EPA/ ACS DL	ISO CRV	ISO MDV	ASTM SL-IDE
Chloroform	524.2		0.036	0.027	0.021	0.225
Chloromethane	502.2	ELCD	0.070	0.049	0.130	0.250
Chloromethane	524.2		0.045	0.036	0.065	0.253
Chromium	1620		0.310	0.254	0.386	0.496
Chromium	200.8		0.073	0.062	0.125	0.408
Cis-1,2-dce+2,2-dcp	502.2	ELCD	0.013	0.009	0.016	0.055
Cis-1,2-dichloroethene	524.2		0.040	0.033	-0.023	0.234
Cis-1,3-dichloropropene	502.2	ELCD	0.007	0.002	0.004	0.074
Cis-1,3-dichloropropene	502.2	PID	0.057	0.048	0.099	0.082
Cis-1,3-dichloropropene	524.2		0.038	0.024	-0.004	0.173
Cobalt	1620		9.820	4.017	8.094	16.463
Cobalt	200.8		0.001	0.001	-0.067	0.074
Copper	1620		6.046	4.990	10.512	21.189
Copper	200.8		0.037	0.027	0.053	0.798
Dibromochloromethane	502.2	ELCD	0.009	0.006	0.011	0.436
Dibromochloromethane	524.2		0.051	0.031	0.004	0.287
Dibromomethane	502.2	ELCD	0.007	0.005	0.010	0.460
Dibromomethane	524.2		0.102	0.082	0.112	0.388
Dichlorodifluoromethane	502.2	ELCD	0.009	0.003	-0.020	0.240
Dichlorodifluoromethane	524.2		0.083	0.054	0.037	0.560
Diethyl Ether	524.2		0.120	0.114	0.163	0.376
Ethyl Methacrylate	524.2		0.045	0.031	0.013	0.273
Ethylbenzene	502.2	PID	0.021	0.015	0.035	0.078
Ethylbenzene	524.2		0.033	0.028	-0.024	0.198
Hardness [1]	130.2		0.828	0.554	1.152	2.258
Hexachlorobutadiene	502.2	ELCD	0.043	0.010	0.021	0.094
Hexachlorobutadiene	524.2		0.068	0.035	-0.031	0.308
Hexachloroethane	524.2		0.056	0.049	0.038	0.288
Hexchlobutadiene+naphthalene	502.2	PID	0.649	0.143	0.321	0.597
Iron	1620		90.409	270.433	472.249	373.590
Isopropylbenzene	502.2	PID	0.020	0.015	0.035	0.060
Isopropylbenzene	524.2		0.011	0.010	0.010	0.120

Table 2. Comparison of Detection Limits (µg/L except where footnoted) for the Episode 6000 Dataset

Analyte	Method	Procedure	EPA/ ACS DL	ISO CRV	ISO MDV	ASTM SL-IDE
Lead	1620		1.647	1.186	1.965	2.423
Lead	200.8		0.655	0.061	0.120	0.204
M+p Xylene	502.2	PID	0.090	0.012	0.026	0.121
M+p Xylene	524.2		0.013	0.008	0.004	0.142
Magnesium	1620		103.033	88.729	175.316	105.998
Manganese	1620		6.856	1.081	2.591	6.808
Manganese	200.8		0.031	0.030	0.049	0.109
Mercury	200.8		0.004	0.003	-0.018	0.027
Methacrylonitrile	524.2		0.356	0.228	0.362	0.718
Methyl Iodide	524.2		0.025	0.023	-0.013	0.193
Methyl Tert-butyl Ether	524.2		0.026	0.016	-0.033	0.225
Methyla crylate	524.2		0.220	0.202	0.353	0.601
Methylene Chloride	502.2	ELCD	0.128	1.835	4.917	2.841
Methylene Chloride	524.2		0.082	0.072	0.093	0.314
Methylm ethacryl ate	524.2		0.225	0.085	0.117	0.535
Molybdenum	1620		2.455	1.714	3.787	3.034
Molybdenum	200.8		0.004	0.003	0.000	0.271
N-butylbenzene	502.2	PID	0.030	0.023	0.049	0.141
N-butylbenzene	524.2		0.016	0.014	0.026	0.152
N-propylbenzene	502.2	PID	0.040	0.022	0.049	0.092
N-propylbenzene	524.2		0.038	0.026	-0.053	0.284
Naphthalene	524.2		0.048	0.040	0.044	0.186
Nickel	1620		20.219	13.262	25.697	25.560
Nickel	200.8		0.146	0.058	0.107	0.083
o-xylene	524.2		0.018	0.015	-0.032	0.198
o-xylene+styrene	502.2	PID	0.059	0.037	0.082	0.116
P-isopropol+1,4-dcb	502.2	PID	0.073	0.056	0.123	0.159
Pentachloroethane	524.2		0.553	0.019	-0.100	0.408
Sec-butylbenzene	502.2	PID	0.055	0.032	0.075	0.081
Sec-butylbenzene	524.2		0.014	0.011	-0.012	0.140
Selenium	1620		0.849	0.619	1.493	1.975
Selenium	200.8		0.192	0.156	0.302	0.416

Table 2. Comparison of Detection Limits (µg/L except where footnoted) for the Episode 6000 Dataset

Analyte	Method	Procedure	EPA/ ACS DL	ISO CRV	ISO MDV	ASTM SL-IDE
Silver	1620		4.907	3.588	6.495	10.668
Silver	200.8		0.004	0.002	0.004	0.012
Sodium	1620		69.530	49.595	97.649	138.768
Styrene	524.2		0.014	0.011	0.010	0.141
Tert-butylbenzene	502.2	PID	0.029	0.020	0.047	0.074
Tert-butylbenzene	524.2		0.022	0.012	0.023	0.186
Tetrachloroethene	502.2	ELCD	0.018	0.014	0.029	0.061
Tetrachloroethene	502.2	PID	0.062	0.040	0.094	0.156
Tetrachloroethene	524.2		0.085	0.084	0.047	0.469
Thallium	1620		0.512	0.651	1.406	1.153
Thallium	200.8		0.000	0.000	0.001	0.001
Thorium	200.8		0.001	0.001	-0.005	0.001
Tin	1620		3.670	2.019	3.143	3.932
Titanium	1620		4.777	4.453	8.050	5.376
Toluene	502.2	PID	0.070	0.028	0.063	0.064
Toluene	524.2		0.020	0.006	-0.004	0.146
Total Phosphorus [1]	365.2		0.006	0.005	0.009	0.013
Total Suspended Solids [1]	160.2		1.170	0.948	1.945	3.005
Trans-1,2-dichloroethene	502.2	ELCD	0.041	0.041	0.090	0.081
Trans-1,2-dichloroethene	524.2		0.038	0.032	-0.016	0.300
Trans-1,3-dichloroPropene	502.2	ELCD	0.012	0.003	0.005	0.098
Trans-1,3-dichloropropene	502.2	PID	0.058	0.045	0.095	0.092
Trans-1,3-dichloropropene	524.2		0.051	0.025	-0.007	0.223
Trans-1,4-dichloro-2-butene	524.2		0.512	0.348	0.576	1.250
Trichloroethene	502.2	ELCD	0.012	0.001	0.003	0.059
Trichloroethene	502.2	PID	0.027	0.018	0.042	0.097
Trichloroethene	524.2		0.061	0.058	0.056	0.332
Trichlorofluoromethane	502.2	ELCD	0.108	0.249	0.612	2.079
Trichlorofluoromethane	524.2		0.087	0.075	0.038	0.384
Uranium	200.8		0.000	0.000	0.000	0.000
Vanadium	1620		7.344	4.207	8.359	10.630
Vanadium	200.8		0.555	0.512	0.994	0.864

Table 2. Comparison of Detection Limits (µg/L except where footnoted) for the Episode 6000 Dataset

Analyte	Method	Procedure	EPA/ ACS DL	ISO CRV	ISO MDV	ASTM SL-IDE
Vinyl Chloride	502.2	ELCD	0.270	0.039	0.077	3.672
Vinyl Chloride	524.2		0.043	0.031	-0.007	0.365
WAD Cyanide	1677		0.572	0.169	0.319	0.701
Xylene (Total)	524.2		0.009	0.005	0.007	0.128
Yttrium	1620		1.923	1.370	2.518	3.247
Zinc	1620		2.597	2.301	3.697	4.500
Zinc	200.8		0.900	0.461	0.806	1.598

[1]Results reported as mg/L

Note: ELCD or PID in the Procedure column indicates the photo-ionization detector (PID) or electrolytic conductivity detector (ELCD) in EPA Method 502.2

Table 3. Percent Differences of Detection Limits to the EPA/ACS DL for the Episode 6000 Dataset

Analyte	Method	Procedure	ISO CRV/ MDL	ISO MDV/ MDL	SL-IDE/ MDL
1,1,1,2-tetrachloroethane	502.2	ELCD	-159.2%	-131.0%	-20.3%
1,1,1,2-tetrachloroethane	524.2		-28.9%	-4142.5%	129.8%
1,1,1-trichloroethane	502.2	ELCD	-34.4%	32.1%	108.8%
1,1,1-trichloroethane	524.2		-89.4%	-177.7%	139.1%
1,1,2,2-tce+1,2,3-tcp	502.2	ELCD	-29.7%	29.9%	94.7%
1,1,2,2-tetrachloroethane	524.2		-0.6%	-3.4%	107.0%
1,1,2-trichloroethane	502.2	ELCD	-146.2%	-116.9%	27.6%
1,1,2-trichloroethane	524.2		-53.2%	-60.4%	124.0%
1,1-dichloroethane	502.2	ELCD	-40.1%	31.0%	156.8%
1,1-dichloroethane	524.2		-50.5%	-70.3%	150.0%
1,1-dichloroethene	502.2	ELCD	-25.4%	61.8%	143.5%
1,1-dichloroethene	524.2		-42.8%	-1080.2%	144.1%
1,1-dichloropropanone	524.2		-48.9%	8.3%	20.6%
1,1-dichloropropene	524.2		-117.1%	-1021.1%	146.2%
1,2,3-trichlorobenzene	502.2	ELCD	-34.9%	30.2%	94.9%
1,2,3-trichlorobenzene	502.2	PID	-29.4%	42.0%	67.0%
1,2,3-trichlorobenzene	524.2		-53.5%	-76.9%	119.2%
1,2,3-trichloropropane	524.2		-197.5%	-198.2%	-141.2%

Table 3. Percent Differences of
Detection Limits to the EPA/ACS DL for the Episode 6000 Dataset

Analyte	Method	Procedure	ISO CRV/ MDL	ISO MDV/ MDL	SL-IDE/ MDL
1,2,4-trichlorobenzene	502.2	ELCD	-39.7%	31.4%	121.2%
1,2,4-trichlorobenzene	502.2	PID	-59.9%	13.5%	55.5%
1,2,4-trichlorobenzene	524.2		-5.0%	-1.3%	123.6%
1,2,4-trimethylbenzene	502.2	PID	-55.5%	23.0%	28.0%
1,2,4-trimethylbenzene	524.2		-25.8%	33.0%	168.6%
1,2-dibromo-3-chloropropane	524.2		-115.4%	-70.1%	18.2%
1,2-dibromoethane	502.2	ELCD	-172.1%	-150.8%	52.9%
1,2-dibromoethane	524.2		-8.6%	28.7%	87.8%
1,2-dichlorobenzene	502.2	ELCD	-12.4%	53.6%	59.5%
1,2-dichlorobenzene	502.2	PID	-30.7%	48.9%	127.6%
1,2-dichlorobenzene	524.2		-28.0%	-655.2%	125.1%
1,2-dichloroethane	502.2	ELCD	-140.1%	-106.3%	83.9%
1,2-dichloroethane	524.2		-48.6%	-98.0%	147.5%
1,2-dichloropropane	502.2	ELCD	-45.0%	22.4%	61.1%
1,2-dichloropropane	524.2		-59.7%	-75.2%	125.7%
1,3,5-tmb+4-chlorotoluene	502.2	PID	-39.6%	39.4%	51.0%
1,3,5-trimethylbenzene	524.2		-33.9%	-28.8%	169.3%
1,3-dichlorobenzene	502.2	ELCD	-151.2%	-112.2%	108.7%
1,3-dichlorobenzene	502.2	PID	-19.1%	58.3%	30.0%
1,3-dichlorobenzene	524.2		-35.5%	-754.8%	144.1%
1,3-dichloropropane	502.2	ELCD	-63.5%	-2.1%	100.1%
1,3-dichloropropane	524.2		-45.7%	-457.8%	136.4%
1,4-dichlorobenzene	502.2	ELCD	-136.9%	-94.1%	80.6%
1,4-dichlorobenzene	524.2		-33.3%	654.4%	142.5%
1-chlorobutane	524.2		-24.0%	-11.7%	166.8%
2,2-dichloropropane	524.2		-183.3%	-228.6%	-109.9%
2-butanone	524.2		-33.5%	20.2%	66.6%
2-chlorotoluene	502.2	ELCD	-116.2%	-64.0%	47.7%
2-chlorotoluene	502.2	PID	-55.5%	23.6%	-3.6%
2-chlorotoluene	524.2		-54.7%	-165.4%	158.1%
2-hexanone	524.2		-159.6%	-140.3%	-37.3%
2-nitropropane	524.2		-106.6%	-66.3%	18.2%

Table 3. Percent Differences of
Detection Limits to the EPA/ACS DL for the Episode 6000 Dataset

Analyte	Method	Procedure	ISO CRV/ MDL	ISO MDV/ MDL	SL-IDE/ MDL
4-chlorotoluene	502.2	ELCD	-119.9%	-74.8%	30.5%
4-chlorotoluene	524.2		-21.8%	-26.2%	170.8%
4-isopropyltoluene	524.2		-18.2%	-95.8%	169.2%
4-methyl-2-pentanone	524.2		-51.4%	-10.3%	38.1%
Acetone	524.2		-64.5%	-6.6%	84.7%
Acrylonitrile	524.2		-64.0%	-27.7%	42.9%
Allyl Chloride	524.2		-19.8%	-150.4%	150.6%
Aluminum	1620		-65.1%	-3.1%	150.0%
Aluminum	200.8		-167.6%	-137.5%	-40.1%
Ammonia as Nitrogen	350.3		-39.8%	30.4%	31.7%
Antimony	1620		-63.8%	12.2%	93.2%
Antimony	200.8		-193.1%	-185.9%	-161.5%
Arsenic	1620		-14.9%	25.4%	27.9%
Arsenic	200.8		-49.1%	18.7%	47.5%
Barium	1620		-24.0%	7.3%	7.6%
Barium	200.8		-12.2%	59.9%	87.9%
Benzene	502.2	PID	-2.5%	76.2%	89.5%
Benzene	524.2		-1.9%	57.8%	158.7%
Beryllium	1620		-43.8%	-25.6%	-16.5%
Beryllium	200.8		-55.8%	-16.7%	109.7%
Boron	1620		-39.1%	14.5%	31.6%
Bromobenzene	502.2	ELCD	-33.8%	34.8%	141.6%
Bromobenzene	502.2	PID	-31.6%	44.4%	121.7%
Bromobenzene	524.2		-18.1%	1274.8%	131.5%
Bromochloromethane	502.2	ELCD	-11.8%	55.9%	189.2%
Bromochloromethane	524.2		-10.3%	23.8%	93.6%
Bromodichloromethane	502.2	ELCD	-35.9%	27.1%	178.8%
Bromodichloromethane	524.2		-47.8%	-76.2%	130.5%
Bromoform	502.2	ELCD	-64.7%	-129.0%	198.4%
Bromoform	524.2		-62.6%	-120.5%	105.6%
Bromomethane	502.2	ELCD	-19.7%	29.2%	185.9%
Bromomethane	524.2		-21.0%	-19.6%	122.1%

Table 3. Percent Differences of
Detection Limits to the EPA/ACS DL for the Episode 6000 Dataset

Analyte	Method	Procedure	ISO CRV/ MDL	ISO MDV/ MDL	SL-IDE/ MDL
Cadmium	1620		-47.1%	5.5%	40.1%
Cadmium	200.8		55.5%	99.6%	138.8%
Calcium	1620		-2.5%	65.4%	11.9%
Carbon Disulfide	524.2		-52.8%	990.7%	160.0%
Carbon Tetrachloride	524.2		-33.8%	10302.8%	156.6%
Carbontet+1,1-dcp	502.2	ELCD	-110.8%	-55.2%	85.6%
Chloroacetonitrile	524.2		-17.3%	49.7%	52.3%
Chlorobenzene	502.2	ELCD	-11.3%	61.5%	190.3%
Chlorobenzene	502.2	PID	-19.2%	58.8%	71.4%
Chlorobenzene	524.2		-12.7%	-66.3%	137.5%
Chloroethane	502.2	ELCD	-171.0%	-169.4%	184.1%
Chloroethane	524.2		-47.0%	-53.1%	142.6%
Chloroform	502.2	ELCD	-150.5%	-129.4%	-27.3%
Chloroform	524.2		-29.2%	-51.9%	144.5%
Chloromethane	502.2	ELCD	-34.7%	60.2%	112.8%
Chloromethane	524.2		-21.8%	37.1%	139.8%
Chromium	1620		-20.0%	21.9%	46.3%
Chromium	200.8		-16.5%	52.5%	139.3%
Cis-1,2-dce+2,2-dcp	502.2	ELCD	-39.4%	21.8%	124.0%
Cis-1,2-dichloroethene	524.2		-19.1%	-760.6%	141.9%
Cis-1,3-dichloropropene	502.2	ELCD	-101.0%	-61.1%	164.6%
Cis-1,3-dichloropropene	502.2	PID	-17.5%	54.1%	36.0%
Cis-1,3-dichloropropene	524.2		-47.6%	-251.5%	127.2%
Cobalt	1620		-83.9%	-19.3%	50.6%
Cobalt	200.8		-23.4%	206.3%	194.5%
Copper	1620		-19.1%	53.9%	111.2%
Copper	200.8		-33.0%	35.3%	182.2%
Dibromochloromethane	502.2	ELCD	-46.7%	17.2%	191.8%
Dibromochloromethane	524.2		-49.9%	-168.4%	139.6%
Dibromomethane	502.2	ELCD	-21.1%	38.8%	194.4%
Dibromomethane	524.2		-21.5%	9.2%	116.8%
Dichlorodifluoromethane	502.2	ELCD	-91.4%	511.1%	185.7%

Table 3. Percent Differences of Detection Limits to the EPA/ACS DL for the Episode 6000 Dataset

Analyte	Method	Procedure	ISO CRV/ MDL	ISO MDV/ MDL	SL-IDE/ MDL
Dichlorodifluoromethane	524.2		-42.3%	-76.4%	148.1%
Diethyl Ether	524.2		-4.9%	30.4%	103.3%
Ethyl Methacrylate	524.2		-37.9%	-108.3%	143.1%
Ethylbenzene	502.2	PID	-35.2%	46.8%	113.8%
Ethylbenzene	524.2		-18.3%	-1245.1%	142.3%
Hardness	130.2		-39.6%	32.7%	92.6%
Hexachlorobutadiene	502.2	ELCD	-123.8%	-69.6%	74.3%
Hexachlorobutadiene	524.2		-63.3%	-528.0%	127.6%
Hexachloroethane	524.2		-12.4%	-38.6%	134.9%
Hexchlobutadiene+naphthalene	502.2	PID	-127.7%	-67.8%	-8.4 %
Iron	1620		99.8%	135.7%	122.1%
Isopropylbenzene	502.2	PID	-30.2%	53.0%	98.7%
Isopropylbenzene	524.2		-8.3%	-3.3%	167.1%
Lead	1620		-32.6%	17.6%	38.1%
Lead	200.8		-165.8%	-138.0%	-105.1%
M+p Xylene	502.2	PID	-154.5%	-109.6%	28.6%
M+p Xylene	524.2		-51.9%	-100.0%	166.8%
Magnesium	1620		-14.9%	51.9%	2.8 %
Manganese	1620		-145.5%	-90.3%	-0.7%
Manganese	200.8		-2.7%	45.4%	112.6%
Mercury	200.8		-22.3%	331.3%	145.0%
Methacrylonitrile	524.2		-43.7%	1.8%	67.4%
Methyl Iodide	524.2		-7.9%	-613.8%	153.7%
Methyl Tert-butyl Ether	524.2		-45.4%	1591.3%	158.7%
Methylacrylate	524.2		-8.6%	46.5%	92.9%
Methylene Chloride	502.2	ELCD	173.9%	189.8%	182.7%
Methylene Chloride	524.2		-13.4%	12.6%	117.2%
Methylmethacrylate	524.2		-90.7%	-63.2%	81.6%
Molybdenum	1620		-35.5%	42.7%	21.1%
Molybdenum	200.8		-25.1%	-195.0%	194.5%
N-butylbenzene	502.2	PID	-26.9%	49.2%	130.0%
N-butylbenzene	524.2		-11.7%	50.0%	162.5%

Table 3. Percent Differences of
Detection Limits to the EPA/ACS DL for the Episode 6000 Dataset

Analyte	Method	Procedure	ISO CRV/ MDL	ISO MDV/ MDL	SL-IDE/ MDL
N-propylbenzene	502.2	PID	-58.0%	20.6%	77.9%
N-propylbenzene	524.2		-38.7%	1215.0%	152.9%
Naphthalene	524.2		-19.7%	-8.6%	117.7%
Nickel	1620		-41.6%	23.9%	23.3%
Nickel	200.8		-86.4%	-30.4%	-55.2%
o-xylene	524.2		-22.0%	735.4%	166.0%
o-xylene+styrene	502.2	PID	-46.2%	32.4%	65.1%
P-isopropitol+1,4-dcb	502.2	PID	-25.1%	51.8%	74.3%
Pentachloroethane	524.2		-186.5%	-288.7%	-30.2%
Sec-butylbenzene	502.2	PID	-52.4%	29.7%	37.9%
Sec-butylbenzene	524.2		-27.1%	2196.0%	163.9%
Selenium	1620		-31.3%	55.0%	79.8%
Selenium	200.8		-20.4%	44.8%	73.8%
Silver	1620		-31.1%	27.9%	74.0%
Silver	200.8		-77.6%	-5.4%	102.6%
Sodium	1620		-33.5%	33.6%	66.5%
Styrene	524.2		-22.6%	-31.1%	163.6%
Tert-butylbenzene	502.2	PID	-36.4%	48.6%	88.6%
Tert-butylbenzene	524.2		-60.7%	5.5%	157.8%
Tetrachloroethene	502.2	ELCD	-26.2%	47.3%	109.0%
Tetrachloroethene	502.2	PID	-42.6%	41.5%	86.4%
Tetrachloroethene	524.2		-0.3%	-57.5%	138.9%
Thallium	1620		24.0%	93.3%	77.0%
Thallium	200.8		-18.1%	44.5%	67.0%
Thorium	200.8		-17.9%	270.2%	50.1%
Tin	1620		-58.1%	-15.5%	6.9%
Titanium	1620		-7.0%	51.0%	11.8%
Toluene	502.2	PID	-85.5%	-11.0%	-8.1%
Toluene	524.2		-112.6%	-290.2%	152.2%
Total Phosphorus	365.2		-25.1%	44.5%	77.5%
Total Suspended Solids	160.2		-21.0%	49.7%	87.9%
Trans-1,2-dichloroethene	502.2	ELCD	1.2%	75.2%	66.8%

Table 3. Percent Differences of
Detection Limits to the EPA/ACS DL for the Episode 6000 Dataset

Analyte	Method	Procedure	ISO CRV/ MDL	ISO MDV/ MDL	SL-IDE/ MDL
Trans-1,2-dichloroethene	524.2		-18.1%	-495.6%	154.9%
Trans-1,3-dichloropropene	502.2	ELCD	-117.4%	-79.8%	157.3%
Trans-1,3-dichloropropene	502.2	PID	-26.6%	47.3%	44.8%
Trans-1,3-dichloropropene	524.2		-69.2%	-260.7%	125.8%
Trans-1,4-dichloro-2-butene	524.2		-38.0%	11.8%	83.8%
Trichloroethene	502.2	ELCD	-156.0%	-127.8%	133.2%
Trichloroethene	502.2	PID	-38.3%	42.8%	112.7%
Trichloroethene	524.2		-4.9%	-9.7%	137.6%
Trichlorofluoromethane	502.2	ELCD	78.9%	140.1%	180.3%
Trichlorofluoromethane	524.2		-15.3%	-78.4%	125.9%
Uranium	200.8		-75.4%	-32.9%	27.6%
Vanadium	1620		-54.3%	12.9%	36.6%
Vanadium	200.8		-8.0%	56.7%	43.6%
Vinyl Chloride	502.2	ELCD	-149.6%	-111.4%	172.7%
Vinyl Chloride	524.2		-32.6%	-274.6%	157.7%
WAD Cyanide	1677		-108.6%	-56.8%	20.2%
Xylene (Total)	524.2		-54.0%	-20.8%	174.0%
Yttrium	1620		-33.6%	26.8%	51.2%
Zinc	1620		-12.1%	34.9%	53.6%
Zinc	200.8		-64.6%	-11.0%	55.8%

Note: ELCD or PID in the Procedure column indicates the photoionization detector (PID) or electrolytic conductivity detector (ELCD) in EPA Method 502.2

Summary Statistics for Table 3

	ISO CRV/ EPA/ACS DL % Difference	ISO MDV/ EPA/ACS DL % Difference	SL-IDE/ EPA/ACS DL % Difference
Minimum	-197.5%	-4142.5%	-161.5%
25th percentile	-60.5%	-76.4%	51.0%
Median	-35.7%	8.8%	108.7%
75th percentile	-19.9%	44.5%	144.1%
Maximum	173.9%	10302.8%	198.4%

	Median % Difference	p-value for % difference=0	
CRV vs. DL	-35.7%	<0.0001	
MDV vs. DL	8.8%	0.164	
SL-IDE vs. DL	1087%	<0.0001	

Table 4. Comparison Quantitation Limits for the Episode 6000 Dataset (µg/L except where footnoted)

Analyte	Method	Procedure	EPA/ ACS QL	ISO/ IUPAC LOQ	ASTM SL-IQE
1,1,1,2-tetrachloroethane	502.2	ELCD	0.2	0.023	0.030
1,1,1,2-tetrachloroethane	524.2		0.2	0.183	0.181
1,1,1-trichloroethane	502.2	ELCD	0.05	0.044	0.830
1,1,1-trichloroethane	524.2		0.2	0.102	0.240
1,1,2,2-tce+1,2,3-tcp	502.2	ELCD	0.2	0.227	5.514
1,1,2,2-tetrachloroethane	524.2		0.5	0.597	0.569
1,1,2-trichloroethane	502.2	ELCD	0.1	0.018	0.060
1,1,2-trichloroethane	524.2		0.2	0.212	0.290
1,1-dichloroethane	502.2	ELCD	0.05	0.037	0.527
1,1-dichloroethane	524.2		0.1	0.099	0.115
1,1-dichloroethene	502.2	ELCD	0.1	0.191	3.796
1,1-dichloroethene	524.2		0.2	0.159	0.129
1,1-dichloropropanone	524.2		20	15.409	12.705
1,1-dichloropropene	524.2		0.2	0.057	0.180
1,2,3-trichlorobenzene	502.2	ELCD	0.2	0.168	0.851
1,2,3-trichlorobenzene	502.2	PID	0.2	0.226	0.248

Table 4. Comparison Quantitation Limits for the Episode 6000 Dataset
(µg/L except where footnoted)

Analyte	Method	Procedure	EPA/ ACS QL	ISO/ IUPAC LOQ	ASTM SL-IQE
1,2,3-trichlorobenzene	524.2		0.2	0.192	0.216
1,2,3-trichloropropane	524.2		20	0.268	11.316
1,2,4-trichlorobenzene	502.2	ELCD	0.1	0.078	0.401
1,2,4-trichlorobenzene	502.2	PID	0.2	0.208	0.439
1,2,4-trichlorobenzene	524.2		0.2	0.231	0.141
1,2,4-trimethylbenzene	502.2	PID	0.5	0.307	0.653
1,2,4-trimethylbenzene	524.2		0.05	0.050	20.896
1,2-dibromo-3-chloropropane	524.2		5	1.842	71.182[6]
1,2-dibromoethane	502.2	ELCD	0.5	0.037	0.592
1,2-dibromoethane	524.2		0.5	0.560	0.417
1,2-dichlorobenzene	502.2	ELCD	0.1	0.158	0.183
1,2-dichlorobenzene	502.2	PID	0.1	0.139	0.346
1,2-dichlorobenzene	524.2		0.1	0.101	0.085
1,2-dichloroethane	502.2	ELCD	0.05	0.015	0.065
1,2-dichloroethane	524.2		0.1	0.122	0.222
1,2-dichloropropane	502.2	ELCD	0.1	0.075	0.102
1,2-dichloropropane	524.2		0.2	0.148	0.196
1,3,5-tmb+4-chlorotoluene	502.2	PID	0.2	0.259	0.189
1,3,5-trimethylbenzene	524.2		0.05	0.044	23.744
1,3-dichlorobenzene	502.2	ELCD	0.1	0.027	0.936
1,3-dichlorobenzene	502.2	PID	0.2	0.438	0.465
1,3-dichlorobenzene	524.2		0.1	0.080	0.076
1,3-dichloropropane	502.2	ELCD	0.05	0.040	0.054
1,3-dichloropropane	524.2		0.1	0.114	0.139
1,4-dichlorobenzene	502.2	ELCD	0.1	0.025	0.101
1,4-dichlorobenzene	524.2		0.1	0.069	0.078
1-chlorobutane	524.2		0.05	0.082	29.943
2,2-dichloropropane	524.2		10	0.572	38.009
2-butanone	524.2		2	1.416	0.893
2-chlorotoluene	502.2	ELCD	0.5	0.145	0.493
2-chlorotoluene	502.2	PID	1	0.781	0.849
2-chlorotoluene	524.2		0.05	0.046	0.053

Table 4. Comparison Quantitation Limits for the Episode 6000 Dataset
(µg/L except where footnoted)

Analyte	Method	Procedure	EPA/ ACS QL	ISO/ IUPAC LOQ	ASTM SL-IQE
2-hexanone	524.2		5	0.669	0.442
2-nitropropane	524.2		2	1.280	0.590
4-chlorotoluene	502.2	ELCD	0.5	0.132	0.142 [1]
4-chlorotoluene	524.2		0.05	0.037	23.810
4-isopropyltoluene	524.2		0.05	0.043	0.016
4-methyl-2-pentanone	524.2		2	2.066	1.785
Acetone	524.2		2	2.114	2.741
Acrylonitrile	524.2		2	1.816	28.056
Allyl Chloride	524.2		0.1	0.129	29.674
Aluminum	1620		100	76.242	464.069
Aluminum	200.8		50	9.418	29.684
Ammonia as Nitrogen [2]	350.3		0.05	0.037	0.035
Antimony	1620		5	4.784	9.551
Antimony	200.8		0.5	0.017	0.034
Arsenic	1620		5	3.684	3.097
Arsenic	200.8		1	0.720	0.798
Barium	1620		5	4.722	4.118
Barium	200.8		0.1	0.161	0.211
Benzene	502.2	PID	0.1	0.173	0.182
Benzene	524.2		0.05	0.075	0.044
Beryllium	1620		2	1.055	0.980
Beryllium	200.8		0.02	0.018	0.044
Boron	1620		50	46.040	51.134
Bromobenzene	502.2	ELCD	0.5	0.599	3.529
Bromobenzene	502.2	PID	0.05	0.050	0.100
Bromobenzene	524.2		0.2	0.167	0.140
Bromochloromethane	502.2	ELCD	0.05	0.065	1.598
Bromochloromethane	524.2		0.5	0.549	0.368
Bromodichloromethane	502.2	ELCD	0.02	0.015	0.424
Bromodichloromethane	524.2		0.2	0.135	0.128
Bromoform	502.2	ELCD	0.02	0.018	3.393
Bromoform	524.2		0.5	0.287	0.482

Table 4. Comparison Quantitation Limits for the Episode 6000 Dataset
(µg/L except where footnoted)

Analyte	Method	Procedure	EPA/ ACS QL	ISO/ IUPAC LOQ	ASTM SL-IQE
Bromomethane	502.2	ELCD	1	undefined [3]	16.351
Bromomethane	524.2		0.2	0.252	0.226
Cadmium	1620		0.5	0.346	0.410
Cadmium	200.8		0.02	0.046	0.063
Calcium	1620		100	186.530	99.975
Carbon Disulfide	524.2		0.1	0.077	0.101
Carbon Tetrachloride	524.2		0.1	0.127	0.140
Carbontet+1,1-dcp	502.2	ELCD	0.1	0.046	0.069
Chloroacetonitrile	524.2		2	4.170	3.310
Chlorobenzene	502.2	ELCD	0.05	0.058	1.766
Chlorobenzene	502.2	PID	0.1	0.143	0.119
Chlorobenzene	524.2		0.1	0.108	0.059
Chloroethane	502.2	ELCD	0.5	0.053	5.826
Chloroethane	524.2		0.2	0.185	0.255
Chloroform	502.2	ELCD	0.2	0.029	0.025
Chloroform	524.2		0.1	0.138	0.121
Chloromethane	502.2	ELCD	0.2	0.342	1.734
Chloromethane	524.2		0.2	0.181	0.141
Chromium	1620		1	0.993	1.259
Chromium	200.8		0.2	0.331	1.028
Cis-1,2-dce+2,2-dcp	502.2	ELCD	0.05	0.045	0.039
Cis-1,2-dichloroethene	524.2		0.1	0.154	0.144
Cis-1,3-dichloropropene	502.2	ELCD	0.02	0.013	0.415
Cis-1,3-dichloropropene	502.2	PID	0.2	0.254	0.017 [1]
Cis-1,3-dichloropropene	524.2		0.1	0.117	0.141
Cobalt	1620		50	20.916	40.837
Cobalt	200.8		0.005	undefined [3]	undefined [4]
Copper	1620		20	27.513	47.509
Copper	200.8		0.1	0.142	1.825
Dibromochloromethane	502.2	ELCD	0.02	0.030	1.252
Dibromochloromethane	524.2		0.2	0.149	0.288
Dibromomethane	502.2	ELCD	0.02	0.028	1.395

Table 4. Comparison Quantitation Limits for the Episode 6000 Dataset
(µg/L except where footnoted)

Analyte	Method	Procedure	EPA/ ACS QL	ISO/ IUPAC LOQ	ASTM SL-IQE
Dibromomethane	524.2		0.5	0.400	0.460
Dichlorodifluoromethane	502.2	ELCD	0.02	0.012	1.091 [5]
Dichlorodifluoromethane	524.2		0.2	0.290	0.480
Diethyl Ether	524.2		0.5	0.563	0.404
Ethyl Methacrylate	524.2		0.2	0.139	0.183
Ethylbenzene	502.2	PID	0.1	0.089	0.157
Ethylbenzene	524.2		0.1	0.123	0.077
Hardness [2]	130.2		2	2.973	5.465
Hexachlorobutadiene	502.2	ELCD	0.2	0.054	0.243
Hexachlorobutadiene	524.2		0.2	0.160	0.228
Hexachloroethane	524.2		0.2	0.232	0.167
Hexchlobutadiene+naphthalene	502.2	PID	2	0.834	1.542
Iron	1620		200	1490.589	996.565 [5]
Isopropylbenzene	502.2	PID	0.1	0.090	0.129
Isopropylbenzene	524.2		0.05	0.056	25.592
Lead	1620		5	5.062	5.698
Lead	200.8		2	0.318	0.685
M+p Xylene	502.2	PID	0.2	0.068	0.222
M+p Xylene	524.2		0.05	0.042	24.651
Magnesium	1620		500	454.043	267.199
Manganese	1620		20	7.948	15.264
Manganese	200.8		0.1	0.133	0.245
Mercury	200.8		0.02	0.056	0.039
Methacrylonitrile	524.2		1	1.066	19.062
Methyl Iodide	524.2		0.1	0.108	0.083
Methyl Tert-butyl Ether	524.2		0.1	0.073	0.122
Methylacrylate	524.2		1	0.966	0.727
Methylene Chloride	502.2	ELCD	0.5	undefined [3]	6.033
Methylene Chloride	524.2		0.2	0.354	0.433
Methylmethacrylate	524.2		1	0.381	20.773
Molybdenum	1620		10	9.752	7.597
Molybdenum	200.8		0.01	0.052	0.608

Table 4. Comparison Quantitation Limits for the Episode 6000 Dataset
(µg/L except where footnoted)

Analyte	Method	Procedure	EPA/ ACS QL	ISO/ IUPAC LOQ	ASTM SL-IQE
N-butylbenzene	502.2	PID	0.1	0.128	0.745
N-butylbenzene	524.2		0.05	0.077	0.067
N-propylbenzene	502.2	PID	0.2	0.128	0.186
N-propylbenzene	524.2		0.1	0.110	29.878
Naphthalene	524.2		0.2	0.184	0.108
Nickel	1620		100	66.486	67.206
Nickel	200.8		0.5	0.287	0.183
o-xylene	524.2		0.05	0.062	0.040
o-xylene+styrene	502.2	PID	0.2	0.210	0.181
P-isopropitol+1,4-dcb	502.2	PID	0.2	0.318	0.456
Pentachloroethane	524.2		2	0.086	0.551
Sec-butylbenzene	502.2	PID	0.2	0.193	0.157
Sec-butylbenzene	524.2		0.05	0.063	0.047
Selenium	1620		2	3.859	5.235
Selenium	200.8		0.5	0.805	1.045
Silver	1620		20	16.734	25.842
Silver	200.8		0.02	0.011	0.056
Sodium	1620		200	251.546	337.755
Styrene	524.2		0.05	0.054	0.041
Tert-butylbenzene	502.2	PID	0.1	0.121	0.203
Tert-butylbenzene	524.2		0.1	0.063	0.073
Tetrachloroethene	502.2	ELCD	0.05	0.076	0.122
Tetrachloroethene	502.2	PID	0.2	0.244	0.750
Tetrachloroethene	524.2		0.2	0.378	30.554[6]
Thallium	1620		2	3.748	2.799
Thallium	200.8		0.002	0.002	0.002
Thorium	200.8		0.002	0.005	0.004
Tin	1620		10	9.237	9.406
Titanium	1620		20	20.807	14.236
Toluene	502.2	PID	0.2	0.162	0.194
Toluene	524.2		0.05	0.028	0.046
Total Phosphorus[2]	365.2		0.02	0.024	0.030

Table 4. Comparison Quantitation Limits for the Episode 6000 Dataset
(µg/L except where footnoted)

Analyte	Method	Procedure	EPA/ ACS QL	ISO/ IUPAC LOQ	ASTM SL-IQE
Total Suspended Solids [2]	160.2		5	5.011	6.729
Trans-1,2-dichloroethene	502.2	ELCD	0.2	0.234	0.191
Trans-1,2-dichloroethene	524.2		0.1	0.141	0.153
Trans-1,3-dichloroPropene	502.2	ELCD	0.05	0.016	0.729
Trans-1,3-dichloropropene	502.2	PID	0.2	0.244	0.175
Trans-1,3-dichloropropene	524.2		0.2	0.121	0.218
Trans-1,4-dichloro-2-butene	524.2		2	1.803	30.108
Trichloroethene	502.2	ELCD	0.05	0.008	3.169
Trichloroethene	502.2	PID	0.1	0.108	0.401
Trichloroethene	524.2		0.2	0.284	0.167
Trichlorofluoromethane	502.2	ELCD	0.5	1.612	4.662
Trichlorofluoromethane	524.2		0.2	0.279	42.490 [6]
Uranium	200.8		0.001	0.001	0.001
Vanadium	1620		20	21.586	24.338
Vanadium	200.8		2	2.627	1.933
Vinyl Chloride	502.2	ELCD	1	0.264	8.234
Vinyl Chloride	524.2		0.2	0.139	0.219
WAD Cyanide	1677		2	0.852	1.624
Xylene (Total)	524.2		0.02	0.027	23.520
Yttrium	1620		5	6.571	8.962
Zinc	1620		10	9.575	10.452
Zinc	200.8		2	2.147	7.024

[1] IQE 10% undefined, IQE 20% reported

[2] Results reported as mg/L

[3] No LOQ could be calculated due to a square root of a negative number in the formula

[4] IQE 10%, IQE 20% and IQE 30% all negative based on chosen model (linear)

[5] IQE 10% and IQE 20% both negative, IQE 30% reported

[6] Hybrid model selected but did not converge; IQE 10% based on constant model instead

Note: ELCD or PID in the Procedure column indicates the photo-ionization detector (PID) or electrolytic conductivity detector (ELCD) in EPA Method 502.2

Table 5. Percent Differences of Quantitation Limits to the EPA/ACS QL for the Episode 6000 Dataset

Analyte	Method	Procedure	ISO LOQ/ML	SL-IQE/ML
1,1,1,2-tetrachloroethane	502.2	ELCD	-158.7%	-147.3%
1,1,1,2-tetrachloroethane	524.2		-8.7%	-9.8%
1,1,1-trichloroethane	502.2	ELCD	-11.8%	177.3%
1,1,1-trichloroethane	524.2		-65.1%	18.0%
1,1,2,2-tce+1,2,3-tcp	502.2	ELCD	12.8%	186.0%
1,1,2,2-tetrachloroethane	524.2		17.6%	12.9%
1,1,2-trichloroethane	502.2	ELCD	-138.2%	-49.6%
1,1,2-trichloroethane	524.2		5.9%	36.6%
1,1-dichloroethane	502.2	ELCD	-28.6%	165.3%
1,1-dichloroethane	524.2		-0.7%	13.7%
1,1-dichloroethene	502.2	ELCD	62.5%	189.7%
1,1-dichloroethene	524.2		-22.8%	-43.3%
1,1-dichloropropanone	524.2		-25.9%	-44.6%
1,1-dichloropropene	524.2		-111.1%	-10.5%
1,2,3-trichlorobenzene	502.2	ELCD	-17.6%	123.9%
1,2,3-trichlorobenzene	502.2	PID	12.2%	21.3%
1,2,3-trichlorobenzene	524.2		-4.2%	7.7%
1,2,3-trichloropropane	524.2		-194.7%	-55.5%
1,2,4-trichlorobenzene	502.2	ELCD	-25.2%	120.2%
1,2,4-trichlorobenzene	502.2	PID	3.8%	74.9%
1,2,4-trichlorobenzene	524.2		14.5%	-34.9%
1,2,4-trimethylbenzene	502.2	PID	-47.8%	26.5%
1,2,4-trimethylbenzene	524.2		0.5%	199.0%
1,2-dibromo-3-chloropropane	524.2		-92.3%	173.7%
1,2-dibromoethane	502.2	ELCD	-172.7%	16.9%
1,2-dibromoethane	524.2		11.3%	-18.1%
1,2-dichlorobenzene	502.2	ELCD	45.1%	58.8%
1,2-dichlorobenzene	502.2	PID	32.9%	110.2%
1,2-dichlorobenzene	524.2		0.6%	-16.5%
1,2-dichloroethane	502.2	ELCD	-108.1%	26.0%
1,2-dichloroethane	524.2		19.7%	75.6%
1,2-dichloropropane	502.2	ELCD	-28.5%	2.3%