United States
Environmental Protection
Agency

Targeted National Sewage Sludge Survey

Statistical Analysis Report

January 2009

U.S. Environmental Protection Agency
Office of Water (4301T)
1200 Pennsylvania Avenue, NW
Washington, DC 20460

EPA-822-R-08-018

Acknowledgments and Disclaimer

This report has been reviewed and approved for publication by the Office of Science and Technology. This report was prepared with the support of Battelle and Computer Sciences Corporation, under the direction and review of the Office of Science and Technology.

TABLE OF CONTENTS

FIGURES

TABLES

APPENDICES

EXECUTIVE SUMMARY

This document provides technical background, statistical methods, and resulting estimates of pollutant concentrations in treated sewage sludge ("biosolids") that represent Publicly Owned Treatment Works (POTWs) in the contiguous United States with flow rates of at least 1 million gallons per day (MGD). Estimates were produced using data from a national probability sample of 74 POTWs that statistically represent 3,337 POTWs that met the study criteria. This sampling effort is known as the Targeted National Sewage Sludge Survey (TNSSS). Estimates presented in this document are generally exploratory, and they provide important input to EPA's efforts to evaluate biosolids generated by the nation's POTWs. The results also may support the development of pollutant limitations, regulatory impact analysis (RIA), and aggregate risk analysis related to biosolids under Part 503 of 40 CFR.

This report presents the concentrations for 145 analytes, including metals, classicals, organics, polybrominated diphenyl ethers (PBDEs), pharmaceuticals, steroids, and hormones. For 34 of the analytes measured in this survey, including eight "target" analytes, this report discusses an in-depth statistical analysis that yielded nationally-representative estimates. For all other analytes, Appendix B.3 provides preliminary summaries and national estimates derived from the concentration data. Because EPA has not performed an in-depth statistical analysis on the 111 analytes listed in Appendix B.3, the reader should exercise caution when interpreting the preliminary summaries. If information becomes available at some later time that warrants further evaluation of these analytes, or if other analytes become the basis for any decision-making activities, EPA will perform in-depth statistical analyses and possibly revise the preliminary results.

For each of the 34 analytes, Table ES-1 presents nationally-representative estimates of the 50^{th} percentile (i.e., median) of the underlying distribution of measurements across POTWs, as well as the 90^{th}, 95^{th}, 98^{th}, and 99^{th} percentiles. Table ES-2 provides selected nationally-representative estimates of the mean and standard deviation, along with the minimum and maximum measurements that were encountered among the samples collected in this survey. For 33 of the 34 analytes (i.e., all but nitrate/nitrite), EPA's statistical approach assumed an underlying lognormal distribution for the measurements. Because lognormality was a poor fit to the observed distribution of nitrate/nitrite data, EPA used a distribution-free nonparametric approach to generate its estimates.

Table ES-1. Nationally-Representative Estimates for 34 Analytes -- Estimates Statistically Adjusted to Represent 3,337 POTWs (>1 MGD)

Analyte	Observed Values		Estimates							
			Percentiles					Summary Statistics		
	Minimum	Maximum	99th	98th	95th	90th	50th	Mean	Standard Deviation	Percent POTWs with Detected Conc
Metals (mg/kg)										
Barium	77	2,117	2,230	1,848	1,396	1,088	452	572	443	100
Beryllium	0.04	2.34	1.81	1.45	1.04	0.77	0.27	0.38	0.37	98.5
Manganese	35	14,900	9,700	6,904	4,156	2,648	540	1,165	2,231	100
Molybdenum	2.51	86.4	68.7	55.6	40.5	30.6	11.4	15.3	13.8	100
Silver*	2	195	105	82	57	42	13	20	22	100
Organics (ug/kg)										
4-Chloroaniline	51	5,900	12,013	8,288	4,762	2,912	513	1,284	2,946	74.4
Fluoranthene	45	12,000	13,173	9,112	5,256	3,226	575	1,421	3,211	89.5
Pyrene	44	14,000	15,918	10,894	6,184	3,742	634	1,654	3,981	84.9
Classicals (mg/kg)										
Nitrate/Nitrite	2	6,120	6,120	2,750	960	463	14	219	828	100
PBDEs (ng/kg)										
BDE-47 (2,2',4,4'-tetrabromodiphenyl)	73,000	5,000,000	2,650,430	2,212,077	1,688,881	1,329,167	570,448	709,174	523,791	100
BDE-99 (2,2',4,4',5-pentabromodiphenyl)	64,000	4,000,000	2,696,928	2,248,181	1,713,370	1,346,295	574,559	716,362	533,447	100
BDE-153 (2,2',4,4',5,5'-hexabromodiphenyl)	9,100	410,000	265,395	220,098	166,454	129,902	54,117	68,334	52,685	100
BDE-209 (decabromodiphenyl)	150,000	17,000,000	15,836,435	11,645,502	7,360,103	4,898,034	1,162,523	2,181,237	3,462,942	98.5
Pharmaceuticals (ug/kg)										
4-Epitetracycline (ETC)	41	4,380	8,026	5,937	3,787	2,540	620	1,135	1,741	96.0
Azithromycin	8	5,205	8,717	5,811	3,172	1,853	278	831	2,342	96.0
Carbamazepine	9	6,030	1,234	856	497	306	55	135	298	96.0
Cimetidine*	4	8,330	19,128	10,975	4,789	2,294	171	1,332	10,314	89.9

Table ES-1 (cont.)

Analyte	Observed Values		Estimates							
			Percentiles					Summary Statistics		
	Minimum	Maximum	99th	98th	95th	90th	50th	Mean	Standard Deviation	Percent POTWs with Detected Conc
Pharmaceuticals (ug/kg) (cont.)										
Ciprofloxacin	75	40,800	79,636	57,975	36,095	23,703	5,367	10,501	17,658	100
Diphenhydramine	37	5,730	5,255	4,021	2,696	1,891	541	871	1,101	100
Doxycycline	34	5,090	7,021	5,046	3,082	1,989	424	877	1,588	92.8
Erythromycin-Total	2	180	264	194	123	82	19	36	58	92.9
Fluoxetine	10	3,130	1,555	1,178	778	539	147	245	329	96.1
Miconazole	7	9,210	16,931	10,083	4,652	2,341	207	1,239	7,311	95.8
Ofloxacin	25	58,100	85,562	57,929	32,363	19,304	3,113	8,573	21,998	98.5
Tetracycline (TC)	38	5,270	10,042	7,250	4,458	2,895	630	1,278	2,255	97.5
Triclocarban	187	441,000	276,708	205,043	131,079	88,120	21,677	39,433	59,924	100
Triclosan	334	133,000	197,288	124,176	62,217	33,693	3,862	16,097	65,135	92.4
Steroids and Hormones (ug/kg)										
Beta Stigmastanol	3,440	1,330,000	1,651,188	1,123,256	632,009	379,365	62,547	168,079	419,232	98.5
Campesterol	2,840	524,000	842,112	598,919	360,119	229,283	46,547	100,879	193,964	100
Cholestanol	3,860	4,590,000	7,874,368	5,071,045	2,629,149	1,467,636	187,244	680,046	2,374,369	100
Cholesterol	2,340	5,390,000	13,376,891	8,538,884	4,369,111	2,410,541	295,092	1,129,268	4,171,366	96.9
Coprostanol	7,720	43,700,000	57,794,254	35,060,035	16,626,022	8,574,467	827,108	4,366,714	22,636,715	100
Epicoprostanol	868	6,030,000	25,579,800	13,441,281	5,143,938	2,193,143	108,028	1,702,708	26,783,520	98.5
Stigmasterol	455	568,500	4,606,900	2,646,615	1,157,099	555,217	41,513	321,199	2,464,383	90.1

* The calculation of these estimates excludes one sample whose concentration was considered a statistical outlier (silver) or whose concentration was missing (cimetidine).

1.0: INTRODUCTION

Biosolids are the nutrient-rich solid, semisolid, or liquid organic materials that result from the treatment of domestic wastewater by municipal wastewater treatment plants, also known as Publicly Owned Treatment Works (POTWs). Local municipalities typically decide how best to manage the treated sewage sludge ("biosolids") that their POTWs generate, such as to recycle them as a fertilizer, incinerate them, or bury them in a landfill. The U.S. Environmental Protection Agency (EPA) is responsible for providing the public with educational information, based on the best science, on the safe recycling and disposal of these biosolids. Furthermore, Section 405(d) of the Clean Water Act requires EPA to identify and regulate toxic pollutants that may be present in biosolids at levels that may negatively impact public health and the environment.

In 1988, EPA conducted the National Sewage Sludge Survey (NSSS) to obtain information on pollutant levels in treated biosolids (USEPA, 1992). EPA used information collected in this survey when promulgating Round 1 of regulations in 1993, which established standards for the final use and disposal of biosolids. EPA completed Round 2 of regulations, which focused on land-applied biosolids containing dioxin and dioxin-like compounds, in October 2003.

Following these first two rounds of regulations, EPA performed a screening assessment of chemical pollutants in biosolids. From this assessment, EPA identified a subset of pollutants for possible regulation. However, additional data were needed for these pollutants. In addition, EPA and other organizations, such as the National Research Council (NRC, 2002), recognized the need for collecting data on other non-regulated analytes that had not been previously assessed. Examples of such analytes included polybrominated diphenyl ethers (PBDEs), pharmaceuticals, steroids, and hormones. To obtain these data, EPA initiated a new survey called the Targeted National Sewage Sludge Survey (TNSSS). In this survey, EPA collected physical samples of biosolids from a statistically representative subset of the nation's POTWs and analyzed these samples for a series of environmental pollutants and contaminants.

This report presents statistical methodology and evaluations related to the data collected in the TNSSS. For selected analytes, it provides estimates of concentrations in biosolids that are representative of the nation's largest 3,337 POTWs. A companion report describes the sampling and chemical analyses (USEPA, 2008: *Sampling and Analysis Report for the Targeted National Sewage Sludge Survey*, EPA-822-R-08-016).

This report about the statistical methodology has six chapters. This first chapter provides background and organization of the report. Chapter 2 provides a summary of the selected analytes for the in-depth statistical evaluation, the target population, and selection of facilities for the survey. Chapter 3 provides an overview of the statistical methodology used to derive the survey weights and model the concentration data. (Appendix C provides the statistical equations and derivations.) Chapter 4 presents the results of the statistical analyses. Finally, Chapter 5 presents a summary of the results and conclusions with Chapter 6 providing references.

2.0: DATA COLLECTION

This chapter lists the set of analytes (i.e., pollutants and contaminants) for which EPA collected concentration measurements within the TNSSS. Then, it provides an overview of the survey design. Section 2.2 defines the target population of POTWs; and Section 2.3 describes the plan for selecting a statistically representative sample of facilities from this target population and deviations during the study. Finally, Section 2.4 describes the biosolids collection and the numbers of biosolids samples collected at each sampled facility.

2.1 Selection of Analytes for In-Depth Statistical Analysis

This report evaluates the concentrations for 145 analytes, including metals, classicals, organics, polybrominated diphenyl ethers (PBDEs), pharmaceuticals, steroids, and hormones. Table 2-1 identifies the 145 analytes measured in TNSSS, and asterisks denote the 34 analytes with in-depth statistical evaluation presented in Chapter 4. This section describes EPA criteria for selecting specific analytes for the in-depth evaluation of the statistical results. As a result of performing the in-depth evaluations, EPA verified or modified distributional assumptions and data selections as described in this document. Although EPA presents preliminary summaries of the remaining 111 analytes in Appendix B.3, it has not thoroughly reviewed these summaries to determine if distributional assumptions are appropriate or statistical outliers are present. As a consequence, the reader should exercise caution when interpreting these preliminary summaries. If information becomes available at some later time that warrants further evaluation of certain analytes, or if other analytes become the basis for any decision-making activities, EPA will perform in-depth statistical analyses and possibly revise the preliminary results at that time.

Table 2-1. Analytes With Reported Data for Biosolids Samples in the TNSSS

Metals	Aluminum Antimony Arsenic Barium* Beryllium* Boron Cadmium Calcium Chromium Cobalt	Copper Iron Lead Magnesium Manganese* Mercury Molybdenum* Nickel Phosphorus	Selenium Silver* Sodium Thallium Tin Titanium Vanadium Yttrium Zinc
Organics	2-Methylnaphthalene 4-Chloroaniline*	Benzo(a)pyrene Bis(2-ethylhexyl) phthalate	Fluoranthene* Pyrene*
Classicals (inorganic ions)	Fluoride Nitrate/Nitrite*	Water-Extractable Phosphorus	
PBDEs	BDE-28 BDE-47* BDE-66 BDE-85	BDE-99* BDE-100 BDE-138 BDE-153*	BDE-154 BDE-183 BDE-209*
Steroids and Hormones	17 Alpha-Dihydroequilin 17 Alpha-Estradiol 17 Alpha-Ethinyl-Estradiol 17 Beta-Estradiol Androstenedione Androsterone Beta Stigmastanol*	Campesterol* Cholestanol* Cholesterol* Coprostanol* Desmosterol Epicoprostanol* Equilenin	Ergosterol Estriol Estrone Norethindrone Norgestrel Progesterone Stigmasterol*

Table 2-1. (continued)

	Beta-Estradiol 3-Benzoate Beta-Sitosterol	Equilin	Testosterone
Pharmaceuticals	1,7-Dimethylxanthine 4-Epianhydrochlortetracycline (EACTC) 4-Epianhydrotetracycline (EATC) 4-Epichlortetracycline (ECTC) 4-Epioxytetracycline (EOTC) 4-Epitetracycline (ETC)* Acetaminophen Albuterol Anhydrochlortetracycline (ACTC) Anhydrotetracycline (ATC) Azithromycin* Caffeine Carbadox Carbamazepine* Cefotaxime Chlortetracycline (CTC) Cimetidine* Ciprofloxacin* Clarithromycin Clinafloxacin Cloxacillin Codeine Cotinine Dehydronifedipine	Demeclocycline Digoxigenin Digoxin Diltiazem Diphenhydramine* Doxycycline* Enrofloxacin Erythromycin-Total* Flumequine Fluoxetine* Gemfibrozil Ibuprofen Isochlortetracycline (ICTC) Lincomycin Lomefloxacin Metformin Miconazole* Minocycline Naproxen Norfloxacin Norgestimate Ofloxacin* Ormetoprim Oxacillin	Oxolinic Acid Oxytetracycline (OTC) Penicillin G Penicillin V Ranitidine Roxithromycin Sarafloxacin Sulfachloropyridazine Sulfadiazine Sulfadimethoxine Sulfamerazine Sulfamethazine Sulfamethizole Sulfamethoxazole Sulfanilamide Sulfathiazole Tetracycline (TC)* Thiabendazole Triclocarban* Triclosan* Trimethoprim Tylosin Virginiamycin Warfarin

* Analytes for which EPA performed in-depth statistical analyses of the survey data.

Eight "target" analytes were identified as an outgrowth of the December 2003 review of biosolids regulations (68 *FR* 75531) where EPA identified 15 toxic pollutants as warranting additional evaluation of potential risks using more up-to-date sludge concentration and occurrence data. For these pollutants, EPA conducted an exposure and hazard assessment using available sewage sludge data (USEPA, 2004). EPA concluded that a new survey, the TNSSS, would be needed to collect more data for eight of the analytes. Table 2-2 identifies the eight target analytes: four metals, three organics, and one classical.[1,2] Because of the importance of these analytes to the study, EPA determined that in-depth statistical analyses were appropriate.

[1] In this document, nitrate/nitrite is counted as one analyte because the chemical analysis generated one value for the combined analytes. In other documents, EPA often refers to them as two separate analytes. Nitrate and nitrite can be analyzed separately in wastewater, but only on a very short holding time (24-48 hours). This was essentially impossible for the survey without raising the shipping and analytical costs dramatically. The two species can undergo transformations back and forth in environmental samples, with nitrate reduced to nitrite under certain conditions, and nitrite oxidized to nitrate under others. It is difficult to look for the two separately in sludge since the process of leaching the sludge with water to make measurements is likely to lead to some conversion of nitrite to nitrate.

[2] EPA has used the term "classicals" to refer to nitrate/nitrite, fluoride, and water-extractable phosphorous. In other documentation for this study, EPA has referred to these analytes as "inorganic ions."

Table 2-2. The Eight Target Analytes Within the TNSSS

Metals	Barium Beryllium	Manganese Silver
Organics	4-Chloroaniline Fluoranthene	Pyrene
Classicals	Nitrate/Nitrite	

EPA selected an additional metal, molybdenum, for the in-depth statistical evaluation. EPA is currently re-evaluating this metal using updated information to determine the need for a revised numeric standard in land applied biosolids.

Four PBDEs were identified for in-depth statistical evaluation because they are most prevalent in various environmental media and acceptable human health benchmarks exist that may be useful for any future risk assessment purposes. The four PBDEs are BDE-47 (2,2',4,4'- tetrabromodiphenyl), BDE-99 2,2',4,4',5- pentabromodiphenyl), BDE-153 (2,2',4,4',5,5'-hexabromodiphenyl), and BDE-209 (decabromodophenyl).

The pharmaceuticals, steroids, and hormones (including some that are naturally occurring) were measured using new chemical analytical methods that were recently developed to monitor POTWs. The data from this method should be considered to be tentative results, pending further study of the chemical analytical method. Consequently, the statistical analyses of these analytes presented in this report should be considered to be exploratory in nature. For the group of 97 pharmaceuticals, steroids, and hormones, EPA used the survey data to estimate the percentage of POTWs, nationally, with detectable levels of the analytes.[3] EPA then conducted an in-depth statistical review of the analytes estimated to be detected at 90 percent[4] or more of the POTWs in the target population.

Table 2-3. Pharmaceuticals, Steroids, and Hormones Selected for In-Depth Statistical Evaluation

Analyte	**Percent of POTWs Nationally Estimated to have Detected Concentrations**
Azithromycin	96.0
Beta Stigmastanol	98.5
Campesterol	100
Carbamazepine	96.0
Cholestanol	100
Cholesterol	96.9
Cimetidine	89.9
Ciprofloxacin	100
Coprostanol	100
Diphenhydramine	100
Doxycycline	92.8

[3] The detection limit is generally considered to be the smallest quantity of the analyte that can be reliably measured with that particular method. Thus, detection is related to the sensitivity of the chemical analytical method, rather than a determination of the presence or absence of a particular analyte.

[4] When rounded to 90 percent.

Table 2-3. (continued)

Analyte	Percent of POTWs Nationally Estimated to have Detected Concentrations
4-Epitetracycline (ETC)	96.0
Epicoprostanol	98.5
Erythromycin-Total	92.9
Fluoxetine	96.1
Miconazole	95.8
Ofloxacin	98.5
Stigmasterol	90.1
Tetracycline (TC)	97.5
Triclocarban	100
Triclosan	92.4

2.2 Target Population and Sample Frame

The target population for the statistical analysis of biosolids data from the TNSSS consisted of all POTWs that:

- were in full operation in 2002 and/or 2004,
- had flow rates greater than 1 million gallons per day (MGD),
- employed a minimum of secondary treatment[5], and
- were located in the contiguous United States.

The target population excluded privately-owned, non-publicly owned, and Tribal facilities. The sample design in Appendix E describes EPA's rationale for focusing the survey on POTWs that met these criteria. For example, EPA excluded POTWs with less than 1 MGD because they collectively contribute only about six percent of the total flow among all POTWs in the nation, suggesting that their potential impact to the environment is minor.

A principal task in the development of a sample survey design is establishing a sample frame that identifies the entities within the target population. EPA's sample frame consisted of 3,337 facilities which it identified from one of two sources: the 2004 Clean Watersheds Needs Survey (CWNS)[6] and the 2002 version of the Permits Compliance System (PCS).[7] Within this sample frame, EPA uniquely identified all members of the target population, and each member had a known chance of being included in the sample (Kish, 1965). EPA then used statistical survey sampling techniques to select a sample of POTWs from the sample frame that would be representative of the entire target population. By applying

[5] At a POTW, all wastewater first must go through the primary treatment process, which involves screening and settling out large particles. The wastewater then moves on to the secondary treatment process, during which organic matter is removed by allowing bacteria to break down the pollutants.

[6] CWNS is a joint EPA-State survey that collected information on water quality programs and projects that may be eligible for funding under the Clean Water State Revolving Fund (CWSRF).

[7] PCS is EPA's computerized information management system that tracks permit issuance, permit limits, monitoring data, and other data pertaining to facilities regulated under the National Pollutant Discharge Elimination System (NPDES).

appropriate statistical techniques, EPA was able to generate statistical estimates from data collected from this sample which could be extrapolated to cover the entire target population.

In the original sample design for the TNSSS (see Appendix E), EPA considered a target population that differed slightly from the final definition. EPA had originally excluded 46 facilities from the target population that utilized either partial treatment (17) or wastewater treatment ponds (29) as the final stage of treatment. This was done because such facilities would have either not produced final biosolids, or the biosolids would have been too difficult to sample. Because EPA expected that the sample frame had some imperfections, as most do, EPA incorporated a slight upward adjustment to the sample size (of approximately four percent) to account for the possibility that some facilities in the sample would actually be ineligible for the target population. In fact, EPA encountered more ineligible facilities in its sample than it had initially anticipated. Thirteen facilities (16 percent) were initially ineligible because they utilized either partial treatment (Section 2.3.3) or wastewater treatment ponds (Section 2.3.4). For reasons discussed in Sections 2.3.3 and 2.3.4, EPA later redefined the target population to include facilities that utilized partial treatment or ponds as the final stage of treatment.

2.3 Selection of Facilities

From the sample frame, EPA used statistical sampling techniques to select 80 facilities from which to collect biosolids samples within the TNSSS. To ensure that the sampled facilities covered the entire range of flow rates, the sampling design divided the sample frame into three flow groups (or strata):

- Facilities exceeding 100 MGD (>100 MGD);
- Facilities exceeding 10 MGD but no higher than 100 MGD (10 to 100 MGD);
- Facilities exceeding 1 MGD but no higher than 10 MGD (1 to 10 MGD).

Most POTWs are located in the eastern part of the country. To ensure that the sample contained POTWs from all parts of the nation, EPA selected the sample according to the following two-step process:

1. The facilities were sorted within each stratum by EPA Region (e.g., Region 1, Region 2, etc.), then by state name within each Region.

2. A systematic sample of facilities was selected within each stratum. If N denotes the size of the stratum and n denotes the stratum's target sample size, systematic sampling involves dividing the stratum into n equal-sized subgroups, generating a random number k between 1 and N/n, and selecting the k^{th} facility within each of the n subgroups

The following sections describe the original sample size selected for the study, the actual number of facilities selected, and deviations from the original target population definition.

2.3.1 Number of Facilities (Sample Size). To determine an appropriate number of facilities to sample, EPA employed a sample design that was based upon the binomial distribution. The binomial distribution applies to situations in which only two outcomes are possible (e.g., yes or no), and it is of interest to estimate the percentage of the target population achieving the outcome of interest. In determining a sample size, EPA assumed that the true value of this percentage was 50 percent (e.g., pyrene was detected in the biosolids samples at 50 percent of the facilities). This assumption yields the maximum sample size necessary to achieve the following two precision criteria:

- ***Overall Criteria***: If the true value of the percentage is 50 percent, then a 90% confidence interval on the percentage is no more than +/- 10% (i.e., the estimated value will be within the

range of 40% to 60%). In other words, the sample size must ensure that the unknown percentage for the target population is estimated to within 20% of its true value with 90% confidence.

- ***Within Stratum Criteria***: If the true value of the percentage was 50 percent, then a 90% confidence interval on the percentage is no more than +/- 30% (i.e., the estimated value will be within the range of 20% to 80%). In other words, the sample size must ensure that an unknown stratum-specific percentage is estimated to within 60% of its true value with 90% confidence. (EPA recognizes that this level of precision is not sufficient to produce stratum-level estimates, but it helps ensure that certain facilities within each stratum are represented within the sample.)

To achieve both precision criteria, EPA determined that it needed to sample a minimum of 74 facilities. EPA increased this by four percent (to 80 facilities total) in anticipation of possible ineligible facilities within the sample. (The size of this upward adjustment was determined by the number of ineligible facilities encountered in the NSSS sample.) Table 2-4 summarizes the original and final sample sizes by strata. Appendix A.1 lists the plant IDs assigned to the selected facilities, the strata in which they belonged, and the geographic region in which they were located.

As is relatively common in sampling, EPA's sample contained some facilities that did not fall within the survey's initial definition of the target population. As a result, EPA replaced some facilities with others; these situations and the replacement facilities are noted within Table 2-4 and Appendix A.1. As first noted in Section 2.2, selected facilities that were outside of the target population were one of two types: facilities that conducted only partial treatment, and facilities that utilized wastewater treatment ponds as final treatment. The following subsections describe how EPA handled these two types of "ineligible" facilities within its sample.

Table 2-4. Original and Final Sample Sizes for the TNSSS

Stratum	Stratum Size	Original Sample Size	No. of Ineligibles	No. of Replace-ments	Final Sample Size	% Change from the Original Sample Size
>100 MGD	51	8	3	3	8	0%
10 to 100 MGD	543	12	0	0	12	0%
1 to 10 MGD	2,743	60	8	2	54	-10.0%
TOTAL	3,337	80	11	5	74	-7.5%

2.3.2 Final Sample Size. Field sampling involved visiting each of the selected POTWs and collecting a single sample of treated biosolids. Separate documents exist on the procedures that were used in the TNSSS to contact the selected POTWs, to arrange for a field visit, and to collect biosolids samples from these POTWs during the field visit. Although EPA had adjusted the sample size upwards for ineligible facilities selected from the sample frame (e.g., partial treatment, ponds as final treatment), early contacts with the facilities indicated that the 4% adjustment was an underestimate. To maintain the target sample size of 74, EPA made the following changes to the facility selection criteria:

- Because the objective was to obtain pollutant concentrations in final biosolids, EPA reevaluated its decision to consider partial treatment as ineligible. Attempts were made to "follow the

sludge," or to replace the facility with the facility receiving the sludge for final treatment as described in Section 2.3.3.

- Because the sample design incorporated locality (through the sorting of the sample frame), EPA reevaluated its decision to eliminate a location because the selected facility used a wastewater treatment pond or lagoon. If another facility within the same system and/or locality produced biosolids on a regular basis, EPA collected biosolids samples at that location as described in Section 2.3.4.

2.3.3 Partial Treatment Facilities. Within the original sample of 80 facilities, five utilized partial treatment. For four of these five facilities, EPA collected biosolids samples from "replacement" facilities. This section describes EPA decisions about replacements for these facilities.

Upon encountering the first facility found to use partial treatment (ID 84), EPA replaced it with another facility in the same municipality (ID 53). Compared to the original facility, the replacement facility shared the same geographic location and management, had similar hydraulic capacity, and was approximately the same size (in terms of flow). Although EPA did not select this replacement using a probabilistic sampling approach, it considers the replacement reasonable because of the similarities between the two facilities.

As it encountered a greater number of partial treatment facilities in its original sample, EPA re-evaluated its earlier decision to exclude such facilities from its target population. As a result of biosolids regulations and other factors, it is possible that partial treatment is more common than it was during the 1988 NSSS. Because the objective was to characterize final biosolids, EPA was concerned that it might be excluding a growing treatment practice, and thus, determined that it would be appropriate to "follow" the partially treated biosolids to the facility that applied full treatment, and then collect a biosolids sample from that facility. In reaching this conclusion, EPA also considered whether the co-mingling of wastes from other facilities would provide misleading results. However, co-mingling is part of the treatment process for the partially treated biosolids. Thus, because the study objective was to measure the pollutant concentrations present in the final treated biosolids at each facility, EPA concluded that concentrations at the facility applying full treatment would appropriately represent treated biosolids for each partial treatment facility. Therefore, EPA replaced three facilities employing partial treatment (IDs 81, 83, and 99) with the facilities that provided final treatment of the biosolids.

EPA did not replace one facility that performed partial treatment (ID 77). This was due to the inability to schedule sample collection at a replacement facility without incurring additional study costs.

Table 2-5 summarizes the partial treatment facilities, the replacements, and the selection criteria.

Table 2-5. Facilities in the Original Sample that Employed Partial Treatment, and Their Replacement Facilities

Original Facility		Replacement Facility		Type of Replacement Facility
ID	**Stratum**	**ID**	**Stratum**	
84	>100 MGD	53	>100 MGD	Similar facility within the same system
81	1 to 10 MGD	31	1 to 10 MGD	Facility that performed final treatment of the biosolids
83	>100 MGD	73	10 to 100 MGD	
99	>100 MGD	61	MGD>100	Facility that performed final treatment of the biosolids
77	1 to 10 MGD	--		Not replaced

2.3.4 Facilities with Wastewater Treatment Ponds. Eight of the original sample of 80 facilities utilized wastewater treatment ponds (or lagoons) in its treatment process. The bottom layer of a pond includes deposits of biosolids and supports anaerobic organisms. Facilities remove solids from their ponds only when they consider the solids to be fully treated. Because the removal process is extensive, facilities tend to perform it infrequently. For example, ID 5 removes its treated biosolids once every five years.

EPA and the wastewater treatment industry have long recognized ponds as an effective method for treating biosolids. However, in the design stage of this survey, EPA felt that coordinating sample collection with the facility's scheduled removal of biosolids from the ponds would be too difficult. For this reason, EPA excluded facilities utilizing ponds from the target population. However, EPA re-evaluated this decision after encountering a greater number of facilities with ponds than expected in the sample. Because ponds provide an effective final treatment of biosolids, EPA decided that it should attempt to collect samples of biosolids from these facilities whenever possible.

As noted in Table 2-6, EPA collected biosolids from two facilities that utilized ponds (IDs 3 and 5). This was possible because EPA was able to schedule physical sampling activities at these facilities as they were recovering the treated biosolids from the ponds. EPA replaced one facility (ID 82) with its "sister" facility (ID 21) in the same system. Compared to the original facility, the replacement facility shared the same geographic location and management, and it was of the same approximate size (in terms of flow). Although EPA did not select this replacement using a probabilistic sampling approach, it considers the replacement to be reasonable because of the similarities between the two facilities. The remaining five facilities with ponds were neither sampled nor replaced, and generally were dropped early in the study before EPA had reevaluated the eligibility requirements. It also was not feasible to return to the general area to sample a replacement. To reduce sampling costs, the contractor had grouped its site visits by region. Consequently, when EPA reevaluated the eligibility requirements, it would have increased the study costs substantially to incorporate several new sampling trips to the affected regions.

Table 2-6. Facilities from the Original Sample that Employed Wastewater Treatment Ponds, and How EPA Handled These Facilities

ID	Stratum	Final Outcome
3	1 to 10 MGD	Biosolids sampled as planned
5	10 to 100 MGD	
75	1 to 10 MGD	Excluded from the study and not replaced
76		
78		
79		
80		
82	1 to 10 MGD	Replaced by ID 21, its sister facility

2.4 Biosolids Collection

EPA collected grab samples of biosolids from the 74 facilities between August 2006 and March 2007. Because EPA collected samples during one day in a relatively short period of time, the concentration data associated with these samples allow EPA to evaluate levels in biosolids at a fixed point in time, rather than to examine trends over time.

EPA collected a single biosolids grab sample from most facilities. However, at ten facilities, EPA collected two grab samples. This was done either when the facility had more than one treatment system (implying different types of biosolids generated), or for quality assurance purposes. The following two sections describe each situation in more detail; and the third section describes how the multiple measurements were used in the statistical analyses. Table 2-7 summarizes all situations where two samples were collected at a given POTW.

Table 2-7. Summary of Situations Where Multiple Samples Were Selected at POTWs

Stratum	Final Sample Size	# POTWs Having Field Duplicates Sampled	# POTWs Having Solid and Liquid Products Sampled from Different Treatment Systems	# POTW Having Different Locations Sampled from Different Treatment Systems	Total No. POTWs With Multiple Samples Collected
MGD > 100	8	0	1 (ID 53)	1 (ID 18)	**2**
10 < MGD < 100	12	1 (ID 49)	0	0	**1**
1 < MGD < 10	54	5 (ID 2, 11, 19*, 28, and 32)	1 (ID 74)	1 (ID 48)	**7**
TOTALS	**74**	**6**	**2**	**2**	**10**

*Data for the duplicate sample for ID 19 were excluded from statistical analyses for classicals, anions, and metals (see Section 4.3.3).

2.4.1 Multiple Biosolids Treatment Systems. Most sampled facilities had only one treatment system that produced biosolids. However, four facilities in the sample utilized two treatment systems. These facilities are represented in columns 4 and 5 of Table 2-7. Because analyte concentrations have the potential to differ among biosolids generated by different systems, EPA obtained a grab sample from both systems within each of these four facilities.

Two facilities (IDs 53, 74) produced biosolids in both liquid and solid forms, both of which were sampled. The other two facilities produced biosolids in solid form from both of their systems.

2.4.2 Field Duplicates. Within its sample, EPA randomly selected eight facilities (10 percent) for the collection of duplicate grab samples. Field duplicates allow EPA to assess sampling procedures as part of its field quality assurance evaluations. EPA does not consider any decisions about field duplicates to affect the conclusions from the study.

While EPA had planned to sample a field duplicate from eight facilities, field personnel were successful in collecting field duplicate samples from six facilities (as noted in the third column of Table 2-7). Table 2-8 lists each of the eight facilities and when field duplicate samples were successfully obtained at each.

EPA had excluded one of the eight facilities (ID 75) from the study because it utilized a pond (Table 2-6). At another facility (ID 18), EPA collected samples of two types of biosolids (Section 2.4.1) rather than a field duplicate sample, without identifying another facility from which to sample a field duplicate. While the number of facilities with field duplicate samples collected was less than planned, EPA determined that this number was sufficient to meet its quality assurance objectives.

Table 2-8. Original Set of Eight Facilities Selected for Field Duplicate Sampling

ID	Stratum	Final Outcome of Field Duplicate Sampling
2	1 to 10 MGD	Duplicate was collected as planned.
11	1 to 10 MGD	
18	>100 MGD	Samples were collected from each of two treatment systems at the facility, rather than a field duplicate sample.
19	1 to 10 MGD	Duplicate was collected as planned.
28		
32		
46	10 to 100 MGD	Duplicate was collected at another facility instead (ID 49).
75	1 to 10 MGD	Excluded from the study and not replaced (Table 2-6).

EPA often grouped facilities in nearby locations into a single sampling trip for convenience. On one trip, the field team discovered that the first facility visited (ID 48) produced two types of biosolids. As a result, they collected two samples from this facility, one of each biosolids type. To allow for the additional sampling, the team used equipment that had been designated for collecting a field duplicate sample at another facility (ID 46) to be visited later in the trip. Because the field team did not expect to receive the replacement equipment until after visiting ID 46, EPA collected the field duplicate sample from another facility that the team visited after visiting ID 46. This facility (ID 49) was in the same flow group and geographic area as the originally selected facility.

2.4.3 Aggregating Data Across Multiple Samples. When a facility had two biosolids samples collected, either for quality control purposes or because the facility generated two types of biosolids products. EPA investigated whether the two data values for a given analyte could be aggregated into a single value prior to performing the data review and analysis. This was done to achieve the objective of characterizing a facility's average analyte concentration within its final treated biosolids at any single point in time.

Aggregation of field duplicate measurements within a facility: For each analyte in each chemical classification, EPA aggregated the data values within a facility when a field duplicate was collected with the regular sample (Section 2.4.2). The aggregation involved calculating a simple arithmetic average of the two data values for each analyte. If one or both samples contained non-detected levels of the given analyte, then the sample-specific detection limit entered into the calculation of this average. EPA classified the aggregated (average) result as "detected" or "not detected" as specified in Table 2-9.

Table 2-9. Determining the Classification of Aggregated Measurements as Detected or Not Detected

If the two sample data values are …	The aggregated value is calculated as the …	This result is labeled …
Both detected	Arithmetic average of the measured values	Detected
Both not-detected	Arithmetic average of the sample-specific detection limits	Not detected
A mixture of detected and not-detected samples	Arithmetic average of the measured value (for the detected sample) and sample-specific detection limit (for the not-detected sample)	Detected

Averaging the duplicates might artificially dampen the variability seen in the concentrations. The effects of averaging would be most pronounced if the duplicates were taken under different wastewater treatment conditions. Because EPA's objective was to characterize a facility's average concentration at a single point in time, the duplicates were collected on the same day from biosolids that were treated under the same process. The effects of averaging the duplicates, therefore, are minimal. EPA determined that the average concentration best represented the concentration of the POTW at that point in time.

Aggregation of measurements for multiple treatment systems within a facility: When multiple samples were collected at a facility having multiple treatment systems (Section 2.4.1), the data from these samples were aggregated for some analytes, but not for others. For analytes within the classicals, metals, and organics classifications, EPA aggregated the two measurements in the same way as field duplicates. However, for the remaining analytes (i.e., PBDEs, pharmaceuticals, steroids, and hormones), measurements often differed considerably between the two biosolids samples generated by different systems. This difference was especially apparent between solid and liquid samples. (These differences can be seen in the data listings and reviews presented in the appendices and Chapter 4.) Therefore, for the statistical analyses, EPA did not aggregate the measurements for PBDEs, pharmaceuticals, steroids, and hormones. That is, the individual sample measurements were included in the analysis as reported, rather than their average. EPA assigned one-half of the facility's assigned survey weight to each sample measurement within the statistical analysis.

3.0: OVERVIEW OF STATISTICAL METHODOLOGY

This chapter describes the statistical techniques which EPA applied to the collected sample measurements for selected analytes. Section 3.1 describes the derivation of the survey weights assigned to the selected facilities. Section 3.2 describes the distributional assumptions used to estimate the means and summary statistics. Section 3.3 describes the quality assurance aspects associated with this report.

3.1 Survey Weights

Each POTW in the sampling frame had a nonzero probability of being selected for the sample. However, as a result of the stratified sampling design, some POTWs had a different probability of being selected than others. Therefore, EPA assigned a *survey weight* to each POTW contributing a biosolids sample to the survey. The survey weight corresponds to the total number of POTWs in the sampling frame that the selected POTW represents. The sum of all survey weights equals the total number of POTWs in the sampling frame. By incorporating survey weights in the statistical analysis, EPA obtained estimates that represented the entire target population.

As a first step in assigning survey weights, EPA assigned an initial "base weight" to each stratum. Because each POTW within a stratum had an equal probability of being selected for the sample, each selected POTW in a stratum received the same base weight. Because stratum and sample sizes differed among the strata, different strata had different base weights.

Once all field sampling was completed, EPA calculated a final set of survey weights. This involved adjusting the base survey weights to account for deviation between EPA's original and final sample of POTWs (Table 2-4). The final weights should be used, rather than the base weights, when analyzing data from this survey.

As noted in Table 2-4, the final sample size for one stratum (1 to 10 MGD) differed from its original targeted sample size. The sample size was reduced by six facilities. This required an adjustment to the base survey weight for this stratum. Because EPA considered the six excluded facilities to fall within the survey's target population, this adjustment corresponded to dividing the stratum size by the actual sample size.

The replacement of five POTWs with other facilities had no effect on the final survey weights for the three strata. In each incidence that a replacement occurred, EPA determined that the biosolids sampled by the replacement POTW were representative of the biosolids generated by the POTW that it replaced. Thus, the replacement had no net impact on the sample size. To each replacement POTW, EPA assigned the survey weight associated with the stratum for the POTW that it replaced.

Table 3-1 provides the final set of survey weights for each stratum. Within a given stratum, EPA assigned the final survey weight to each POTW that contributed one or more biosolids samples to the survey.

As detailed in Section 4.4.3.1, the final statistical analysis for silver excluded one plant from the "10<MGD<100" stratum whose measurement was deemed excessively large compared to measurements from other plants. In this single situation, we adjusted the survey weight further for this stratum to account for using measurements for 11 plants rather than 12 (i.e., the final survey weight was 543/11 = 49.36).

Table 3-1. Final Set of Survey Weights

Stratum	Stratum Size	Original Sample Size	Base Weight	Final Sample Size	Final Weight[b]
>100 MGD	51	8	51/8 = 6.375	8[a]	51/8 = 6.375
10 to 100 MGD	543	12	543/12 = 45.25	12	543/12 = 45.25
1 to 10 MGD	2,743	60	2,743/60 = 45.7167	54	2,743/54 = 50.80

[a] One of the eight POTWs performed final treatment of the partially-treated biosolids of a facility originally selected from the "> 100 MGD" stratum. Thus, this replacement facility was assigned the final weight for the ">100 MGD" stratum.

[b] Assigned to each POTW within the final sample. The final weight, rather than the base weight, is utilized in all statistical analyses. For silver, the final weight for the "10 to 100 MGD" stratum was 543/11 = 49.36.

3.2 Statistical Analysis Approaches

As noted in Section 2.1, EPA applied an in-depth statistical analysis to concentration data for 34 analytes, including the survey's eight target analytes. The primary objective of the statistical analysis was to generate national estimates of the mean, standard deviation, and selected percentiles of analyte concentrations (i.e., 50^{th}, 90^{th}, 95^{th}, 98^{th}, 99^{th} percentiles). EPA used one of two statistical approaches for obtaining these national estimates: a lognormal-based approach and a nonparametric approach that did not assume any underlying distributional form in the data. To decide which statistical approach was more appropriate, EPA performed a preliminary investigation of the data as described in Section 4.2. In general, EPA selected the lognormal approach unless it was clear for a particular analyte that its data were not consistent with lognormality. An overview of each approach is given in the following subsections, with details provided in Appendix C.

3.2.1 Lognormal Approach. This was EPA's primary statistical approach. The lognormal approach assumed that, for a given analyte, average concentrations of biosolids among the nation's POTWs follow a *lognormal distribution*. This is equivalent to assuming that the log-transformed concentrations follow a normal distribution. Experience has shown that for a variety of environmental media and POTW-generated discharges, including biosolids, concentrations at a given point in time generally follow a lognormal distribution. The NSSS found that a lognormal distribution was a reasonable assumption for concentrations of target pollutants in biosolids.

The lognormal approach takes into account the stratified sample design and the survey weights assigned to each facility. The approach uses established equations associated with the lognormal distribution to obtain stratum-specific estimates of the mean, standard deviation, and percentiles. These equations are provided in Section C.1 of Appendix C.

The lognormal approach treats non-detects as observations that are censored at the sample-specific detection limit. Appendix C notes how the approach is modified in the presence of non-detects.

3.2.2 Nonparametric (Distribution Free) Approach. As an alternative to the lognormal approach, the nonparametric approach does not assume that the data follow any particular function. The estimates of the mean, standard deviation, and percentiles are determined solely from the observed data, while taking into account the stratified sample design and the survey weights. The mathematical formulas for estimating these statistics are provided in Section C.2 of Appendix C.

3.3 Quality Assurance

While performing the statistical analyses presented in this report, we adhered to all procedures specified in a formal EPA-approved Quality Assurance Project Plan. The statistical analysis approach followed an analysis plan that EPA approved prior to implementation.

For the statistical summaries and analyses presented in this report, we downloaded and utilized the final version of the TNSSS data in SAS dataset format from EPA's mainframe computer without modification. EPA had previously performed a comprehensive review of the laboratory data packages for data completeness and compliance with project and method specifications. The overall objective of the data review process was to identify any limitations apparent in the results that might affect their end use. This information was encoded in the database through a series of qualifiers. In a few instances, EPA determined that the laboratory results were so seriously flawed that no reasonable use could be made of the concentration values. In these instances, EPA excluded the concentration values, but the dataset includes the qualifiers that led to its exclusion. In all other cases, the database included the concentration values and any qualifiers. Appendix B.1.1 identifies and defines the qualifier codes used in the database.

Prior to statistical analysis, we further assessed the quality and integrity of the survey data relative to their acceptability for use for the analysis. This assessment included:

- Performing exploratory analyses in which definitions of the data variables are reviewed, their appropriate units of measure were noted, and any known relationships were assessed.
- Utilizing statistical and graphical techniques to characterize the data distribution, noting presence of missing data, identifying outliers, and influential data points, assessing the type and degree of censoring in the data and any censoring patterns, and determining deviation relative to assumed underlying distributions (i.e., lognormal).

Chapter 4 documents the outcome of these assessments.

All data analysis programs were written and tested using good programming practices. Programs were constructed to be modular, to include sufficient comments, and to include code which performs interim validation steps on the summaries and analyses.

4.0: FINDINGS FROM IN-DEPTH STATISTICAL ANALYSES

As noted in Section 2.1, EPA identified eight target analytes, along with molybdenum, 4 PBDEs, 14 pharmaceuticals, and 7 steroids and hormones, for in-depth statistical evaluation. Table 4-1 lists these 34 analytes. This chapter presents the results of the in-depth statistical analyses, including an evaluation of detection frequency (Section 4.1); statistical graphics (Section 4.2); data review of outliers and distributional assumptions (Section 4.3); national estimates of means and percentiles (Section 4.4); and comparisons to current standards for land application under 40 CFR 503 and to previous survey data (Sections 4.5 and 4.6).

Table 4-1. 34 Analytes Considered for In-Depth Statistical Analysis

Metals*	Barium Beryllium Manganese	Molybdenum Silver
Organics*	4-Chloroaniline Fluoranthene	Pyrene
Classicals*	Nitrate/Nitrite	
PBDEs	BDE-47 (2,2',4,4'- tetrabromodiphenyl) BDE-99 (2,2',4,4',5- pentabromodiphenyl)	BDE-153 (2,2',4,4',5,5'-hexabromodiphenyl) BDE-209 (decabromodiphenyl)
Pharmaceuticals	4-Epitetracycline (ETC) Azithromycin Carbamazepine Cimetidine Ciprofloxacin Diphenhydramine Doxycycline	Erythromycin-Total Fluoxetine Miconazole Ofloxacin Tetracycline (TC) Triclocarban Triclosan
Steroids and Hormones	Beta Stigmastanol Campesterol Cholestanol Cholesterol	Coprostanol Epicoprostanol Stigmasterol

* With the exception of molybdenum, the analytes listed for these chemical classes represent the survey's target analytes.

Concentration measurements for biosolids samples collected in the TNSSS are found in six SAS datasets, one for each analyte classification (shaded column of Table 4-1). For a given analyte, the SAS dataset contained one record for each biosolids sample. If a sample result met EPA's quality requirements, its data record specified either a measured value ("detected") or a sample-specific detection limit ("non-detected"). The datasets also include any qualifier flags that EPA assigned to the sample measurements in its quality assurance review of the laboratory data packages (see Section 3.3). Appendices A.2 through A.6 list the sample measurements for each analyte along with detection indicators and qualifier flags. Appendix B.1.1 defines the qualifier flags that EPA assigned during the data quality review. Because the SAS datasets reported the concentration values on a dry-weight basis, the data are directly comparable across facilities without the need to consider the percentage of solids present in each sample.

4.1 National Estimates of Detection Percentages

The detection limit is generally considered to be the smallest quantity of the analyte that can be reliably measured with that particular method. Thus, detection is related to the sensitivity of the chemical

analytical method, rather than a determination of the presence or absence of a particular analyte. EPA is sometimes interested in this aspect of the data. Consequently, for each analyte, EPA used the survey data to estimate the percentage of POTWs nationally that had detectable concentrations. Because these estimates take into account the final survey weights, they are representative of detection percentages for biosolids generated by POTWs within EPA's target population. Table 4-2 provides estimates of these percentages for each of the 34 analytes listed in Table 4-1. Table 4-3 provides these estimates for the remaining analytes.

Table 4-2 shows that all eight target analytes had detection percentages of at least 74 percent, with four achieving 100 percent. Nine of the 11 PBDEs had detection rates of 100 percent; BDE-209 had a detection rate of 98.5 percent (Table 4-2), while BDE-138 had a detection rate of 65.5 percent (Table 4-3). All of the 21 pharmaceuticals, steroids, and hormones subject to in-depth statistical analysis had estimated detection percentages of at least 90 percent (when rounded).

Table 4-2. Nationally-Representative Estimates of Detection Percentages in Biosolids for Analytes Included in the In-Depth Statistical Analysis

	Analytes	Detection Percentage
Metals	Barium	100%
	Beryllium	98.5%
	Manganese	100%
	Molybdenum	100%
	Silver	100%
Organics	4-Chloroaniline	74.4%
	Fluoranthene	89.5%
	Pyrene	84.9%
Classicals	Nitrate/Nitrite	100%
PBDEs	BDE-47	100%
	BDE-99	100%
	BDE-153	100%
	BDE-209	98.5%
Pharmaceuticals	4-Epitetracycline (ETC)	96.0%
	Azithromycin	96.0%
	Carbamazepine	96.0%
	Cimetidine	89.9%
	Ciprofloxacin	100%
	Diphenhydramine	100%
	Doxycycline	92.8%
	Erythromycin-Total	92.9%
	Fluoxetine	96.1%
	Miconazole	95.8%
	Ofloxacin	98.5%
	Tetracycline (TC)	97.5%
	Triclocarban	100%
	Triclosan	92.4%
Steroids and Hormones	Beta Stigmastanol	98.5%
	Campesterol	100%
	Cholestanol	100%
	Cholesterol	96.9%
	Coprostanol	100%
	Epicoprostanol	98.5%
	Stigmasterol	90.1%

Table 4-3. Nationally-Representative Estimates of Detection Percentages in Biosolids for Analytes Not Included in the In-Depth Statistical Analysis

	Analytes	Detection Percentage	Analytes	Detection Percentage
Metals	Aluminum	100%	Mercury	100%
	Antimony	87.8%	Nickel	100%
	Arsenic	100%	Phosphorus	100%
	Boron	97.1%	Selenium	100%
	Cadmium	100%	Sodium	100%
	Calcium	100%	Thallium	94.1%
	Chromium	100%	Tin	94.1%
	Cobalt	100%	Titanium	98.5%
	Copper	100%	Vanadium	100%
	Iron	100%	Yttrium	100%
	Lead	100%	Zinc	100%
	Magnesium	100%		
Organics	2-Methylnaphthalene	40.9%	Bis(2-ethylhexyl) phthalate	100%
	Benzo(a)pyrene	77.1%		
Classicals	Fluoride	100%	Water-Extractable Phosphorus	100%
PBDEs	BDE-28	100%	BDE-138	65.5%
	BDE-66	100%	BDE-154	100%
	BDE-85	100%	BDE-183	100%
	BDE-100	100%		
Pharmaceuticals	1,7-Dimethylxanthine	4.7%	Lomefloxacin	2.9%
	4-EACTC	0%	Metformin	6.5%
	4-EATC	38.8%	Minocycline	48.2%
	4-ECTC	1.4%	Naproxen	50.5%
	4-EOTC	11.3%	Norfloxacin	36.2%
	Acetaminophen	3.0%	Norgestimate	0%
	Albuterol	1.5%	Ormetoprim	1.5%
	ACTC	1.5%	Oxacillin	0%
	Anhydrotetracycline (ATC)	64.9%	Oxolinic Acid	0.2%
	Caffeine	47.4%	Oxytetracycline (OTC)	38.2%
	Carbadox	0%	Penicillin G	0%
	Cefotaxime	0%	Penicillin V	0%
	Chlortetracycline (CTC)	1.4%	Ranitidine	60.6%
	Clarithromycin	54.8%	Roxithromycin	3.0%
	Clinafloxacin	0%	Sarafloxacin	2.9%
	Cloxacillin	0%	Sulfachloropyridazine	3.1%
	Codeine	23.3%	Sulfadiazine	4.5%
	Cotinine	47.4%	Sulfadimethoxine	7.0%
	Dehydronifedipine	23.0%	Sulfamerazine	0.1%
	Demeclocycline	4.6%	Sulfamethazine	2.8%
	Digoxigenin	0%	Sulfamethizole	0%
	Digoxin	0%	Sulfamethoxazole	40.8%
	Diltiazem	83.1%	Sulfanilamide	12.0%
	Enrofloxacin	15.8%	Sulfathiazole	0.1%
	Flumequine	0%	Thiabendazole	71.7%
	Gemfibrozil	87.8%	Trimethoprim	27.3%
	Ibuprofen	64.4%	Tylosin	0%
	Isochlortetracycline (ICTC)	1.4%	Virginiamycin	18.9%
	Lincomycin	4.6%	Warfarin	0%

Table 4-3. Nationally-Representative Estimates of Detection Percentages in Biosolids for Analytes Not Included in the In-Depth Statistical Analysis (Continued)

	Analytes	Detection Percentage	Analytes	Detection Percentage
Steroids and Hormones	17 Alpha-Dihydroequilin	1.6%	Equilenin	1.6%
	17 Alpha-Estradiol	7.2%	Equilin	15.5%
	17 Alpha-Ethinyl-Estradiol	0%	Ergosterol	61.3%
	17 Beta-Estradiol	10.6%	Estriol	23.1%
	Androstenedione	41.5%	Estrone	75.0%
	Androsterone	65.4%	Norethindrone	6.3%
	Beta-Estradiol 3-Benzoate	24.6%	Norgestrel	4.7%
	Beta-Sitosterol	85.5%	Progesterone	20.2%
	Desmosterol	65.9%	Testosterone	21.6%

4.2 Statistical Graphics: Bar Charts and Box Plots

To provide an initial view of how the survey measurements for the 34 analytes were distributed, this section presents two types of statistical graphical data displays: bar charts and boxplots. Both display data without considering survey weights. As a result, these displays portray only the distribution of analyte concentrations among the survey samples. EPA did not apply survey weights when preparing these graphics because their purpose was to explore the distributional properties of the actual data collected. As such, they provide insight into which statistical methods should be used. Once the appropriate methods are decided upon, the survey weights were appropriately applied within each approach to obtain national estimates.

In these evaluations, EPA has assumed that non-detects have the same value as the sample-specific detection limit (or more correctly, sample-specific reporting limit). The detection limit is generally considered to be the smallest quantity of the analyte that can be reliably measured with that particular method. If the value could be measured with more specificity, it would have a value between zero and the detection limit. However, for convenience, EPA has assumed the upper bound for every non-detected value. Thus, for datasets with many non-detected values, the results from the graphical displays should be viewed with caution.

Bar charts partition the observed range of values into groups and use vertical bars to express the number (and percentage) of values within each group. The bar charts consider aggregated data within a POTW, as described in Section 2.4. To distinguish between them, the bars have gray diagonals for the measured (detected) portion and black for non-detected portion at each value range.

Within each bar chart, the horizontal axis represents the range of observed concentrations. The axis is logarithmic, with powers of 10 equally spaced along the axis. Thus, the observed shape of the bar chart is actually associated with the log-transformed data. If the bar chart resembles a symmetric, bell-shaped curve, this suggests a lognormal distribution assumption is appropriate. Because the purpose of the graphical analysis is to view the general shape of the concentrations, the vertical axis does not indicate the number of values associated with each bar.

Boxplots provide a visual summary of the key parameters of the data distribution. The boxplots display the sample-specific measurements as originally reported (i.e., without aggregation), with non-detects represented by their detection limits. One boxplot represents each analyte and is interpreted as follows:

- The length of the box represents the interquartile range of the observed log-transformed data (i.e., the distance between the 25^{th} and the 75^{th} percentiles).

- The asterisk represents the mean of the observed log-transformed data.

- The horizontal line within the box represents the median of the observed log-transformed data.

- The vertical lines (or "whiskers") extending from both ends of the box extend to the most extreme data value in that direction that falls within 1.5 interquartile ranges from the end of the box.

- The open circles denote data values that exceed 1.5 interquartile ranges from the end of the box. (While they suggest possible extreme values, they are not necessarily statistical outliers unless they fall quite far from the end of the vertical line.)

Each boxplot is plotted along a logarithmic vertical axis. Thus, like the bar charts, the distribution represented in each boxplot represents log-transformed data. If data originate from a lognormal distribution, their boxplot would have the following properties:

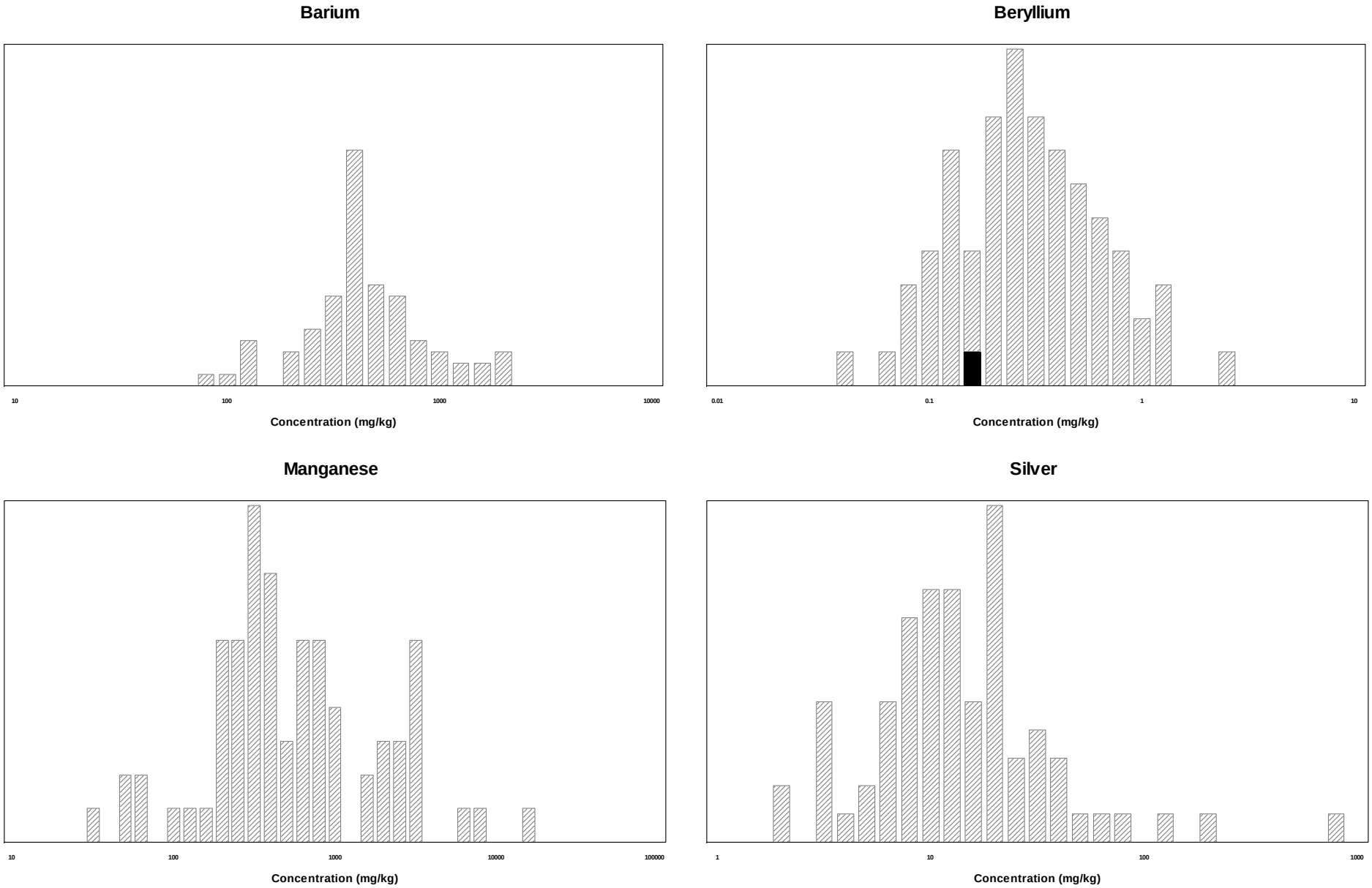
- The asterisk (mean) and horizontal line within the box (median) would be plotted on top of each other, midway through the vertical length of the box.

- The "whiskers" would be of equal length on each side of the box.

- The number of any open circles would be very limited and distributed equally on both sides of the box.

4.2.1 Metals. Figure 4-1a contains bar charts for each of the four metals among EPA's target analytes, along with molybdenum. The bar chart for silver contains an isolated bar at the far right end of the chart, suggesting a possible statistical outlier. Otherwise, each bar chart is relatively symmetric and unimodal. This suggests that a lognormal assumption is plausible for these metals.

Figure 4-1b includes boxplots for the five metals. The boxplot for barium suggests the data are tightly clustered around the mean and median. Thus, while its distribution may resemble a lognormal distribution in shape, its tails (i.e., and lower and upper ends of the curve) may be "skinnier" than what is typical for a lognormal distribution. This suggests that lognormal-based estimates for upper percentiles may be slightly lower than what would be estimated from the observed data alone. The boxplot for silver indicates that a few large values could influence the calculation of upper percentiles.



Note: Sample-specific detection limits are noted in solid black for samples associated with not detected outcomes.

Figure 4-1a. Bar Charts for Metals

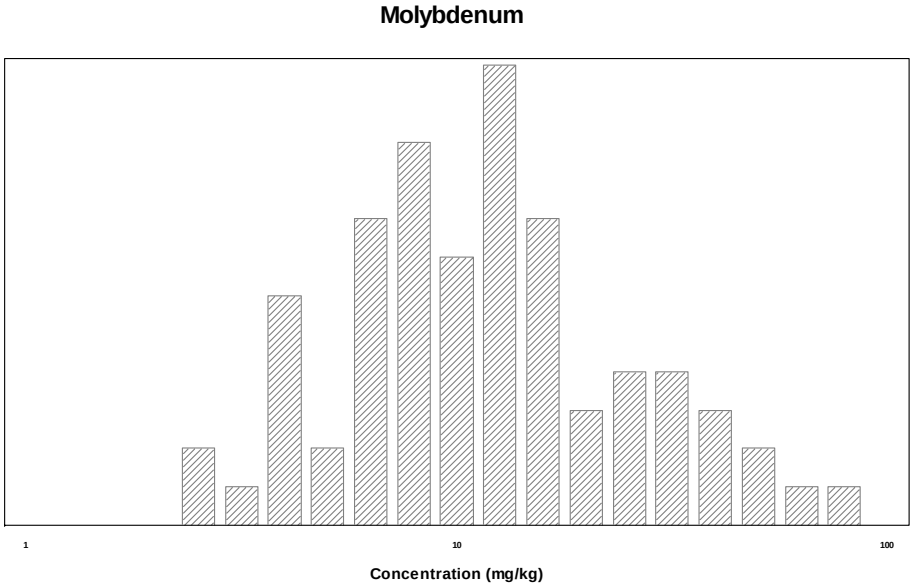


Figure 4-1a. Bar Charts for Metals (Continued)

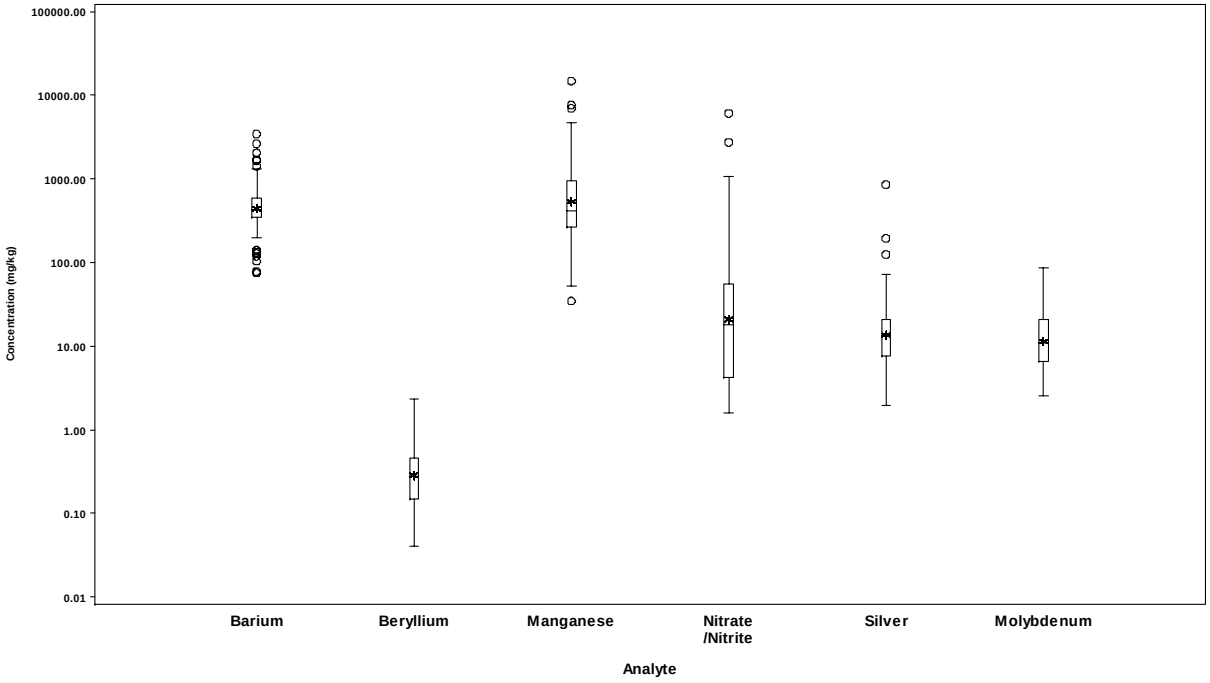
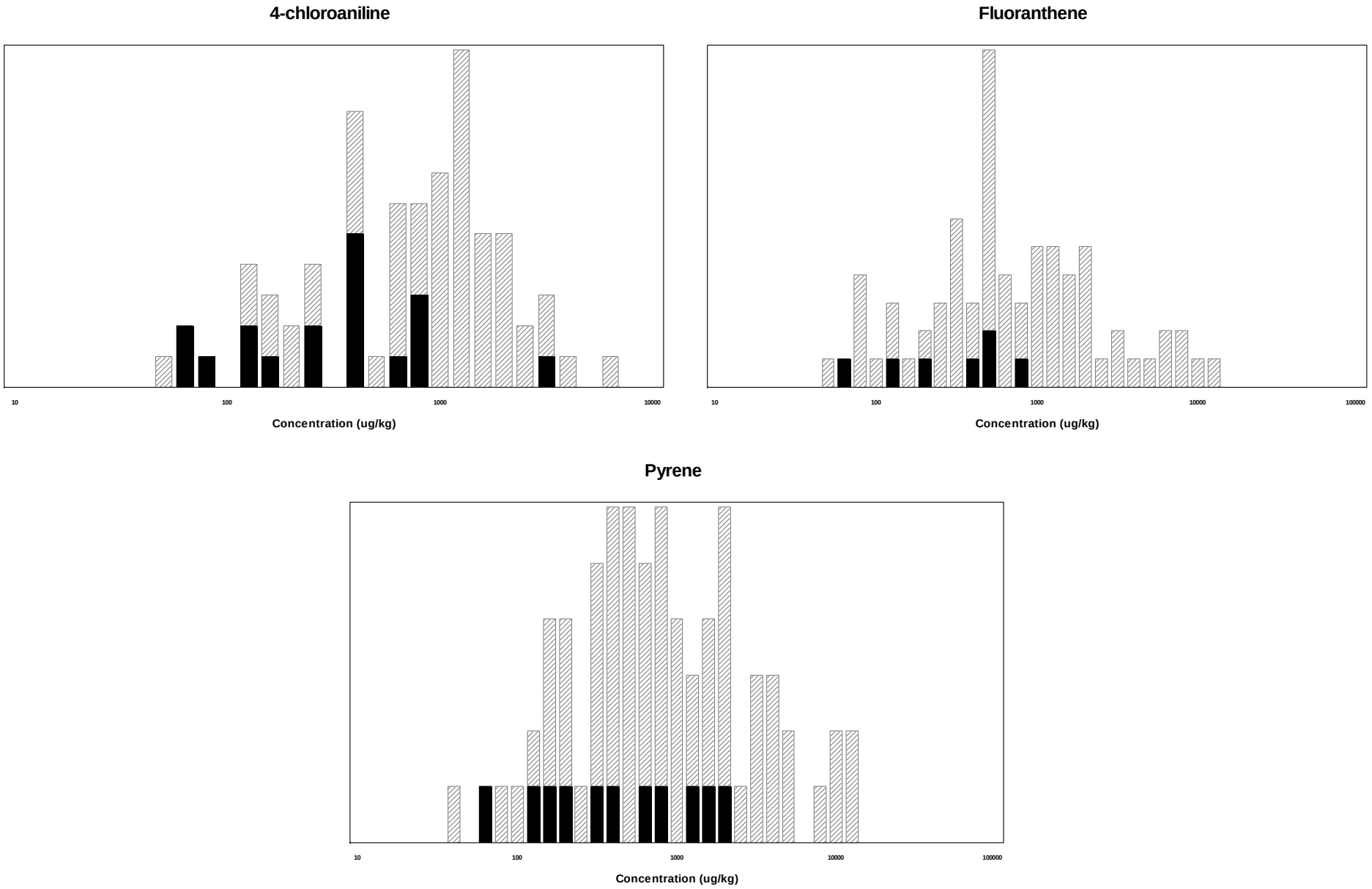


Figure 4-1b. Box Plots for Metals and Classicals

4.2.2 Organics. Figure 4-2a includes bar charts for the three organics that are among the target analytes, while Figure 4-2b contains the boxplots for these analytes. They suggest that there are no apparent outliers. However, because of the presence of a fair number of non-detected values, it is difficult to definitively draw conclusions about distributional assumptions, especially for 4-chloroaniline with 25.6 percent of the samples being classified as non-detected. Because the shape of the detected values tend to support lognormality, EPA concluded that this distributional assumption was likely to be appropriate.



Note: Sample-specific detection limits are noted in solid black for samples associated with not detected outcomes.

Figure 4-2a. Bar Charts for Organics

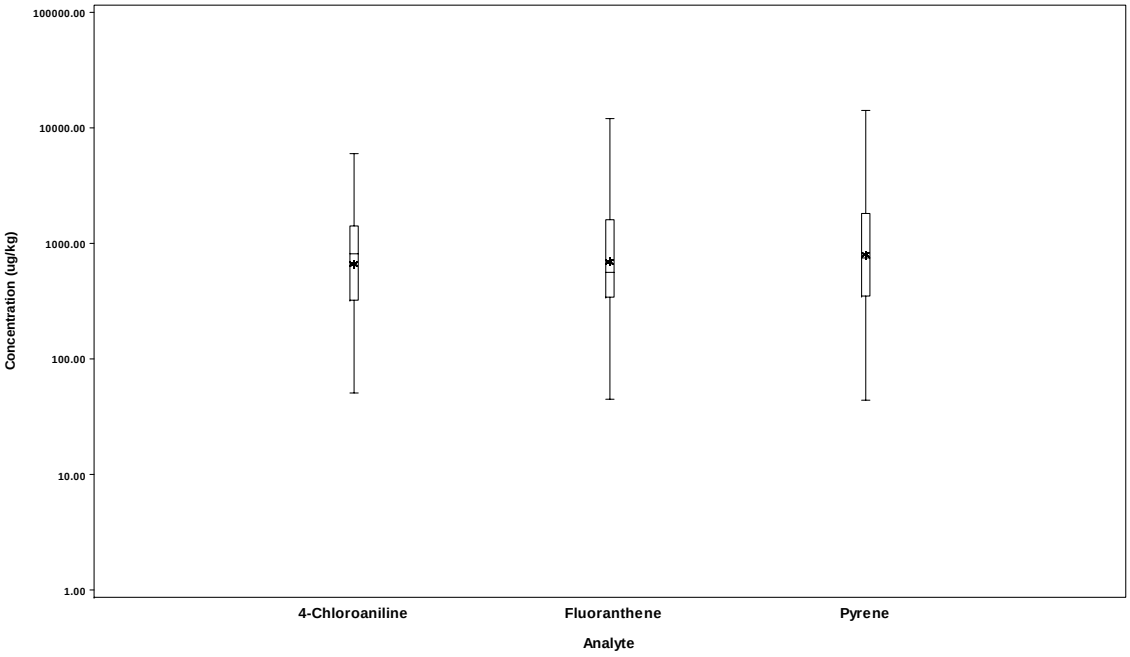


Figure 4-2b. Box Plots for Organics

4.2.3 Classicals. The target analytes include one classical compound: nitrate/nitrite. Its bar chart appears in Figure 4-3. Its boxplot was included with the metals in Figure 4-1b. These plots demonstrate how measurements for nitrate/nitrite are spread out along the entire range of measurements. Furthermore, the distribution tends to be bimodal (i.e., has two distinct peaks). A few large data values appear separated from the other values at the rightmost end of the bar chart. The boxplot suggests that the distribution is skewed, with a few large values extending beyond the other values. Thus, the nitrate/nitrite data do not appear to resemble a lognormal distribution.

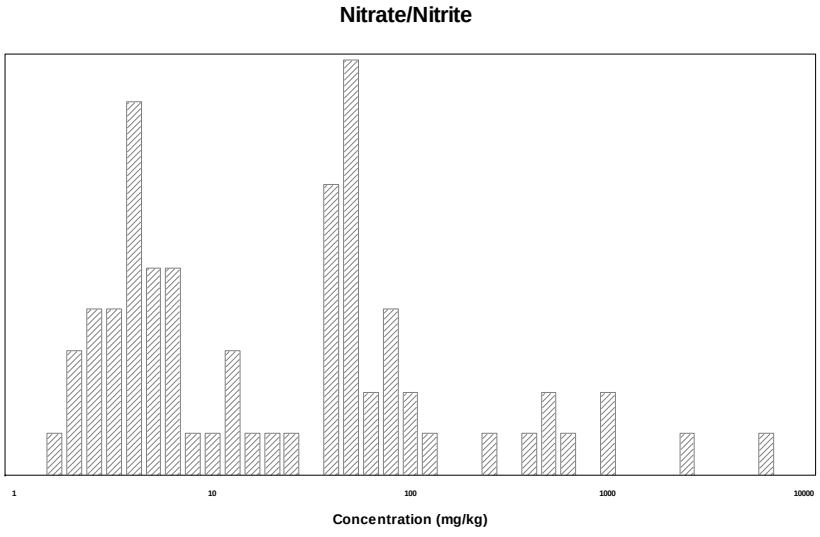
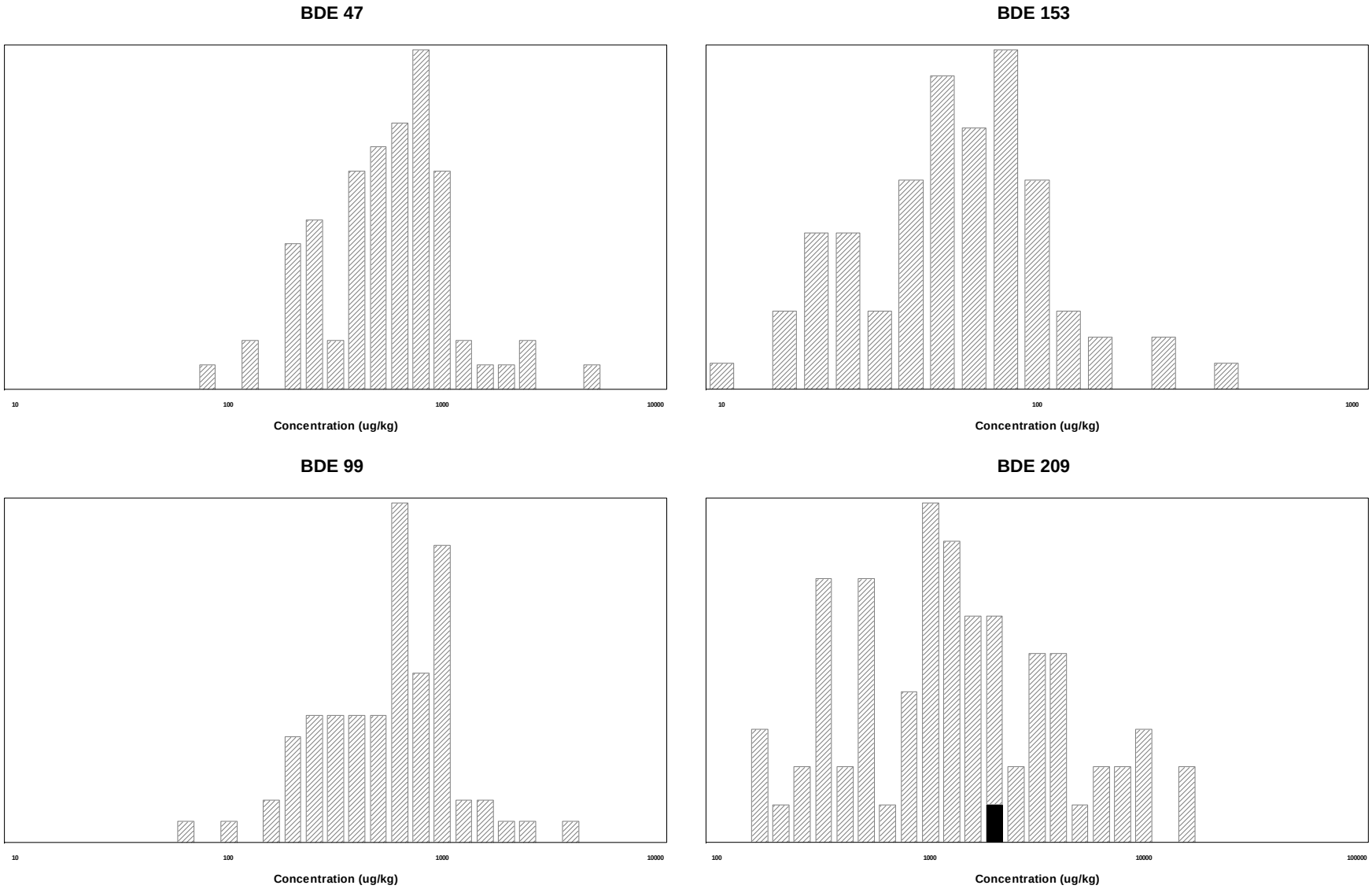


Figure 4-3. Bar Chart for Classicals (Nitrate/Nitrite)

4.2.4 PBDEs. Figure 4-4a contains bar charts for the four PBDEs on which EPA performed in-depth statistical analysis, while Figure 4-4b contains boxplots. Each bar chart has a unimodal, symmetric shape with no obvious outliers. This suggests that lognormality is a reasonable assumption for these PBDEs.

The boxplots in Figure 4-4b also provide strong evidence for lognormality. For each PBDE, the mean and median log-concentrations hold similar values, while the "whiskers" on each end of the box appear similar in length. Furthermore, measurements represented by open circles are not extreme.



Note: The bar charts summarize n=78 measurements. Each of the six POTWs having field duplicate samples collected is represented by the average of the two sample measurements. Sample-specific detection limits are noted in solid black for POTWs associated with not-detected outcomes

Figure 4-4a. Bar Charts for PBDEs

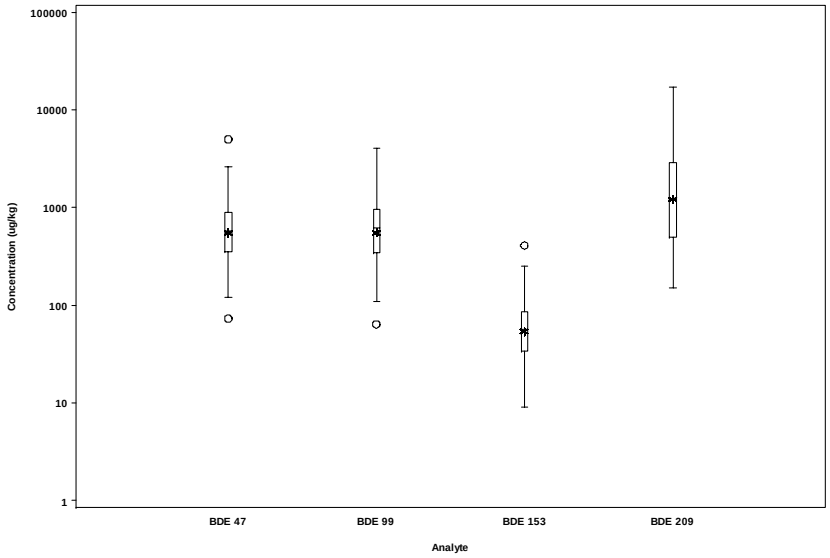


Figure 4-4b. Box Plots for PBDEs

4.2.5 Pharmaceuticals, Steroids, and Hormones. Figures 4-5a and 4-5b contain bar charts for the 14 pharmaceuticals and the seven steroids and hormones, respectively, that were included among the analytes on which in-depth statistical analyses were performed. Figure 4-5c presents boxplots for these 21 pharmaceuticals, steroids, and hormones.

Triclocarban is the only analyte having extreme measured (detected) values on the low end of its distribution range. Other analytes occasionally have one or two measurements that are high compared to the others, but they do not appear to be overly extreme. Overall, considering the shapes for the detected values in the bar charts, the lognormal assumption seems plausible for this set of 21 analytes.

In reviewing the shapes of the bar charts, EPA noted a different pattern for this set of 21 analytes than it generally had noted for the other analytes. For the other analytes, the sample-specific detection limits and detected values were consistent. In contrast, for some analytes in this set, the histograms show the sample-specific detection limits to be clustered separately from the detected values.

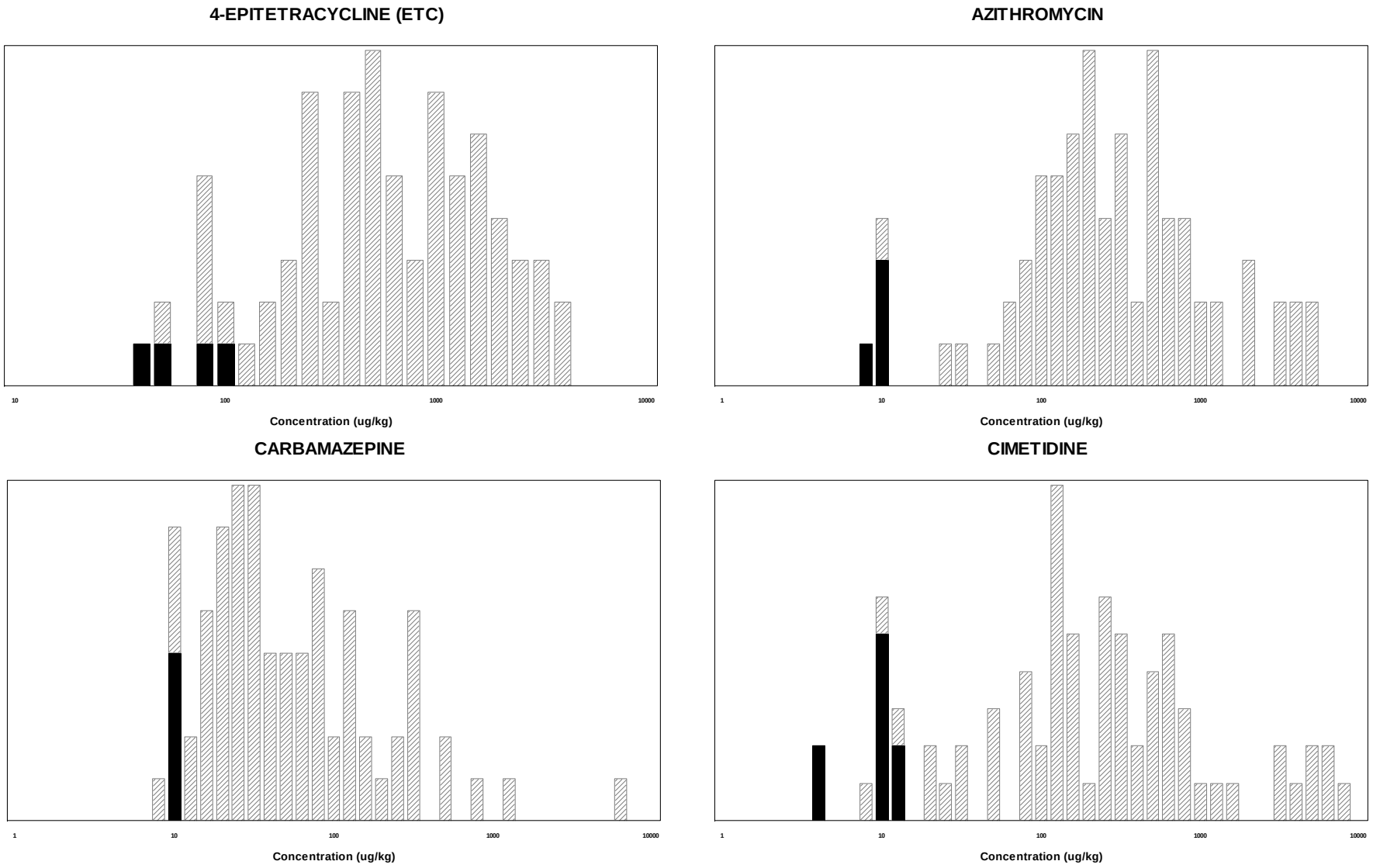


Figure 4-5a. Bar Charts for Pharmaceuticals

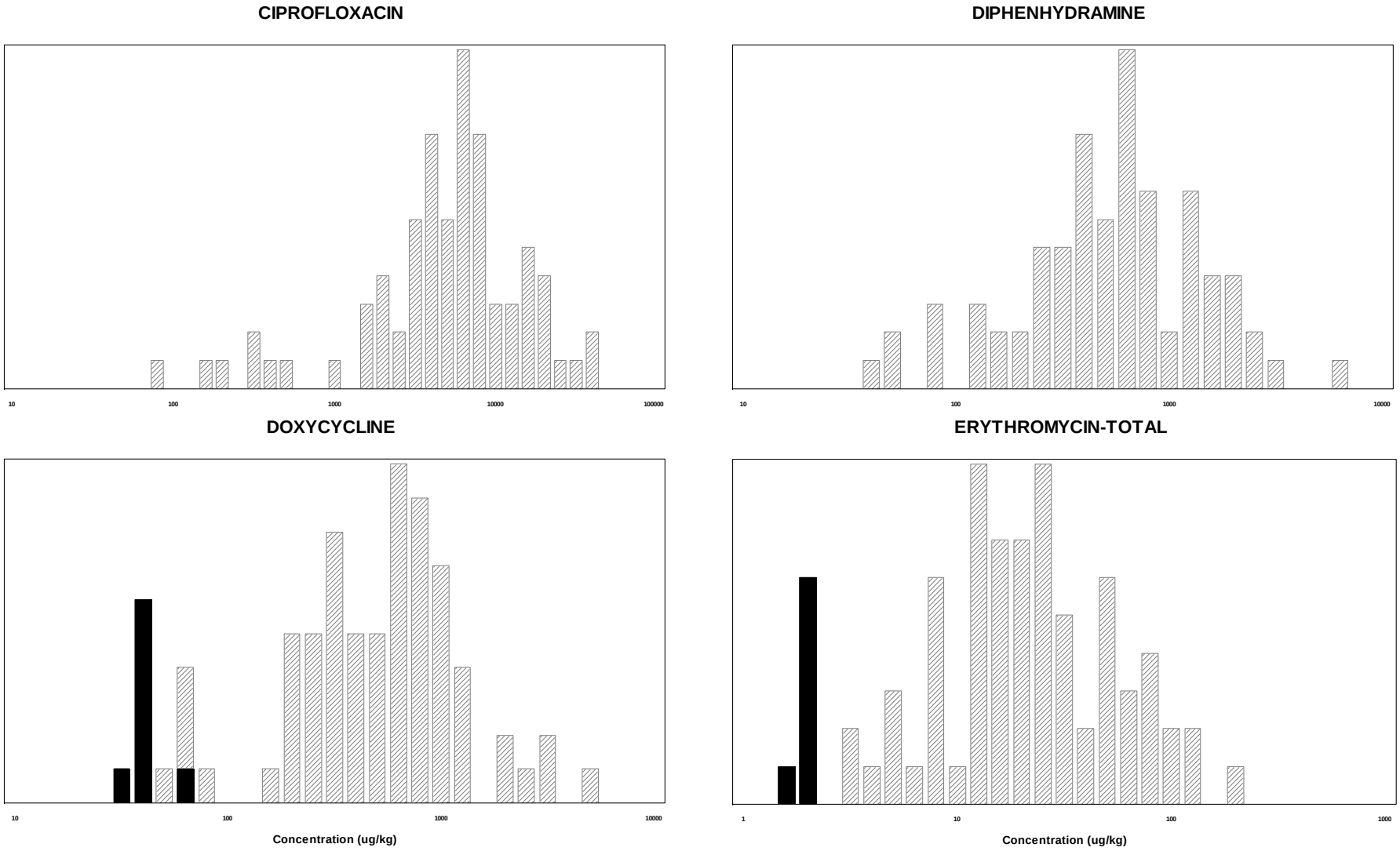


Figure 4-5a. Bar Charts for Pharmaceuticals (continued)

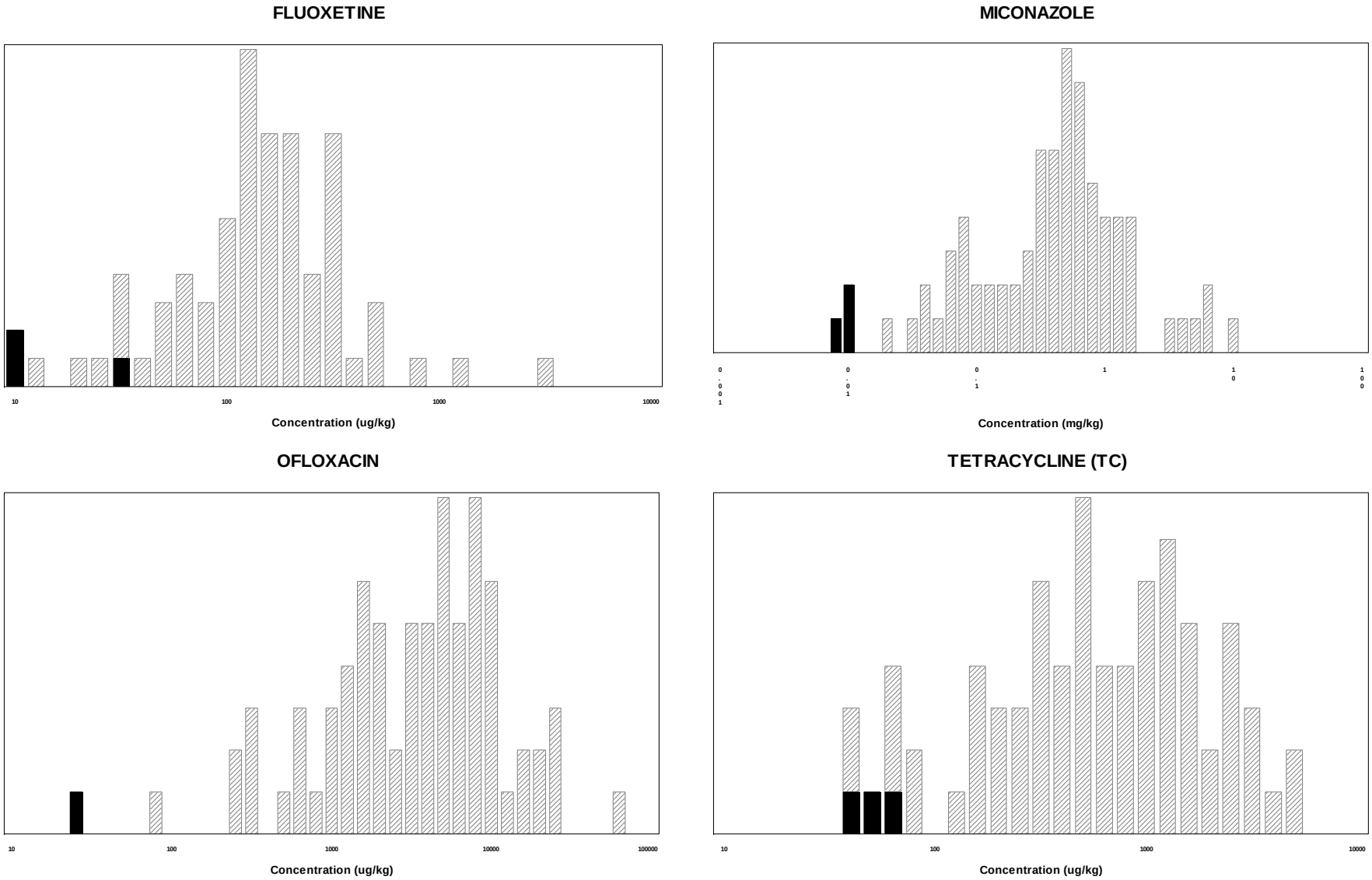


Figure 4-5a. Bar Charts for Pharmaceuticals (continued)

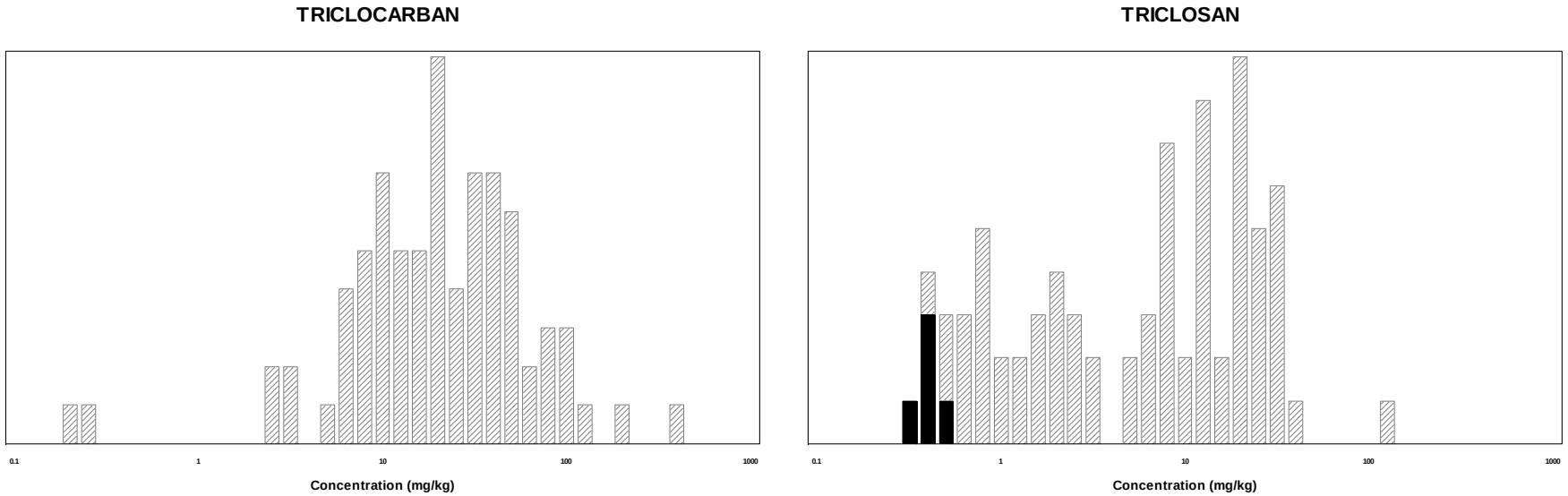


Figure 4-5a. Bar Charts for Pharmaceuticals (continued)

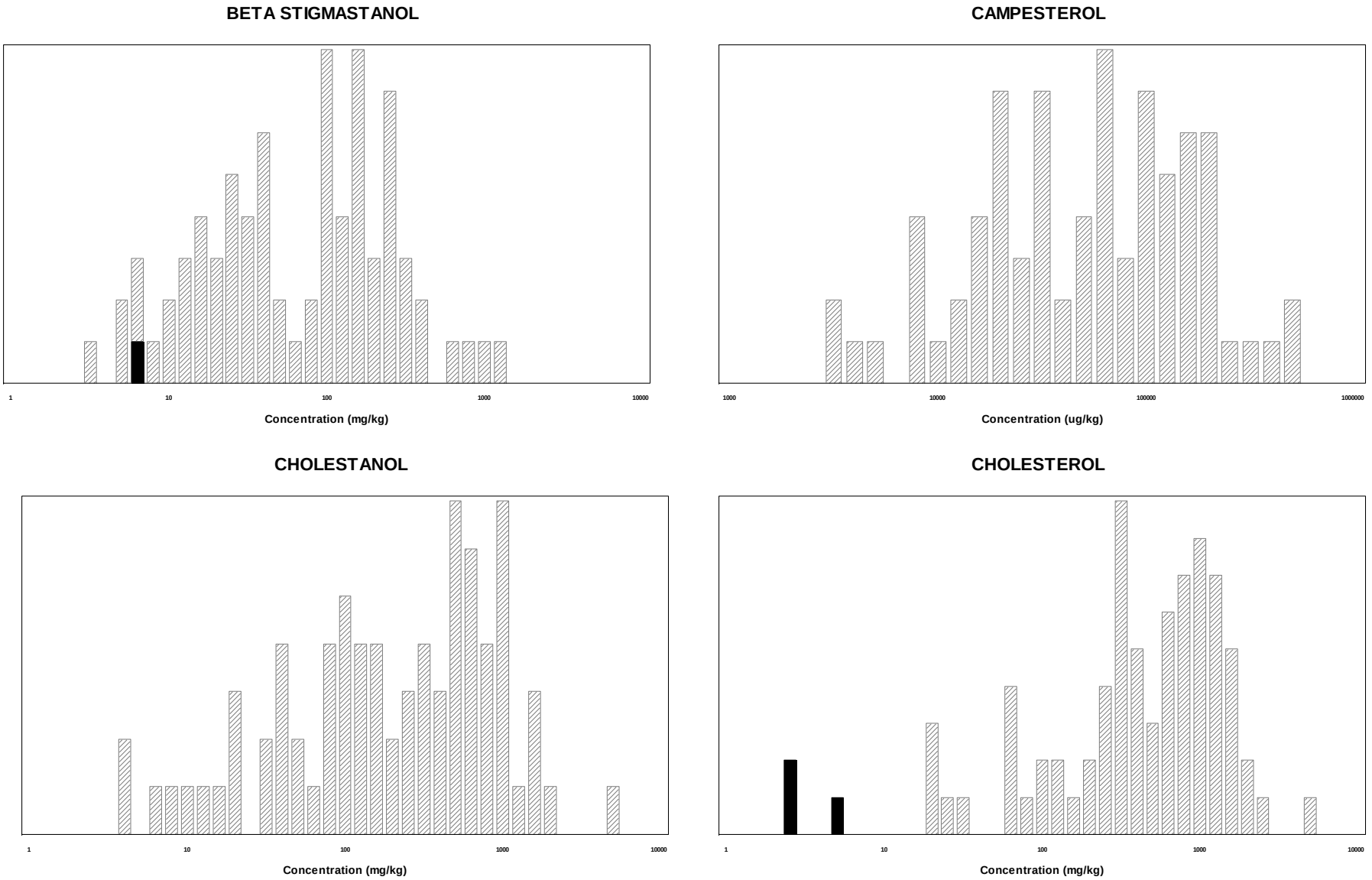
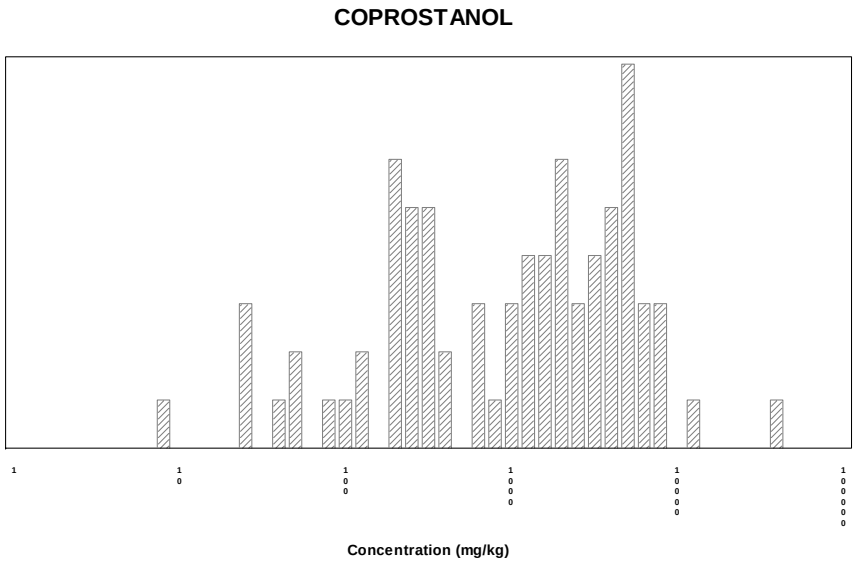
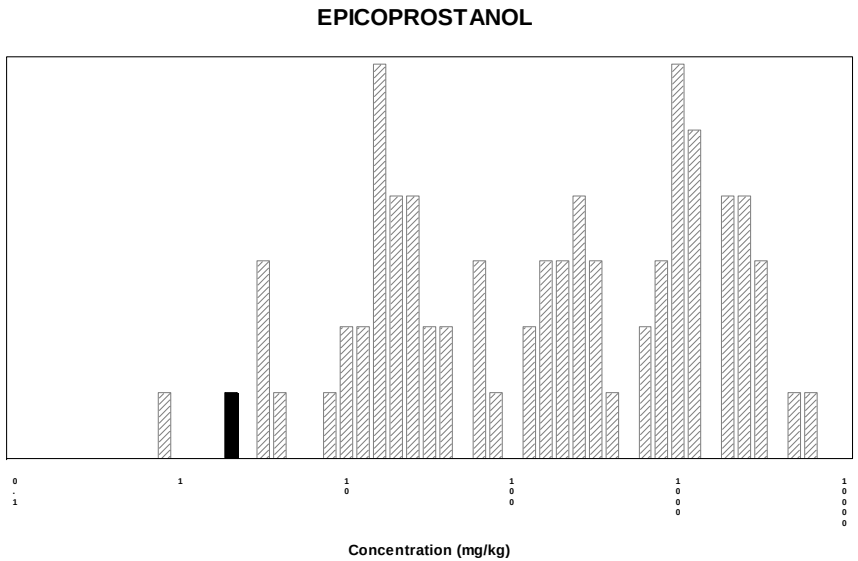


Figure 4-5b. Bar Charts for Steroids and Hormones





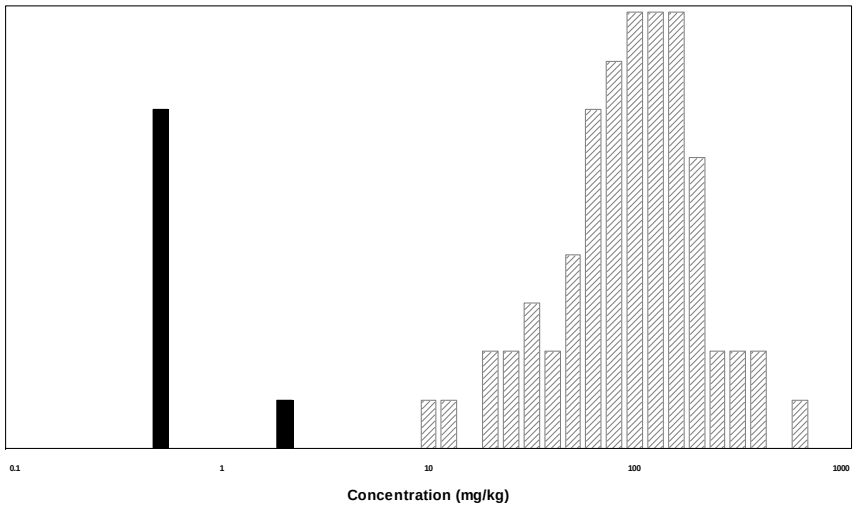


Figure 4-5b. Bar Charts for Steroids and Hormones (continued)

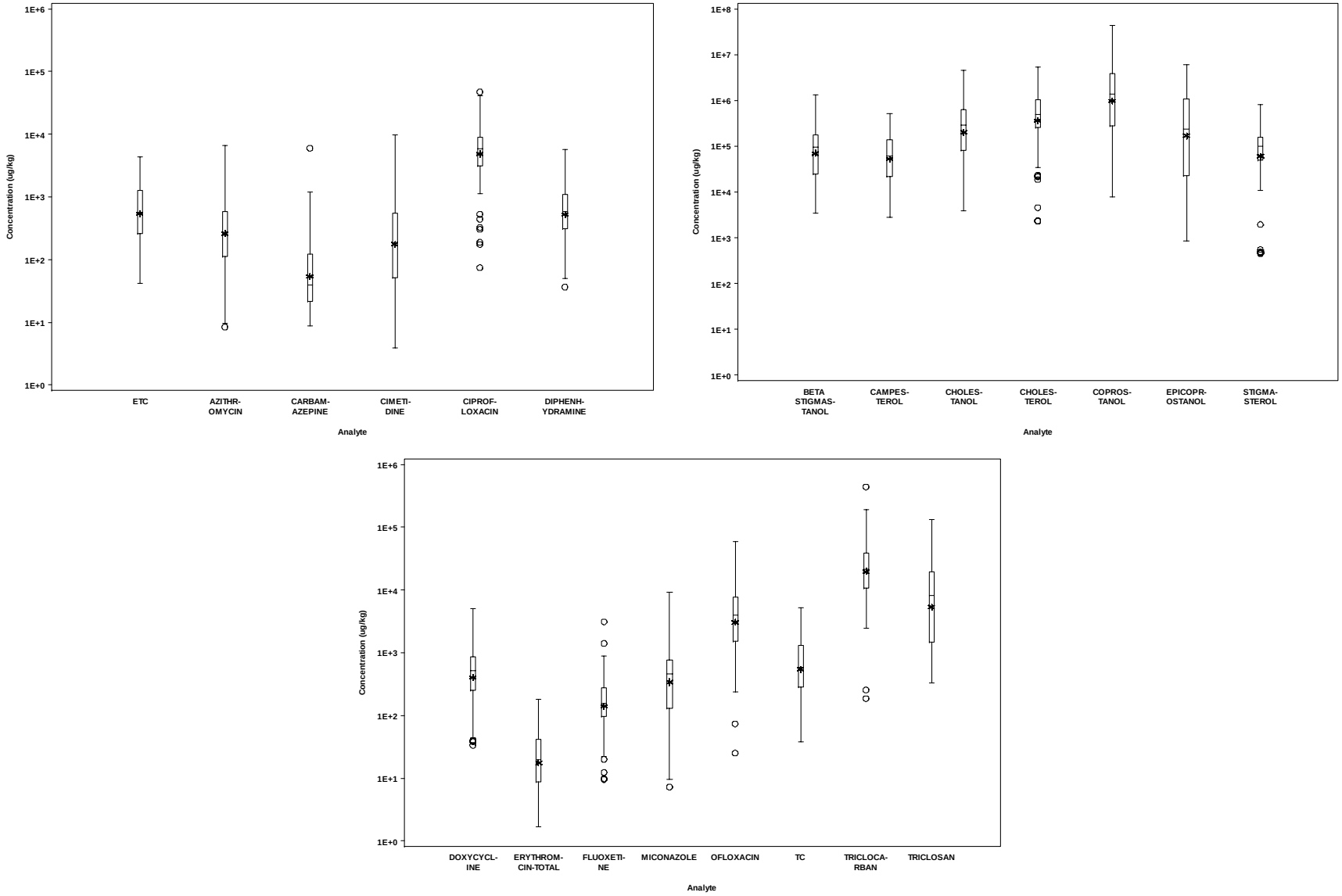


Figure 4-5c. Box Plots for Pharmaceuticals, Steroids, and Hormones

4.3 Data Review

In addition to reviewing the measurement data visually through bar charts and boxplots, we applied statistical techniques to review data for the 34 analytes in Table 4-1. The primary objectives of this review were 1) to identify statistical outliers and evaluate whether to include them in the statistical analysis, and 2) to decide whether to take a lognormal or nonparametric approach to estimate means and percentiles for a particular analyte (Section 2.6 and Appendix C).

4.3.1 Statistical Outliers. We applied two statistical techniques to identify the presence of statistical outliers among detected measurements: the "generalized extreme-Studentized deviate (ESD) many-outlier" procedure (Rosner, 1983), and analysis of variance (ANOVA) modeling approach. Both assumed a lognormal distribution to the data. The Rosner test was capable of identify multiple outliers simultaneously among the observed measurements. The ANOVA model expressed average log-transformed measurements as a linear function of the stratum (flow group) and geographic region in which the POTW was classified. This yielded a "studentized residual" for each measurement. The studentized residual equaled the difference between the observed log-transformed measurement and what the model predicts for this value, divided by the estimated standard error of this difference. A measurement could be considered extreme if its studentized residual exceeded three in absolute value. The ANOVA model approach was applied to both unaggregated and aggregated measurements, as well as to aggregated measurements that were weighted by the survey weights.

Not all measurements flagged by one of these statistical approaches appeared to be extreme. We assessed these findings with information from the bar charts and boxplots (Section 4.2) and data lists (Appendices A.2 through A.6). In each case, no analytical concerns existed that would suggest excluding these values from the statistical analysis. However, if the validity of an extreme sample measurement was brought into question, such as when a given sample had extremely high measurements for multiple analytes, this could lead to excluding the measurement(s). Table 4-4 lists extreme data values that had the greatest potential of being highly influential to the outcome of the statistical analyses, with more details beginning in Section 4.3.3.

When considering only samples collected from plants with flow rates between 1 and 10 MGD, the liquid sample from ID 74 contained the largest concentrations of the PBDEs subject to in-depth analysis. Within this sample, the values for three of these PBDEs (BDE-47, BDE-99, and BDE-153) were highest among all samples and were extreme enough to be detected as statistical outliers in Table 4-4. The fourth PBDE (BDE-209) had a concentration of 15,000,000 ng/kg, which was relatively closer in value to the next largest value (11,000,000 ng/kg) among plants with flow rates between 1 and 10 MGD, and it was not extreme enough to be detected as an outlier (e.g., one sample had a higher concentration).

Table 4-4. Listing of Detected Measurements Labeled as Statistical Outliers for Analytes Subject to In-Depth Statistical Analysis

Analyte	Plant ID	Flow Group	Amount	Units	High/ Low	Range of Detected Results, Excluding the Outliers
Barium	74	1<MGD<10	3,460	mg/kg	High	75.10 to 2650
Silver	27	10<MGD<100	856	mg/kg	High	1.94 to 195
BDE 47	74	1<MGD<10	5,000,000	ng/kg	High	73,000 to 2,600,000
BDE 99	74	1<MGD<10	4,000,000	ng/kg	High	64,000 to 2,500,000
BDE 153	74	1<MGD<10	410,000	ng/kg	High	9,100 to 250,000
Azithromycin	69	1<MGD<10	10.2	ug/kg	Low	26.5 to 6,530
Carbamazepine	74	1<MGD<10	6,030	ug/kg	High	8.74 to 1,190
Ciprofloxacin	21	1<MGD<10	74.5	ug/kg	Low	189 to 47,500
	23	1<MGD<10	176	ug/kg	Low	
Fluoxetine	21	1<MGD<10	20.1	ug/kg	Low	12.4 to 1410
	70	10<MGD<100	3,130	ug/kg	high	
Tetracycline (TC)	21	1<MGD<10	38.3	ug/kg	Low	39.7 to 5,270
Triclocarban	20	1<MGD<10	441,000	ug/kg	High	2,470 to 189,000
	48	1<MGD<10	256	ug/kg	Low	
	61	MGD>100	187	ug/kg	Low	
Cholesterol	74	1<MGD<10	21,900	ug/kg	Low	18,700 to 5,390,000

4.3.2 Statistical Tests for Lognormality. In addition to evaluating outliers, we evaluated whether the lognormal distribution was appropriate for modeling the data. For this evaluation, we used the Shapiro-Wilk test for normality (Shapiro and Wilk, 1965). We applied this test to the logarithms of the reported values. If the reported values are lognormally distributed, then their logarithms will be normally distributed. We start this test by hypothesizing that the log-transformed concentrations are normally distributed, then either reject or fail to reject this hypothesis based on available information in the observed data. We reject the hypothesis of normality if the test's reported p-value is below 0.05. Table 4-5 lists the p-values for the Shapiro-Wilk test when applied to both aggregated and unaggregated log-transformed concentrations for the 34 analytes.

We considered the bar charts and boxplots in addition to p-values from the Shapiro-Wilk test in deciding whether to take the lognormal statistical approach for a given analyte. After reviewing the three outputs (Shapiro-Wilk, bar charts, boxplots), we decided that only nitrate/nitrite deviated substantially from lognormality assumptions. For this one analyte, we used a nonparametric approach to estimate the statistical parameters, including the percentiles.

Table 4-5. Results of Shapiro-Wilk Tests for Normality of Log-Transformed Biosolids Data for Analytes Subject to In-Depth Statistical Analysis

Analyte	P-value of Shapiro-Wilk Test	
	Performed on log-transformed unaggregated detected data	Performed on all log-transformed aggregated data
Metals		
Barium	0.0152*	0.0291*
Beryllium	0.7796	0.9844
Manganese	0.1489	0.1207
Molybdenum	0.6203	0.7402
Silver	0.0010*	0.0007*
Organics		
4-Chloroaniline	0.0068*	0.0433*
Fluoranthene	0.3604	0.4850
Pyrene	0.6515	0.7069
Classicals		
Nitrate/Nitrite	<0.0001*	0.0001*
PBDEs		
BDE 47	0.2600	0.2504
BDE 99	0.2130	0.2678
BDE 153	0.6190	0.7709
BDE 209	0.2650	0.4056
Pharmaceuticals		
4-Epitetracycline (ETC)	0.4439	0.2019
Azithromycin	0.1810	0.0747
Carbamazepine	0.0006*	0.0004*
Cimetidine	0.3294	0.1466
Ciprofloxacin	<0.0001*	<0.0001*
Diphenhydramine	0.2147	0.2827
Doxycycline	0.1092	0.0024*
Erythromycin-total	0.5804	0.0636
Fluoxetine	0.0738	0.0210*
Miconazole	0.1279	0.0445*
Ofloxacin	0.2360	0.0176*
Tetracycline (TC)	0.1900	0.0584
Triclocarban	0.0002*	0.0004*
Triclosan	0.0002*	0.0004*
Steroids and Hormones		
Beta Stigmastanol	0.1737	0.2594
Campesterol	0.1834	0.3016
Cholestanol	0.0091*	0.0244*
Cholesterol	0.0008*	<0.0001*
Coprostanol	0.0033*	0.0127*
Epicoprostanol	0.0023*	0.0058*
Stigmasterol	0.3952	<0.0001*

* P-value is below 0.05, indicating that the hypothesis of normality in log-transformed data can be rejected at the 0.05 level.

4.3.3 Findings from the Data Review. The following sections provide general findings and conclusions made from data reviews for the 34 analytes, along with decisions made on the statistical analysis approach.

4.3.3.1 Metals. For silver, we identified one extreme data value (856 mg/kg; ID 27) that was over four times larger than the next largest value of 195 mg/kg. All tests identified this as a statistical outlier, and it is clearly extreme within the bar chart and boxplot for silver. Upon excluding this value, the p-value reported from the Shapiro-Wilk test was 0.20, suggesting that lognormality appeared reasonable. Thus, we initially applied the lognormal approach to the silver data both with and without this extreme value. EPA performed an additional evaluation of the laboratory's analytical data and documentation, including results from a similar chemical method (Method 200.7) that supported a value close to 856 mg/kg in the database from Method 200.8. EPA then contacted the POTW to ask if they had ever had high silver results in their sludge before, and was told no. Their only major industrial contributors are a dog food plant and a tire manufacturer. (The latter has a pretreatment system.) EPA would not expect either industry to contribute much silver to the plant. Because photo processors are known to contain silver in effluents, EPA also asked if any large photo processors discharged on the system, and the POTW was not aware of any. Because the extreme value of 856 mg/kg appears to be an anomaly that may not reflect normal operations at the POTW, we excluded the value from the final statistical analyses.

In addition to silver, we decided to take the lognormal approach for each of the other three metals:

- Although the Shapiro-Wilk test formally rejected the hypothesis of lognormality for barium at the 0.05 level, any observed deviation from lognormality appeared to be minor. This deviation was not significant enough to warrant taking a nonparametric approach.
- For beryllium and manganese, the assumption of normality in the log-transformed data was reasonable.

A field duplicate collected at one plant (#68357, collected at ID 19) had high concentration values for several metals, organics, and classical compounds, especially when compared to the other sample from the plant with which it was paired. In particular, for all metals, the concentrations associated with this sample were two to four times higher than its paired sample and were frequently high compared to samples from other POTWs. EPA performed an additional evaluation of the laboratory's analytical data and documentation, and they appeared to be acceptable for both samples. The documentation indicated that the samples were liquid sludge products collected from a large storage tank. In addition, the data for total solids suggest that the liquid product was not particularly homogeneous, and the sampling procedures used for this facility did not result in true duplicate samples. Because of sizeable differences from the results of its paired sample and the data from other POTWs, the representativeness of sample #68357 was put into question. As a result, EPA decided to exclude data from this field duplicate sample from the statistical analyses for all metals (as well as for anions and classicals, as noted below). Thus, ID 19 was represented in the statistical analysis by one sample result rather than two.

While we identified one statistical outlier for barium (3,460 mg/kg; ID 74), it was one of two samples collected at this POTW. Upon averaging this measurement with the other sample result for this POTW, it lowered its influence on the distribution. In other words, the average value did not appear to be a statistical outlier, and thus, was retained for the statistical analyses.

4.3.3.2 Organics. As was done with the metals, EPA determined that all data associated with the field duplicate sample from ID 19 (#68357) would be excluded from the statistical analyses for the three organics.

No statistical outliers were identified among the other samples for the three organics.

We selected the lognormal approach for each of the organics. Although the p-value of the Shapiro-Wilk test for 4-chloroaniline was slightly below 0.05, its observed distribution for detected values was similar to that for the other organics, for which lognormality appeared to be sufficient.

4.3.3.3 Classicals. The distribution of nitrate/nitrite concentrations appeared to deviate considerably from the symmetric bell-shaped curve that signifies a lognormal distribution. The bar chart exhibited two peaks in the data and a long right tail. The outcome of the Shapiro-Wilk test verified the lack of lognormality. Therefore, we selected a nonparametric approach for nitrate/nitrite data analysis. This compound may have a different distribution from others, because it is the combination of two analytes (nitrate and nitrite).

As was done for the metals and organics, EPA excluded data associated with the field duplicate sample (#68357) for one plant (ID 19) from the statistical analysis applied to the nitrate/nitrite data. While we identified no other statistical outliers among the nitrate/nitrite data, this could be the result of high variability in the data.

4.3.3.4 PBDEs. Within the four facilities having samples collected from different treatment systems (Section 2.4.1), we observed that variability in measurement values within a facility appeared to be a significant component of total variability for PBDEs. For several PBDEs, one sample's measurement was two to six times higher than the other sample for that facility. If these paired measurements were averaged, the statistical analysis would have ignored this potentially significant source of variability. Therefore, for each of these four facilities, we included the measurements for both samples in the statistical analyses without averaging them together. We assigned a weight to each sample result equal to the facility's survey weight divided by two. (Deviations of this magnitude were considerably less prevalent among the six facilities having regular and field duplicate samples collected. Thus, measurements for the paired samples within these facilities were averaged.)

- For one of these facilities (ID 74), its liquid sample consistently had the highest concentrations among all samples in the survey for each PBDE. For three of the four PBDEs included in the in-depth statistical analysis, these measurements were flagged as statistical outliers (Table 4-4). However, none of these measurements had data qualifiers assigned to them that suggested validity concerns. Thus, none were excluded from the statistical analysis.

Data for each of the four PBDEs included in the in-depth statistical analysis well-resembled a lognormal distribution. Thus, we took the lognormal-based approach for each of these PBDEs.

4.3.3.5 Pharmaceuticals, Steroids, and Hormones. Among the set of 21 pharmaceuticals, steroids, and hormones included in the in-depth statistical analyses, one analyte (cimetidine) had no reported measurement for ID 34 because the laboratory result for this facility did not meet EPA's quality assurance criteria. As noted in the data listings within Appendix A.6.2, EPA excluded at least one measurement for 28 other pharmaceuticals, steroids, and/or hormones that were not selected for the in-depth statistical analyses, also for quality assurance reasons.

For the six facilities with field duplicate samples collected, the measurement for one sample was frequently no more than two times the other. Some exceptions occurred, such as for total erythromycin and miconazole at ID 32, and for several steroids and hormones at ID 19. Even with the exceptions, EPA

considered the results to be reasonable, and continued to average field duplicate measurements within a facility for pharmaceuticals, steroids, and hormones.

For some of the pharmaceuticals, steroids, and hormones, within a facility with two treatment systems, the extent of differences in measurements between the two samples was similar to what we observed with the PBDEs. For six analytes (azithromycin, cholestanol, cholesterol, ciprofloxacin, diphenhydramine, ofloxacin), each of these four facilities had one sample measurement that exceeded twice the value of the other sample, while all facilities with field duplicate samples had smaller deviations. As a result, we did not average measurements within these facilities. Instead, we used the measurements individually in the statistical analysis. Like the PBDEs, we assigned a weight to each sample result equal to the facility's survey weight divided by two.

We chose the lognormal approach for each of the 21 pharmaceuticals, steroids, and hormones for the in-depth statistical analysis. Although the p-values from the Shapiro-Wilk test were occasionally below 0.05 for some of these analytes, the bar charts and boxplots suggested that any deviation from lognormality tended to be minor.

Although our statistical tests identified some potential outliers among seven of the 21 analytes (Table 4-4; Appendix B.1.2), there was not sufficient evidence to warrant exclusion of any of these values from the statistical analysis. Unlike the other analyte classifications, the outliers associated with this set of 21 analytes were on both the low and high side. No single facility was the primary source of these outliers.

4.4 National Estimates

By applying the statistical approach specified in Section 4.3, we obtained estimates of the mean, standard deviation, and selected percentiles (99^{th}, 98^{th}, 95^{th}, 90^{th}, and 50^{th} percentiles) for each of the 34 analytes specified in Table 4-1. Appendix C provides details on how these estimates were calculated within each approach (lognormal and nonparametric). Each method incorporated the final survey weights assigned to the POTWs. Therefore, these estimates are representative of the distribution of concentrations in biosolids for the entire target population (i.e., they represent "national" estimates).

Table 4-6 summarizes the statistical estimates for each of the 34 analytes. We list the number of data points used in the analysis within the column labeled 'n'. The 'n' column lists three different values: 74 when multiple measurements at some POTWs were averaged, 73 under the same conditions with one value excluded during the chemical quality assurance review, and 78 when some multiple measurements were used separately with half of the survey weight. (See Section 2.4.3.) We also provide the estimated number of POTWs in the target population which they represent, within the column labeled 'Est. N.' These values equal the sum of the survey weights. When we applied the nonparametric approach (for nitrate/nitrite only), we represented non-detects by one-half of the sample-specific detection limit.

For each of the 34 analytes, we produced two sets of estimates by applying both statistical approaches. This was done to investigate how the estimates may differ if a different approach was taken. Both sets of estimates are presented in Appendix D. However, EPA considers the set of estimates presented in Table 4-6 as the final set for each analyte.

For nitrate/nitrite, whose underlying data distribution did not appear to be lognormal, the estimates presented in Table 4-6 for the mean and percentiles are higher than the estimates generated under the lognormal approach. The standard deviation estimates for nitrate/nitrite, however, were similar between the two methods. Because our sample size was less than 100, the nonparametric approach sets the 99^{th} percentile for a given analyte to the largest reported measurement.

Table 4-6. Nationally Representative Estimates of the Mean, Standard Deviation, and Selected Upper Percentiles of the Distribution of Concentrations for 34 Analytes in the TNSSS

Analyte	Observed Values		Estimates							
			Percentiles					Summary Statistics		
	Minimum	Maximum	99th	98th	95th	90th	50th	Mean	Standard Deviation	Percent POTWs with Detected Conc
Metals (mg/kg)										
Barium	77	2,117	2,230	1,848	1,396	1,088	452	572	443	100
Beryllium	0.04	2.34	1.81	1.45	1.04	0.77	0.27	0.38	0.37	98.5
Manganese	35	14,900	9,700	6,904	4,156	2,648	540	1,165	2,231	100
Molybdenum	2.51	86.4	68.7	55.6	40.5	30.6	11.4	15.3	13.8	100
Silver*	2	195	105	82	57	42	13	20	22	100
Organics (ug/kg)										
4-Chloroaniline	51	5,900	12,013	8,288	4,762	2,912	513	1,284	2,946	74.4
Fluoranthene	45	12,000	13,173	9,112	5,256	3,226	575	1,421	3,211	89.5
Pyrene	44	14,000	15,918	10,894	6,184	3,742	634	1,654	3,981	84.9
Classicals (mg/kg)										
Nitrate/Nitrite	2	6,120	6,120	2,750	960	463	14	219	828	100
PBDEs (ng/kg)										
BDE-47 (2,2',4,4'-tetrabromodiphenyl)	73,000	5,000,000	2,650,430	2,212,077	1,688,881	1,329,167	570,448	709,174	523,791	100
BDE-99 (2,2',4,4',5-pentabromodiphenyl)	64,000	4,000,000	2,696,928	2,248,181	1,713,370	1,346,295	574,559	716,362	533,447	100
BDE-153 (2,2',4,4',5,5'-hexabromodiphenyl)	9,100	410,000	265,395	220,098	166,454	129,902	54,117	68,334	52,685	100
BDE-209 (decabromodiphenyl)	150,000	17,000,000	15,836,435	11,645,502	7,360,103	4,898,034	1,162,523	2,181,237	3,462,942	98.5
Pharmaceuticals (ug/kg)										
4-Epitetracycline (ETC)	41	4,380	8,026	5,937	3,787	2,540	620	1,135	1,741	96.0
Azithromycin	8	5,205	8,717	5,811	3,172	1,853	278	831	2,342	96.0
Carbamazepine	9	6,030	1,234	856	497	306	55	135	298	96.0
Cimetidine*	4	8,330	19,128	10,975	4,789	2,294	171	1,332	10,314	89.9

Table 4-6. Nationally Representative Estimates of the Mean, Standard Deviation, and Selected Upper Percentiles of the Distribution of Concentrations for 34 Analytes in the TNSSS (Continued)

Analyte	Observed Values		Estimates							
			Percentiles					Summary Statistics		
	Minimum	Maximum	99th	98th	95th	90th	50th	Mean	Standard Deviation	Percent POTWs with Detected Conc
Pharmaceuticals (ug/kg) (cont.)										
Ciprofloxacin	75	40,800	79,636	57,975	36,095	23,703	5,367	10,501	17,658	100
Diphenhydramine	37	5,730	5,255	4,021	2,696	1,891	541	871	1,101	100
Doxycycline	34	5,090	7,021	5,046	3,082	1,989	424	877	1,588	92.8
Erythromycin-Total	2	180	264	194	123	82	19	36	58	92.9
Fluoxetine	10	3,130	1,555	1,178	778	539	147	245	329	96.1
Miconazole	7	9,210	16,931	10,083	4,652	2,341	207	1,239	7,311	95.8
Ofloxacin	25	58,100	85,562	57,929	32,363	19,304	3,113	8,573	21,998	98.5
Tetracycline (TC)	38	5,270	10,042	7,250	4,458	2,895	630	1,278	2,255	97.5
Triclocarban	187	441,000	276,708	205,043	131,079	88,120	21,677	39,433	59,924	100
Triclosan	334	133,000	197,288	124,176	62,217	33,693	3,862	16,097	65,135	92.4
Steroids and Hormones (ug/kg)										
Beta Stigmastanol	3,440	1,330,000	1,651,188	1,123,256	632,009	379,365	62,547	168,079	419,232	98.5
Campesterol	2,840	524,000	842,112	598,919	360,119	229,283	46,547	100,879	193,964	100
Cholestanol	3,860	4,590,000	7,874,368	5,071,045	2,629,149	1,467,636	187,244	680,046	2,374,369	100
Cholesterol	2,340	5,390,000	13,376,891	8,538,884	4,369,111	2,410,541	295,092	1,129,268	4,171,366	96.9
Coprostanol	7,720	43,700,000	57,794,254	35,060,035	16,626,022	8,574,467	827,108	4,366,714	22,636,715	100
Epicoprostanol	868	6,030,000	25,579,800	13,441,281	5,143,938	2,193,143	108,028	1,702,708	26,783,520	98.5
Stigmasterol	455	568,500	4,606,900	2,646,615	1,157,099	555,217	41,513	321,199	2,464,383	90.1

* Nitrate/nitrite estimates were estimated using the non-parametric model with not-detected values replaced with one-half of the sample specific detection limit. All other estimates were calculated using the lognormal model.
** Outlier removed for ID 27.

4.5 Comparison of Metals to Current Standards

EPA established the current standards for land application (40 CFR 503) as a ceiling (i.e., upper limit) on the dry-weight concentrations for nine distinct metals. Table 4-7 documents these nine metals and their land application ceiling standards, along with the maximum observed concentrations among samples collected in the TNSSS for these nine metals. The maximum concentrations are calculated by considering both the individual sample results (presented under "unweighted statistics") and after averaging results within each POTW when data values for multiple samples were reported (presented under "weighted statistics"). The number of POTWs in the sample with data values exceeding the ceiling, as well as an estimate of the total number of POTWs in the target population that exceed the ceiling, are reported.

As noted in the table, only three metals have maximum observed concentrations exceeding their respective land application ceiling concentrations: molybdenum, nickel, and zinc. The maximum observed concentration for all other metals in this table are well below their respective land application regulatory limits.

Table 4-7 also shows the number of POTWs in the survey with concentrations exceeding the specified land application ceiling. After excluding sample #68357 for ID 19 as explained in Section 4.3.3.1, only four samples in this study had concentrations that were greater than the land application ceiling concentrations. Two of the samples were from a single POTW (ID 71), which exceeded the limits for both molybdenum and nickel, while the other two samples were from other POTWs (ID 2 exceeded the nickel standard, and ID 57 exceeded the zinc standard). When we apply the survey weights to these POTWs to obtain national estimates, we determine that less than three percent of POTWs in the survey's target population might be expected to exceed the land application standards for any of these three metals. EPA notes that three percent is likely to be an overestimate, because the regulations apply only to land application, and many facilities use other methods of disposal.

Of the POTWs observed exceeding these standards in the survey, one incinerated its treated biosolids on site, while the others sent their biosolids to landfills. Thus, results from this survey indicate that POTWs were generally complying with the existing land application standards for metals.

Table 4-7. Land Application Ceiling Standards for Nine Metals, and Maximum Concentrations As Observed in Samples Collected in the TNSSS

Analyte	CAS No.	Land Application Ceiling (mg/kg)	Max.. Conc. of Individual Samples (mg/kg)[a]	Max. Conc., After Averaging (mg/kg)	# Sampled POTWs Over the Ceiling	Estimated # POTWs Nationally Over the Ceiling	Estimated % POTWs Nationally Over Ceiling
Arsenic	7440382	75	49.2	49.2	0	0	0.0
Cadmium	7440439	85	11.8	11.8	0	0	0.0
Copper	7440508	4300	2580.0	1720.0	0	0	0.0
Lead	7439921	540	450.0	350.0	0	0	0.0
Mercury	7439976	57	8.3	7.5	0	0	0.0
Molybdenum	7439987	75	86.4*	86.4*	1	45	1.4
Nickel	7440020	420	526.0*	526.0*	2	96	2.9
Selenium	7782492	100	24.7	24.2	0	0	0.0
Zinc	7440666	7500	8550.0*	8550.0*	1	51	1.5

* Exceeds the land application ceiling.

4.6 Comparison of Metals, Organics, and Classicals to NSSS Results

EPA conducted the 1988 National Sewage Sludge Survey (NSSS) to obtain national estimates of over 400 pollutants in biosolids that POTWs have treated and prepared for disposal or some other use (e.g., land application). EPA used the data collected in the NSSS to support its development of pollutant limitations, regulatory impact analysis, and aggregate risk analysis in the Final Standards for the Use or Disposal of Sewage Sludge (40 CFR 503).

While it is computationally possible to compare distributional estimates between the NSSS and the TNSSS, it is not entirely appropriate to do so. The two studies were designed with different statistical objectives and target populations. More importantly, the TNSSS was not designed in a manner that would allow for statistical inferences about the differences between the two studies. As a result, any observed differences between the two sets of estimates do not necessary imply that levels have changed from the time that the NSSS occurred. The following differences between how EPA designed the NSSS and how the Agency designed the TNSSS result in limitations on how estimates obtained from the two surveys can be compared:

- The target population sizes were considerably different. In 1988, EPA identified 11,307 POTWs from the 1986 NEEDS survey, contrasted with the 3,337 POTWs for the TNSSS.
    - The NSSS included an additional stratum that represented POTWs with flow rates below 1.0 MGD. One-quarter of the sample for this survey (46 out of a total of 185 POTWs sampled) came from this stratum. The TNSSS did not represent such POTWs at all within its sample, and therefore, in its results. Because neither survey was designed to report stratum-specific results, distributional estimates from the NSSS cannot be obtained from existing documents for only those POTWs with flow rates above 1.0 MGD.
    - The NSSS excluded POTWs using lagoons and partial treatment from its sample, but the TNSSS considered such POTWs as eligible.
- The TNSSS required laboratories to measure percent solids first, then adjust the aliquot of wet sludge used to get both a consistent and a manageable amount of dry solids to extract. In the NSSS, laboratories analyzed a standard aliquot volume, and a dry weight concentration was obtained by mathematically adjusting the analytical measurement using the sample's percent solids.
- The TNSSS achieved greater sensitivity in measuring organic concentrations. In addition, TNSSS also required the laboratories to run a specific cleanup technique, called gel permeation chromatography (GPC), to remove much of the lipid content from the raw sample extracts before they were concentrated. It was very effective at removing the lipids as well as other interferences. GPC was used on some samples in the NSSS, but at the discretion of the laboratory. For the TNSSS, it was required for all samples.
- The two studies primarily focused on different sets of analytes. NSSS evaluated 412 analytes, including many organics. TNSSS evaluated 145 analytes which largely consisted of PBDEs and pharmaceuticals, steroids, and hormones that were not measured in the NSSS. It also included a few analytes that had previously been evaluated in NSSS.

Despite its concerns about the validity of comparisons between the two studies, EPA has presented a comparison of the two surveys in Table 4-8. One set of estimates is based on data from the NSSS (SAIC,

2003). The second set of estimates originates from data collected in the TNSSS. Of the eight target analytes (i.e., barium, beryllium, manganese, silver, fluoranthene, pyrene, 4-chloroaniline, and nitrate/nitrite), only beryllium had lognormal model-based estimates reported in USEPA (1992) for its distributional parameters, which were calculated from 1988 NSSS data. As a result, Table 4-8 presents estimates derived from nonparametric (distribution-free) approaches. These numbers were taken from Appendix D for TNSSS and Appendix F for NSSS, with the sample-specific detection limits substituted for non-detected outcomes.

Table 4-8. Comparison of Distributional Parameter Estimates Between the NSSS and the TNSSS, Obtained Using Nonparametric (Distribution-Free) Approaches

Analyte	Survey	n	N	% Det.	Est. Mean	Est. S.D.	Estimated Percentiles				
							99th	98th	95th	90th	50th
Metals (mg/kg)											
Barium	NSSS	176	7,750	100%	673	840	3000	2370	1730	1230	499
	TNSSS	74	3,337	100%	575	454	2117	2060	1700	1240	426
Beryllium	NSSS	176	7,750	22%	1.84	2.43	8.56	8.33	6.00	5.00	0.56
	TNSSS	74	3,337	99%	0.386	0.374	2.34	1.23	1.17	0.89	0.27
Manganese	NSSS	176	7,750	100%	538	1040	4060	3720	1620	929	276
	TNSSS	74	3,337	100%	1247	2228	14900	7690	3430	3020	449
Silver	NSSS	176	7,750	84%	48.2	112	546	218	128	75.8	25.5
	TNSSS	74	3,337	100%	32.3	101	856	195	71.6	42.3	13.5
Organics (ug/kg)											
4-Chloroaniline	NSSS	176	7,750	5%	8640	13800	46700	43000	33300	28800	4760
	TNSSS	74	3,337	76%	1099	1051	5900	3700	3200	2500	865
Fluoranthene	NSSS	176	7,750	5%	8950	13400	46700	43000	32800	27700	4760
	TNSSS	74	3,337	91%	1420	2247	12000	9700	6700	3500	550
Pyrene	NSSS	176	7,750	5%	8850	13400	46700	43000	33000	28000	4760
	TNSSS	74	3,337	85%	1647	2593	14000	10000	8700	3900	620
Classicals (mg/kg)											
Nitrate	NSSS	176	7,750	95%	1420	5040	26500	15500	5020	1890	96.5
Nitrite	NSSS	176	7,750	83%	201	1210	2920	2910	462	215	12.9
Nitrate/Nitrite	TNSSS	74	3,337	100%	219	828	6120	2750	960	463	13.8

Source of estimates from NSSS: SAIC (2003). Some numbers are approximate.

5.0: CONCLUSIONS

For a variety of targeted chemicals, the primary goal of the TNSSS was to characterize mean concentration levels and selected percentiles of analytes in biosolids generated by the nation's POTWs (having flow rates of at least 1.0 MGD). EPA successfully collected 84 biosolids samples from its targeted sample size of 74 POTWs. EPA collected these samples from 69 of the 80 POTWs in its original sample, and from five POTWs that served as replacement facilities. While EPA had anticipated that it would find some ineligible facilities in its sample, it encountered more ineligible facilities than initially anticipated. As a result, EPA redefined its target population to include wastewater treatment ponds as a final treatment stage. Also, if a facility utilized partial treatment, it "followed" the biosolids to the facility that applied a final treatment stage and sampled from that facility instead. EPA determined that the biosolids generated by the five replacement POTWs in this survey were characteristic of final treated biosolids for the POTWs that they replaced. Although EPA had excluded facilities known to use ponds or to conduct only partial treatment from the sample frame, the sample design report (Appendix E) notes that only 46 facilities with flow rates exceeding 1.0 MGD were excluded as a result (29 facilities with ponds, and 17 facilities conducting partial treatment). This is less than 1.5 percent of the 3,337 facilities that EPA included within the sample frame. Even if these 46 facilities had been included in the sample frame and one or more had been selected for the sample, it is expected that the results presented in this report would have been impacted in only a very minor way, if at all. Therefore, EPA has concluded that the estimates generated from the survey data retain the original statistical properties of the original sample design.

The TNSSS database, available from EPA, contains concentration measurements for 145 different analytes. These analytes included three classicals, 28 metals, six organics (PAHs and semivolatiles), 11 PBDEs, 72 pharmaceuticals and 25 steroids and hormones. As necessary, the database qualifies each reported measurement by noting any quality-control issues associated with the laboratory analysis. None of these qualifications were severe enough to warrant excluding any measurement from EPA's statistical analyses that was reported in the database.

At 64 of the 74 facilities, EPA collected a single biosolids grab sample. At the other ten facilities, EPA collected two grab samples on a single day. The second sample represented either a field duplicate (at six facilities) or a sample collected from a second treatment system or product generated by the facility (at four facilities). In general, EPA found that a field duplicate sample's measurements were similar to those of the regular sample collected at the same facility, with one measurement seldom more than twice the other. (One exception occurred, however, at ID 19, with the field duplicate measurement being considerably higher than its paired sample measurement.) However, measurements could differ considerably between the two samples collected from different treatment systems or products within a facility. This was especially true for PBDEs, pharmaceuticals, steroids, and hormones.

EPA performed an in-depth statistical analysis on data for 34 analytes. The statistical analysis produced national estimates by incorporating survey weights based upon the statistical sample design. Thus, the estimates represent analyte concentrations from the target population of POTWs. Except for nitrate/nitrite, EPA assumed that lognormality procedures were appropriate for percentile estimates, based upon its review of statistical graphics and goodness of fit tests. Nitrate/nitrite had an observed data distribution that deviated considerably from what would be expected under lognormality. For this analyte, EPA applied a nonparametric (distribution-free) approach to deriving estimates. The percentiles are presented in Table ES-2 and Table 4-6 for the 34 analytes selected for in-depth review. Although EPA was less interested in the remaining analytes and did not perform in-depth evaluation of the statistical results for them, it has provided preliminary data summaries in Appendix B.3. The reader should exercise caution in interpreting these preliminary summaries.

6.0: REFERENCES

Kish, L. (1965). *Survey Sampling*. New York: John Wiley and Sons, Inc.

NRC (2002). *Biosolids Applied to Land: Advancing Standards and Practices*. Report by the National Research Council. Washington, DC: National Academy Press. July 2002.

Rosner, B. (1983). "Percentage Points for a Generalized ESD Many-outlier Procedure." *Technometrics* 25(2):165-172.

SAIC (2003). National Pollutant Concentration Percentile Estimates from the National Sewage Sludge Survey: Candidate Pollutants for Round Two Regulations. Facsimile to Henry Kahn from T. Stralka. January 30.

Shapiro, SS, and Wilk, MB. (1965). "An Analysis of Variance Test for Normality (Complete Samples)." *Biometrika.* 52(3,4):591-607.

United States Federal Register. 68 FR 75531. Wednesday December 31, 2003, pp. 75531-75552.

USEPA (2008). *Sampling and Analysis Report for the Targeted National Sewage Sludge Survey*. Final Report, Office of Science and Technology, U.S. Environmental Protection Agency. EPA-822-R-08-016.

USEPA (2004). *Technical Background Document: Refined Biosolids Exposure and Hazard Assessment Interim Final*. Contract No. 68-C-04-006. Office of Water, Washington, DC. October 8, 2004.

USEPA (2002a). *Statistical Support Document for the Development of Round 2 Biosolids Use or Disposal Regulations.* Final Report, Office of Science and Technology, U.S. Environmental Protection Agency. EPA-822-R-02-034. April 2002.

USEPA (2002b). *Wastewater Technology Fact Sheet: Facultative Lagoons*. U.S. Environmental Protection Agency. Retrieved from http://www.epa.gov/OW-OWM.html/mtb/faclagon.pdf. December 2008

USEPA (1992). *Statistical Support Document for the 40 CFR, Part 503: Final Standards for the Use or Disposal of Sewage Sludge. Volumes I and II.* Final Report, Engineering and Analysis Division, Office of Science and Technology, U.S. Environmental Protection Agency. 11 November 1992.