Department of Defense
Environmental Data Quality Workgroup

DEVELOPMENT OF DEPARTMENT OF DEFENSE LABORATORY CONTROL SAMPLE CONTROL LIMITS

Final
May 2004

Acknowledgments

The study was overseen by the Quality Assurance Authors/Task Action Team (QAA/TAT) of the Department of Defense (DoD) Environmental Data Quality Workgroup (EDQW), which was composed of environmental chemists from each of the three DoD components (represented by Naval Sea Systems Command [NAVSEA], Air Force Center for Environmental Excellence [AFCEE], and the Army Corps of Engineers [USACE]). The study was performed with the cooperation of the American Council of Independent Laboratories (ACIL). The study team was composed of the QAA/TAT, ACIL members from the environmental testing laboratory community, and a contractor support team from Versar, Inc.

ACIL solicited and collected data from over 20 laboratories, then provided the data volunteered by the laboratories to DoD. DoD made all final decisions; however, ACIL provided input on the methodology and policy for setting and applying limits. This study would not have been possible without their support. In particular, the QAA/TAT would like to thank Richard Burrows, Charles Carter, Jack Farrell, Debra Henderer, Deb Loring, Tony Pagliaro, Jerry Parr, Michael Shepherd, John Webb, and Chuck Wibby

Representatives from EPA, DoD, and ACIL membership conducted an informal peer review on the LCS Study.

DoD LCS Study Team

Quality Assurance Authors/Task Action Team (current members):

William H. Batschelet, Ph.D., AFCEE
Chung-Rei Mao, Ph.D., U.S. Army Corps of Engineers
Fred McLean, NAVSEA
Pati Moreno, Naval Facilities Engineering Service Center
Jackie Sample, NAVSEA

Quality Assurance Authors/Task Action Team (former members):

Cheryl Groenjes, U.S. Army Corps of Engineers
Maj. W. Kevin Kuhn, AFCEE

Versar, Inc. (support team):

Sharon McCarthy, Ph.D.
Clem Rastatter
Nicole Weymouth

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF FIGURES

EXECUTIVE SUMMARY

In 1999 the Department of Defense (DoD) Environmental Data Quality Workgroup (EDQW) initiated a study of laboratory control samples (LCSs) from commercial environmental laboratories that have shown good performance on work done for DoD.

The objectives of the study were twofold:

- To develop and publish LCS control limits (LCS-CLs) based on empirical data, which must be used by laboratories doing work for DoD.
- To establish objective benchmarks for analytical method performance to assist in evaluating the suitability of alternative methods.

The DoD LCS study focused on nine different analytical methods published in *Test Methods for Evaluating Solid Waste* (SW-846): semivolatiles 8270C, volatiles 8260B, herbicides 8151A, polynuclear aromatic hydrocarbons (PAHs) 8310, explosives 8330, pesticides 8081A, polychlorinated biphenyls (PCBs) 8082, metals 6010B, and mercury 7470A/7471A.

This report presents the outcome of the study and is organized into four major sections:

1. Purpose (Section 1.0): Briefly identifies the reasons DoD initiated the LCS study.
2. Background (Section 2.0): Describes the current use of LCSs in laboratories and DoD's goals and requirements for the study.
3. DoD LCS-CLs Development (Section 3.0): Presents the process DoD went through in developing the LCS-CLs, including a detailed description of the methodology, study findings, and analysis of policy issues.
4. DoD LCS-CLs Implementation (Section 4.0): Describes the final LCS-CL policy developed by DoD and presents the data tables. These tables are also published as quality requirements in the *Quality Systems Manual for Environmental Laboratories* (QSM) Version 2 (June 2002).

This page intentionally left blank.

DEVELOPMENT OF DEPARTMENT OF DEFENSE LABORATORY CONTROL SAMPLE CONTROL LIMITS

1.0 PURPOSE

As part of its charter to develop and coordinate environmental sampling and testing policy for the Department of Defense (DoD), the DoD Environmental Data Quality Workgroup (EDQW) developed the *DoD Quality Systems Manual for Environmental Laboratories* (QSM), of which Version 1 (October 2000) and Version 2 (June 2002) have now been published. As part of that work, the EDQW recognized the need for minimum objective standards against which laboratory and analytical method performance can be judged. They focused on the use of a particular quality control sample, the laboratory control sample (LCS), to provide a measure of analytical performance. DoD wished to set realistic and scientifically defensible targets for LCS recoveries based on the routine performance of commonly used methods. Their objectives were twofold:

- To develop and publish LCS control limits (LCS-CLs) based on empirical data, which must be used by laboratories doing work for DoD.
- To establish objective benchmarks for analytical method performance to assist in evaluating the suitability of alternative methods.

2.0 BACKGROUND

LCSs are used as quality control (QC) measures to establish and track intra-laboratory performance of the analytical system. The percent recovery of each spiked compound is compared with a range of acceptable recoveries (control limits) that are typically statistically calculated. Laboratories should establish in-house LCS-CLs annually. The control limits capture both systematic and random errors and serve as benchmarks against which analyst and instrument performance are measured. If the LCS recovery for any analyte in a particular batch of samples is outside the established limits for that analyte and method, then the batch results may be considered unacceptable, triggering corrective action as appropriate (e.g., reanalysis may be required). Unacceptable LCS recovery (i.e., LCS failure) is of great concern to both laboratories and DoD because of the cost and time associated with reanalysis. As currently implemented, the failure of a single compound in an LCS can constitute failure of the entire analytical batch.

Laboratory Control Samples are clean matrices (e.g., reagent water or a clean solid such as sand, glass beads, or sodium sulfate) that have been spiked with a known quantity of a compound or group of compounds and are processed with every analytical batch of environmental samples. The percentage of the compound that is recovered in the analysis provides a measure of method accuracy. When analysis of the LCS is repeated, the standard deviation provides a measure of analytical precision.

2.1 Calculation of LCS Control Limits

According to the widely used *Test Methods for Evaluating Solid Waste* (SW-846 methods, Chapter 1, Section 4.4.2), analyte-specific control limits are calculated as 3 standard deviations around the mean.

$$CL = \bar{\chi} \pm 3SD$$

where:

CL = control limit

$\bar{\chi}$ = mean recovery of data set

SD = standard deviation of data set

Method 8000B of SW-846, *Determinative Chromatographic Separations*, suggests that the control limits should be generated from an LCS data set consisting of at least 15 to 20 data points for each analyte.

Prior to the DoD LCS study, laboratories generally either set their own control limits or met limits published in the *AFCEE Quality Assurance Project Plan* or in the method. Since most of the AFCEE QAPP limits and the method limits are based on a limited amount of data from a single laboratory, some laboratories voiced concerns that the limits do not reflect the true capabilities of the methods to recover analytes. Failure to meet these limits was costly to the laboratories (due to reanalysis) and to DoD (due to increased costs from laboratories for reanalysis and time delays).

2.2 DoD Goals and Data Requirements for the Study

DoD's goal when initiating the LCS study was to establish a consistent set of default LCS control limits to be used DoD-wide, in the absence of project-specific requirements. Key criteria for *developing* the LCS-CLs were that the limits be:

- Scientifically valid and statistically defensible.
- Based on actual laboratory data from laboratories that performed satisfactory work for DoD.
- Able to accommodate the variability that exists in the ways laboratories execute the methods.
- Based on SW-846 methods, since those methods are commonly used by DoD for the two largest programs that require the collection of analytical data — the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), and the Resource Conservation and Recovery Act (RCRA).

Key requirements for *implementing* the LCS-CLs included the following:

- The default LCS-CLs would not take the place of project-specific limits that were based on site-specific information.
- The use of SW-846 methods as the basis for this study would not limit the use of alternative analytical methods, as appropriate. Instead, the LCS-CLs would provide objective benchmarks against which the adequacy of an alternative method could, in part, be evaluated.
- The complexity of implementation by the laboratories would be taken into account (e.g., no requirement for the bench chemist to manage multiple sets of limits that vary by analyte).

The DoD LCS study purposely included data from multiple laboratories. This approach was considered necessary in order to calculate control limits that encompassed the method-allowed variations in procedures routinely used by different laboratories. The goal was to establish LCS-

CLs that reflected routine performance by laboratories that performed well according to method specifications. Environmental laboratories that had passed an audit by one or more of the DoD components within the past 18 to 24 months were deemed to be “good performing,” and data submitted by those laboratories were considered to reflect routine method performance for good laboratories.

3.0 DoD LCS-CLs DEVELOPMENT

Development of DoD LCS-CLs involved establishing the statistical methodology, analyzing the results, and evaluating the policy implications.

3.1 Methodology

The study was conducted in two phases. During the pilot study, or Phase I, two different statistical methodologies for generating control limits were tested using multi-laboratory data for a single analytical method (SW-846 Method 8270C for semivolatile organic compounds). During Phase II, the selected statistical methodology was applied to multi-laboratory data for eight other SW-846 analytical methods, including volatile organic compounds 8260B, chlorinated herbicides 8151A, polynuclear aromatic hydrocarbons (PAHs) 8310, explosives 8330, organochlorine pesticides 8081A, polychlorinated biphenyls (PCBs) 8082, metals 6010B, and mercury 7470A/7471A.

Laboratories voluntarily provided data for the study according to data submittal instructions placed on the DoD DENIX website and distributed by ACIL (see Attachments 1 and 2 of Appendix A). Data were requested for target analytes routinely reported for DoD compliance and restoration programs (target analyte lists are found in the DoD QSM Version 2, Appendix DoD-C). Information submitted by each laboratory included the LCS sample ID number, analyte name, matrix type (solid or water), preparation/extraction methods, spike concentrations, and percent recovery.

For Phase I, 17 laboratories submitted data for 77 semivolatile target analytes. Data sets ranged from 74 to 435 data points per analyte. For Phase II, 16 laboratories submitted data for at least one of the eight methods. Data sets for the 162 total analytes ranged from 91 to 396 data points per analyte.

During Phase I of the study, the team divided the data into two groups, by laboratory: a test group, which went through every step of the proposed methodologies, and a control group. After the statistical methodology was selected, the data sets from both the test group and the control group were then compared with the control limits generated from the test group data. The comparisons demonstrated that overall failure rates were similar, without significant differences between the control group data and test group data. Therefore, the EDQW decided to use consolidated data sets and generate a single set of LCS-CLs for each analyte using the selected statistical methodology for both the 8270C method and the Phase II analytical methods.

The final statistical methodology used by the study team included analysis of variances (ANOVA) between different method-specific parameters, identification of outliers, calculation of mean and standard deviation, and calculation of control limits. Figure 1 presents a flow chart of the general methodology. The methodology is described in detail in Appendix A to this report; Attachment 3 to Appendix A presents the original methodology strategy for the study.

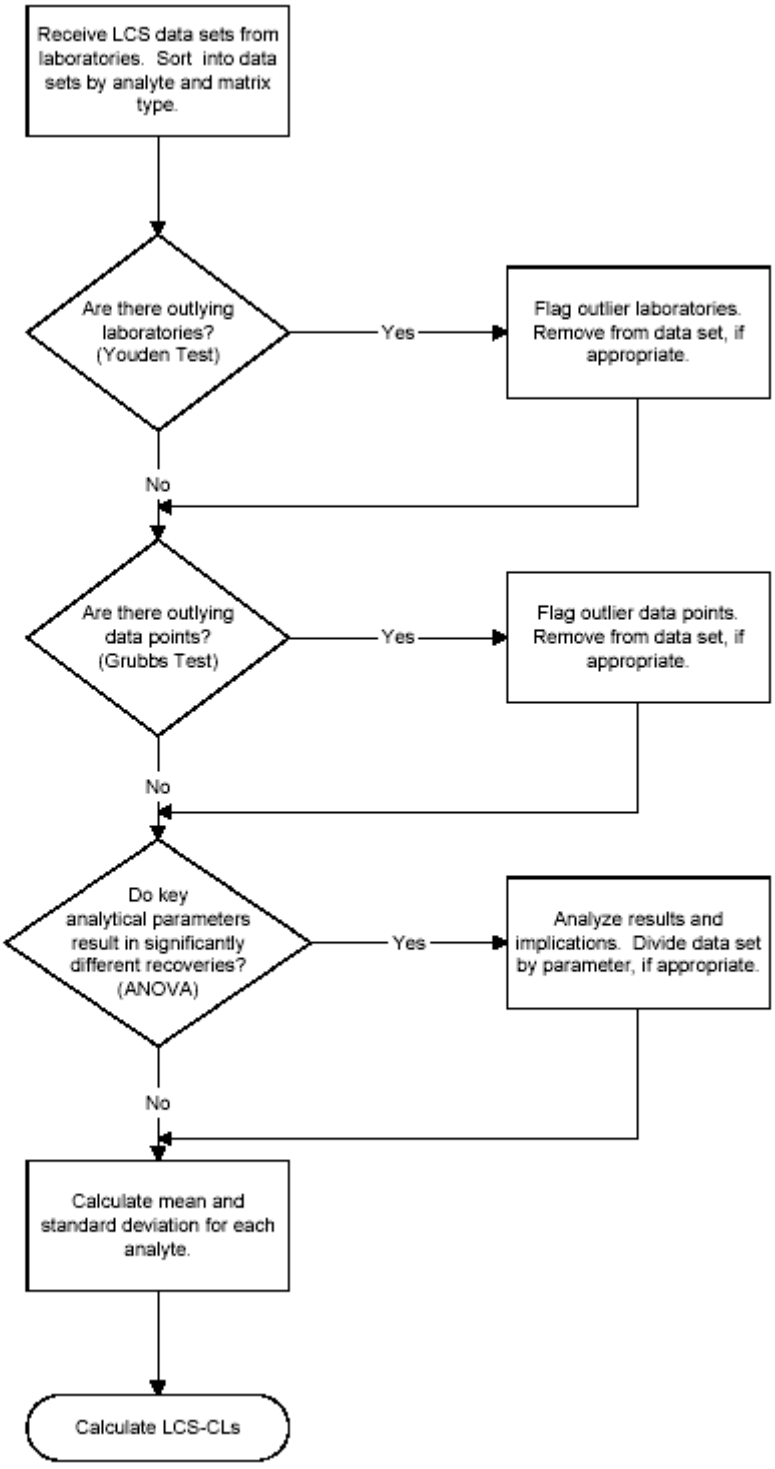

Figure 1. Statistical Methodology Flow Chart

3.2 Findings

This section presents the results of the primary study analyses and an evaluation of the effects of applying the calculated control limits to the data.

3.2.1 Summary of Findings

The study involved nine analytical methods, for both solid and water matrices, which resulted in more than 450 different analyte data sets. (Note: Data for solid and water matrices for the same analyte are counted as two different data sets.) The following is a summary of the limits generated using the selected methodology and an analysis of quantitative results:

- In general, mean recoveries were high, greater than 70% recovery for the majority (93%) of 454 total analytes.
- For organics, LCS recoveries were more variable, yielding higher standard deviations and, therefore, a high level of uncertainty.
- Not surprisingly, inorganics produced much better results. Means were near 100% with low standard deviations.

Figures 2 and 3 present the range of means, across analytes, for each of the nine analytical methods (solid and water matrix data, respectively). Mean recoveries are typically between 70 and 100%. The relative standard deviation (RSD) charts (Figures 4 and 5) demonstrate the varied precisions of the different methods across analytes. (Note: For the purposes of this report, high precision is defined by low RSD.) The figures use bar graphs that represent all of the data for a given method. The graphs are color-coded to show the percentage of compounds within that method that have low, medium, or high precision. Metals (methods 6010B and 7470A/7471A) have low RSD and high levels of precision, and herbicides (method 8151A) have medium to high RSD, therefore less precision. The mean, standard deviation, and lower and upper control limits for each analyte can be found in the tables at the end of this report.

General findings from further analysis of the data and methodology include the following:

- The outlier methodology (Youden/Grubbs), in almost all cases, lowered the standard deviation. In addition, outliers were typically biased high. Therefore, removing outliers lowered the resulting upper control limit by lowering both the mean and the standard deviation.
- Occasionally, significant differences were identified by the ANOVA test; however, the differences did not have a material effect on the calculation of the LCS-CLs with the exception of explosives method 8330 in water. In some cases, not enough data were available to conduct ANOVA, since many laboratories use the same parameter (e.g., extraction methods).
- The analysis of certain analytes by a specific analytical method resulted in such inconsistent performance that high standard deviations established lower control limits at or below 10%. These compounds were defined as poor performing analytes by DoD.
- The LCS-CLs were evaluated by comparing them with existing acceptance limits from alternative sources (benchmarks; see Section 3.2.5). This comparison demonstrated that the limits calculated in the study were comparable to or more stringent than most existing limits.

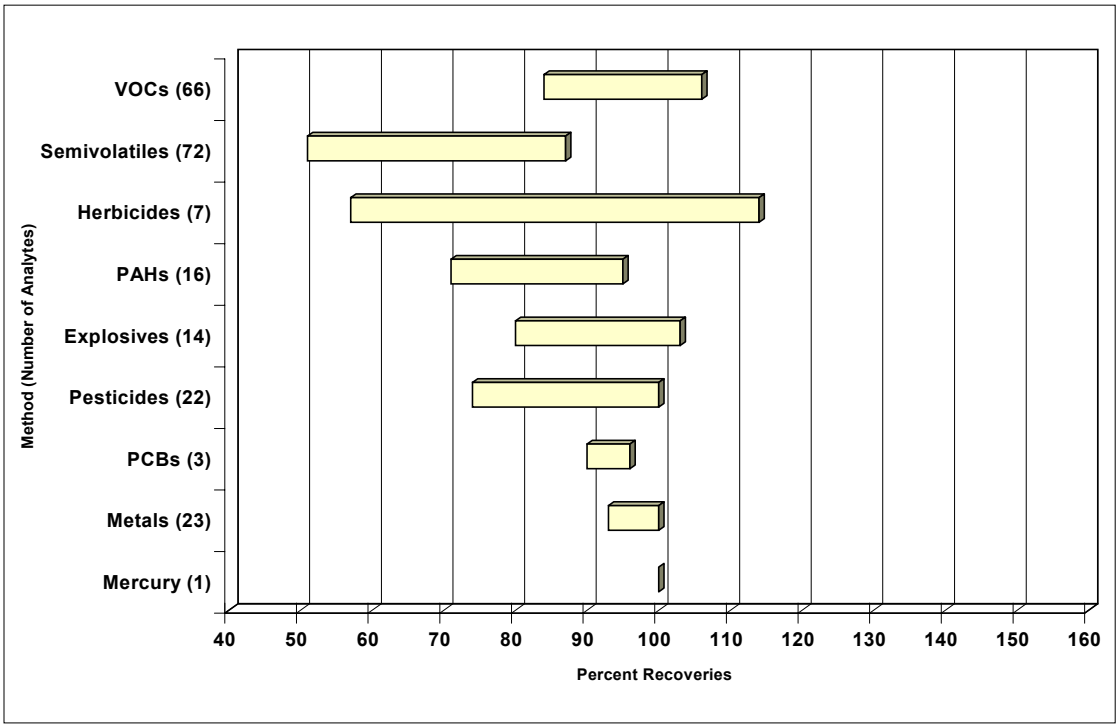

Figure 2. Range of Mean Recoveries in Solid*

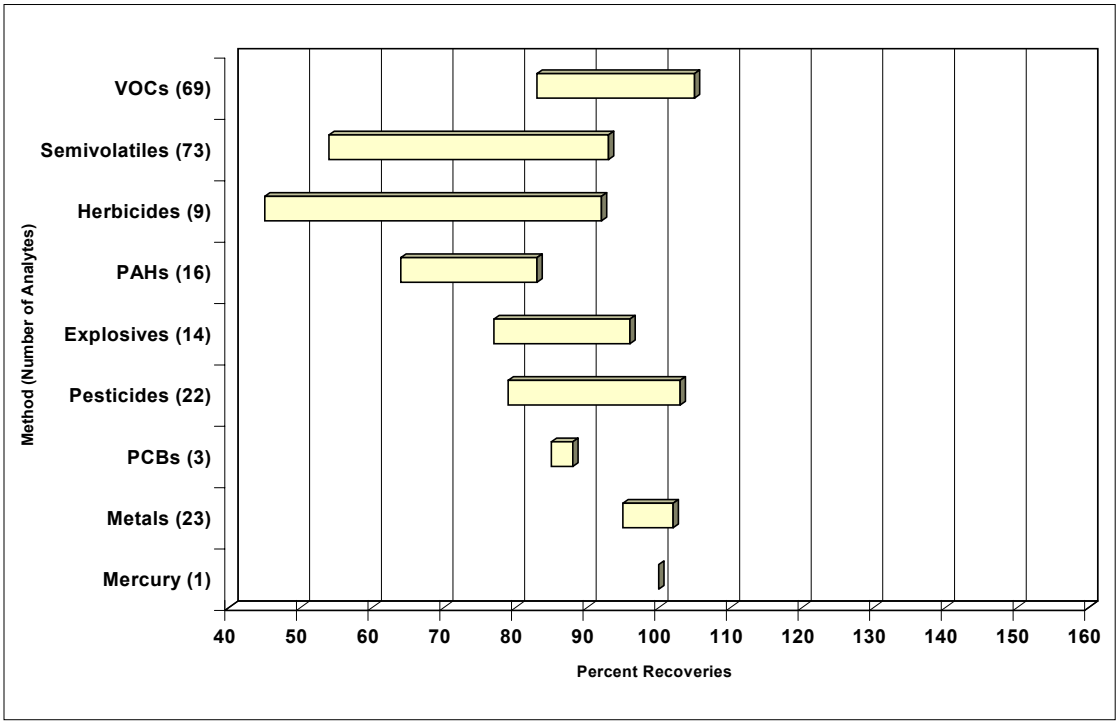

Figure 3. Range of Mean Recoveries in Water*

* The number of analytes varies between the solid and water matrices because of differences in the amount of data received from laboratories.

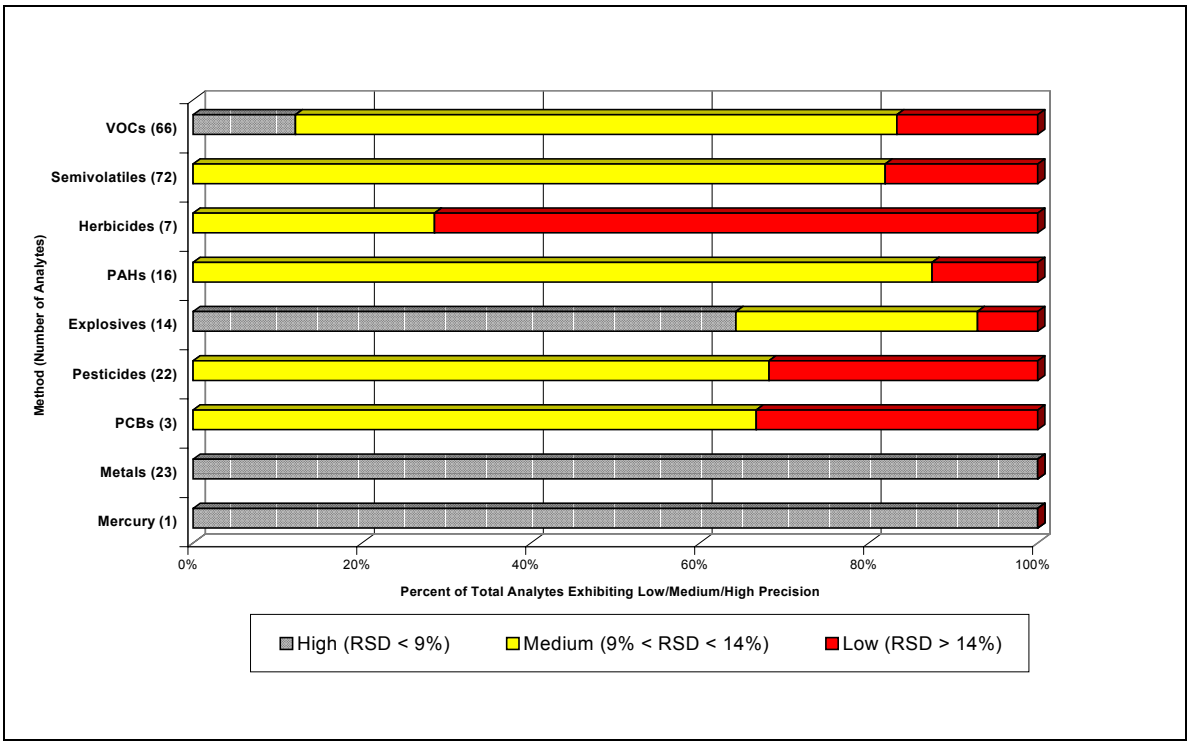

Figure 4. Precision of Methods in Solid*

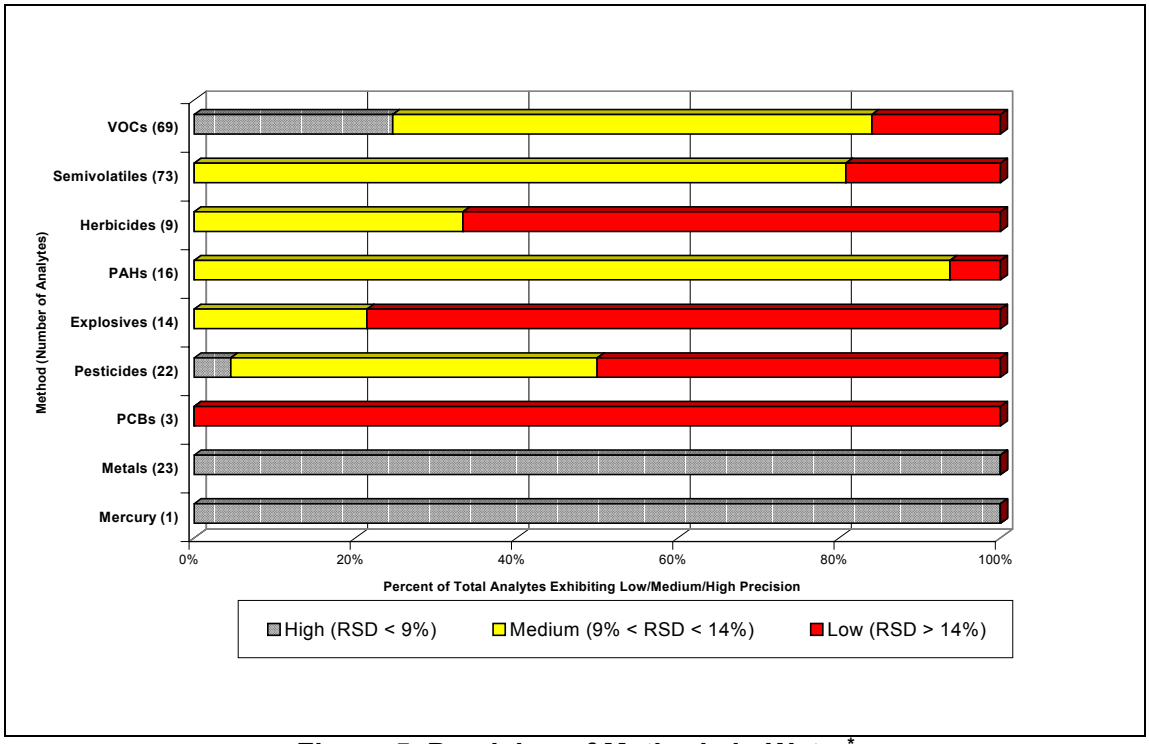

Figure 5. Precision of Methods in Water*

* The number of analytes varies between the solid and water matrices because of differences in the amount of data received from laboratories.

- Calculation of estimated failure rates where one or more of the analytes were outside the LCS-CLs demonstrated that failure was more likely at the upper limit.
- Estimated failure rates showed that LCS failure is statistically more likely with longer lists of analytes.

3.2.2 Effects of Outlier Removal

The selected methodology called for identification of outliers using the Youden test and the Grubbs test. Most outliers were identified using the Youden test. Because the Youden test excluded data for a particular analyte from an entire laboratory, often the study identified a large number of data points as outliers. A laboratory's data set was identified as a Youden outlier for various reasons. In some cases the outlier laboratories had consistently higher or lower recoveries than the other laboratories. In other cases the outlier laboratories' recoveries were more tightly clustered than the other laboratories'.

Analysis of the effect of outlier removal on the LCS-CLs led to the conclusion that lower standard deviations were usually achieved when outliers were removed. In two-thirds of the cases, a lower mean also resulted (often only one or two points). However, the effect on control limits of a change in standard deviation was 6 times as great as a change in mean (i.e., the standard deviation is multiplied by 3 on both the upper and lower ends). Therefore, the slightly lower means were considered acceptable by DoD, since the overall effect of outlier removal was tighter control limits.

3.2.3 ANOVA Results

Analysis of the variance in method parameters could result in several outcomes:

- Multiple sets of LCS-CLs based on a particular parameter (e.g., spiking level, preparation/extraction method).
- LCS-CLs based only on the parameter that produced a "better" result (e.g., higher and tighter recoveries).
- LCS-CLs based on all data if no significant difference in recoveries was identified.

As the ANOVA results were being reviewed, it became apparent that compelling evidence was needed to justify the creation of multiple control limits for the same analyte. First of all, LCS recoveries may not be indicative of the performance of the parameter in environmental samples. Second, multiple sets of control limits for a single analytical method would be too confusing for laboratories to manage at the bench. Third, the methods allow laboratories to make choices in implementation. These choices may have cost and time implications or may be appropriate for achieving the level of data quality necessary for decision-making (e.g., selection of a particular preparation method). Finally, the identification of significant differences in ANOVA results may not always lead to significant differences in the generated LCS-CLs. DoD did not want to limit its or the laboratory's choices unless there was a significant benefit; therefore, although the use of different parameters resulted in some findings of statistically different recoveries, the LCS-CLs were calculated using entire data sets. The only exception was for explosives method 8330 in water.

For explosives method 8330, water matrix only, the ANOVA results demonstrated that there was a significant difference in recovery depending on the extraction method used. Solid phase extraction (SPE) using acetonitrile elution produced higher mean recoveries and considerably

lower standard deviations than those of the alternative salting out extraction method. In addition, SPE is less expensive, cumbersome, and time and labor intensive than the alternative. As a result, the EDQW chose to set LCS-CLs for method 8330 (water matrix) using SPE data only. Because of the small number of laboratories in that data set (approximately 4, depending on the analyte), no outliers were removed prior to calculating the limits. This approach ensured that a reasonably sized, representative data set was used to generate the control limits. (Note: Laboratories may use any extraction method they feel is appropriate; however, the LCS recoveries must fall within the LCS-CLs generated with the SPE data.)

3.2.4 Poor Performing Analytes

After running all the data through the statistical methodology, the study team identified analytes that did not perform well with specific methods. DoD felt those analytes needed to be addressed because of the high level of uncertainty in their results. DoD defined those poor performing analytes as analytes with lower LCS-CLs of 10% or less. They typically have low mean recoveries and high standard deviations, resulting in wide LCS-CLs. (Note: Although the term "poor performing analytes" is used, DoD is aware that this is a reflection of the analytical system as routinely implemented and not an indictment of the laboratories' performance.)

The EDQW discussed extensively the options for defining poor performing analytes (e.g., lower limit less than 10 or 20%, mean less than 70%). They looked at scatter plots and found that the poor performing analytes had high variability both within a given laboratory as well as across laboratories. As described in Section 3.2.6, estimation of failure rates after various adjustments to the limits demonstrated that raising lower limits above 10% increased failure rates (sometimes significantly). Raising the cutoff to 20% or higher would significantly increase the number of poor performing analytes, thereby eliminating from regular evaluation compounds frequently found at DoD sites. Eventually, a compromise was reached, and poor performing analytes were defined as those analytes with a statistically generated lower control limit of 10% or less.

The decision to use 10% was a means of letting the data speak for themselves and not accepting extremely low recoveries. The purpose of the LCS study was to evaluate routinely achievable performance, not optimize performance for a particular problematic analyte or group of analytes. DoD did not want to penalize the laboratories or itself for the poor performance of the methods. In many cases the lower limit published in the SW-846 methods for the poor performing analytes was lower than 10% (sometimes nondetect or zero). However, DoD did not feel that extremely low recoveries should be considered acceptable and felt the issue should be addressed in some way.

Table 1 presents the poor performing analytes, as identified by a lower control limit of 10% or less. See Section 4.3 for an explanation of DoD's policy on addressing poor performing analytes.

Table 1. Poor Performing Analytes

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
8270C Water:				
4-Nitrophenol	54.3	23.0	0	123
Benzoic acid	54.9	24.0	0	127
Phenol	55.9	19.9	0	116
Phenol-d5/d6 (surrogate)	62.6	18.0	9	117
8270C Solid:				
3,3'-Dichlorobenzidine	68.9	19.6	10	128
4-Chloroaniline	51.0	14.2	8	94
Benzoic acid	55.7	18.7	0	112
8151A Solid:				
Dinoseb	57.3	50.9	0	210
8330 Solid:				
Methyl-2,4,6-trinitrophenylnitramine (Tetryl)	80.2	23.3	10	150

3.2.5 Comparison with Benchmarks

One step in analyzing the effects of the methodology on the calculated control limits was to compare the LCS-CLs that were statistically generated in this study with a variety of benchmarks, including the following:

- The laboratory's in-house limits (as provided by the laboratories that submitted data for the study)
- The method limits (when available)
- AFCEE published limits
- Proficiency testing (PT) acceptance limits for water (calculated using regression constants from EPA's National Standards for Water Proficiency Testing Studies)
- Limits from the USACE Quality Assurance Laboratory in Omaha, Nebraska

The findings of this comparison varied by method, but in the majority of cases the upper control limits generated in this study were more stringent (i.e., lower) than the benchmark upper limits. The comparison of lower control limits produced mixed results. For the most part, the lower limits in this study were more stringent than the PT limits; however, in only half the cases were they more stringent than the limits published in the methods. Since method limits were calculated using extremely limited data (i.e., from a single laboratory), the LCS study data was considered more typical of laboratory performance and therefore more appropriate to use.

3.2.6 Estimation of Failure Rates

The EDQW was concerned about the effects of the new control limits on laboratory LCS failure rates. They approached the study with an understanding that several key factors drive the capabilities of the analytical system:

- The methods themselves are far from perfect. As documented in many of the published methods, the anticipated lower control limits for LCS recoveries of certain analytes approach zero percent.
- LCS failure can occur as a result of both random and systematic problems. When analyzing a list of analytes, there is a statistical probability that one or more of the

analytes will fail to meet acceptance criteria due to random errors that are beyond the control of the laboratory. Although they raise a level of concern, these random failures do not reflect the laboratory's implementation of the method.

- A significant increase in the failure rate beyond what already occurs under the existing approaches to LCSs would have negative cost implications for both the laboratory and DoD.

To test the limits' effect on laboratory LCS failures, the LCS-CLs were applied to the individual LCS results submitted for the study. If one or more analytes exceeded the LCS-CLs (less than the lower limit or greater than the upper limit), the LCS failed and corrective action would be required for the batch of environmental samples. Estimated failure rates were first calculated using the limits generated in the study and the definition of failure described above. Table 2 presents total failure rates for all laboratories, as well as failure rates when the lower and upper limits were considered separately. (Note: It is possible for a single LCS to fail as a result of separate analytes failing the lower limit and the upper limit. Consequently, the sum of the lower limit and upper limit failures may be greater than the total number of failures.)

Table 2. Baseline LCS Failure Rates

	Failure Rates – Solid Matrix			Failure Rates – Water Matrix		
Method	**Total (%)**	**Lower Limit (%)**	**Upper Limit (%)**	**Total (%)**	**Lower Limit (%)**	**Upper Limit (%)**
Semivolatiles (8270C)	18	4	15	28	14	14
Volatiles (8260B)	22	6	19	15	6	10
Herbicides (8151A)	24	6	19	16	6	10
PAHs (8310)	28	5	23	9	2	8
Explosives (8330)	13	9	6	14	7	6
Pesticides (8081A)	24	14	14	18	11	8
PCBs (8082)	9	6	3	5	3	2
Metals (6010B)	21	8	15	11	7	4
Mercury (7470A/7471A)	2	1	1	1	0	1

Analysis of the baseline failure rates demonstrated that more LCS failures were caused by exceedance of the upper control limits than exceedance of the lower control limits and therefore are more likely to result in unnecessary actions (false positives) than in not enough action. Failure rates sometimes varied significantly by laboratory. For some methods, only a handful of laboratories accounted for most of the failures. Failure rates for in-house control limits (each laboratory's data compared with the in-house limits it provided for the study) showed that laboratories were generally less likely to fail using their own limits than using the limits generated by the study. This is not surprising considering that the in-house limits should be generated using historical data from that laboratory. Laboratories having more variability in LCS recoveries generate wide limits within which their data could fall. However, not all laboratories that submitted data for the study generated in-house limits using historical data. Some laboratories appear to have adopted AFCEE published limits or arbitrarily set limits, such as 80 to 120% for all analytes.

After determining the baseline failure rates, the study team performed numerous additional analyses to evaluate the effects of modifying the manner in which LCS limits were set and applied. This analysis varied (1) the manner in which the LCS limits were set (e.g., lower limits raised to 10 or 20%) and (2) the manner in which LCS limits were applied (e.g., the definition of failure of an LCS to allow for sporadic marginal exceedances of limits).

Some of the adjustments of the limits included raising the lower limits (to 10, 20, and 50%), raising the upper limits (to 100, 110, and 120%), and setting limits at 2 standard deviations around the mean instead of three. Adjusting the definition of failure included multiple variations of the marginal exceedance approach (allowing a certain number of analytes to marginally exceed the LCS-CLs based on the total number of analytes spiked in the LCS). Failure rates increased when raising the lower limit and decreased when raising the upper limit. Adjusting the upper limit usually had less effect on failure rates than adjusting the lower limit, since failures of the upper limit tended to be by larger amounts (i.e., greater than 120%). However, adjusting the upper limits affected more compounds. Modifying the definition of failure always decreased failure rates from the baseline, since more than one failed analyte was allowed. The amount of change in failure rates varied depending on the number of allowances. Section 4.0 discusses adjustments in how the limits were set and the final approach for determining failures.

3.3 Establishing DoD LCS-CLs

When DoD initiated the project to establish LCS-CLs, it determined that any decisions to come out of the study would be based on sound science. However, since these final decisions represent DoD-wide policy, they had to be tempered with scientific insight. With that in mind, once the statistically generated limits were determined, a number of issues were considered as to how the LCS-CLs would be both set and applied. These factors reflect the following considerations:

- The LCS-CLs should be used to identify blunders and generally not to penalize laboratories for random out-of-control events.
- Given the variability within the laboratory community and the fact that the data reflect analytical practice at a given point in time, the study results are not necessarily predictive of future laboratory performance. However, understanding potential LCS and analytical batch failure rates is critically important to policy development. Unwarranted increases in failure rates (i.e., those associated with random failures) could lead to excessively penalizing the laboratory and DoD for factors out of their control.
- Failure rates based on the application of LCS-CLs to default lists of analytes may be different from those resulting from the application of LCS-CLs to individual analytes identified as project-specific target analytes.
- High levels of variability (as measured by wide standard deviations) can be associated with entire methods or specific analytes.
- Implementation of the DoD-wide LCS-CLs by commercial laboratories should minimize complexity.
- The LCS policy should encourage laboratories to maintain or improve performance beyond the default limits.

3.3.1 Statistical Probabilities of Random and Nonrandom Failures

Random error during laboratory analysis is inevitable. Given the complexity of the analytical methods, there is a finite probability that an LCS result will fall outside the LCS-CLs as a result of random error. By spiking multiple analytes in a single LCS, the probability of LCS failures due to random error is compounded, and the chance that one or more of the analytes will not meet acceptance criteria increases. DoD does not accept the results of an analytical batch when its associated LCS has failed; however, DoD does not want to penalize laboratories for random events beyond their control. At the same time it seeks to minimize the acceptance of LCSs that

reflect systematic problems over which the laboratory should have control. This issue can be framed by two questions:

1. What is the likelihood that failure of the LCS is due to the random occurrences that are out of the laboratory's control?
2. What is the likelihood that failure is due to nonrandom events (e.g., systematic errors or blunders) that the laboratory does have some control over?

Probability theory using a binomial distribution indicates that the chance for a random event increases as the number of trials increases. For an LCS with multiple analytes, each analyte would be considered a separate trial. The Army Corps of Engineers has a system in place for allowing a certain number of analytes to fail based on the number of analytes in the LCS. The EDQW agreed with the concept and performed multiple statistical analyses to determine the maximum allowable number of failed analytes.

After analyzing the results, DoD chose to set the allowable number of failures at 5% of the total number of analytes. This is a straightforward yet still conservative approach that is based on professional judgment. Table 4 in Section 4.2 presents the final number of allowable failures versus the number of analytes in the LCS.

3.3.2 Evaluation of Adjustments to Limits and Application

As described in Section 3.2.6, the study team calculated LCS failure rates for a baseline scenario (statistically generated limits using the standard definition of failure) as well as for a variety of scenarios involving modifications to how the limits were set and how failure was defined.

3.3.2.1 Setting the Limits

Adjustments to the limits reflected the following concerns:

- Excessively low lower control limits could result in a low bias and lead to false negatives (and potential risks to human health and the environment). This concern was addressed by the poor performing analyte concept discussed in Section 3.2.4.
- Excessively high upper control limits could allow a high bias and lead to false positives (and thus unnecessary expense to DoD); however, high bias was generally not a problem in this study.
- Control limits in which the upper limit was less than 100% could in effect penalize laboratories for good performance. Producing the correct recovery (100%) would result in failure of the LCS.
- There was no benefit in requiring laboratories to achieve LCS acceptance criteria that were more stringent than method-defined acceptance criteria, if the method limits were already sufficiently stringent.

The EDQW discussed the advantages and disadvantages of adjusting the statistically generated limits. They considered whether the limits should be arbitrarily modified or whether the data should be allowed to speak for themselves, thereby identifying where improvements in the methods need to be made. Ultimately the EDQW struck a balance by:

- Generally keeping the LCS-CLs close to those generated by the statistical methodology; allowing exceptions only if supported by sound scientific rationale.

- Noting that if a project-specific analyte of concern has a level of variability and resulting LCS-CLs that are inadequate for the use of the data, the client should be contacted about the need for potential method optimization.
- Identifying certain analytes that are poor performing analytes, and noting that the client should be contacted about method optimization if data suggest that those analytes may be present at the site.

For herbicides method 8151A (both water and solid matrix) the intra-laboratory variability in recoveries was large for almost every analyte. The standard deviations were high, resulting in extremely wide control limits. Scatter plots for every compound were reviewed to confirm this variability. The EDQW chose to set control limits for method 8151A using nonparametric statistics. The control limits were based on 5th and 95th percentiles for each analyte (no outliers were removed). As described in Section 3.2.3, LCS-CLs for explosives method 8330 in water were based only on data that used solid phase extraction (SPE).

The EDQW decided to define poor performers as analytes with lower control limits of 10% or less and treat those analytes separately on a project-specific basis. They felt that it was inappropriate to control batch acceptance on analytes with lower control limits of 10% or less. However, artificially raising the lower limits from the statistically generated level did not address the problems of a method that produced extremely low or variable recoveries.

For inorganic compounds, the limits were adjusted to be at least 80 to 120%, consistent with the allowable acceptance criteria in proposed method 6010C.

All limits were rounded to the nearest 5% for ease of implementation.

3.3.2.2 Applying the Limits: Sporadic Marginal Exceedances

The study team also considered options for applying the LCS-CLs (i.e., defining LCS failure), recognizing that larger lists of analytes result in higher rates of random failures. Simple probability calculations (binomial statistics) predict that there is a finite chance that random errors will cause an analyte to fall outside the LCS-CLs, and that the chance will increase with the number of analytes. Thus, laboratories that include a long list of analytes in the LCS spike can be penalized in terms of higher LCS failure rates and the associated costs of repreparing and reanalyzing the samples. After evaluating the failure rates, the study team developed a marginal exceedance approach for calculating failure for methods with longer lists of analytes (see Section 4.2 for a complete explanation).

Allowing a certain number of analytes to exceed the control limits on the basis of analyte list length lessens the likelihood that laboratories will fail an LCS because of random error, while still maintaining acceptable data quality. Calculating failure rates using this approach resulted in lower failure rates than with the standard approach, with the greatest effect being on the methods with long lists of analytes (e.g., methods 8270C and 8260B). Table 3 summarizes the failure rates for each method using the final limits and final definition of failure: rounding the limits to the nearest 5%, adjusting limits to be at least as wide as 80 to 120% for inorganics, applying the marginal exceedance approach, and excluding poor performing analytes. (Note: The final policy specifies that project-specific requirements supersede all DoD-specified limits. In addition, the marginal exceedance policy cannot be used for any analytes specifically identified as project-specific analytes of concern.)

Table 3. LCS Failure Rates Using Final LCS Policy

Method	Failure Rates – Solid Matrix			Failure Rates – Water Matrix		
	Total (%)	Lower Limit (%)	Upper Limit (%)	Total (%)	Lower Limit (%)	Upper Limit (%)
Semivolatiles (8270C)	9	2	7	18	10	8
Volatiles (8260B)	13	2	11	8	4	5
Herbicides (8151A)	28	14	15	34	19	16
PAHs (8310)	19	4	15	5	1	4
Explosives (8330)	13	9	5	3*	3*	0*
Pesticides (8081A)	20	12	11	10	7	5
PCBs (8082)	9	6	3	5	3	2
Metals (6010B)	6	3	3	1	1	0
Mercury (7470A/7471A)	0	0	0	0	0	0

* Included only laboratories that used SPE preparatory method.

A comparison of the final failure rates with the baseline failure rates found that the total rate of expected failures decreased between 0 and 15 percentage points under the final policy, depending on the method. The one exception to this decrease was herbicides method 8151A, in which a nonparametric methodology was used to generate limits. The nonparametric methodology produced more stringent control limits than the standard methodology; therefore, it was more likely that recoveries would fall outside the limits (see Section 3.3.2.1). Failure rates actually increased from the baseline by 4 percentage points for solid and 18 percentage points for water. There was no change in failure rate for PCBs method 8082 because short analyte lists do not benefit from the marginal exceedance allowance, and failure rates for mercury method 7470A/7471A decreased only 2 and 1 percentage points (solid and water matrix, respectively) because of the widening of the limits to 80 to 120%.

Failure rates for in-house laboratory limits were generally comparable to the final policy rates. The most significant exception was for herbicides, where failure rates increased significantly under the final policy as a result of the more stringent limits from the nonparametric methodology.

4.0 DoD LCS-CLs IMPLEMENTATION

The EDQW developed a final approach regarding the setting and applying of LCS-CLs after substantial input from a variety of stakeholders. This approach is described in Appendix DoD-D of the QSM Version 2 and is summarized in this section.

4.1 Setting the Limits

The general approach to setting the control limits used 3 standard deviations around the mean, calculated after outliers had been removed. Limits were then rounded to the nearest 5% for ease of use. LCS-CLs for metals method 6010B and mercury methods 7470A/7471A were set at 80 to 120% if the statistically generated limits were within that range. If the statistically generated limits were outside 80 to 120% (e.g., silver in the solid matrix has a lower LCS-CL of 75%), the control limit remained at the statistically generated value. These values are consistent with the allowable LCS acceptance criteria in proposed method 6010C.

4.2 Applying the Limits: Allowance for Sporadic Marginal Exceedances

DoD redefined LCS failure in order to allow a number of sporadic marginal exceedances of the LCS-CLs. This policy reflects DoD's desire to not penalize laboratories for small random errors, while still identifying significant systematic errors. The number of exceedances is based on the total number of analytes spiked in the LCS. The number of allowable marginal exceedances is based on a policy decision that no more than 5% of the total number of analytes spiked in the LCS may exceed the DoD limits. This is a simple and conservative approach. Table 4 presents the allowable number of marginal exceedances for a given number of analytes in the LCS. The marginal exceedance limits were set at 4 standard deviations around the mean with a lower limit of at least 10%.

Table 4. Number of Marginal Exceedances

Number of Analytes in LCS	Allowable Number of Marginal Exceedances of LCS-CLs
> 90	5
71 – 90	4
51 – 70	3
31 – 50	2
11 – 30	1
< 11	0

A *marginal exceedance* is defined as beyond the LCS-CL but still within the marginal exceedance limits of 4 standard deviations around the mean. This outside boundary prevents a grossly out-of-control LCS from passing. Marginal exceedances are not allowed for analytes that are project-specific analytes of concern. DoD also requires that the marginal exceedances be sporadic (i.e., random). If the same analyte repeatedly exceeds the LCS-CL (e.g., 2 out of 3 consecutive LCSs), that is an indication that the problem is systematic and something is wrong with the measurement system. The source of error should be located and the appropriate corrective action taken.

Under this policy, failure of the LCS can occur several ways:

- Exceedance of an LCS-CL by any project-specific analyte of concern
- Marginal exceedance of the LCS-CLs by more than the allowable number of analytes
- Exceedance of the marginal exceedance limits by one or more analytes

4.3 Addressing Poor Performing Analytes

Laboratories are required to include all target analytes in the calibration standards, including the poor performing analytes. However, they should not apply LCS-CLs to the poor performing analytes when determining LCS acceptance. If one of the poor performing analytes identified in Table 1 is a project-specific analyte of concern, or if it is detected in the project samples, the laboratory should contact the client (DoD), who will then work with the laboratory on an appropriate course of action. Ideally, DoD and the laboratory will use an alternative method to test for the analyte (one that is known to produce higher recoveries) or else modify the original method to optimize conditions for the poor performing analyte. The lower control limit for alternative or modified methods must be greater than 10% to be considered acceptable. The LCS-CLs for the poor performing analytes generated in this study are provided as a benchmark

against which laboratories may measure the effectiveness of alternative methods or modifications to the current methods.

4.4 Maintaining In-house LCS Limits

In keeping with current, accepted practices, laboratories should continue to maintain their own in-house LCS limits. These in-house limits must be consistent with the limits produced in the LCS study, where available. The laboratory should calculate in-house limits from its historical LCS data and monitor its performance through the use of control charts.

The laboratory's in-house limits should be used for several purposes:

- As part of the laboratory's quality control system, to evaluate trends and monitor and improve performance.
- To evaluate the effects of laboratory performance on environmental data quality, on a batch-specific basis. When a laboratory's in-house limits are outside the DoD control limits (upper or lower), the laboratory must include its in-house limits in the laboratory report, even if the LCS associated with the batch was within the DoD limits.
- To enable DoD to determine acceptability of a laboratory's overall performance. DoD may review the laboratory in-house limits and the associated trends reflected in control charts. If DoD deems the performance unacceptable, they may use the in-house limits as a basis for deciding to not use the laboratory until substantial improvement has occurred.

4.5 LCS-CLs

The LCS study used real-world data to demonstrate current method performance by environmental laboratories. The EDQW expects that laboratories will be able to routinely achieve the LCS-CLs. Project managers should incorporate the LCS-CLs in their quality assurance project plans, and laboratories can use the limits to benchmark alternative methods as part of a performance-based approach.

Tables 5 through 20 present the mean (or median), standard deviation, and control limits as generated by the DoD LCS policy (excluding rounding to the nearest 5%). Refer to Appendix DoD-D of the QSM Version 2 for the rounded LCS-CLs and marginal exceedance limits.

Table 5. LCS Control Limits for Volatile Organic Compounds
SW-846 Method 8260B Water Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
1,1,1,2-Tetrachloroethane	104.7	8.0	81	129
1,1,1-Trichloroethane	99.7	10.8	67	132
1,1,2,2-Tetrachloroethane	95.6	10.7	63	128
1,1,2-Trichloroethane	100.0	8.4	75	125
1,1-Dichloroethane	100.8	10.7	69	133
1,1-Dichloroethene	98.6	10.3	68	130
1,1-Dichloropropene	102.3	9.9	73	132
1,2,3-Trichlorobenzene	99.3	14.1	57	142
1,2,3-Trichloropropane	98.2	8.5	73	124
1,2,4-Trichlorobenzene	99.9	11.4	66	134
1,2,4-Trimethylbenzene	102.9	9.7	74	132
1,2-Dibromo-3-chloropropane	91.3	13.7	50	132
1,2-Dibromoethane	100.4	6.7	80	121
1,2-Dichlorobenzene	96.5	8.5	71	122
1,2-Dichloroethane	100.1	10.5	69	132
1,2-Dichloroethane-d4 (surrogate)	95.2	7.8	72	119
1,2-Dichloropropane	100.2	8.3	75	125
1,3,5-Trimethylbenzene	102.3	9.5	74	131
1,3-Dichlorobenzene	99.6	8.1	75	124
1,3-Dichloropropane	99.6	8.9	73	126
1,4-Dichlorobenzene	98.8	8.1	74	123
2,2-Dichloropropane	102.9	11.2	69	137
2-Butanone	91.0	19.7	32	150
2-Chlorotoluene	99.5	9.0	73	126
2-Hexanone	92.4	12.0	56	128
4-Bromofluorobenzene (surrogate)	97.6	7.1	76	119
4-Chlorotoluene	101.0	8.9	74	128
4-Methyl-2-pentanone	96.0	12.7	58	134
Acetone	90.7	17.2	39	142
Benzene	101.7	6.9	81	122
Bromobenzene	100.0	7.9	76	124
Bromochloromethane	97.3	10.6	65	129
Bromodichloromethane	98.2	7.5	76	121
Bromoform	98.6	9.9	69	128
Bromomethane	88.0	19.5	30	146
Carbon disulfide	99.7	20.8	37	162
Carbon tetrachloride	101.9	12.0	66	138
Chlorobenzene	101.8	6.9	81	122
Chlorodibromomethane	95.7	12.5	58	133
Chloroethane	98.6	12.1	62	135
Chloroform	99.6	12.2	63	136
Chloromethane	83.2	14.6	39	127

Table 5. LCS Control Limits for Volatile Organic Compounds SW-846 Method 8260B Water Matrix (continued)

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
cis-1,2-Dichloroethene	98.6	9.0	72	126
cis-1,3-Dichloropropene	100.3	10.3	69	131
Dibromofluoromethane (surrogate)	99.9	5.1	85	115
Dibromomethane	100.6	8.3	76	125
Dichlorodifluoromethane	93.0	20.6	31	155
Ethylbenzene	100.2	9.1	73	127
Hexachlorobutadiene	96.9	15.2	51	142
Isopropylbenzene	101.1	8.8	75	127
m,p-Xylene	102.3	8.7	76	128
Methyl tert-butyl ether	94.0	9.7	65	123
Methylene chloride	96.4	14.4	53	140
Naphthalene	96.1	14.0	54	138
n-Butylbenzene	102.6	11.3	69	137
n-Propylbenzene	100.5	9.4	72	129
o-Xylene	100.3	6.8	80	121
p-Isopropyltoluene	101.7	9.7	73	131
sec-Butylbenzene	99.6	9.2	72	127
Styrene	99.8	11.5	65	134
tert-Butylbenzene	99.4	9.8	70	129
Tetrachloroethene	96.3	17.6	44	149
Toluene	99.8	7.5	77	122
Toluene-d8 (surrogate)	101.6	6.1	83	120
trans-1,2-Dichloroethene	99.3	13.3	60	139
trans-1,3-Dichloropropene	97.7	14.8	53	142
Trichloroethene	98.7	9.4	70	127
Trichlorofluoromethane	102.7	14.6	59	146
Vinyl chloride	98.9	16.1	50	147

Table 6. LCS Control Limits for Volatile Organic Compounds
SW-846 Method 8260B Solid Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
1,1,1,2-Tetrachloroethane	99.7	8.6	74	125
1,1,1-Trichloroethane	100.5	10.9	68	133
1,1,2,2-Tetrachloroethane	92.5	13.0	54	131
1,1,2-Trichloroethane	94.9	10.9	62	127
1,1-Dichloroethane	99.0	8.7	73	125
1,1-Dichloroethene	100.2	11.8	65	136
1,1-Dichloropropene	102.2	10.8	70	135
1,2,3-Trichlorobenzene	97.5	11.7	62	133
1,2,3-Trichloropropane	96.7	11.2	63	130
1,2,4-Trichlorobenzene	97.6	11.0	65	131
1,2,4-Trimethylbenzene	100.0	11.8	65	135
1,2-Dibromo-3-chloropropane	87.4	15.7	40	135
1,2-Dibromoethane	97.1	9.1	70	124
1,2-Dichlorobenzene	96.6	7.4	74	119
1,2-Dichloroethane	104.3	10.8	72	137
1,2-Dichloropropane	95.0	8.1	71	119
1,3,5-Trimethylbenzene	98.9	11.4	65	133
1,3-Dichlorobenzene	98.1	8.7	72	124
1,3-Dichloropropane	99.8	7.8	76	123
1,4-Dichlorobenzene	98.5	8.9	72	125
2,2-Dichloropropane	100.6	11.3	67	134
2-Butanone	94.0	21.6	29	159
2-Chlorotoluene	98.5	9.9	69	128
2-Hexanone	96.7	16.4	47	146
4-Bromofluorobenzene (surrogate)	101.3	5.6	84	118
4-Chlorotoluene	99.8	8.8	73	126
4-Methyl-2-pentanone	97.2	16.6	47	147
Acetone	88.2	23.1	19	158
Benzene	99.4	8.8	73	126
Bromobenzene*	93.4	9.3	66	121
Bromochloromethane	99.4	9.3	71	127
Bromodichloromethane	99.8	9.4	72	128
Bromoform	96.5	13.4	56	137
Bromomethane	95.0	21.3	31	159
Carbon disulfide	102.7	18.7	47	159
Carbon tetrachloride	99.7	11.0	67	133
Chlorobenzene	98.9	8.1	75	123
Chlorodibromomethane	98.0	10.5	66	130
Chloroethane	98.3	19.6	39	157
Chloroform	98.0	8.7	72	124
Chloromethane	89.8	13.0	51	129

*Provisional limits – outlier analyses during the LCS study resulted in LCS-CLs generated with data from fewer than four laboratories.

Table 6. LCS Control Limits for Volatile Organic Compounds SW-846 Method 8260B Solid Matrix (continued)

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
cis-1,2-Dichloroethene	96.2	9.7	67	125
cis-1,3-Dichloropropene	98.8	8.9	72	126
Dibromomethane	100.4	9.2	73	128
Dichlorodifluoromethane*	84.7	17.0	34	136
Ethylbenzene	100.5	8.8	74	127
Hexachlorobutadiene	97.8	14.9	53	142
Isopropylbenzene	103.0	8.8	77	129
m,p-Xylene	102.4	7.9	79	126
Methylene chloride	97.4	14.4	54	141
Naphthalene	83.5	14.4	40	127
n-Butylbenzene	101.1	12.2	65	138
n-Propylbenzene	99.0	11.9	63	135
o-Xylene	101.4	8.0	77	125
p-Isopropyltoluene	103.6	9.6	75	133
sec-Butylbenzene	97.2	11.5	63	132
Styrene	100.7	9.1	74	128
tert-Butylbenzene	98.8	11.1	65	132
Tetrachloroethene	103.0	11.9	67	139
Toluene	98.9	9.2	71	127
Toluene-d8 (surrogate)	100.3	5.3	84	116
trans-1,2-Dichloroethene	100.1	11.3	65	135
trans-1,3-Dichloropropene	95.8	10.4	65	125
Trichloroethene	100.5	7.8	77	124
Trichlorofluoromethane	105.6	26.9	25	186
Vinyl chloride	92.1	11.4	58	126

*Provisional limits – outlier analyses during the LCS study resulted in LCS-CLs generated with data from fewer than four laboratories.

Table 7. LCS Control Limits for Semivolatile Organic Compounds
SW-846 Method 8270C Water Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
1,2,4-Trichlorobenzene	71.7	11.6	37	107
1,2-Dichlorobenzene	67.3	11.4	33	102
1,2-Diphenylhydrazine	84.8	9.4	57	113
1,3-Dichlorobenzene	64.8	10.9	32	98
1,4-Dichlorobenzene	64.8	10.9	32	98
2,4,5-Trichlorophenol	79.7	10.3	49	111
2,4,6-Tribromophenol (surrogate)	82.9	13.6	42	124
2,4,6-Trichlorophenol	80.7	10.7	49	113
2,4-Dichlorophenol	76.3	9.6	48	105
2,4-Dimethylphenol	68.8	13.5	28	109
2,4-Dinitrophenol	75.8	20.6	14	138
2,4-Dinitrotoluene	84.3	11.2	51	118
2,6-Dinitrotoluene	82.7	11.3	49	117
2-Chloronaphthalene	76.5	9.3	49	104
2-Chlorophenol	71.3	11.4	37	106
2-Fluorobiphenyl (surrogate)	79.9	10.6	48	112
2-Fluorophenol (surrogate)	63.7	14.8	19	108
2-Methylnaphthalene	75.0	9.5	46	104
2-Methylphenol	73.3	11.7	38	109
2-Nitroaniline	81.8	11.2	48	115
2-Nitrophenol	75.8	12.4	39	113
3,3'-Dichlorobenzidine	65.2	15.3	19	111
3-Methylphenol/4-Methylphenol	71.3	13.0	32	110
3-Nitroaniline	72.6	17.7	19	126
4,6-Dinitro-2-methylphenol	84.9	15.0	40	130
4-Bromophenyl phenyl ether	82.9	10.2	52	113
4-Chloro-3-methylphenol	78.6	10.7	47	111
4-Chloroaniline	62.2	15.6	15	109
4-Chlorophenyl phenyl ether	80.6	10.3	50	111
4-Nitroaniline	77.2	13.7	36	118
Acenaphthene	77.6	10.1	47	108
Acenaphthylene	78.5	9.4	50	107
Anthracene	83.0	9.7	54	112
Benz(a)anthracene	82.7	8.9	56	109
Benzo(a)pyrene	81.3	9.5	53	110
Benzo(b)fluoranthene	81.8	12.1	45	118
Benzo(g,h,i)perylene	80.5	14.1	38	123
Benzo(k)fluoranthene	84.6	13.2	45	124
Benzyl alcohol	71.0	13.8	30	112
Bis(2-chlorethoxy)methane	76.2	10.2	46	107
Bis(2-chloroethyl)ether	73.3	12.3	37	110

Table 7. LCS Control Limits for Semivolatile Organic Compounds SW-846 Method 8270C Water Matrix (continued)

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
Bis(2-chloroisopropyl) ether	78.2	17.5	26	131
Bis(2-ethylhexyl) phthalate	84.2	14.0	42	126
Butyl benzyl phthalate	81.1	11.7	46	116
Carbazole	82.5	11.4	48	117
Chrysene	82.1	8.9	55	109
Dibenz(a,h)anthracene	84.7	14.1	42	127
Dibenzofuran	80.3	8.8	54	107
Diethyl phthalate	79.2	12.9	41	118
Dimethyl phthalate	75.9	16.9	25	127
Di-n-butyl phthalate	84.8	10.3	54	116
Di-n-octyl phthalate	87.4	16.6	37	137
Fluoranthene	85.2	10.4	54	116
Fluorene	80.6	10.3	50	112
Hexachlorobenzene	82.3	10.0	52	112
Hexachlorobutadiene	65.2	12.6	27	103
Hexachloroethane	60.9	11.1	28	94
Indeno(1,2,3-cd)pyrene	84.3	13.6	43	125
Isophorone	81.0	10.5	50	112
Naphthalene	70.8	10.5	39	102
Nitrobenzene	76.8	10.8	44	109
Nitrobenzene-d5 (surrogate)	76.0	11.8	41	111
N-Nitrosodimethylamine	67.9	14.1	26	110
N-Nitrosodi-n-propylamine	80.9	15.7	34	128
N-Nitrosodiphenylamine	79.6	10.6	48	111
Pentachlorophenol	77.6	13.3	38	117
Phenanthrene	84.0	11.0	51	117
Pyrene	88.6	13.2	49	128
Terphenyl-d14 (surrogate)	92.7	14.0	51	135

Table 8. LCS Control Limits for Semivolatile Organic Compounds SW-846 Method 8270C Solid Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
1,2,4-Trichlorobenzene	77.4	11.2	44	111
1,2-Dichlorobenzene	70.9	8.7	45	97
1,3-Dichlorobenzene	69.7	10.3	39	100
1,4-Dichlorobenzene	69.0	11.4	35	103
2,4,5-Trichlorophenol	80.1	10.4	49	111
2,4,6-Tribromophenol (surrogate)	80.9	15.1	36	126
2,4,6-Trichlorophenol	76.3	11.0	43	109
2,4-Dichlorophenol	77.2	10.9	45	110
2,4-Dimethylphenol	67.3	11.9	32	103
2,4-Dinitrophenol	72.6	20.0	13	132
2,4-Dinitrotoluene	82.0	11.4	48	116
2,6-Dinitrotoluene	80.2	10.7	48	112
2-Chloronaphthalene	75.2	9.9	45	105
2-Chlorophenol	74.7	10.3	44	106
2-Fluorobiphenyl (surrogate)	72.8	10.0	43	103
2-Fluorophenol (surrogate)	70.6	11.1	37	104
2-Methylnaphthalene	77.3	10.0	47	107
2-Methylphenol	71.7	10.6	40	104
2-Nitroaniline	81.0	12.2	44	118
2-Nitrophenol	76.2	11.5	42	111
3-Methylphenol/4-Methylphenol	73.9	10.9	41	107
3-Nitroaniline	68.8	13.8	27	110
4,6-Dinitro-2-methylphenol	83.1	18.0	29	137
4-Bromophenyl phenyl ether	81.7	11.8	46	117
4-Chloro-3-methylphenol	79.5	11.1	46	113
4-Chlorophenyl phenyl ether	79.6	10.7	47	112
4-Nitroaniline	73.6	13.1	34	113
4-Nitrophenol	77.0	20.2	17	138
Acenaphthene	77.3	10.3	46	108
Acenaphthylene	75.7	10.4	44	107
Anthracene	79.9	9.0	53	107
Benz(a)anthracene	81.6	9.8	52	111
Benzo(a)pyrene	80.7	10.3	50	111
Benzo(b)fluoranthene	79.7	11.4	45	114
Benzo(g,h,i)perylene	81.8	14.7	38	126
Benzo(k)fluoranthene	83.8	12.9	45	123
Benzyl alcohol	70.9	17.4	19	123
Bis(2-chlorethoxy)methane	75.5	10.9	43	108
Bis(2-chloroethyl) ether	71.1	11.2	38	105
Bis(2-chloroisopropyl) ether	68.4	15.7	21	115
Bis(2-ethylhexyl) phthalate	87.4	13.3	47	127
Butyl benzyl phthalate	86.4	12.3	49	123

Table 8. LCS Control Limits for Semivolatile Organic Compounds SW-846 Method 8270C Solid Matrix (continued)

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
Carbazole	80.4	12.3	44	117
Chrysene	82.6	9.9	53	112
Dibenz(a,h)anthracene	82.9	13.9	41	125
Dibenzofuran	77.1	8.8	51	103
Diethyl phthalate	82.2	10.6	50	114
Dimethyl phthalate	79.6	10.2	49	110
Di-n-butyl phthalate	83.2	9.1	56	110
Di-n-octyl phthalate	86.4	15.2	41	132
Fluoranthene	83.9	10.1	54	114
Fluorene	78.3	9.8	49	108
Hexachlorobenzene	82.5	11.7	47	118
Hexachlorobutadiene	78.2	12.9	40	117
Hexachloroethane	71.9	12.6	34	110
Indeno(1,2,3-cd)pyrene	79.7	13.8	38	121
Isophorone	77.0	11.4	43	111
Naphthalene	73.4	11.1	40	107
Nitrobenzene	77.2	11.9	41	113
Nitrobenzene-d5 (surrogate)	69.5	10.7	37	102
N-Nitrosodimethylamine	66.1	15.9	18	114
N-Nitrosodi-n-propylamine	76.8	12.3	40	114
N-Nitrosodiphenylamine	82.4	11.1	49	116
Pentachlorophenol	71.9	15.6	25	119
Phenanthrene	80.1	10.0	50	110
Phenol	69.7	10.2	39	100
Phenol-d5/d6 (surrogate)	71.0	10.2	40	102
Pyrene	84.4	12.8	46	123
Terphenyl-d14 (surrogate)	78.8	15.5	32	125

Table 9. LCS Control Limits for Chlorinated Herbicides SW-846 Method 8151A Water Matrix*

Analyte	Median	Lower Control Limit	Upper Control Limit
2,4-D	88	35	113
2,4-DB	99	44	132
2,4,5-T	83	34	112
2,4,5-TP (Silvex)	87	49	116
Dalapon	62	40	108
Dicamba	86	60	112
Dichloroprop	91	68	122
Dinoseb	65	21	97
MCPA	93	62	144

*LCS-CLs were generated using nonparametric statistics (see Section 3.3.2.1 for further explanation).

Table 10. LCS Control Limits for Chlorinated Herbicides SW-846 Method 8151A Solid Matrix*

Analyte	Median	Lower Control Limit	Upper Control Limit
2,4-D	88	36	144
2,4-DB	108	52	157
2,4,5-T	86	43	137
2,4,5-TP (Silvex)	90	46	125
Dicamba	90	56	110
Dichloroprop	99	77	138

*LCS-CLs were generated using nonparametric statistics (see Section 3.3.2.1 for further explanation).

Table 11. LCS Control Limits for Polynuclear Aromatic Hydrocarbons SW-846 Method 8310 Water Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
Acenaphthene	69.5	11.5	35	104
Acenaphthylene	73.7	13.2	34	113
Anthracene	76.9	11.8	41	112
Benzo(a)anthracene	80.7	10.5	49	112
Benzo(a)pyrene	79.4	11.3	45	113
Benzo(b)fluoranthene	81.6	10.3	51	112
Benzo(g,h,i)perylene	76.6	14.1	34	119
Benzo(k)fluoranthene	79.3	10.4	48	110
Chrysene	83.3	10.9	50	116
Dibenzo(a,h)anthracene	64.2	15.5	18	111
Fluoranthene	82.1	11.3	48	116
Fluorene	69.1	11.3	35	103
Indeno(1,2,3-cd)pyrene	79.6	10.8	47	112
Naphthalene	68.1	11.8	33	104
Phenanthrene	80.2	13.4	40	120
Pyrene	80.0	9.3	52	108

Table 12. LCS Control Limits for Polynuclear Aromatic Hydrocarbons SW-846 Method 8310 Solid Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
Acenaphthene	70.6	12.4	33	108
Acenaphthylene	72.8	13.4	33	113
Anthracene	86.1	13.0	47	125
Benzo(a)anthracene	78.0	9.3	50	106
Benzo(a)pyrene	86.5	15.4	40	133
Benzo(b)fluoranthene	89.3	10.7	57	121
Benzo(g,h,i)perylene*	84.6	10.4	53	116
Benzo(k)fluoranthene	84.5	12.2	48	121
Chrysene	87.0	10.7	55	119
Dibenzo(a,h)anthracene	80.8	11.4	47	115
Fluoranthene	88.2	15.6	41	135
Fluorene	76.4	10.1	46	107
Indeno(1,2,3-cd)pyrene	94.9	13.0	56	134
Naphthalene	79.9	10.5	48	111
Phenanthrene	91.2	11.5	57	126
Pyrene	82.3	11.0	49	115

* Provisional limits – outlier analyses during LCS study resulted in LCS-CLs generated with data from fewer than four laboratories.

Table 13. LCS Control Limits for Explosives
SW-846 Method 8330 Water Matrix*

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
1,3,5-Trinitrobenzene	101.5	12.6	64	139
1,3-Dinitrobenzene	102.5	18.4	47	158
2,4-Dinitrotoluene	97.6	12.3	61	135
2,6-Dinitrotoluene	98.5	12.7	60	137
2,4,6-Trinitrotoluene (TNT)	97.8	15.2	52	143
2-Amino-4,6-dinitrotoluene**	101.2	17.1	50	153
2-Nitrotoluene	88.1	15.0	43	133
3-Nitrotoluene	89.9	14.1	48	132
4-Amino-2,6-dinitrotoluene**	104.3	16.5	55	154
4-Nitrotoluene	90.2	14.0	48	132
Hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX)	106.3	18.3	51	161
Methyl-2,4,6-trinitrophyenylnitramine (Tetryl)**	97.9	25.2	22	174
Nitrobenzene	93.6	14.7	49	138
Octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (HMX)	98.8	5.8	81	116

*LCS-CLs were generated with data using solid phase extraction with acetonitrile only, without removing outliers from the data set (see Section 3.2.3 for further explanation).
**Provisional limits – LCS-CLs were generated with data from fewer than four laboratories.

Table 14. LCS Control Limits for Explosives
SW-846 Method 8330 Solid Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
1,3,5-Trinitrobenzene	99.0	8.5	73	125
1,3-Dinitrobenzene	102.3	7.8	79	126
2,4-Dinitrotoluene	101.9	7.3	80	124
2,6-Dinitrotoluene	100.2	7.3	78	122
2,4,6-Trinitrotoluene (TNT)	98.5	13.8	57	140
2-Amino-4,6-dinitrotoluene	102.0	7.0	80	124
2-Nitrotoluene	101.2	7.2	80	123
3-Nitrotoluene	99.9	7.5	77	122
4-Amino-2,6-dinitrotoluene	101.0	7.0	79	124
4-Nitrotoluene	100.6	8.1	76	125
Hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX)	103.0	10.0	72	134
Nitrobenzene	100.4	7.8	77	124
Octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (HMX)	100.0	9.0	74	126

Table 15. LCS Control Limits for Organochlorine Pesticides
SW-846 Method 8081A Water Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
4,4'-DDD	88.1	20.4	27	149
4,4'-DDE	86.7	17.8	33	140
4,4'-DDT	92.5	15.0	47	138
Aldrin	82.8	18.6	27	138
alpha-BHC	94.1	11.4	60	128
alpha-Chlordane	93.1	10.0	63	123
beta-BHC	96.1	10.0	66	126
Decachlorobiphenyl (surrogate)	83.3	17.2	32	135
delta-BHC	90.9	15.0	46	136
Dieldrin	95.5	11.0	62	129
Endosulfan I*	80.1	10.4	49	111
Endosulfan II	79.2	17.1	28	130
Endosulfan sulfate	95.8	13.9	54	137
Endrin	95.2	13.0	56	134
Endrin aldehyde	96.4	13.6	56	137
Endrin ketone	102.1	8.2	77	127
gamma-BHC	81.9	18.3	27	137
gamma-Chlordane	93.8	10.7	62	126
Heptachlor	86.6	14.8	42	131
Heptachlor epoxide	96.4	11.5	62	131
Methoxychlor	103.0	15.5	56	150
TCMX (surrogate)	81.4	18.8	25	138

*Provisional limits – outlier analyses during the LCS study resulted in LCS-CLs generated with data from fewer than four laboratories.

Table 16. LCS Control Limits for Organochlorine Pesticides
SW-846 Method 8081A Solid Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
4,4'-DDD	81.3	17.9	28	135
4,4'-DDE	97.1	9.7	68	126
4,4'-DDT	92.3	15.8	45	140
Aldrin	93.3	15.6	47	140
alpha-BHC	93.4	10.5	62	125
alpha-Chlordane	92.1	9.7	63	121
beta-BHC	94.5	10.7	62	127
Decachlorobiphenyl (surrogate)	93.9	12.6	56	132
delta-BHC	93.6	12.3	57	130
Dieldrin	96.0	9.7	67	125
Endosulfan I	73.7	19.8	14	133
Endosulfan II	88.9	17.3	37	141
Endosulfan sulfate	98.6	12.2	62	135
Endrin	96.9	12.1	61	133
Endrin aldehyde	92.0	18.4	37	147
Endrin ketone	99.7	11.3	66	134
Gamma-BHC	90.5	10.7	59	123
Gamma-Chlordane	96.4	10.0	66	126
Heptachlor	95.6	14.9	51	140
Heptachlor epoxide	98.0	10.6	66	130
Methoxychlor	100.0	14.2	57	143
TCMX (surrogate)	96.6	9.1	69	124

Table 17. LCS Control Limits for Polychlorinated Biphenyls
SW-846 Method 8082 Water Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
Aroclor 1016	84.6	19.8	25	144
Aroclor 1260	87.5	19.2	30	145
Decachlorobiphenyl (surrogate)	87.5	15.1	42	133

Table 18. LCS Control Limits for Polychlorinated Biphenyls SW-846 Method 8082 Solid Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
Aroclor 1016	89.5	16.1	41	138
Aroclor 1260	96.0	11.6	61	131
Decachlorobiphenyl (surrogate)	91.4	11.2	58	125

Table 19. LCS Control Limits for Metals SW-846 Methods 6010B and 7470A Water Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
Aluminum	97.2	4.6	83	111
Antimony	98.0	4.1	86	110
Arsenic	97.9	4.3	85	111
Barium	99.4	3.8	88	111
Beryllium	99.2	4.0	87	111
Cadmium	99.5	4.2	87	112
Calcium	98.4	3.8	87	110
Chromium	99.9	4.1	88	112
Cobalt	98.7	3.1	89	108
Copper	99.0	3.4	89	109
Iron	101.6	4.0	90	113
Lead	98.9	4.0	87	111
Magnesium	98.4	3.6	88	109
Manganese	100.1	3.9	88	112
Mercury	100.2	5.0	85	115
Molybdenum	94.9	5.2	79	111
Nickel	100.2	4.4	87	113
Potassium	97.7	4.3	85	111
Selenium	98.1	6.0	80	116
Silver	97.3	5.3	82	113
Sodium	99.1	4.0	87	111
Thallium	97.1	3.8	86	109
Vanadium	99.4	4.0	88	111
Zinc	99.7	4.5	86	113

Table 20. LCS Control Limits for Metals
SW-846 Methods 6010B and 7471A Solid Matrix

Analyte	Mean	Standard Deviation	Lower Control Limit	Upper Control Limit
Aluminum	95.1	5.5	79	112
Antimony	96.1	4.7	82	110
Arsenic	95.1	3.9	84	107
Barium	98.4	3.4	88	108
Beryllium	99.1	3.5	89	110
Cadmium	96.8	4.4	83	110
Calcium	96.6	4.1	84	109
Chromium	98.7	4.5	85	112
Cobalt	97.8	4.1	86	110
Copper	96.9	3.1	88	106
Iron	100.3	4.2	88	113
Lead	94.9	4.1	83	107
Magnesium	96.5	3.3	87	106
Manganese	97.4	4.0	85	109
Mercury	100.3	5.9	83	118
Molybdenum	95.5	5.2	80	111
Nickel	97.5	3.9	86	109
Potassium	95.7	4.1	83	108
Selenium	92.8	4.3	80	106
Silver	96.4	7.2	75	118
Sodium	95.6	4.4	82	109
Thallium	94.5	4.2	82	107
Vanadium	98.7	3.4	89	109
Zinc	95.2	5.1	80	110

REFERENCES

Department of Defense. 2002. Quality Systems Manual for Environmental Laboratories, Version 2.

Kafadar, Karen. 1982. Using Biweight M-Estimates in Two-Sample Problems. Part 1: Symmetric Populations. Commun. Statistics – Theory Method 11(17): 1883-1901.

Kafadar, Karen. 1983. The Efficiency of the Biweight as a Robust Estimator of Location. J of Research of the National Bureau of Standards. 88(2).

Taylor, John K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers.

This page intentionally left blank.

Appendix A
Statistical Approach Used to Develop DoD LCS-CLs

This page intentionally left blank.

Appendix A: Statistical Approach Used to Develop DoD LCS-CLs

1.0 Introduction

The DoD Environmental Data Quality Workgroup (EDQW) established DoD-wide control limits for laboratory control samples (LCS-CLs) using empirical data from commercial laboratories that perform work for DoD. The EDQW consulted chemists, statisticians, laboratory representatives, and quality assurance personnel to establish a statistical methodology that would produce reasonable and defensible results. The strategy developed for the study included two phases: In the pilot phase the study team tested the methodology; in the second phase the study team incorporated professional judgment and cost and time implications to arrive at the final outcome. This appendix provides details on the statistical methodology and the initial raw data results.

2.0 Description of the Data Set

The LCS study depended on commercial laboratories to voluntarily submit LCS data to DoD. The American Council of Independent Laboratories (ACIL) assisted in efforts to collect data (see Attachments 1 and 2 for data submittal instructions provided to the laboratories). Ultimately 17 laboratories submitted data for the Phase I analyte group (semivolatiles using SW-846 method 8270C), and 16 laboratories submitted data for at least one analyte group in Phase II. Table A-1 presents the number of laboratories that submitted data for each Phase II analyte group, by matrix.

Table A-1. Phase II Data Received

	Number of Laboratories	
Analyte Group (SW-846 method)	**Water**	**Solid**
Volatile Organic Compounds (8260B)	15	13
Chlorinated Herbicides (8151A)	12	9
Polynuclear Aromatic Hydrocarbons (8310)	10	10
Explosives (8330)	10	10
Organochlorine Pesticides (8081A)	15	15
Polychlorinated Biphenyls (8082)	12	12
Metals (6010B)	12	11
Mercury (7470A/7471A)	10	10

Laboratories do not necessarily perform all of the nine methods analyzed in the study for both solid and water matrices. In addition, the analyte list for a given method will likely vary slightly by laboratory. As a result the number of available data points in the LCS study varied by analyte – from a minimum of 91 points submitted for dichloroprop using chlorinated herbicides method 8151A in solid matrix to a maximum of 396 data points for benzene using volatiles method 8260B in water. Section 4.0 of this appendix provides a detailed summary of all the data received.

3.0 Description of Approach

This section describes the assumptions of the statistical approach for establishing the LCS-CLs, followed by a detailed description of each step of the approach.

3.1 Assumptions

The study approach used the following primary assumptions to develop the LCS-CLs:

- *The laboratories responding to the request for data are a representative sample of the population of "good performing" laboratories.* At the time the study began, a total of 81 laboratories met the criteria for good performing laboratories (i.e., passed an audit by one or more of the DoD components within the past 18 to 24 months). Seventeen laboratories responded to Phase I, and 16 laboratories responded to Phase II.

- *The LCS data submitted by the laboratories was the result of analytical processes that were "in control."* This assumption was met by requiring that LCS data be from batches that passed both initial calibration verification and continuing calibration verification tests.

- *The LCS-CLs developed for each analyte/matrix combination were calculated from data sets that were representative of the capabilities of good performing laboratories.* This assumption was met, first, by requiring that data for an analyte/matrix combination be available from a minimum of five laboratories before the LCS-CL would be calculated. Data sets were tested for the presence of outlying laboratories and individual data points. Analysis of variance (ANOVA) was performed to determine whether differences in laboratory execution of the subject methods (e.g., differences in extraction methods used) resulted in significantly different performance. Finally, the resulting LCS-CLs were benchmarked against in-house control limits from individual laboratories.

3.2 LCS-CL Development Process

During Phase I of the study, the team tested and finalized the process used to develop the DoD LCS-CLs. The study team divided the data set into a test group and a control group. They applied control group data to the control limits that were generated using the test group data to analyze the effect on failure rates. In addition, the team compared two different outlier methodologies and performed extensive analysis of variance and carefully assessed the results. The original study strategy is presented in Attachment 3 to this appendix.

During Phase II of the study, the methodology consisted of identifying outlier laboratories using the Youden test, identifying outlier data points using the Grubbs test, determining significantly different recoveries between key parameters in the analytical method using ANOVA, and calculating the mean and standard deviation of the final data set. The LCS-CLs were calculated at 3 times the standard deviation around the mean. The statistical methodologies used for each step are described below.

3.2.1 Test for Outlying Laboratories

A rank-sum test, called the Youden test (Taylor, 1987), was used to check each analyte data set for outlying laboratories. The test was implemented as follows:

1. The data set was sorted by laboratory.
2. If more than 15 laboratories submitted data for the analyte, the analyte data set was divided into two groups, with laboratories randomly assigned to each group.
3. Fifteen data points were randomly selected for each laboratory.

4. The first data points selected for the laboratories were assigned ranks based on their relative magnitudes, with the largest value assigned a rank of 1, the next largest a rank of 2, etc.
5. Step 4 was repeated for each of the 15 data points.
6. The 15 ranks for each laboratory were summed, and those scores were compared with reference values based on the number of laboratories and number of data points being tested. Laboratories with scores outside of the range of reference values were flagged as possible outliers.
7. Steps 2 through 6 were repeated two more times, to mitigate the possibility that test results were biased either by the division of the laboratories into two groups, or by the 15 randomly selected data points. Laboratories that were flagged as outliers all three times were then identified as potential outliers for the analyte.

The reference values used for the Youden test provide a 95% confidence that non-outlying laboratories will be correctly identified as such (in other words, there is a 5% chance that the test will identify a laboratory as an outlier when it is not). The test assumes that the sources of variation within the data for each laboratory are the same, although it is possible that individual laboratories implemented the analytical methods in different ways. Therefore, the results of the Youden test were examined in conjunction with the results for the ANOVA before a decision was made to exclude a flagged laboratory from the analyte data set.

The Youden test identified at least one laboratory as an outlier for almost all analyte data sets. In most cases DoD chose to remove the Youden outlier data (except in cases where their removal left fewer than four laboratories for a given data set). Since each laboratory had approximately 15 data points per analyte, the removal of Youden outliers had a significant impact on the results. DoD reviewed scatter plots for many data sets to understand how the outlier data points were distributed compared with the rest of the data. The Youden test identified as outliers those laboratories that had consistently higher or lower recoveries than the other laboratories or those with more tightly clustered recoveries.

3.2.2 Test for Outlying Data Points

The Grubbs test was applied to each data set to identify outlier data points. In the Grubbs test, the mean and standard deviation of the entire data set were calculated and the minimum and maximum data points in the data set were identified. Next, the T-values for the minimum and maximum data points were calculated as follows:

$$T = (X_{av} - X_{min})/S \qquad \text{or} \qquad T = (X_{max} - X_{av})/S$$

X_{av} = mean of the data set
X_{min} = minimum value of the data set
X_{max} = maximum value of the data set
S = standard deviation of the data set

The T-values were compared with reference values (Taylor, 1987) using a 5% false rejection rate. This means that there is a 5% chance that a non-outlier would be falsely rejected as an outlier. The reference values depend on both the risk factor and size of the data set. If the T-value is larger than the reference value, the maximum or minimum data point is identified as an outlier. For this study, the Grubbs test was applied to a maximum of 100 data points. If a data set consisted of more than 100 data points for a particular analyte, the program randomly

divided the data set into the appropriate number of groups, each with 100 points or less. The Grubbs test was then performed on each group.

The Grubbs test identified outlier data points at both the low and high end equally. Since the test identified only single data points as outliers, the removal had little effect on the results (except in cases in which the outlier was an order of magnitude higher than the rest of the data).

3.2.3 Analysis of Variance

The analytical methods published in SW-846 allow for variations in their implementation. For example, specific methods may allow variations in the following parameters:

- LCS spike concentrations
- Type of extraction or preparatory method
- LCS matrix
- Sample cleanup method
- Type of chromatography column
- Injection volume

The study used one-way ANOVA to evaluate the effect of these variations on mean LCS recovery results. The ANOVA identified statistically significant differences in mean LCS recoveries for data using opposing method-allowed parameters. The effects of specific variations were evaluated only if the laboratories provided sufficient data to make a valid comparison. For ANOVA results to be considered valid in this study, each parameter (e.g., extraction method) had to have data from at least two different laboratories and a total of more than 30 data points. The amount of data often varied from analyte to analyte within a given method; therefore, ANOVA was not conducted in all cases.

The ANOVA tests were applied both with and without the outliers removed. The ANOVA results were examined in conjunction with the Youden test results and the scientific basis of differing results were considered. The team used the results to decided whether to exclude outlying laboratories or divide the analyte data set by parameter.

When evaluating the ANOVA results and their implications for each method, there were indications that the data should not be divided according to the parameter of interest. In one case, although there was a statistically significant difference in the means, the difference was not enough to have a real effect on the limits (i.e., no practical difference in absolute numbers). For example, the ANOVA on data for metals method 6010B in water showed significant difference in recoveries between extraction methods 3005 and 3010. However, the difference in the means was often less than 4 percentage points. Because the calculation of control limits is driven by the standard deviation (it is multiplied by 3 for both the lower and upper limits), a minor difference in means did not result in significantly different limits.

Another circumstance showed a lack of consistency across analytes in a given method. For the 22 analytes analyzed using method 8081A in water, 8 showed significantly higher recoveries using a *narrow-bore* GC column; however, another 8 showed significantly higher recoveries using a *wide-bore* GC column. The remaining 6 analytes showed no significant difference based on column width. The absence of a consistent trend in results represented a problem with implementation, since it would require two variations in methodology when analyzing a single LCS.

If only one laboratory submitted data for a particular parameter, ANOVA was not performed within the given method. It was not reasonable to make assessments about the effects of certain parameters on an analytical method when the data for one parameter came from a single laboratory. In such a case there could be no certainty whether the differences were truly significant or were due to an outlier laboratory. For instance, PAH method 8310 and semivolatile method 8270C (both solid and water matrices) had only one laboratory that performed a cleanup method. All others did not indicate that cleanup was performed. Similarly, data for volatile method 8260B in water indicated that all but one laboratory used the same extraction method (5030) and same purge temperature (ambient). No ANOVA was performed on these data sets.

In several circumstances, an appropriate amount of data demonstrated noticeable trends; however, after much discussion the EDQW chose to keep the LCS-CLs as they were and not separate the data set by parameters. For example, in Phase I of the study for method 8270C in water, ANOVA tests indicated that extraction method 3520 produced significantly higher recoveries than extraction method 3510. The DoD chemists involved in the study felt that LCS recoveries may not be indicative of the quality of performance of the extraction methods on environmental samples. Opposite trends concerning those same extraction methods were observed in Phase II of the study for PAH method 8310 in water and pesticide method 8081A in water (3510 produced higher recoveries than 3520), but these differences were not pursued for the same reasons.

For several methods in the solid matrix, differences in means were observed between matrix materials (e.g., Ottawa sand and sodium sulfate). However, since they are all clean matrices, none of the materials can truly predict the performance of the analytical method on environmental samples. Although means were often higher using sodium sulfate, DoD chose not to indicate a preference for matrix material by modifying the control limits. Similarly, differences in mean recovery based on spiking concentration did not result in generation of alternative control limits. ANOVA indicated that in some cases, higher spiking concentrations produced higher means; however, the choice of spiking concentration is often a project-specific decision and should not be broadly dictated.

DoD did choose to set LCS-CLs based on ANOVA results for explosives method 8330 in water. There were higher mean recoveries and lower standard deviations for LCS using solid phase extraction (SPE) with acetonitrile elution than for those using the salting out extraction method.

4.0 Raw Data Results

The following tables provide information on the data received (number of laboratories and data points) and the effects of the outlier and ANOVA tests, on an analyte-by-analyte basis.

Results for Method 8260B – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
1,1,1,2-Tetrachloroethane	9	219	103.5	11.7	178	104.7	8.0	
1,1,1-Trichloroethane	10	257	101.3	12.9	175	99.7	10.8	Injection volume 5 mL > 25 mL
1,1,2,2-Tetrachloroethane	10	236	96.4	13.8	173	95.6	10.7	Injection volume 5 mL > 25 mL
1,1,2-Trichloroethane	10	257	97.5	14.5	235	100.0	8.4	
1,1-Dichloroethane	10	257	100.1	12.9	255	100.8	10.7	
1,1-Dichloroethene	14	343	101.2	12.2	247	98.6	10.3	Injection volume 5 mL > 25 mL
1,1-Dichloropropene	9	211	103.6	13.9	189	102.3	9.9	Injection volume 5 mL > 25 mL
1,2,3-Trichlorobenzene	9	208	100.3	16.0	192	99.3	14.1	
1,2,3-Trichloropropane	9	222	98.8	18.5	220	98.2	8.5	Injection volume 5 mL > 25 mL
1,2,4-Trichlorobenzene	9	210	100.6	14.2	188	99.9	11.4	
1,2,4-Trimethylbenzene	9	209	102.8	12.2	187	102.9	9.7	Injection volume 5 mL > 25 mL
1,2-Dibromo-3-chloropropane	9	207	94.1	13.1	147	91.3	13.7	Injection volume 5 mL > 25 mL
1,2-Dibromoethane	9	232	101.0	10.4	170	100.4	6.7	Injection volume 5 mL > 25 mL
1,2-Dichlorobenzene	9	203	99.3	10.4	81	96.5	8.5	Injection volume 5 mL > 25 mL
1,2-Dichloroethane	11	297	98.2	15.8	252	100.1	10.5	
1,2-Dichloroethane-d4 (surrogate)	4	100	99.8	14.1	79	95.2	7.8	
1,2-Dichloropropane	10	257	98.5	12.1	235	100.2	8.3	
1,3,5-Trimethylbenzene	9	209	102.0	12.0	184	102.3	9.5	Injection volume 5 mL > 25 mL
1,3-Dichlorobenzene	9	203	99.6	10.7	161	99.6	8.1	Injection volume 5 mL > 25 mL
1,3-Dichloropropane	8	188	99.7	12.0	168	99.6	8.9	Injection volume 5 mL > 25 mL
1,4-Dichlorobenzene	10	203	99.0	10.4	138	98.8	8.1	Injection volume 5 mL > 25 mL
2,2-Dichloropropane	9	211	103.2	15.9	206	102.9	11.2	Injection volume 5 mL > 25 mL
2-Butanone	9	244	92.3	21.4	222	91.0	19.7	
2-Chlorotoluene	9	206	99.6	11.7	184	99.5	9.0	Injection volume 5 mL > 25 mL
2-Hexanone	9	236	96.9	24.4	192	92.4	12.0	
4-Bromofluorobenzene (surrogate)	7	160	100.9	11.3	140	97.6	7.1	
4-Chlorotoluene	9	206	100.9	11.5	184	101.0	8.9	Injection volume 5 mL > 25 mL
4-Methyl-2-pentanone	9	204	93.7	20.3	162	96.0	12.7	

Results for Method 8260B – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Acetone	9	236	91.4	24.7	194	90.7	17.2	
Benzene	14	356	101.0	8.6	335	101.7	6.9	
Bromobenzene	9	210	99.7	10.7	188	100.0	7.9	Injection volume 5 mL > 25 mL
Bromochloromethane	10	229	97.5	13.1	207	97.3	10.6	
Bromodichloromethane	9	227	100.4	11.6	165	98.2	7.5	Injection volume 5 mL > 25 mL
Bromoform	10	256	97.2	16.4	174	98.6	9.9	
Bromomethane	10	247	93.1	20.6	167	88.0	19.5	
Carbon disulfide	9	237	99.6	20.5	176	99.7	20.8	High spiking > low
Carbon tetrachloride	11	279	100.5	16.2	234	101.9	12.0	
Chlorobenzene	14	352	101.1	10.0	251	101.8	6.9	
Chlorodibromomethane	9	227	100.3	16.1	125	95.7	12.5	Injection volume 5 mL > 25 mL
Chloroethane	10	257	94.7	14.5	201	98.6	12.1	Low spiking > high
Chloroform	11	277	98.7	15.2	274	99.6	12.2	
Chloromethane	10	247	93.4	21.6	147	83.2	14.6	
cis-1,2-Dichloroethene	10	194	98.7	13.7	128	98.6	9.0	Injection volume 5 mL > 25 mL
cis-1,3-Dichloropropene	10	216	99.3	14.9	173	100.3	10.3	
Dibromofluoromethane (surrogate)	5	100	103.1	11.5	60	99.9	5.1	
Dibromomethane	9	208	100.3	11.0	166	100.6	8.3	Injection volume 5 mL > 25 mL
Dichlorodifluoromethane	8	201	90.3	24.1	140	93.0	20.6	
Ethylbenzene	11	252	101.1	12.0	197	100.2	9.1	
Hexachlorobutadiene	9	207	96.8	17.7	203	96.9	15.2	
Isopropylbenzene	9	210	101.5	11.9	172	101.1	8.8	High spiking > low; Injection vol 5 mL > 25 mL
m,p-Xylene	8	113	102.3	8.7	113	102.3	8.7	High spiking > low; Injection vol 5 mL > 25 mL
Methyl tert-butyl ether	4	92	95.5	9.6	72	94.0	9.7	
Methylene chloride	10	250	98.4	17.0	192	96.4	14.4	
Naphthalene	9	210	98.1	15.7	169	96.1	14.0	Injection volume 5 mL > 25 mL
n-Butylbenzene	9	172	102.4	14.0	150	102.6	11.3	
n-Propylbenzene	9	172	100.5	12.4	147	100.5	9.4	Injection volume 5 mL > 25 mL

Results for Method 8260B – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
o-Xylene	10	169	101.0	12.5	131	100.3	6.8	Injection volume 5 mL > 25 mL
p-Isopropyltoluene	9	176	100.8	12.6	174	101.7	9.7	Injection volume 5 mL > 25 mL
sec-Butylbenzene	9	169	100.6	13.1	147	99.6	9.2	Injection volume 5 mL > 25 mL
Styrene	10	249	100.8	13.2	207	99.8	11.5	
tert-Butylbenzene	9	169	100.2	12.7	131	99.4	9.8	Injection volume 5 mL > 25 mL
Tetrachloroethene	11	260	101.4	15.9	132	96.3	17.6	Injection volume 5 mL > 25 mL
Toluene	14	349	101.3	9.1	268	99.8	7.5	
Toluene-d8 (surrogate)	6	140	104.0	11.5	100	101.6	6.1	
trans-1,2-Dichloroethene	10	187	100.1	14.8	165	99.3	13.3	Injection volume 5 mL > 25 mL
trans-1,3-Dichloropropene	10	196	97.5	17.6	173	97.7	14.8	
Trichloroethene	14	343	100.9	9.8	188	98.7	9.4	Injection vol 5 mL > 25 mL; Purge temp 40 deg C > ambient
Trichlorofluoroethane	10	257	97.2	19.3	202	102.7	14.6	
Vinyl chloride	11	277	94.6	17.8	222	98.9	16.1	

Results for Method 8260B – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
1,1,1,2-Tetrachloroethane	6	143	99.1	8.9	105	99.7	8.6	
1,1,1-Trichloroethane	8	180	100.5	10.9	180	100.5	10.9	
1,1,2,2-Tetrachloroethane	8	180	94.5	14.3	158	92.5	13.0	
1,1,2-Trichloroethane	8	180	96.1	10.6	113	94.9	10.9	
1,1-Dichloroethane	8	181	100.5	10.1	114	99.0	8.7	
1,1-Dichloroethene	13	362	101.0	50.6	294	100.2	11.8	Low spiking > high
1,1-Dichloropropene	6	143	101.1	10.7	105	102.2	10.8	
1,2,3-Trichlorobenzene	6	143	94.8	15.4	122	97.5	11.7	

Results for Method 8260B – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
1,2,3-Trichloropropane	6	143	93.4	14.3	123	96.7	11.2	
1,2,4-Trichlorobenzene	6	143	96.4	14.5	103	97.6	11.0	
1,2,4-Trimethylbenzene	6	143	100.0	11.8	143	100.0	11.8	
1,2-Dibromo-3-chloropropane	6	133	89.7	15.7	113	87.4	15.7	
1,2-Dibromoethane	7	138	97.1	9.1	138	97.1	9.1	
1,2-Dichlorobenzene	6	133	95.7	9.9	93	96.6	7.4	
1,2-Dichloroethane	9	232	101.2	12.9	194	104.3	10.8	Ambient purge temp > 40 deg C
1,2-Dichloropropane	8	180	98.0	9.3	131	95.0	8.1	
1,3,5-Trimethylbenzene	6	143	98.9	11.4	143	98.9	11.4	
1,3-Dichlorobenzene	6	133	96.9	10.7	93	98.1	8.7	
1,3-Dichloropropane	6	143	97.7	9.5	125	99.8	7.8	
1,4-Dichlorobenzene	8	182	96.7	10.5	162	98.5	8.9	
2,2-Dichloropropane	6	143	101.6	13.7	105	100.6	11.3	
2-Butanone	9	179	116.2	86.4	159	94.0	21.6	
2-Chlorotoluene		143	97.2	11.0	103	98.5	9.9	
2-Hexanone	8	169	96.7	19.0	167	96.7	16.4	
4-Bromofluorobenzene (surrogate)	6	173	101.1	5.9	172	101.3	5.6	
4-Chlorotoluene	6	143	97.7	10.7	123	99.8	8.8	
4-Methyl-2-pentanone	8	157	96.7	17.7	156	97.2	16.6	
Acetone	8	175	92.9	24.2	125	88.2	23.1	
Benzene	13	360	105.0	97.8	289	99.4	8.8	
Bromobenzene	6	144	96.5	10.5	55	93.4	9.3	
Bromochloromethane	7	163	97.7	10.4	145	99.4	9.3	
Bromodichloromethane	7	151	99.8	9.4	151	99.8	9.4	
Bromoform	8	181	95.8	13.5	143	96.5	13.4	
Bromomethane	8	170	96.3	25.2	100	95.0	21.3	
Carbon disulfide	8	177	111.6	35.0	138	102.7	18.7	
Carbon tetrachloride	9	232	101.2	12.2	212	99.7	11.0	
Chlorobenzene	13	364	104.4	92.4	323	98.9	8.1	

Results for Method 8260B – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Chlorodibromomethane	8	175	98.0	10.5	175	98.0	10.5	
Chloroethane	8	180	98.8	21.7	134	98.3	19.6	
Chloroform	9	232	102.5	59.2	212	98.0	8.7	
Chloromethane	8	170	92.3	15.4	149	89.8	13.0	
cis-1,2-Dichloroethene	7	162	99.3	10.2	113	96.2	9.7	
cis-1,3-Dichloropropene	8	176	97.3	10.5	138	98.8	8.9	
Dibromomethane	6	142	100.4	9.2	142	100.4	9.2	
Dichlorodifluoromethane	6	142	90.3	28.7	55	84.7	17.0	
Ethylbenzene	9	202	101.4	9.1	182	100.5	8.8	
Hexachlorobutadiene	6	143	95.2	16.5	123	97.8	14.9	
Isopropylbenzene	6	144	101.4	9.9	124	103.0	8.8	
m,p-Xylene	7	160	100.9	9.2	140	102.4	7.9	
Methylene chloride	8	181	100.9	14.8	131	97.4	14.4	
Naphthalene	7	146	92.6	14.9	56	83.5	14.4	
n-Butylbenzene	6	143	100.2	13.4	103	101.1	12.2	
n-Propylbenzene	6	143	99.0	11.9	143	99.0	11.9	
o-Xylene	7	164	100.9	8.9	124	101.4	8.0	
p-Isopropyltoluene	5	127	100.9	11.8	107	103.6	9.6	
sec-Butylbenzene	6	144	98.8	12.6	125	97.2	11.5	
Styrene	8	192	100.7	9.1	192	100.7	9.1	
tert-Butylbenzene	6	143	98.8	11.1	143	98.8	11.1	
Tetrachloroethene	9	209	100.6	13.0	168	103.0	11.9	Low spiking > high
Toluene	13	380	103.3	86.5	379	98.9	9.2	
Toluene-d8 (surrogate)	5	147	100.8	5.2	127	100.3	5.3	
trans-1,2-Dichloroethene	7	162	100.1	11.3	162	100.1	11.3	
trans-1,3-Dichloropropene	8	177	95.6	10.6	138	95.8	10.4	
Trichloroethene	13	362	105.9	97.0	321	100.5	7.8	
Trichlorofluoromethane	7	172	106.3	28.7	171	105.6	26.9	
Vinyl chloride	9	227	93.6	12.3	207	92.1	11.4	Low spiking > high

Results for Method 8270C – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
1,2,4-Trichlorobenzene	17	418	73.6	16.1	274	71.7	11.6	Extraction 3520 > 3510*
1,2-Dichlorobenzene	11	302	70.8	16.8	215	67.3	11.4	
1,2-Diphenylhydrazine	6	115	86.6	11.9	70	84.8	9.4	
1,3-Dichlorobenzene	11	301	69.2	18.5	213	64.8	10.9	
1,4-Dichlorobenzene	16	401	69.5	16.8	294	64.8	10.9	Extraction 3520 > 3510*
2,4,5-Trichlorophenol	11	291	84.8	14.8	185	79.7	10.3	Extraction 3520 > 3510*
2,4,6-Tribromophenol (surrogate)	7	207	89.4	16.6	139	82.9	13.6	Extraction 3520 > 3510*; High spiking > low**
2,4,6-Trichlorophenol	12	318	85.0	15.1	187	80.7	10.7	
2,4-Dichlorophenol	12	318	81.4	15.0	167	76.3	9.6	Extraction 3520 > 3510*
2,4-Dimethylphenol	12	320	66.9	17.6	255	68.8	13.5	
2,4-Dinitrophenol	12	318	82.6	25.1	231	75.8	20.6	Extraction 3520 > 3510*
2,4-Dinitrotoluene	17	434	88.1	15.6	344	84.3	11.2	Extraction 3520 > 3510*
2,6-Dinitrotoluene	11	297	87.8	13.6	206	82.7	11.3	Extraction 3520 > 3510*
2-Chloronaphthalene	12	314	80.8	14.3	203	76.5	9.3	Extraction 3520 > 3510*
2-Chlorophenol	17	411	76.3	17.2	261	71.3	11.4	Extraction 3520 > 3510*
2-Fluorobiphenyl (surrogate)	7	230	82.4	13.7	142	79.9	10.6	Extraction 3520 > 3510*
2-Fluorophenol (surrogate)	7	208	67.7	22.7	61	63.7	14.8	High spiking > low**
2-Methylnaphthalene	11	291	78.9	16.2	164	75.0	9.5	Extraction 3520 > 3510*
2-Methylphenol	10	281	74.3	16.8	167	73.3	11.7	Extraction 3520 > 3510*; High spiking > low**
2-Nitroaniline	10	292	87.0	14.9	225	81.8	11.2	Extraction 3520 > 3510*
2-Nitrophenol	11	301	81.7	18.6	189	75.8	12.4	Extraction 3520 > 3510*
3,3'-Dichlorobenzidine	12	312	75.7	26.6	184	65.2	15.3	High spiking > low**
3-Methylphenol/4-Methylphenol	10	284	75.5	21.0	150	71.3	13.0	Extraction 3520 > 3510*
3-Nitroaniline	9	259	79.4	20.4	192	72.6	17.7	
4,6-Dinitro-2-methylphenol	11	301	90.3	20.6	213	84.9	15.0	
4-Bromophenyl phenyl ether	12	313	86.0	13.5	154	82.9	10.2	

Results for Method 8270C – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
4-Chloro-3-methylphenol	16	403	82.2	14.9	274	78.6	10.7	Extraction 3520 > 3510*
4-Chloroaniline	10	276	69.9	18.5	189	62.2	15.6	
4-Chlorophenyl phenyl ether	12	313	84.6	12.4	203	80.6	10.3	Extraction 3520 > 3510*
4-Nitroaniline	10	278	81.1	15.1	211	77.2	13.7	High spiking > low**
4-Nitrophenol	17	417	64.3	29.9	291	54.3	23.0	Extraction 3520 > 3510*
Acenaphthene	17	436	80.3	12.3	331	77.6	10.1	Extraction 3520 > 3510*
Acenaphthylene	12	334	80.8	12.3	202	78.5	9.4	Extraction 3520 > 3510*; High spiking > low**
Anthracene	12	333	83.7	10.2	308	83.0	9.7	Extraction 3520 > 3510*; High spiking > low**
Benz(a)anthracene	11	325	86.4	11.1	233	82.7	8.9	Extraction 3520 > 3510*; High spiking > low**
Benzo(a)pyrene	12	337	85.6	12.2	245	81.3	9.5	Extraction 3520 > 3510*
Benzo(b)fluoranthene	12	334	84.9	13.3	266	81.8	12.1	
Benzo(g,h,i)perylene	12	323	87.3	16.8	232	80.5	14.1	
Benzo(k)fluoranthene	12	330	87.0	13.0	220	84.6	13.2	Extraction 3520 > 3510*; High spiking > low**
Benzoic acid	10	234	59.5	36.2	108	54.9	24.1	Extraction 3520 > 3510*
Benzyl alcohol	10	248	77.0	21.7	123	71.0	13.8	Extraction 3520 > 3510*
Bis(2-chlorethoxy)methane	12	312	82.9	16.3	201	76.2	10.2	Extraction 3520 > 3510*; High spiking > low**
Bis(2-chloroethyl) ether	12	312	77.6	14.6	202	73.3	12.3	Extraction 3520 > 3510*
Bis(2-chloroisopropyl) ether	10	290	82.1	20.5	177	78.2	17.5	
Bis(2-ethylhexyl) phthalate	12	320	90.6	27.0	231	84.2	14.0	
Butyl benzyl phthalate	12	313	87.0	15.3	226	81.1	11.7	Extraction 3520 > 3510*; High spiking > low**
Carbazole	8	174	84.8	14.5	153	82.5	11.4	High spiking > low**
Chrysene	12	334	86.3	11.3	243	82.1	8.9	Extraction 3520 > 3510*; High spiking > low**
Dibenz(a,h)anthracene	11	323	87.6	14.8	236	84.7	14.1	
Dibenzofuran	11	287	82.4	12.0	180	80.3	8.8	Extraction 3520 > 3510*
Diethyl phthalate	12	314	82.5	14.4	246	79.2	12.9	Extraction 3520 > 3510*; High spiking > low**
Dimethyl phthalate	12	314	77.6	21.5	183	75.9	16.9	High spiking > low**
Di-n-butyl phthalate	11	304	84.8	10.3	304	84.8	10.3	High spiking > low**
Di-n-octyl phthalate	12	314	89.0	18.6	288	87.4	16.6	High spiking > low**
Fluoranthene	12	331	85.8	10.7	306	85.2	10.4	Extraction 3520 > 3510*; High spiking > low**
Fluorene	12	331	84.3	11.4	241	80.6	10.3	Extraction 3520 > 3510*; High spiking > low**

Results for Method 8270C – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Hexachlorobenzene	12	314	85.6	11.9	203	82.3	10.0	Extraction 3520 > 3510*
Hexachlorobutadiene	12	313	70.7	18.8	206	65.2	12.6	Extraction 3520 > 3510*
Hexachloroethane	12	311	67.8	19.9	203	60.9	11.1	Extraction 3520 > 3510*
Indeno(1,2,3-cd)pyrene	12	334	86.1	15.4	225	84.3	13.6	
Isophorone	11	293	83.4	13.5	197	81.0	10.5	Extraction 3520 > 3510*
Naphthalene	12	328	74.9	14.3	218	70.8	10.5	Extraction 3520 > 3510*
Nitrobenzene	12	315	80.3	15.5	175	76.8	10.8	Extraction 3520 > 3510*
Nitrobenzene-d5 (surrogate)	7	229	81.9	16.2	142	76.0	11.8	Extraction 3520 > 3510*
N-Nitrosodimethylamine	9	238	73.6	27.0	132	67.9	14.1	Extraction 3520 > 3510*
N-Nitrosodi-n-propylamine	17	418	80.4	16.0	360	80.9	15.7	Extraction 3520 > 3510*
N-Nitrosodiphenylamine	9	198	81.3	12.0	173	79.6	10.6	
Pentachlorophenol	17	410	81.3	18.3	322	77.6	13.3	Extraction 3520 > 3510*
Phenanthrene	12	331	84.8	11.1	307	84.0	11.0	Extraction 3520 > 3510*; High spiking > low**
Phenol	17	416	62.2	27.1	234	55.9	19.9	Extraction 3520 > 3510*
Phenol-d5/d6 (surrogate)	7	209	65.6	29.0	77	62.6	18.0	High spiking > low**
Pyrene	17	431	88.2	14.2	409	88.6	13.2	
Terphenyl-d14 (surrogate)	7	227	88.8	23.1	180	92.7	14.0	Extraction 3510 > 3520*; High spiking > low**

* Controlled for higher spiking level.

** Controlled for extraction method 3520.

Results for Method 8270C – Solid Matrix								
			All Data		Outliers Removed			
Analyte	# of Labs	Total # of points	Mean	Std Dev.	Total # of points	Mean	Std Dev.	ANOVA
1,2,4-Trichlorobenzene	17	408	76.6	12.6	312	77.4	11.2	Extraction 3540 > 3550 (SS)
1,2-Dichlorobenzene	11	261	73.2	12.6	131	70.9	8.7	Extraction 3540 > 3550 (SS)
1,3-Dichlorobenzene	10	259	72.4	13.8	166	69.7	10.3	Extraction 3540 > 3550 (SS)
1,4-Dichlorobenzene	16	435	70.9	13.2	398	69.0	11.4	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS)
2,4,5-Trichlorophenol	11	258	82.3	13.1	154	80.1	10.4	Extraction 3540 > 3550
2,4,6-Tribromophenol (surrogate)	7	189	85.1	17.1	152	80.9	15.1	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS); Extraction 3550 > 3540 (Ottawa)
2,4,6-Trichlorophenol	12	282	80.9	13.2	177	76.3	11.0	Extraction 3540 > 3550 (SS)
2,4-Dichlorophenol	12	281	79.1	12.8	185	77.2	10.9	Extraction 3540 > 3550 (SS)
2,4-Dimethylphenol	11	272	67.6	14.2	184	67.3	11.9	SS > Ottawa (extraction 3550)*; Extraction 3550 > 3540 (Ottawa)
2,4-Dinitrophenol	12	282	74.1	24.6	173	72.6	20.0	Extraction 3540 > 3550 (SS)
2,4-Dinitrotoluene	17	409	84.0	15.3	297	82.0	11.4	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS)
2,6-Dinitrotoluene	11	271	83.1	12.3	197	80.2	10.7	Extraction 3540 > 3550 (SS)
2-Chloronaphthalene	11	271	78.2	12.1	197	75.2	9.9	Extraction 3540 > 3550 (SS)
2-Chlorophenol	17	409	75.2	12.6	313	74.7	10.3	Extraction 3540 > 3550 (SS)
2-Fluorobiphenyl (surrogate)	7	203	76.2	11.9	167	72.8	10.0	Extraction 3540 > 3550 (SS)
2-Fluorophenol (surrogate)	7	193	73.1	13.6	135	70.6	11.1	Extraction 3540 > 3550 (SS); Extraction 3550 > 3540 (Ottawa)
2-Methylnaphthalene	11	256	78.8	13.2	135	77.3	10.0	Extraction 3540 > 3550
2-Methylphenol	10	251	74.0	11.5	215	71.7	10.6	Extraction 3540 > 3550 (SS)
2-Nitroaniline	9	240	83.0	13.5	168	81.0	12.2	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550
2-Nitrophenol	10	259	77.8	14.1	146	76.2	11.5	Extraction 3540 > 3550 (SS)
3,3'-Dichlorobenzidine	11	270	70.7	22.2	166	68.9	19.6	SS > Ottawa (extraction 3550)*; Extraction 3540 > 3550 (SS)
3-Methylphenol/4-Methylphenol	10	249	76.3	13.1	196	73.9	10.9	Extraction 3540 > 3550 (SS)
3-Nitroaniline	9	240	71.5	17.8	156	68.8	13.8	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS)
4,6-Dinitro-2-methylphenol	10	259	85.0	21.2	186	83.1	18.0	SS > Ottawa (extraction 3550)*; Extraction 3540 > 3550 (SS)
4-Bromophenyl phenyl ether	11	271	83.7	13.1	170	81.7	11.8	Extraction 3540 > 3550 (SS)
4-Chloro-3-methylphenol	16	400	81.9	12.2	304	79.5	11.1	Extraction 3540 > 3550
4-Chloroaniline	9	239	58.3	20.7	155	51.0	14.2	Extraction 3540 > 3550
4-Chlorophenyl phenyl ether	11	271	81.9	11.4	190	79.6	10.7	Extraction 3540 > 3550

Results for Method 8270C – Solid Matrix								
			All Data		Outliers Removed			
Analyte	# of Labs	Total # of points	Mean	Std Dev.	Total # of points	Mean	Std Dev.	ANOVA
4-Nitroaniline	9	240	77.3	15.0	204	73.6	13.1	Extraction 3540 > 3550
4-Nitrophenol	17	409	81.3	21.8	353	77.0	20.2	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550
Acenaphthene	17	422	78.5	10.7	386	77.3	10.3	Extraction 3540 > 3550
Acenaphthylene	12	290	78.2	10.9	209	75.7	10.4	Extraction 3540 > 3550
Anthracene	12	290	81.4	9.4	241	79.9	9.0	Extraction 3540 > 3550
Benz(a)anthracene	11	282	83.8	10.5	201	81.6	9.8	Extraction 3540 > 3550 (SS)
Benzo(a)pyrene	12	293	83.8	11.6	233	80.7	10.3	Extraction 3540 > 3550 (SS)
Benzo(b)fluoranthene	12	293	83.0	12.4	229	79.7	11.4	Extraction 3540 > 3550 (SS)
Benzo(g,h,i)perylene	12	281	84.4	16.7	230	81.8	14.7	Extraction 3540 > 3550 (SS)
Benzo(k)fluoranthene	11	280	85.5	13.0	244	83.8	12.9	Extraction 3540 > 3550 (SS)
Benzoic acid	8	177	58.2	24.7	140	55.7	18.7	SS > Ottawa (extraction 3550)*; Extraction 3550 > 3540
Benzyl alcohol	8	187	78.7	22.0	117	70.9	17.4	
Bis(2-chlorethoxy)methane	11	270	77.7	14.5	197	75.5	10.9	Extraction 3540 > 3550 (SS)
Bis(2-chloroethyl) ether	11	271	73.5	12.7	196	71.1	11.2	Extraction 3540 > 3550 (SS)
Bis(2-chloroisopropyl) ether	10	254	73.1	21.7	178	68.4	15.7	Extraction 3540 > 3550 (SS); Extraction 3550 > 3540 (Ottawa)
Bis(2-ethylhexyl) phthalate	11	274	86.5	13.6	257	87.4	13.3	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS)
Butyl benzyl phthalate	11	271	86.1	13.5	186	86.4	12.3	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS)
Carbazole	8	184	81.6	12.9	167	80.4	12.3	Extraction 3540 > 3550 (Ottawa)
Chrysene	12	293	83.8	10.9	238	82.6	9.9	Extraction 3540 > 3550 (SS)
Dibenz(a,h)anthracene	11	285	84.9	14.2	249	82.9	13.9	Extraction 3540 > 3550 (SS)
Dibenzofuran	11	253	78.6	12.6	155	77.1	8.8	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550
Diethyl phthalate	11	274	83.7	11.2	165	82.2	10.6	Extraction 3540 > 3550
Dimethyl phthalate	11	271	81.7	11.3	197	79.6	10.2	Extraction 3540 > 3550 (SS)
Di-n-butyl phthalate	11	265	84.3	10.3	198	83.2	9.1	Extraction 3540 > 3550 (SS)
Di-n-octyl phthalate	11	271	87.8	16.1	249	86.4	15.2	Extraction 3540 > 3550 (SS)
Fluoranthene	12	290	83.0	10.4	271	83.9	10.1	Extraction 3540 > 3550
Fluorene	12	289	81.1	10.7	195	78.3	9.8	Extraction 3540 > 3550
Hexachlorobenzene	11	275	83.2	11.8	222	82.5	11.7	Extraction 3540 > 3550 (SS)
Hexachlorobutadiene	11	275	78.0	14.9	162	78.2	12.9	Extraction 3540 > 3550

Results for Method 8270C – Solid Matrix								
			All Data		Outliers Removed			
Analyte	# of Labs	Total # of points	Mean	Std Dev.	Total # of points	Mean	Std Dev.	ANOVA
Hexachloroethane	11	272	73.3	15.0	199	71.9	12.6	Extraction 3540 > 3550 (SS)
Indeno(1,2,3-cd)pyrene	12	293	83.9	15.2	229	79.7	13.8	Extraction 3540 > 3550
Isophorone	11	271	78.5	13.8	158	77.0	11.4	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS)
Naphthalene	12	293	74.7	11.8	237	73.4	11.1	Extraction 3540 > 3550
Nitrobenzene	11	273	76.1	14.0	168	77.2	11.9	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS)
Nitrobenzene-d5 (surrogate)	7	202	74.6	14.7	166	69.5	10.7	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550
N-Nitrosodimethylamine	7	177	74.7	22.8	140	66.1	15.9	
N-Nitrosodi-n-propylamine	17	409	77.1	15.2	301	76.8	12.3	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550
N-Nitrosodiphenylamine	9	192	83.7	12.3	170	82.4	11.1	SS > Ottawa (extraction 3550)*
Pentachlorophenol	17	412	75.5	19.3	322	71.9	15.6	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS)
Phenanthrene	12	292	82.0	10.1	211	80.1	10.0	Extraction 3540 > 3550
Phenol	17	408	73.8	13.3	330	69.7	10.2	Extraction 3540 > 3550 (SS)
Phenol-d5/d6 (surrogate)	7	193	75.4	14.4	155	71.0	10.2	Extraction 3540 > 3550 (SS)
Pyrene	17	420	85.0	13.1	400	84.4	12.8	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS)
Terphenyl-d14 (surrogate)	7	206	83.9	18.0	129	78.8	15.5	Ottawa > SS (extraction 3550)*; Extraction 3540 > 3550 (SS); Extraction 3550 > 3540 (Ottawa)

Notes: Ottawa = Ottawa sand; SS = sodium sulfate
* Controlled for lower spiking level.

Results for Method 8151A – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
2,4,5-T	10	215	81.9	21.7	174	83.0	17.4	Low spiking > high
2,4,5-TP	11	222	86.4	22.9	122	84.4	16.5	Narrow GC column > wide; Low spiking > high
2,4-D	11	235	81.4	23.8	135	80.3	18.6	
2,4-DB	8	160	95.6	26.5	140	91.6	25.7	No cleanup > method 8151
Dichloroprop	7	140	91.5	18.6	98	92.0	11.9	Narrow GC column > wide; No cleanup > method 8151
Dalapon	7	138	68.5	29.2	77	59.6	12.6	No cleanup > method 8151

Results for Method 8151A – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Dicamba	8	153	85.2	17.7	112	86.6	12.9	
Dinoseb	8	150	62.4	24.3	70	44.8	16.3	No cleanup > method 8151
MCPA	7	138	97.7	25.8	78	89.8	15.7	No cleanup > method 8151

Results for Method 8151A – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
2,4,5-T	8	191	85.0	31.8	105	95.1	21.7	Wide GC column > narrow
2,4,5-TP	8	196	88.1	26.2	136	92.5	15.7	Wide GC column > narrow
2,4-D	8	188	88.5	35.0	108	86.3	24.0	
2,4-DB	6	105	112.6	61.9	86	114.0	31.7	Wide GC column > narrow
Dicamba	6	111	88.0	16.4	91	92.7	12.5	High spiking > low
Dichloroprop	5	91	103.1	18.2	52	93.1	12.3	Wide GC column > narrow
Dinoseb	6	115	71.2	62.5	53	57.3	50.9	Wide GC column > narrow

Results for Method 8310 – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Acenaphthene	7	135	73.4	15.8	103	69.5	11.5	Extraction 3510 > 3520
Acenaphthylene	7	135	76.6	13.2	104	73.7	13.2	Extraction 3510 > 3520
Anthracene	8	155	84.3	13.3	91	76.9	11.8	Extraction 3510 > 3520
Benzo(g,h,i)perylene	6	115	82.5	14.4	71	76.6	14.1	Low spiking > high
Benzo(b)fluoranthene	7	135	89.3	14.0	71	81.6	10.3	Extraction 3510 > 3520
Benzo(k)fluoranthene	8	155	87.0	11.4	71	79.3	10.4	Extraction 3510 > 3520; Low injection vol > high*

Results for Method 8310 – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Benzo(a)anthracene	8	145	88.6	11.5	61	80.7	10.5	Extraction 3510 > 3520
Benzo(a)pyrene	8	155	82.1	12.6	131	79.4	11.3	Low injection vol > high*
Chrysene	8	155	88.9	11.4	91	83.3	10.9	
Dibenzo(a,h)anthracene	7	135	74.0	19.8	91	64.2	15.5	
Fluoranthene	7	135	88.1	13.3	91	82.1	11.3	Extraction 3510 > 3520
Fluorene	7	135	77.1	14.4	80	69.1	11.3	
Indeno(1,2,3-cd)pyrene	8	155	87.1	12.0	71	79.6	10.8	
Naphthalene	7	135	70.4	13.5	103	68.1	11.8	Low spiking > high
Phenanthrene	8	155	85.5	13.7	100	80.2	13.4	Extraction 3510 > 3520
Pyrene	8	155	84.9	11.4	111	80.0	9.3	

* Injection volume: Low = 0.01 – 0.06 mL; High = 5 – 20 mL

Results for Method 8310 – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Acenaphthene	8	150	90.0	80.0	94	70.6	12.4	SS > Ottawa; High spiking > medium*
Acenaphthylene	8	150	83.0	19.6	94	72.8	13.4	High spiking > medium; Low spiking > medium*
Anthracene	8	158	85.8	17.9	74	86.1	13.0	
Benzo(g,h,i)perylene	8	158	86.2	17.7	55	84.6	10.4	
Benzo(b)fluoranthene	8	158	89.8	14.6	75	89.3	10.7	
Benzo(k)fluoranthene	8	158	88.0	16.7	93	84.5	12.2	
Benzo(a)anthracene	8	148	89.2	16.9	64	78.0	9.3	
Benzo(a)pyrene	8	158	82.8	19.3	94	86.5	15.4	
Chrysene	8	158	90.3	14.7	94	87.0	10.7	
Dibenzo(a,h)anthracene	8	158	83.9	18.3	95	80.8	11.4	
Fluoranthene	8	158	90.4	17.9	114	88.2	15.6	

Results for Method 8310 – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Fluorene	8	158	82.4	19.8	94	76.4	10.1	
Indeno(1,2,3-cd)pyrene	8	158	90.7	16.7	121	94.9	13.0	SS > Ottawa
Naphthalene	8	153	82.4	23.3	74	79.9	10.5	High spiking > medium*
Phenanthrene	8	158	89.5	17.3	94	91.2	11.5	
Pyrene	8	158	86.8	15.8	94	82.3	11.0	Medium spiking > low*

Notes: Ottawa = Ottawa sand; SS = sodium sulfate

* Spiking level: High = 1,330 – 10,050 ug/kg; Medium = 100 – 999 ug/kg; Low = 3.33 – 99 ug/kg

Results for Method 8330 – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
1,3,5-Trinitrobenzene	9	158	86.8	29.0	131	82.2	29.1	SPE > Salting out; High spiking > low
1,3-Dinitrobenzene	9	157	86.1	29.3	125	81.1	25.6	SPE > Salting out; High spiking > low
2,4,6-Trinitrotoluene (TNT)	9	158	85.9	27.1	108	77.4	28.8	SPE > Salting out; High spiking > low
2,4-Dinitrotoluene	9	157	86.7	25.6	118	83.0	23.6	SPE > Salting out; High spiking > low
2,6-Dinitrotoluene	9	157	86.8	26.2	127	82.2	26.8	
2-Amino-4,6-dinitrotoluene	7	105	95.6	14.5	48	86.7	9.3	High spiking > low
2-Nitrotoluene	9	154	83.9	21.9	104	76.6	22.2	High spiking > low
3-Nitrotoluene	9	153	85.7	21.7	117	80.4	21.8	High spiking > low
4-Amino-2,6,-dinitroluene	7	109	98.9	18.1	92	96.3	13.6	
4-Nitrotoluene	9	153	84.9	21.2	117	79.8	21.2	SPE > Salting out; High spiking > low
Hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX)	8	137	93.1	23.5	108	88.1	16.0	SPE > Salting out
Methyl-2,4,6-trinitrophenylnitramine (Tetryl)	7	123	89.2	24.4	107	85.1	22.8	
Nitrobenzene	9	161	83.9	23.6	132	79.5	23.5	High spiking > low
Octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (HMX)	8	136	91.7	17.0	87	89.4	14.0	

Results for Method 8330 – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
1,3,5-Trinitrobenzene	8	212	94.9	22.2	169	95.1	20.3	
1,3-Dinitrobenzene	8	209	96.6	22.6	159	101.5	7.6	SS > Ottawa
2,4-Dinitrotoluene	8	212	98.6	23.5	169	98.4	20.8	
2,6-Dinitrotoluene	8	207	96.6	23.9	157	99.8	7.4	
2,4,6-Trinitrotoluene (TNT)	8	212	94.6	24.8	192	95.1	25.9	Acetonitrile extraction > ultrasonic
2-Amino-4,6-dinitrotoluene	8	169	101.3	10.1	134	101.7	7.3	SS > Ottawa
2-Nitrotoluene	8	208	95.5	21.5	185	97.2	19.5	
3-Nitrotoluene	8	206	94.5	22.9	204	95.4	21.1	Acetonitrile extraction > ultrasonic; SS > Ottawa
4-Amino-2,6-dinitrotoluene	8	166	102.7	14.0	113	101.5	7.4	
4-Nitrotoluene	8	207	96.0	21.9	197	100.6	7.8	
Hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX)	8	175	100.3	14.0	154	103.2	10.3	
Methyl-2,4,6-trinitrophenylnitramine (Tetryl)	8	171	79.7	24.0	170	80.2	23.3	
Nitrobenzene	8	211	94.5	21.9	167	96.3	19.2	Acetonitrile extraction > ultrasonic
Octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (HMX)	8	172	101.2	11.1	132	100.0	8.5	

Note: Ottawa = Ottawa sand; SS = sodium sulfate

Results for Method 8081A – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
4,4'-DDD	11	215	92.5	20.3	137	88.1	20.4	Narrow GC column > wide
4,4'-DDE	11	215	90.5	21.3	176	86.7	17.8	
4,4'-DDT	14	278	94.6	16.4	186	92.5	15.0	
Aldrin	14	288	84.4	18.9	268	82.8	18.6	Extraction 3520 > 3510; Narrow GC column > wide; High spiking > low
alpha-BHC	11	223	90.2	20.3	140	94.1	11.4	

Results for Method 8081A – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
alpha-Chlordane	9	185	92.5	13.3	142	93.1	10.0	Extraction 3510 > 3520
beta-BHC	11	223	92.4	22.3	160	96.1	10.0	Extraction 3510 > 3520
Decachlorbiphenyl (surrogate)	8	170	76.9	23.1	109	83.3	17.2	
delta-BHC	11	223	90.5	22.4	122	90.9	15.0	
Dieldrin	14	288	95.0	17.6	186	95.5	11.0	Extraction 3510 > 3520
Endosulfan I	9	186	81.5	20.8	58	80.1	10.4	
Endosulfan II	10	206	85.2	20.8	93	79.2	17.1	
Endosulfan sulfate	9	186	94.2	21.5	93	95.8	13.9	Extraction 3510 > 3520
Endrin	14	288	97.5	22.1	184	95.2	13.0	
Endrin aldehyde	10	206	89.7	20.1	164	96.4	13.6	
Endrin ketone	7	150	97.4	16.3	79	102.1	8.2	Extraction 3510 > 3520; Wide GC column > narrow
gamma-BHC	13	258	86.7	19.6	168	81.9	18.3	
gamma-Chlordane	3	186	91.5	12.8	165	93.8	10.7	
Heptachlor	14	288	85.7	16.7	247	86.6	14.8	Narrow GC column > wide
Heptachlor epoxide	10	208	92.5	17.6	145	96.4	11.5	Extraction 3510 > 3520
Methoxychlor	10	208	100.6	17.9	187	103.0	15.5	Extraction 3510 > 3520; Wide GC column > narrow
TCMX (surrogate)	9	190	78.8	23.5	130	81.4	18.8	Narrow GC column > wide; High spiking > low

Results for Method 8081A – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
4,4'-DDD	11	238	94.6	19.4	89	81.3	17.9	
4,4'-DDE	11	237	93.4	19.8	167	97.1	9.7	
4,4'-DDT	14	295	95.7	18.1	222	92.3	15.8	Narrow GC column > wide
Aldrin	14	303	95.3	21.3	182	93.3	15.6	
alpha-BHC	11	248	91.7	17.8	159	93.4	10.5	
alpha-Chlordane	8	188	97.3	16.3	89	92.1	9.7	

Results for Method 8081A – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
beta-BHC	11	248	93.7	18.9	159	94.5	10.7	
Decachlorobiphenyl (surrogate)	8	191	114.8	77.8	150	93.9	12.6	
delta-BHC	11	248	91.2	23.8	158	93.6	12.3	
Dieldrin	13	283	94.0	19.3	191	96.0	9.7	
Endosulfan I	9	208	85.2	26.1	109	73.7	19.8	
Endosulfan II	10	227	87.1	24.0	158	88.9	17.3	High spiking > low
Endosulfan sulfate	9	207	96.2	19.9	138	98.6	12.2	High spiking > low
Endrin	14	303	96.6	21.3	191	96.9	12.1	Low spiking > high
Endrin aldehyde	10	228	88.4	24.8	138	92.0	18.4	
Endrin ketone	7	178	98.4	15.5	129	99.7	11.3	
gamma-BHC	13	274	89.5	17.6	183	90.5	10.7	
gamma-Chlordane	2	188	96.4	14.8	139	96.4	10.0	
Heptachlor epoxide	10	227	95.1	18.4	157	98.0	10.6	High spiking > low
Heptachlor	14	305	93.8	18.4	234	95.6	14.9	Sodium sulfate > Ottawa sand; High spiking > low
Methoxychlor	9	207	102.6	22.6	158	100.0	14.2	
TCMX (surrogate)	9	210	106.6	48.5	150	96.6	9.1	Low spiking > high

Results for Method 8082 – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Aroclor 1016	12	241	88.1	21.4	181	84.6	19.8	
Aroclor 1260	13	261	90.6	19.8	180	87.5	19.2	
Decachlorobiphenyl (surrogate)	6	121	82.2	27.5	99	87.5	15.1	

Results for Method 8082 – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Aroclor 1016	12	236	92.2	21.0	174	89.5	16.1	
Aroclor 1260	13	256	97.6	56.6	194	96.0	11.6	
Decachlorobiphenyl (surrogate)	6	121	87.7	17.5	81	91.4	11.2	

Results for Methods 6010B and 7470A – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Aluminum	12	248	98.3	5.6	206	97.2	4.6	Extraction 3010 > 3005
Antimony	11	227	98.3	4.1	207	98.0	4.1	
Arsenic	13	259	99.4	5.0	205	97.9	4.3	High spiking > low
Barium	13	265	99.7	4.4	204	99.4	3.8	
Beryllium	12	246	100.1	4.4	207	99.2	4.0	Extraction 3005 > 3010
Cadmium	13	259	100.1	4.5	227	99.5	4.2	
Calcium	13	260	99.8	4.9	189	98.4	3.8	
Chromium	13	266	100.3	4.5	206	99.9	4.1	
Cobalt	12	240	99.3	3.8	198	98.7	3.1	
Copper	13	265	98.6	3.7	243	99.0	3.4	
Iron	13	263	102.3	7.3	188	101.6	4.0	Extraction 3010 > 3005
Lead	12	247	99.9	4.6	209	98.9	4.0	High spiking > low
Magnesium	13	258	99.3	3.9	208	98.4	3.6	
Manganese	12	247	100.2	4.0	167	100.1	3.9	
Mercury	12	224	100.5	5.4	210	100.2	5.0	
Molybdenum	10	192	97.4	5.4	118	94.9	5.2	Extraction 3005 > 3010
Nickel	13	264	100.6	4.5	244	100.2	4.4	High spiking > low
Potassium	13	261	96.7	11.0	171	97.7	4.3	
Selenium	13	260	99.6	6.2	206	98.1	6.0	High spiking > low; Extraction 3005 > 3010
Silver	13	266	97.1	9.8	149	97.3	5.3	

Results for Methods 6010B and 7470A – Water Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Sodium	13	261	102.0	47.3	259	99.1	4.0	
Thallium	12	223	98.0	4.3	167	97.1	3.8	High spiking > low
Vanadium	11	230	99.6	4.0	170	99.4	4.0	
Zinc	13	266	100.5	6.2	168	99.7	4.5	

Results for Methods 6010B and 7471A – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Aluminum	12	216	96.4	5.6	155	95.1	5.5	
Antimony	11	230	96.7	7.2	189	96.1	4.7	
Arsenic	12	253	96.3	6.3	188	95.1	3.9	
Barium	12	250	100.1	6.8	138	98.4	3.4	
Beryllium	11	231	98.6	4.8	169	99.1	3.5	
Cadmium	12	252	96.9	5.5	202	96.8	4.4	
Calcium	11	204	98.1	5.3	160	96.6	4.1	Low spiking > high
Chromium	12	250	100.0	5.3	180	98.7	4.5	
Cobalt	11	231	97.7	4.3	191	97.8	4.1	
Copper	12	244	98.3	5.0	158	96.9	3.1	
Iron	12	227	102.2	8.4	142	100.3	4.2	Low spiking > high
Lead	11	233	96.0	4.4	183	94.9	4.1	Low spiking > high; Extraction 3050 > 3051
Magnesium	11	212	97.2	4.0	141	96.5	3.3	
Manganese	11	213	99.6	4.9	130	97.4	4.0	
Mercury	12	240	100.3	6.2	238	100.3	5.9	
Molybdenum	9	140	96.8	5.2	103	95.5	5.2	
Nickel	12	241	98.7	4.4	170	97.5	3.9	Low spiking > high
Potassium	10	181	93.8	6.8	94	95.7	4.1	
Selenium	12	249	93.2	8.0	139	92.8	4.3	Extraction 3051 > 3050

Results for Methods 6010B and 7471A – Solid Matrix								
		All Data			Outliers Removed			
Analyte	# of Labs	Total # of Points	Mean	Std Dev.	Total # of Points	Mean	Std Dev.	ANOVA
Silver	12	250	95.4	10.6	168	96.4	7.2	Low spiking > high
Sodium	11	199	97.5	5.3	149	95.6	4.4	Low spiking > high
Thallium	11	220	95.1	4.6	190	94.5	4.2	
Vanadium	11	231	99.3	4.6	141	98.7	3.4	
Zinc	12	244	98.6	7.1	133	95.2	5.1	

This page intentionally left blank.

ATTACHMENT 1
PILOT STUDY DATA COLLECTION INSTRUCTIONS

Department of Defense
Environmental Data Quality Workgroup
Laboratory Control Sample (LCS) Study Data Submittal Instructions

Please submit electronically all the LCS results for **SW-846 Method 8270C** (see target analyte list below) from the most recent **thirty days, with a minimum of 20 results**. If your lab generates less than 20 results in a month, please extend the time period until 20 data sets can be retrieved. Equivalent data sets are requested for **both solid and water matrices**.

LCS samples must be from batches that passed initial calibration verification (ICV) and continuing calibration verification (CCV) tests. The LCS sample should still be provided if it passed the ICV and CCV tests but is outside your laboratory's LCS limits.

The following is the information required from those who wish to contribute to the LCS study. All the fields are required and most fields are either followed by the required format of the data or a list of acceptable values to be chosen from. If an option for a field is not listed, enter a value in the same form or format as the listed values. The lab-specific information is only required once while the detail file must be repeated for the entire analyte list for every data set being submitted.

Data may be submitted as a Microsoft Excel file or a text delimited file. A variable field length separated by the vertical bar is preferred over comma delimited since many analyte names contain commas.

Lab-Specific

1) Lab Name
2) Small Business (yes or no)
3) SIC Code (if a small business)
4) Were outlier data points removed? (yes or no)

Detail File:

1) Sample Number (if not unique within the data set, please include Time Analyzed (6))
2) SW-846 Method (only **8270C** for this initial pilot study)
3) Matrix (solid or water) (may also submit two separate files, if clearly identified)
4) Date Extracted (**/**/1999)
5) Date Analyzed (**/**/1999)
6) Time Analyzed (**:**) (hour:min)
7) LCS Matrix Material:
   - teflon chips;
   - quartz beads;
   - glass beads;

- sodium sulfate;
- in-house purified solids;
- Ottawa sand; or
- water.

8) Extraction Method:
   - 3540;
   - 3541;
   - 3545;
   - 3550;
   - 3560/3561;
   - 3510;
   - 3520; or
   - 3535.
9) Cleanup Method:
   - 3610;
   - 3611;
   - 3620;
   - 3630;
   - 3640;
   - 3650;
   - 3660; or
   - 3665.
10) Type of Instrument Used (i.e., GC/MS)
11) Lab-specific Instrument ID
12) Analyte Name (see target list)
13) CAS Number or PAR Label (**data will be sorted by this field, please include**)
14) Spiking Level
15) Spiking Level Units
16) Lower In-house LCS Acceptance Limit (%)
17) Upper In-house LCS Acceptance Limit (%)
18) Measured Concentration
19) Measured Concentration Units
20) Actual Recovery (%)

ATTACHMENT 2
PHASE II DATA COLLECTION INSTRUCTIONS

Department of Defense
Environmental Data Quality Workgroup
Laboratory Control Sample (LCS) Study Data Submittal Instructions

Please submit electronically **the most recent 20 LCS results** for each of the following **SW-846 Methods: 8260B, 6010B, 7470A/7471A, 8310, 8081A, 8082, 8330, and 8151A** (see target analyte list). Equivalent data sets are requested for **both soil and water matrices**.

LCSs must be from batches that passed initial calibration verification (ICV) and continuing calibration verification (CCV) tests. The LCS should still be provided if it passed the ICV and CCV tests but is outside your laboratory's LCS limits. Do not exclude outlier data points.

The following is the information required from those who wish to contribute to the LCS study. All the fields are required and most fields are either followed by the required format of the data or a list of acceptable values to be chosen from. If an option for a field is not listed, enter a value in the same form or format as the listed values. The lab-specific information is only required once while the detail file must be repeated for the entire analyte list for every data set being submitted.

Data may be submitted as a Microsoft Excel file or a text delimited file. A variable field length separated by the vertical bar is preferred over comma delimited since many analyte names contain commas.

<u>Lab-Specific</u>
1) Lab Name
2) Small Business (yes or no)
3) SIC Code (if a small business)
4) Were outlier data points removed? (yes or no)

<u>Detail File:</u>
1) Sample Number (if not unique within the data set, please include Time Analyzed [5])
2) SW-846 Method
3) Matrix (soil or water) (may also submit two separate files, if clearly dentified)
4) Date Analyzed (**/**/2000)
5) Time Analyzed (**:**) (hour:min)
6) LCS Matrix Material:
    - teflon chips;
    - quartz beads;
    - glass beads;
    - sodium sulfate;
    - in-house purified soils;

- Ottawa sand; or water.

7) Preparation (Extraction or Digestion) Method:

Analytical Method:	6010B	7470A/ 7471A	8260B	8081A	8082	8151A	8310	8330
Preparation Method:	3005 3010 3015 3020 3050 3051 3052	7470A 7471A 7471A alt.-autoclave	5030 5035 Direct injection	3510 3520 3535 3540 3541 3545 3550	3510 3520 3535 3540 3545 3550	Ultrasonic Shaker Separatory funnel	3510 3520 3540 3541 3545 3550 3561	8330: Salting out (*filtered or unfiltered*) Direct injection (*Acetonitrile or Methanol*) Acetonitrile extraction

8) Extraction Solvent
8081/8082 - solids:
- Hexane:acetone
- Methylene chloride:acetone

9) Is alkaline hydrolysis required? (yes or no) (for 8151A only)

10) Type of esterification? (for 8151A only):
- Diazomethane
- Pentafluorobenzyl Bromide

11) Cleanup Method:

Analytical Method:	6010B	7470A/ 7471A	8260B	8081A	8082	8151A	8310	8330
Cleanup Method:	None specified	None specified	Not applicable	3610 3620 3630 3640 3660	3620 3630 3640 3660 3665	8151	3610 3611 3630 3640 3650	None specified

12) Type of Instrument Used (i.e., GC/MS, HPLC, ICP, etc.)

13) Instrument Configuration (for 8151A, 8081A, and 8082 only)
- primary column with confirmation column
- dual column

14) Type of GC Column (for 8151A, 8081A, and 8082 only)
- Narrow bore
- Wide bore

15) Injection volume (for 8310 and 8260B only)
8260B:
- 5 mL
- 25 mL

16) Purge temperature (for 8260B only):
- ambient
- 40 degrees C
- Other

17) Analyte Name (see attached target analyte list)
18) CAS Number (**data will be sorted by this field, please include**)
19) Spiking Level
20) Spiking Level Units
21) Lower In-house LCS Acceptance Limit (%)
22) Upper In-house LCS Acceptance Limit (%)
23) Actual Recovery (%)

This page intentionally left blank.

ATTACHMENT 3

METHDOLOGY FOR ESTABLISHING DOD-WIDE LABORATORY CONTROL SAMPLE TARGET ACCEPTANCE LIMITS (OCTOBER 1999)

This page intentionally left blank.

FINAL

METHODOLOGY FOR ESTABLISHING DOD-WIDE LABORATORY CONTROL SAMPLE TARGET ACCEPTANCE LIMITS

Submitted to:

DOD Quality Assurance Authors Task Action Team
Environmental Data Quality Workgroup

Submitted by:

Versar, Inc.
6850 Versar Center
Springfield, Virginia 22151

Under Contract No.:
N00174-96-D-0001/0065
Subcontract C048-98-D-18
Delivery Order 3

October 13, 1999

TABLE OF CONTENTS

METHODOLOGY FOR ESTABLISHING DOD-WIDE LABORATORY CONTROL SAMPLE TARGET ACCEPTANCE LIMITS

1.0 Purpose

This paper describes the strategy to develop standardized DOD-wide method specific acceptance limits for laboratory control sample (LCS) recoveries. These limits will be used to identify quantitative target windows that analytical batches processed for the U.S. Department of Defense (DOD) will be expected to achieve. These LCS acceptance limits will be included in an appendix to the Laboratory Quality Systems Manual now under development by a Quality Assurance subgroup of the Environmental Data Quality Workgroup.[1]

The purpose of this paper is to document the methodology for development of acceptance limits for LCS and foster dialogue on the approach with interested parties.

2.0 Overview

The purpose of this study is to establish standardized, routinely achievable, method-specific acceptance limits for LCS recoveries that will ensure high data quality and be applicable DOD-wide. In determining the DOD-wide LCS acceptance limits, both the measurement variability inherent in an analytical method and the inter-laboratory variability must be considered. In this study, the DOD-wide LCS acceptance limits will be determined based on the statistical confidence interval generated from the LCS data sets obtained from multiple laboratories.

The study strategy consists of three elements:

- Obtaining data sets from laboratories for each method (listed in Section 5.1), and variables within the method, for which a Target Acceptance Limit for LCS samples will be established;
- Establishment of the Target Acceptance Limit for the method (or variable within the method) using accepted statistical methodologies, including outlier analyses; and
- Reality testing of the results through comparison to method recommendations, the laboratories' own LCS limits, and experience with recoveries in proficiency testing.

A number of policy issues are posed that are not addressed by this study. Some of these policy issues concern whether the generated LCS limits will be mandatory, how data that is outside the DOD limits but within the laboratories' own limits will be viewed, and the nature of corrective actions required. These and other policy issues will be addressed at later stages in the project. This paper focuses solely on the methodology for developing DOD-wide limits.

[1] The Environmental Data Quality Workgroup (EDQW) is a four service workgroup established by Sherri Wasserman Goodman under the leadership of the U.S. Navy. The EDQW is charged with establishing policies and procedures to improve the management of environmental data throughout DOD.

3.0 Study Phases

The work will be conducted in three phases:

- Phase 1 – Exploration of the methodology and testing of the data collection approach;
- Phase 2 – Pilot testing of the full methodology on one method (SW-846 method 8270C); and
- Phase 3 – Expansion of the study project to additional methods (listed in Section 5.1).

Information from each phase will feed the subsequent phases.

Phase 1 of the project will include:

- Exploring potential sources of LCS data that may have been collected by others and will fit the needs of the project;
- Conducting exploratory discussions to ascertain interest of laboratories in contributing data;
- Creating a database for storage of data from multiple laboratories (further detail presented in Section 6.0);
- Pilot testing statistical methodologies for merging data from multiple laboratories; and
- Finalizing the information collection strategy.

Phase 2 of the project will include:

- Obtaining data from laboratories on the pilot study method (SW-846, method 8270C);
- Developing sample acceptance limits for that method; and
- Examining the generated acceptance limits, and comparison of these limits to other published limits (including method specific limits and recoveries that may have been generated for that method in association with PE samples).

Phase 3:

Phase 3 of the project will include developing LCS levels for the remaining methods. The details of Phase 3 are not spelled out here, because they are so dependent on the outcomes of Phases 1 and 2. The purpose of Phase 3 is to implement a data collection strategy based on the results of the previous two phases.

4.0 Background

The LCS acceptance limits are a statistical measure for the analytic uncertainty resulting from uncontrollable systematic and random errors inherent in an analytic method. They are used to screen measurements for avoidable human errors and instrumental failures during sample analysis. A LCS consists of an aliquot of a clean (control) matrix similar to the sample matrix spiked with standards for selected analytes. The LCS is used to verify that the laboratory can perform the indicated method in a clean matrix.

LCSs measure the percent of a known quantity of chemical injected into a clean matrix that can be seen by the analytical instrument. Typically, a laboratory establishes LCS limits annually, as a range (plus or minus a percent recovery that reflects the mean and standard deviation around that mean) of the amount of the chemical that is identified. At least one LCS is run per analytical batch after the calibration, but before the samples are run. The percent recovery for each batch is benchmarked against the pre-established limits. If the LCS recovery for a particular batch is outside the established limits for that method, then the batch results may be considered to be unacceptable, triggering corrective action as appropriate (e.g., reanalysis may be required).

According to the widely used SW-846 methods, the LCS acceptance limits are defined as the mean recovery ± 3 * the standard deviation, with the mean recovery and standard deviation generated from an LCS data set consisting of 20 data points. A common protocol for establishing laboratory-specific acceptance limits is to take the first 20 consecutive LCS sample results at the beginning of 1 year and calculate the mean recovery and standard deviation of the LCS for each analyte in the sample (U.S. EPA, 1995). The LCS acceptance limits determined are used to control the quality of sample analysis for the whole year. However, in some laboratories, the LCS acceptance limits are continuously updated whenever another set of 20 LCS samples has been analyzed. In still other laboratories, the mean and standard deviation may be calculated with an entire year's worth of data to establish the LCS limits for the following year.

5.0 Study Design

The study design addresses a variety of issues, including the methods for which DOD will calculate LCS limits, the universe of laboratories from which data will be sought, the data required from the laboratories, and the nature of the information about the LCS data sets that will be sought.

5.1 Methods of Concern

The methods for which LCS limits will initially be developed include the following SW-846 methods: 8260B (volatile organics), 8081A (pesticides), 8082 (polychlorinated biphenyls), 8151A (herbicides), 8270C (semivolatile organics), 8330 (explosives), 6010B (metals), 8310 (PAHs), and 7470/7471 (Mercury).

5.2 Obtaining data from laboratories

LCS limits will be set using recent actual LCS data from laboratories working on DOD projects that have demonstrated quality work. The initial strategy will involve providing opportunities for laboratories to voluntarily offer data for participation in the study. Data collection instructions and a description of desired data will be placed on the EDQW web-site. In addition, trade associations such as the American Council of Independent Laboratories (ACIL) will be notified that the EDQW will be accepting historic data from laboratories that perform work for DOD.

In order to provide a clear record of the quality status of the laboratories who are voluntarily contributing to the study, a list has been prepared of laboratories that represent the universe of laboratories currently in good standing for performing work for at least one of the DOD components overseeing this study. In addition to performing the methods that are the subject of this study, the laboratories on this list have passed a laboratory quality audit within the last 18 to 24 months with one of the following agencies: U.S. Navy, U.S. Air Force, U.S. Army (and U.S. Army Corps of Engineers), and Defense Logistics Agency. A total of 81 laboratories have been identified that meet the criteria for the methods that are the subject of this study.

In using the data submitted by the laboratories, the study team will first identify those laboratories that meet the criteria listed above and flag that data in the database that is created. In the data analysis methodology described in Section 7.0, those laboratories will be identified as "Group A" and will provide an initial benchmark against which LCS limits will be developed. A few of these laboratories will also be put into the control group ("Group B"). In addition, the distribution of the laboratories within the population of laboratories that meet the criteria above will be analyzed. If an insufficient data set is obtained from laboratories that meet the criteria, then additional data may be directly solicited from up to nine laboratories selected at random from that portion of the set of 81 laboratories that did not respond.

5.3 Data Required From Laboratories

The DOD workgroup is preparing a Target Analyte List (TAL). This TAL will be listed in the DOD Quality Systems for Laboratories Manual and will be used for a variety of purposes. For the purpose of this study, the TAL will define the specific analytes for a given method that will be the subject of the LCS study. Historic LCS data will be sought from laboratories only for those analytes. However, if the laboratory has gathered historical data on a broader array of analytes, then the study team will accept the full array of data provided by the laboratory to ACIL.[2]

The generation of a statistical confidence interval requires that the mean and standard deviation used must be derived from a data set consisting of a **minimum** of 15 data points for each variable involved so that the whole population of possible LCS values can be well represented (Taylor, 1987). A data set of 20 data points is commonly used in environmental laboratories to determine in house acceptance limits.[3]

As described in Section 4.0, however, laboratories vary in the way they set LCS limits. In order to ensure uniformity, each LCS limit set by this study will use data sets consisting of LCS results from **consecutive analytical runs from the most recent 30 days, with a minimum**

[2] The TAL list is intended to make data collection easier, not harder. The laboratory will be invited to supply data for the list of analytes that is easiest for them.

[3] SW-846 specifies that the average percent recovery and standard deviation(s) for each matrix spike compound are calculated after analysis of 15-20 matrix spike samples of the same matrix. In Quality Assurance of Chemical Measurement (Taylor, 1987), an F test is recommended for calculation of control limits. "It is recommended that each of the s values in question be based on at least 14 degrees of freedom." Fifteen are the minimum number of data points required. SW-846 recommends 15 to 20 data points.

of 20 data points, from each laboratory.[4] If a laboratory performs less than 20 analytical runs in the 30-day period, that lab would extend the time period until 20 LCS results are compiled. The LCS data submitted will be for those recent, consecutive batches that have passed initial calibration (ICV) and continuing calibration verification (CCV) tests. This will ensure that the batches represented in the study are "in control," even if individual LCS values are not within the laboratory's limits. The final LCS acceptance limits for each analyte will then be estimated based on the combined LCS data sets from many laboratories. This final data set may represent hundreds of data points depending on the total number of laboratories participating in the study and the number of LCS results submitted by each lab (e.g., X labs times 20+ data points per analyte).

Because all methods of interest are applicable to both soils and water matrices, at least two sets of acceptance limits, one for soils and the other for water, will be determined. The laboratories will be requested to submit their last 30 days worth of LCS runs for each matrix.

5.4 Information about LCS Data Sets

It is hypothesized that there may be certain variables that affect the final recovery value for the LCS. Such variables include the type of solid matrix, specific preparatory method, or spiking level. If different laboratories address these variables differently, this can lead to significantly different results.

Therefore, for each set of LCS acceptance limits to be determined, the following information on a given analytical method will be requested from the environmental laboratories along with the LCS data set:

- A full list of analytes addressed in the batch;
- Preparatory methods used;
- Description of the material used as a soil blank;
- Spiking levels of analytes in laboratory control samples;
- Cleanup methods used;
- Instruments used to generate LCS data; and
- The LCS acceptance criteria in use by the laboratory for the method and matrix associated with the LCS run supplied.

Statistical tests (described below) will be used to decide if the variables identified above significantly affect the magnitude of the LCS data points provided by the laboratories. Depending on the outcome of this analysis, the study team will determine if additional LCS limits (or collection of additional data sets) are necessary.

[4] Although some laboratories may use a year's worth of data to set in-house LCS limits, use of all of their data could bias the study toward those laboratories' results. If the combined data sets using 30 days worth of data are still dominated by a few laboratories, a weighted adjustment or a random selection of individual data sets will be used to ensure that data from a few laboratories do not dictate results.

6.0 Database

The study team is evaluating the use of a database in StatSoft's STATISTICA statistical analysis software (or MS-Access, if that is not possible) to process the LCS data requested from laboratories. Every effort will be made to collect the data from the laboratories in a common format. This database will be composed of the following three components:

- An input spreadsheet in MS-Excel whose main function is to ensure the electronic data from laboratories is consistent;

- Statistical analysis software (STATISTICA) used to compare and consolidate the data sets for a given analytic method, evaluate the quality of the LCS data, determine the nature of data distribution, and calculate the LCS recovery acceptance limits; and

- An output file for storing the calculated LCS recovery acceptance limits in both numeric and graphical form.

7.0 Pilot Study Work Plan

A pilot study will be performed using data gathered from the complete universe of laboratories identified on only one analytical method (SW-846, method 8270C). This will serve as a way to test the data quality evaluation steps that are proposed before beginning a full-scale study. A flow chart, attached to this document (Figure 1), presents the following methodology. In the following discussion the term 'data set' refers to a set of LCS values for a specific analyte from an individual laboratory.

The objectives of the pilot study are two fold: first, to determine if the chosen software tools are appropriate, and second, to determine if the approach yields the desired outcome. The first objective will be evaluated by the Project Team staff and modifications recommended for the full study as needed. The second objective will be achieved in two ways. First, the laboratories will be divided into two groups: a larger group (group A) to be taken through every step of the data quality evaluation and a smaller group (group B) to be set aside and used as a control. The data sets from group B will then be compared to the pilot study acceptance limits generated from group A. Next, a small peer-review team will be assembled as an outside check on the study methodology and the reasonableness of the acceptance limits. As part of this step, an alternative method to calculating acceptance limits, the biweight approach, will also be used. Thus two sets of acceptance limits will be generated and compared in this phase.

Three different statistical tests will be used to generate the final data set from which the acceptance limits will be calculated. Initially, analysis of variance (ANOVA) will be used to determine if LCS recoveries vary according to any of the descriptor variables (e.g., preparatory method, spiking level). If significant differences are identified, each data set will be tested for outlier data points using the Grubbs test (Section 8.3) as a way to double check that the ANOVA results were not driven by extreme values. Then the data will be subdivided into groups based

on common descriptor variables. Next, the Youden test (Section 8.4) will be used to identify outlier laboratories within any of the subgroups. Data from outlier laboratories will be flagged and not included in the final data set. Lastly, the entire data set will be tested for outlier data points using the Grubbs test. Data points not meeting the test requirements will be flagged and not included in calculation of the acceptance criteria from the final data set.

The LCS acceptance criteria will be based on the 99 percent (%) confidence interval for each analyte calculated using the mean and standard deviation of the final data set (Section 9.0), assuming the final data set is approximately normally distributed. At this point, the data sets in control group B will be tested against the resulting confidence intervals. If 95% of the data sets in group B are within the calculated acceptance criteria, then the group B data sets will go through the previously described steps of data quality evaluation. The data remaining in group B after the evaluation is complete will be integrated with the group A data, and a new overall confidence interval calculated. If 95% of the group B data sets are not within the calculated acceptance criteria, all steps of the statistical analysis should be reviewed and potentially revised.

For the pilot study, the biweight approach to calculate an estimate of the central tendency and spread of the distribution, developed by Karen Kafadar (Kafadar, 1982, 1983), will be run in parallel with the tests mentioned above (Section 8.5). Two of the advantages of this approach are that it does not require the identification and removal of outliers and it does not require the data to be normally distributed. The disadvantage is that computationally it is extremely complex and it is not available in commercial software. The effectiveness of the two approaches will be evaluated by comparing the acceptance criteria by both approaches.

Finally, the acceptance criteria by both approaches will be compared to participating laboratory LCS limits, PE sample LCS limits, or any other source of comparison. If the calculated LCS limits are reasonable, a decision will be made regarding the technique (biweight or traditional) to be used for the entire study. If the results are not reasonable, the entire process will be reviewed and alternative methods developed.

8.0 Data Quality Evaluation

Section 7.0 and the attached flowchart describe the overall approach to data quality evaluation that will be used in the pilot study. This approach may be adjusted as appropriate and as the methodology is proven. The basic approach is to first evaluate the shape of the data sets (e.g., testing the data points for normal distribution) and then combine the data sets if ANOVA indicates that there is no significant difference among them. Second, analyze for outlier laboratories and then outlier data points in the set of data for each analyte. Finally, the combined data set, after being tested again for normality, will be used to calculate the LCS acceptance criteria. These steps and the biweight approach are discussed below.

8.1 Distribution of Data

LCS acceptance limits will be generated based on 99% confidence intervals. This requires that each LCS data set exhibit a normal distribution. In this study, a two-step procedure will be used to test the normality of data for each LCS data set. Distribution tests will be

performed using the Statistica software. This software provides several techniques for distribution fitting. These include skewness and kurtosis as well as two goodness-of-fit tests (Kolmogorov-Smirnov and Chi-square). In addition, the ANOVA procedure includes a test for homogeneity of variance, which is analogous to a distribution-fitting test.

After the outlier tests have been performed, the normality of the final data set will be tested using the procedure just described.

8.2 One Way ANOVA Analysis

Each data set will contain not only the LCS values but also coded information pertaining to the different parameters described in Section 5.4 (e.g., preparatory method, spiking level). The data sets will be evaluated to determine if those parameters affect the LCS recovery values using one-way Analysis of Variance (ANOVA). If a significant difference is observed between data sets due to a certain parameter, the data sets will be sorted according to that parameter, and the LCS recovery limits generated separately for each parameter. For example, one set of LCS acceptance limits might need development for the spiking level of 50 parts per billion (ppb) and another set for the spiking level of 200 ppb. If a majority of the data sets are found to be non-normal, a non-parametric test can be used as an alternative to ANOVA.

8.3 Grubbs Test for Outlying Data Points

The Grubbs test will be conducted on the subgroups identified by the ANOVA analysis (if any) to determine if extreme values are driving ANOVA results. It will be used again on the entire LCS data set (after the Youden test) to identify and flag outlying data points, using the following procedure:

1. Calculate the mean and standard deviation of each LCS data set;
2. Identify minimum and maximum data points in the data set; and
3. Calculate the appropriate values of T for minimum and maximum data points:

$$T = (X_{av} - X_{min})/S \qquad \text{or} \qquad T = (X_{max} - X_{av})/S$$

Where:

X_{av}	=	Mean of the LCS data set
X_{min}	=	Minimum of the LCS data set
X_{max}	=	Maximum of the LCS data set
S	=	Standard deviation of the LCS data set

4. Select the risk factor for false rejection (e.g., 1 or 5%); and
5. Compare T with values tabulated in Appendix B (from Taylor, 1987), depending on the size of LCS data set and acceptable risk. If T is larger than the tabulated values, maximum or minimum data points will be rejected as outliers.

Dixon's Test could also be used as an alternate approach to determine outlier data points but this is more complex and will only be considered if the Grubbs test is not adequate for our purposes.

8.4 Youden Test for Outlying Laboratories

The Youden Test will be conducted to identify LCS data sets or laboratories that consistently report high or low LCS recoveries. The test ranks each data point in LCS data sets, as shown in the following tables. The rankings for all data points in each data set will be summed as cumulative scores. The cumulative scores are compared to the statistical ranges listed in Appendix A (Taylor, 1987). If the scores are not within the range, then the LCS data set is an outlier, consistently lower or higher than other LCS data sets, and should not be used in generating LCS limits. In the example below, seven laboratories reported on five samples; the range is expected to be between 8 to 32, with 95% confidence. Laboratory A is considered to provide results consistently higher than other members of the group and is an outlier.

Youden Test Example: Data Sets Collected from Seven Laboratories

	Data Points				
Laboratory	**1**	**2**	**3**	**4**	**5**
A	10.5	14.2	20.0	18.1	12.3
B	9.9	13.7	19.7	18.2	11.7
C	10.2	14.1	19.9	17.8	12.0
D	9.7	13.9	19.5	17.9	12.2
E	10.4	14.0	19.7	17.5	11.6
F	10.0	13.6	19.4	17.6	11.9
G	10.1	13.8	19.6	17.7	12.1

Youden Test Example: Rankings and Cumulative Scores for Each Laboratory

	Rankings of Data Points					Cumulative Score
Laboratory	**Data Point 1**	**Data Point 2**	**Data Point 3**	**Data Point 4**	**Data Point 5**	
A	1	1	1	2	1	6
B	6	6	3	1	6	22
C	3	2	2	4	4	15
D	7	4	6	3	2	22
E	2	3	4	7	7	23
F	5	7	7	6	5	30
G	4	5	5	5	3	22

The Youden test can only be used when the number of observations for each laboratory is equal. This may not always be the case for this project. Therefore, if a laboratory submits more than the minimum number of 20 LCS data points, 20 points will be randomly selected for the sole purpose of testing for outlier laboratories. All submitted LCS data will be used in the calculation of the acceptance criteria.

8.5 Alternative Pilot Test: Biweight Approach

The biweight approach to identifying outliers is an alternative technique to calculating the central tendency of a population and the variability of the population around the central tendency measure. The approach assigns a zero weight to very extreme values and very small weights

(e.g., 0.1) to samples that are not quite as extreme. Therefore, it does not require the removal of outliers. It utilizes a rather complex iterative approach to calculate the central tendency value starting from the median. These steps have already been programmed by an outside source) and the pilot study data will be processed by that source for this stage of the parallel evaluation.

The biweight approach is effectively a substitute for the two outlier tests described above, and, since the approach does not require a normal distribution, the normality tests are no longer necessary. The one-way ANOVA analysis is still required, however, and will be conducted for this approach in the same manner described in Section 8.2. The LCS recovery acceptance limits for this alternative approach will be generated by the central tendency and variability calculated by the biweight approach.

9.0 Generation of LCS Recovery Acceptance Limits

In this study (for the main statistical approach), to be consistent with common practice, the confidence interval rather than tolerance intervals or other statistical intervals will be used to generate LCS acceptance limits. The acceptance limits will be based on a 99% confidence interval

After the quality of LCS data sets has been examined, the final LCS data sets will be generated based on the ANOVA results (e.g., combined as one data set or split into subgroups). The mean recovery and standard deviation for the final data set (containing potentially hundreds of data points) will be calculated and used to generate LCS acceptance limits according to the following two-sided 99% confidence interval:

$$LCS\ Recovery\ Acceptance\ Limits = Mean\ Recovery \pm t * S\tan dardDeviation$$

The value for t will depend on the level of confidence desired and the number of degrees of freedom (the number of data points minus one) associated with the estimation of the standard deviation. The values for t are provided in Appendix C (Taylor, 1987). If it is determined that data has been collected from the entire population of labs that meet the defined criteria for the population, then the z (or standard normal curve) rather than the t-distribution will be used.

10.0 Assessment of Results of LCS Recovery Acceptance Limits

The LCS Recovery Acceptance Limits generated by both statistical approaches will be compared to one another. Only one final approach and methodology will be used to analyze the remaining methods.

Prior to finalizing the LCS limits that result from the chosen approach, it is desirable to compare these results to standard measures for a "reality check." Several types of standard measures can be used:

- In-house LCS acceptance limits established by and obtained from the selected laboratories;

- LCS limits established by PE providers who may cooperate in the study;
- Comparison of results from available data bases of PE samples; and
- Single or multiple laboratory method performance data published along with the method(s).

The DOD study team will review the various benchmark comparisons, as well as comments from the analytical community to establish final LCS limits.

REFERENCES

Kafadar, Karen. 1982. "Using Biweight M-Estimates in Two-Sample Problems Part 1: Symmetric Populations." Commun. Statistics – Theory Method. 11(17): 1883-1901.

Kafadar, Karen. 1983. "The Efficiency of the Biweight as a Robust Estimator of Location." J. of Research of the National Bureau of Standards. Vol. 88. No. 2.

U.S. Environmental Protection Agency. 1989. *Guidance Document on the Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities*. EPA Contract No. 68-01-7310. MRI Projection No. 8962-78-14.

Taylor, John K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers.

Figure 1
Overview of Pilot Study Methodology

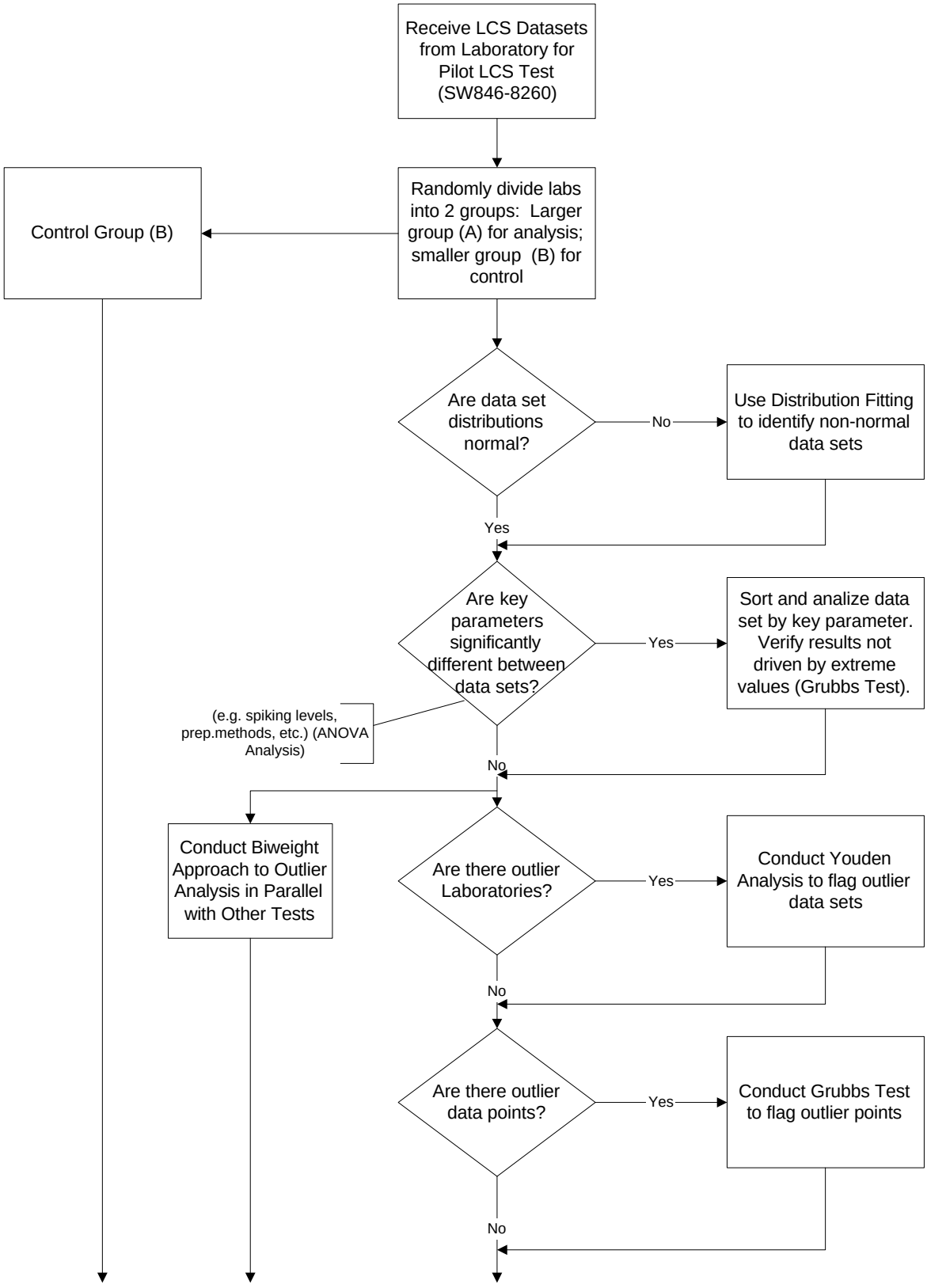

Figure 1 (page 2)
Overview of Pilot Study Methodology