DATA MANIPULATION SUMMARY FOR HIGH PRODUCTION VOLUME (HPV) AND PESTICIDE INERT CHEMICALS

Prepared for:

U.S. Environmental Protection Agency
Office of Science Coordination and Policy

Prepared by:

Eastern Research Group, Inc.
14555 Avion Parkway, Suite 200
Chantilly, Virginia 20151

08 May 2007

EPA Contract No. EP-W-05-014
Work Assignment No. 3-03

Table of Contents

List of Tables

1.0 INTRODUCTION

The 1996 Food Quality Protection Act and 1996 Amendments to the Safe Drinking Water Act required the Environmental Protection Agency (EPA) to develop a screening program to determine whether certain substances have endocrine effects similar to those produced by naturally occurring hormones. Such substances are known as endocrine disruptors.

In 1996, EPA formed the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC). The EDSTAC Report of August 1998 discussed the Committee's recommendations for many aspects of EPA's Endocrine Disruptor Screening Program (EDSP), including priority setting.

On September 27, 2005, EPA issued a Federal Register notice (70 FR 56449) that describes the chemical selection approach that the Agency would use to select an initial list of chemicals to undergo Tier 1 Screening under EDSP. In this Notice, EPA indicated that the initial universe of chemicals to be considered would include: 1) high production volume (HPV) chemicals that are also pesticide inerts (HPV/Pesticide Inerts); and 2) pesticide active ingredients. The focus of this report is on HPV chemicals that are also pesticide inerts. In this Federal Register notice, EPA identified four potential pathways for human exposure HPV chemicals that are also pesticide inerts, including:

- Human biomonitoring (i.e., presence in human tissues);
- Drinking water;
- Ecological biomonitoring; and
- Indoor air.

EPA also identified specific data sources of information related to these exposure pathways, as listed in Table 1-1.

Table 1-1
Data Sources Evaluated

Exposure Pathway	Data Source
Human Biological Monitoring Data	National Health and Nutrition Examination Survey III (NHANES III) - Priority Toxicant Range Study for Volatile Organic Compounds (VOCs)
	CDC's Second and Third National Reports on Human Exposure to Environmental Chemicals (NHANES 1999 - 2002)
	National Human Adipose Tissue Survey (NHATS)
	Total Exposure Assessment Methodology (TEAM) Study Human Biological Monitoring (Breath) Data

Table 1-1 (Continued)

Exposure Pathway	Data Source
Ecological Biological Monitoring Data	National Sediment Inventory (NSI) Tissue Data
	National Fish Tissue Study
	National Water Quality Assessment (NAWQA) Aquatic Tissue Data
Chemicals in Drinking Water Monitoring Data	National Contaminant Occurrence Database (NCOD)
	National Human Exposure Assessment Survey (NHEXAS) Drinking and Tap Water Data
	TEAM Water Data
	National Stream Quality Accounting Network (NASQAN) Surface Water Data
	NASQAN Sediment Data
	NAWQA Ground Water Data
	NAWQA Surface Water Data
	NAWQA Sediment Data
Indoor Air Monitoring Data	EPA/Office of Research and Development (ORD) Journal Articles - Indoor Air Data
	EPA/Office of Research and Development (ORD) Journal Articles - Personal Air Data
	NHEXAS - Indoor Air Data
	NHEXAS - Personal Air Data
	TEAM Air Data

Each of these data sources was reviewed, relevant information was collected, and databases to summarize this information were prepared for HPV chemicals that are also pesticide inerts. Data were aggregated for each data source and exposure pathway into "summary" tables, which include only one record per chemical name/CAS number combination.

This report is organized by exposure pathway and describes each of the data sources evaluated and specific data manipulation activities. The sections of this report include the following:

- Section 2.0 - High Production Volume and Pesticide Inerts List of Chemicals;
- Section 3.0 - Summary Statistics for Data Collection;
- Section 4.0 - Summary of Data Sources and Data Manipulation; and
- Section 5.0 - Integration of Pathway Data.

2.0 HIGH PRODUCTION VOLUME AND PESTICIDE INERTS LIST OF CHEMICALS

EPA's current approach for developing an initial list of chemicals to undergo Tier 1 Screening as part of EDSP was outlined in the September 27, 2005 Federal Register notice (70 FR 56449). In this Notice, EPA indicated that the initial universe of chemicals to be considered would include: 1) high production volume (HPV) chemicals that are also pesticide inerts (HPV/Inerts); and 2) pesticide active ingredients. Data collection efforts for pesticide active ingredients are described in a separate report (*Data Manipulation Summary for Pesticide Active Ingredients*, Document ID: EPA-HQ-OPPT-2004-0109-0012). The pesticide active ingredients data collection and results are described within this report

HPV chemicals are those substances that are produced or imported into the U.S. in amounts greater than or equal to one million pounds per year. Based on direction from the EDSP, the list of HPV chemicals is based on the non-confidential list of 2002 Toxic Substances Control Act (TSCA) Inventory Update Rule (IUR) chemicals (http://www.epa.gov/oppt/iur/tools/data/2002-vol.htm).

Pesticide inert chemicals are defined as any ingredients in pesticide product formulations other than the active ingredient. (7 U.S.C. 136a (m)). The Office of Pesticide Programs (OPP) maintains an inventory of pesticide inert chemicals that are categorized into the following four lists (http://www.epa.gov/opprd001/inerts/lists.html):

- List 1 - Inert Ingredients of Toxicological Concern. Any product containing a List 1 ingredient must include the label statement "This product contains the toxic inert ingredient (name of inert)."

- List 2 - Potentially Toxic Inert Ingredients/High Priority for Testing Inerts. The substances on this list may be structurally similar to chemicals known to be toxic; some have data suggesting a concern.

- List 3 - Inerts of Unknown Toxicity. Inert ingredients on this list have not yet been determined to be of known potential toxicological concern nor have they been determined to be of minimal concern. These substances will continue to be evaluated to determine if they merit reclassification to List 1, 2, or 4.

- List 4 - Inerts of Minimal Concern. This list is subdivided into List 4A (minimal risk inert ingredients) and List 4B (inerts which have sufficient data to substantiate that they can be used safely in pesticide products).

Table 2-1 presents the number of HPV and pesticide inert chemicals and the number of chemicals that are contained on both lists.

Table 2-1
HPV and Pesticide Inert Chemical Counts

Chemical List	Number of Chemicals
High Production Volume Chemicals[1]	2,708
Pesticide Inert Chemicals[2]	2,775
Overlap of HPV/Pesticide Inert Chemicals	643

As shown in Table 2-1, there are a total of 643 chemicals that are both an HPV and pesticide inert chemical. This overlap was identified by matching Chemical Abstract Service (CAS) Registry numbers on each of the lists. Note that the list of pesticide inerts contains 109 chemicals without corresponding CAS numbers. This list of 109 pesticide inert chemicals was reviewed to determine whether any overlap could be identified based on chemical name. Table 2-2 presents chemical matches that were identified based on name, and also includes the CAS number provided on the HPV list. These chemicals shown in Table 2-2 are included in the universe of HPV/pesticide inert overlap chemicals considered for EDSP screening as shown in Table 2-1.

Table 2-2
Additional Chemicals Identified for the HPV/Pesticide Inert Overlap List

HPV CAS Number	HPV Name	Inert Name
67784901	Fatty acids, coco, reaction products with 2-[(2-aminoethyl)amino]ethanol	Fatty acids, coco, reaction products with 2-[(2-aminoethyl)amino]ethanol, alkylation products with methyl acrylate, sodium salts
68442091	Naphthalene sulfonic acid, sodium salt, isopropylated	Naphthalene sulfonic acid, isopropylisohexyl-, sodium salt

[1]Based on the 2002 TSCA IUR.

[2]The number of inert ingredients contained in one or more registered pesticide products as of April 27, 2007. Note that as new products and formulations are registered, and as other products are canceled or reformulated, the number of inert ingredients contained in one or more registered pesticide products can change

3.0 SUMMARY STATISTICS FOR DATA COLLECTED

A summary of data sources and data manipulation is provided in Section 4.0 of this report. A unique list of HPV/pesticide inert chemicals detected in monitoring samples in the data sources evaluated was generated for each exposure pathway. Each list only contains chemicals that are on both HPV and pesticide inert chemical lists. Each of the exposure pathway lists were then combined to generate a unique list of HPV/pesticide inert chemicals represented by all exposure pathways. Table 3-1 below presents the number of unique HPV/pesticide inert chemicals included in each list.

Table 3-1
Summary of the Number of HPV/Pesticide Inert Chemicals on the Pathway Lists

Exposure Pathway	Number of Unique HPV/Pesticide Inert Chemicals
Overall Combined List	62
Human Biological Monitoring Exposure Pathway List	14
Ecological Biological Monitoring Exposure Pathway List	17
Chemicals in Drinking Water Exposure Pathway List	19
Indoor Air Monitoring Exposure Pathway List	48

Table 3-2 presents summary statistics for each data source evaluated. The chemicals identified from each data source are presented in Section 4.0. Note that CAS numbers were not provided for chemicals in some data sources; therefore, CAS numbers were updated from readily available sources. To determine whether a chemical is an HPV or inert, the chemicals were matched by CAS number, followed by chemical name.

Table 3-2
Summary Chemical Counts for Data Sources Evaluated

Exposure Pathway	Data Source	No. HPV/Inerts
Human Biological Monitoring Data	NHANES – VOC	3
	NHANES – CDC	5
	NHATS	8
	TEAM – Breath	1

Table 3-2 (Continued)

Exposure Pathway	Data Source	No. HPV/Inerts
Ecological Biological Monitoring Data	NSI – Tissue	13
	NFTS	10
	NAWQA – Tissue	1
Chemicals in Drinking Water Monitoring Data	NCOD	5
	NHEXAS	0
	TEAM	1
	NASQAN – Surface Water	6
	NASQAN – Sediment	7
	NAWQA– Ground Water	9
	NAWQA – Surface Water	9
	NAWQA – Sediment	8
Indoor Air Monitoring Data	ORD – Indoor Air	48
	ORD – Personal Air	3
	NHEXAS – Indoor Air	1
	NHEXAS – Personal Air	0

4.0 SUMMARY OF DATA SOURCES AND DATA MANIPULATION

Data were obtained from data sources identified in the September 2005 FR Notice for the following four exposure pathways:

- Human Biological Monitoring Exposure Pathway;
- Ecological Biological Monitoring Exposure Pathway;
- Chemicals in Drinking Water Exposure Pathway; and
- Indoor Air Exposure Pathway.

A summary of each of the data sources evaluated and data manipulations that were performed is presented below. Descriptions of the data sources, including limitations and quality control/quality assurance information, were obtained directly from the Web sites and/or reports that are cited in the sections below.

4.1 Human Biological Monitoring Exposure Pathway

Relevant data were extracted from the following data sources to determine the presence of HPV and Inert chemicals in human biological samples:

- National Health and Nutrition Examination Survey III (NHANES III) Priority Toxicant Reference Range Study for Volatile Organic Compounds;

- Centers for Disease Control and Prevention's (CDC) Second and Third National Reports on Human Exposure to Environmental Chemicals (NHANES 1999 to 2002);

- National Human Adipose Tissue Survey (NHATS); and

- Total Exposure Assessment Methodology (TEAM) Breath Study.

4.1.1 National Health and Nutrition Examination Survey III (NHANES III) Priority Toxicant Reference Range Study for Volatile Organic Compounds

NHANES III was conducted between 1988 and 1994 and surveyed 33,994 people. Several studies (e.g., high blood pressure, immunization status, nutritional blood measures, etc.) were conducted under NHANES III. The *Priority Toxicant Reference Range Study* involved the collection of blood and urine samples from a non-random (i.e., not statistically representative of the United States) sample size of approximately 1,300 people to determine the concentration of volatile organic compounds (VOCs) and selected pesticides and their metabolites.

More recent NHANES data are available for samples collected during 1999 and 2002 in the Centers for Disease Control and Prevention's *Third National Report on Human Exposure to Environmental Chemicals* (discussed in Section 4.1.2); however, VOCs were not sampled in

the 1999 to 2002 study. Therefore, VOC data obtained from the earlier *Priority Toxicant Reference Range Study* from NHANES III were included in this analysis, and urinary concentrations of pesticide analytes were obtained from more recent NHANES data (discussed in Section 4.1.2). Note that previously, only a subset of 1988 to 1994 data were available in a journal article by Ashley et al., "Blood Concentrations of Volatile Organic Compounds in a Nonoccupationally Exposed US Population and in Groups with Suspected Exposure" published in *Clinical Chemistry* in 1994. The complete set of 1988 to 1994 data, which was published in January of 2000, were obtained and included in this analysis.

Blood samples were retrieved from 1,018 people (the total number of samples). Results could not be obtained for a certain number of specimens for specific analytes; therefore, these samples were subtracted from the total number of samples for each chemical to arrive at a number of relevant samples. Blood samples were not available for 320 persons in the sample and a urine specimen was not available to measure levels of pesticide metabolites for 360 persons in this sample (i.e., "No" samples).

4.1.1.1 Data Source Location

The NHANES III data are available for download at the following web site: http://www.cdc.gov/nchs/about/major/nhanes/nh3data.htm#nhanes%20iii%20series%2011,%20no.%204a.

4.1.1.2 Data Source Contact

U.S. Department of Health and Human Services (DHHS)
National Center for Health Statistics,
Division of Data Services
Hyattsville, MD 20782
(301) 458-4636

4.1.1.3 Data Format

NHANES III VOC data may be obtained electronically in ASCII format with a SAS readme file.

4.1.1.4 Frequency of Updates to the Data

The last NHANES III survey for VOCs was conducted between 1988 through 1994. No additional updates are expected.

4.1.1.5 Data Quality

No formal statistical sampling procedures were used to recruit volunteers for this study. Because the sample for this study is not statistically based, sample weights cannot be used in analyses of these data. As a result of this, estimates obtained from analyses of these data cannot be weighted; therefore, representative estimates for the U.S. population cannot be calculated. However, these data are a valid cross-sectional survey to assess chemical burden in the population and with appropriate statistical techniques, one can make some comparisons of pesticide levels over time.

4.1.1.6 Data Manipulation

1988-1994 data from the NHANES III study were downloaded from the web site in SAS format and were converted and imported into a database. An electronic version of the *Priority Toxicant Reference Range Study* was also obtained and was used to extract the analyte codes, names, and other information for use in the NHANES III VOC data manipulations.

The following data manipulations were performed on the imported raw NHANES III VOC data:

1. Relevant data elements were identified and included in the raw NHANES III table:

    - Respondent identification number;
    - Analyte code;
    - Analyte name; and
    - Concentration.

2. The number of detects were calculated for each analyte by counting records containing concentration amounts greater than the analyte detection limit (found in the electronic study document) and where the concentration was not a 'blank, but applicable'[3] record.

3. The number of nondetects were calculated for each analyte by counting records containing concentration amounts less than or equal to the analyte detection limit (found in the electronic study document).

4. The number of blank[4] amounts were calculated for each analyte by counting the null records for each analyte code.

---

[3]The code 'blank, but applicable' was used to identify where blood specimens were available but results were not obtained for specific analytes.

[4]Blood measures of volatile organic compounds were not obtained for 320 of the volunteers in this study. For these participants, fields for all blood measures are left blank.

5. The number of ‘blank but applicable’ records were calculated for each analyte by counting records containing a ‘blank, but applicable’ record (i.e., was coded with ‘88888' or longer string of ‘8’'s).

6. The NHANES VOC aggregated data table was created with the following data elements:

   - Detect flag;
   - Analyte code;
   - Analyte name;
   - CAS number;
   - Chemical type (i.e., ‘Urinary Phenol’ or ‘VOC’);
   - Detection limit (ug/L);
   - Square root of the detection limit;
   - Number of detects;
   - Total number of samples;
   - Number of relevant samples;
   - Number of samples below detection limit;
   - Number of “No” samples; and
   - Outside Standard Range (OSR) flag.

   The detection limit fields were populated by extracting the relevant limits for each analyte from the electronic study document.

7. Some of the analyte codes indicated that the analyte contained an outside the standard range (OSR) flag. The OSR flag in the data table was set to “Yes” for each analyte with an OSR flag.

   Some blood level VOC laboratory measures were defined as OSR for that analytical method. These results may have lower quantitative validity than values within the standard range; however, additional laboratory QC testing showed that standard curves continued to be linear at least an order of magnitude higher than the highest measured value, suggesting that results flagged as OSR are valid. OSR records were flagged, as explained above, and data were included in the analysis.

8. The Detect flag in the data table was set to “Yes” for each analyte that was detected at least once in the study (i.e., the number of detects is greater than zero). Analytes that were not detected in any samples (i.e., Detect flag = “No”) and that were urinary phenols (i.e., chemical type = ‘Urinary Phenol’)[5], as indicated in the data table, were not considered for Priority Setting.

[5]Although data were available for urinary phenols, these data were excluded from the NHANES III VOC summary since more recent data is available in the Center for Disease Control’s (CDC’s) *Second and Third National Reports on Human Exposure to Environmental Chemicals*, discussed in Section 4.1.2.

HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

4.1.2 Centers for Disease Control and Prevention's (CDC) Second and Third National Reports on Human Exposure to Environmental Chemicals (NHANES 1999 to 2002)

The CDC published three reports summarizing NHANES sampling data:

- First National Report on Human Exposure to Environmental Chemicals (issued in March 2001);
- Second National Report on Human Exposure to Environmental Chemicals (issued in March 2003); and
- Third National Report on Human Exposure to Environmental Chemicals (issued in July 2005).

Each year's report presents data from prior years, in addition to exposure data collected for current and additional chemicals studied. The First Report presented exposure data for 27 chemicals from NHANES 1999; the Second Report presented exposure data for 116 chemicals (including the 27 in the first Report) during NHANES 1999 and 2000. The Third Report presents exposure data for the U.S. population for 148 environmental chemicals over the period 2001-2002. The Third Report also includes the data from the Second Report.

The First Report measured lead, mercury, cadmium, and other metals; dialkyl phosphate metabolites of organo-phosphate pesticides; cotinine; and phthalates. The Second Report included chemicals from the First Report and also added the following chemical groups:

- Polycyclic aromatic hydrocarbons (PAHs);
- Dioxins, furans, and coplanar polychlorinated biphenyls (PCBs);
- Non-coplanar PCBs;
- Phytoestrogens;
- Selected organophosphate pesticides;
- Organochlorine pesticides;
- Carbamate pesticides;
- Herbicides; and
- Pest repellents and disinfectants.

The Third Report added the following chemical groups:

- Pyrethroid insecticides;
- Additional polycyclic aromatic hydrocarbons (including benzo-[a]-pyrene);
- Aldrin, endrin, dieldrin;
- Additional phthalate metabolites;
- Additional pesticides and herbicides; and

- Additional dioxins, furans, and polychlorinated biphenyls (PCBs).

The 2002 and 2003 reports present tables of descriptive statistics on the distribution of blood, urine, or blood serum levels for each environmental chemical. Statistics include geometric means and percentiles (i.e., 50th, 75th, 90th, and 95th) with confidence intervals. For chemicals measured in urine, two levels are presented: per volume of urine and per gram of creatinine. Levels per gram of creatinine (i.e., creatinine-corrected) adjust for urine dilution. Creatinine is excreted from the body at a relatively constant rate over time, so expressing the result per gram of creatinine helps adjust for the effects of urinary dilution.

Blood serum levels are measured for dioxins, furans, PCBs, and organochlorine pesticides and are presented per gram of total lipid. These compounds are lipophilic and concentrate in the body's lipid stores including the lipid in the serum. Levels for these compounds are expressed as per gram of total lipid in the serum because the chemicals reside in this part of the serum.

Concentrations less than the limit of detection (LOD) were assigned a value equal to the LOD divided by the square root of 2 for calculation of geometric means. Percentile estimates that are less than the LOD for the chemical analysis are reported as < LOD. If the proportion of results below the LOD was greater than 40%, geometric means were not calculated.

For dioxins, furans, PCBs, organochlorine pesticides, and several additional pesticides, each individual sample has its own LOD. For these chemicals the mean LOD with a standard deviation (SD) and a maximum LOD is presented. The maximum LOD was the highest LOD among all the individual samples analyzed.

4.1.2.1 Data Source Location

Reports on NHANES data are available for download at the following web site: http://www.cdc.gov/exposurereport.

4.1.2.2 Data Source Contact

John Osterloah, Chief Medical Officer
Centers for Disease Control and Prevention
National Center for Environmental Health
Division of Laboratory Sciences, Mail Stop F-20
4770 Buford Highway, N.E. Atlanta, Georgia 30341-3724
Telephone (toll-free): 1-866-670-6052
E-mail: ncehdls@cdc.gov

Susan Schrober
Centers for Disease Control and Prevention
National Center for Health Statistics
(301) 458-4484

4.1.2.3 Data Format

NHANES data are contained in a report in portable document format (PDF).

4.1.2.4 Frequency of Updates to the Data

NHANES became a continuous and annual survey beginning in 1999. It is expected that a report will be generated following completion of each subsequent NHANES survey. Future reports are expected to be released every two years. The next report is expected to cover 2003 and 2004.

4.1.2.5 Data Quality

The blood, urine, and blood serum analyses presented in the reports were made at CDC's Environmental Health Laboratory (Division of Laboratory Sciences, National Center for Environmental Health). The analytical methods used for measuring these environmental chemicals or their metabolites in blood and urine were isotope dilution mass spectrometry, inductively coupled plasma mass spectrometry, or graphite furnace atomic absorption spectrometry.

The NHANES sample design is complex; therefore, sample weights must be used to account for the unequal probability of selection into the survey. Sample weights also are used to adjust for possible bias resulting from non-response and are post-stratified to U.S. Census Bureau estimates of the U.S. population. Data were analyzed using the statistical software package from Statistical Analysis System (SAS)® (SAS Institute Inc., 2002) and the statistical software package SUDAAN® (SUDAAN Release 8.0, 2001). SUDAAN uses sample weights and calculates variance estimates that account for the complex survey design.

Laboratory Quality Control

As part of the overall quality assurance process for the survey, all collection materials, vacutainer tubes, and storage containers used for trace elements assays were initially prescreened by the CDC/NCEH, Environmental Health Laboratory Sciences Laboratory for background contamination levels of lead, cadmium, total and speciated mercury. Lead, cadmium, and total and speciated mercury are fairly ubiquitous contaminants; and blood may be collected in red-top tubes after the acceptability of the test tubes has been confirmed. Special lead-free tubes are not required. Ordinary EDTA tubes may similarly be used after prescreening has confirmed no contamination.

The National Center for Health Statistics (NCHS) uses several methods to monitor the quality of the analyses performed by the NHANES contract laboratories. These methods include performing second examinations on previously examined participants and "blind" split samples collected during practice ("dry run") sessions. In addition, contract laboratories randomly perform repeat testing on two percent of all specimens.

4.1.2.6 Data Manipulation

The NHANES Second Report was downloaded and information was data entered into a database. Results contained in the 2003 Report were presented in a table for each chemical. Data were presented as totals and also subdivided by age, gender, and race/ethnicity. Only data representing total results were extracted and entered into the database.

The following data fields were extracted:

- Chemical name;
- CAS number;
- Sample size;
- Geometric mean;
- 10^{th} percentile;
- 25^{th} percentile;
- 50^{th} percentile;
- 75^{th} percentile;
- 90^{th} percentile;
- 95^{th} percentile;
- Media (i.e, blood, urine, or serum);
- Units;
- Level of detection (LOD); and
- LOD flag.

The NHANES Third Report was downloaded and results contained in the Third Report were also presented in a table for each chemical. Data were presented as totals and also subdivided by age, gender, and race/ethnicity. Only data representing total results were extracted for the analysis. The following data fields were extracted:

- Chemical name;
- CAS number;
- Sample size;
- Geometric mean;
- 50^{th} percentile;
- 75^{th} percentile;
- 90^{th} percentile;
- 95^{th} percentile;
- Media (i.e, blood, urine, or serum);
- Units;
- Level of detection (LOD); and
- LOD flag.

Note that in the Third Report, all variance estimates (both 1999-2000 and 2001-2002 data) were calculated using the Taylor series (linearization) method within SUDAAN. In the Second Report, 1999-2000 variance estimates were calculated using the jackknife method. The two methods produce very similar, but not identical, variance estimates. Consequently, some confidence intervals for 1999-2000 presented in the Second Report will differ slightly from

confidence intervals for the same time period presented in the Third Report. The 1999-2000 data extracted for use in priority setting are based on the Second Report.

1. For the purposes of this analysis, a chemical was considered detected if a geometric mean was presented in the report.

2. A detect flag was added to the aggregate data table and was updated to "Yes" for each chemical with a geometric mean presented.

   In some instances metabolites of parent chemicals were measured in participants as surrogates of exposure to the parent chemical. A comparison between parent chemicals and corresponding metabolites measured was created. This table was used to ensure that each parent chemical could be associated with a measured metabolite where applicable.

HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

4.1.3 National Human Adipose Tissue Survey (NHATS)

The National Human Adipose Tissue Survey (NHATS) was the primary activity of the National Human Monitoring Program (NHMP), operated by EPA's Office of Pollution Prevention and Toxics (USEPA/OPPT), until the early 1990s. NHATS analyzed human adipose tissue specimens in order to monitor human exposure to potentially toxic chemicals. A nationwide network of pathologists and medical examiners from 47 metropolitan statistical areas collected tissue specimens from cadavers and surgical patients that were analyzed for certain chemicals. Throughout the 1970s and early 1980s, the chemical residues of primary interest were organochlorine pesticides and polychlorinated biphenyls (PCBs). In 1982, volatile organic compounds (VOCs) and semivolatile organic compounds (SVOCs) were included in the survey.

NHATS results were published in 14 journal articles and reports. NHATS data were previously compiled into a database by EPA. This database was obtained from EPA in September 1999. The NHATS database includes summarized data from 1970 through 1987 and provides information for 134 chemicals.

4.1.3.1 Data Source Location

The NHATS raw data was obtained directly from EPA. The following reports summarizing NHATS data can be obtained from http://www.regulations.gov, Docket Id. No. EPA-HQ-OPPT-2002-0066:

1) U.S. EPA. Chlorinated Dioxins and Furans in the General U.S. Population: NHATS FY87 Results - Executive Summary. EPA-560/5-91-003. May 1991.

2) Cramer, Paul H.; Stanley, John S.; Bauer, Karin; Ayling, Randy E.; Thornburg, Kelly R.; and Schwemberger, John. Brominated Dioxins and Furans in Human Adipose Tissue: Final Report. EPA-560/5-90-005 (NTIS PB91-103507). April 11, 1990.

3) Cramer, Paul H.; Stanley, John S.; and Thornburg, Kelly R. Mass Spectral Confirmation of Chlorinated and Brominated Diphenylethers in Human Adipose Tissues: Final Report. EPA-560/5-90-012 (NTIS PB91-159699). June 15, 1990.

4) Mack, Gregory A. and Mohadjer, Leyla. Baseline Estimates and Time Trends for Beta-benzene hexachloride, Hexachlorobenzene, and Polychlorinated Biphenyls in Human Adipose Tissue 1970-1983. EPA-560/5-85-025. September 30, 1985.

5) Onstot, J.D.; Ayling, R.E.; and Stanley, J.S. Characterization of HRGC/MS Unidentified Peaks from the Analysis of Human Adipose Tissue: Volume I - Technical Approach. EPA-560/5-87-002A (NTIS PB88-100367). May 1987.

6) Onstot, J.D.; Ayling, R.E.; and Stanley, J.S. Characterization of HRGC/MS Unidentified Peaks from the Analysis of Human Adipose Tissue: Volume II - Appendices. EPA-560/5-87-002B (NTIS PB88-100375). May 1987.

7) Onstot, J.D. and Stanley, J.S. Identification of SARA Compounds in Adipose Tissue. EPA-260/5-89-003 (NTIS PB90-132564). August 1989.

8) Orban, John E.; Stanley, John S.; Schwemberger, John G.; and Remmers, Janet C. Dioxins and Dibenzofurans in Adipose Tissue of the General US Population and Selected Subpopulations. American Journal of Public Health. (1994) 84: 439-445.

9) U.S. EPA. Semivolatile Organic Compounds in the General U.S. Population: NHATS FY86 Results - Volume I. EPA-747-R-94-001. July 1994.

10) Stanley, John S. Broad Scan Analysis of the FY82 National Human Adipose Tissue Survey Specimens: Volume I - Executive Summary. EPA-560/5-86-035 (NTIS PB87-177218). December 1986.

11) Stanley, John S. Broad Scan Analysis of the FY82 National Human Adipose Tissue Survey Specimens: Volume II - Volatile Organic Compounds. EPA-560/5-86-036 (NTIS PB87-177226). December 1986.

12) Stanley, John S. Broad Scan Analysis of Human Adipose Tissue: Volume III - Semivolatile Organic Compounds: Final Report. EPA-560/5-86-037 (NTIS PB87-180519). December 1986.

13) Stanley, John S. Broad Scan Analysis of Human Adipose Tissue: Volume IV -

Polychlorinated Dibenzo-p-Dioxins (PCDDs) and Polychlorinated Dibenzofurans (PCDFs): Final Report. EPA-560/5-86-038 (NTIS PB87-177234). December 1986.

14) Stanley, John S. and Stockton, Rodney A. Broad Scan Analysis of the FY82 National Human Adipose Tissue Survey Specimens: Volume V - Trace Elements. EPA-560/5-86-039 (NTIS PB87-180527). December 1986.

4.1.3.2 Data Source Contact

Data may be extracted from the above-referenced reports.

4.1.3.3 Data Format

NHATS data were provided in electronic database format by EPA. EPA obtained NHATS electronic study data files in 1997 from Battelle, Inc. and compiled a subset of these data into database compatible files. Original data are also available in hard copy format in the reports listed above in the "Data Source Location" section.

4.1.3.4 Frequency of Updates to the Data

The last NHATS survey was conducted in 1992. NHATS monitoring of human tissue for chemical contaminants was discontinued due to budget cutbacks. Human tissue monitoring data are available if chemicals were analyzed during the final 1986 survey. No additional updates are expected.

4.1.3.5 Data Quality

Quality control samples, such as method and equipment blank samples, control samples, and spike samples, were collected to evaluate the quality of the sampled data. Data are available for years 1970 through 1987; however, a standard set of summarized data parameters are not available.

4.1.3.6 Data Manipulation

1970-1987 NHATS raw data were obtained from EPA. The original raw data obtained from EPA were extracted and the following data elements were used to create the NHATS aggregated data table:

- Chemical name;
- Number of samples;
- CAS number;

- Proportion detected (%);
- Mean level of detection (ug/g); and
- Year.

The following data manipulations were performed on the imported raw NHATS data:

1. The number of measured samples was added to the aggregated data table and calculated for each record using the following equation:

   Number measured = Total number of samples x (% measurable x 0.01)

2. The detect flag was added to the aggregated data table and updated for each record having a *number measured samples* calculated to be greater than zero (see Step 1 above).

HPV/pesticide inert chemicals that were measured in monitoring samples from this data source were considered for priority setting purposes.

4.1.4 Total Exposure Assessment Methodology (TEAM) Breath Study

The TEAM study was designed to develop methods to measure individual total exposure (exposure through air, food, and water) and resulting body burden of toxic and carcinogenic chemicals. These methods could then be applied within a probability-based sampling framework to estimate the exposures and body burdens of urban populations in several U.S. cities. The TEAM study reports the results of eight monitoring studies performed in five communities during different seasons of the year.

Breath, personal and outdoor air samples, and water samples were collected and subsequently reported in a 1987 four-volume report entitled: *The Total Exposure Assessment Methodology (TEAM) Study*. Two of the four volumes provide data in a form useful for EDSP priority setting activities:

The Total Exposure Assessment Methodology (TEAM) Study: Elizabeth and Bayonne, New Jersey, Devils Lake, North Dakota, and Greensboro, North Carolina: Volume II, Part 2; and

The Total Exposure Assessment Methodology (TEAM) Study: Selected Communities in Northern and Southern California: Volume III.

These volumes include data collected at the following locations and times:

- New Jersey in the fall of 1981, spring/summer of 1982, and winter of 1983;
- North Carolina in May 1982;
- North Dakota in October 1982;
- Los Angeles, California in February and May of 1984; and
- Contra Costa County, California in June of 1984.

4.1.4.1 Data Source Location

The TEAM data were obtained from Volumes II and III of the 1987 *TEAM Study*, as discussed above.

4.1.4.2 Data Source Contact

The citation of the relevant volumes of the *TEAM Study* are provided below. These reports can be ordered from the National Technical Information Service (NTIS) at (800) 553-6847.

- U.S. EPA. The Total Exposure Assessment Methodology (TEAM) Study: Elizabeth and Bayonne, New Jersey, Devils Lake, North Dakota, and Greensboro, North Carolina: Volume II. Part 2. EPA-600/6-87/002b (NTIS PB88–100078). June 1987.

- U.S. EPA. The Total Exposure Assessment Methodology (TEAM) Study: Selected Communities in Northern and Southern California: Volume III. EPA-600/6-87/002c (NTIS PB88–00086). June 1987.

- Wallace, Lance. Project Summary: The Total Exposure Assessment Methodology (TEAM) Study. EPA/600/S6-87/002. September 1987.

4.1.4.3 Data Format

The raw TEAM data are available in hard copy format in the two volumes of the 1987 *TEAM Study* previously discussed. These data were manually extracted from these hard copies and compiled into a spreadsheet.

4.1.4.4 Frequency of Updates to the Data

The TEAM Study was published in 1987. No additional updates are expected.

4.1.4.5 Data Quality

All of the TEAM data appear to be of high quality. Established methods were used for collecting and analyzing data and are well documented throughout the 1987 *TEAM Study*. Quality control and quality assurance samples were collected and analyzed, including reagent blanks, field blanks, duplicate samples, and spiked samples.

In addition, the TEAM data that were manually extracted from the hard copy documents were verified against the original documents to ensure the quality of the electronic data collection used for priority setting purposes.

4.1.4.6 Data Manipulation

Breath data collected from 1981 to 1984 that were contained in two hard-copy volumes of *The TEAM Study* were manually extracted and saved to a spreadsheet.

The raw TEAM breath data were extracted, and the following data elements are included in the TEAM breath aggregated data table:

- Compound;
- CAS number;
- Arithmetic mean of the concentration (ug/m^3);
- Geometric mean of the concentration (measured or quantifiable limit) (ug/m^3);
- Units (ug/m^3);
- Central tendency measurement label (e.g., 'Geometric Mean' or 'Minimum Quantifiable Limit');
- Total number of samples;
- Percent measurable;
- Number of measured samples;
- Site (e.g., Contra Costa County, CA ('CC'), Los Angeles, CA ('LA'), 'NJ', 'NC', or 'ND');
- Season;
- Year;
- Inhalation rate (m^3/day);
- Total load per hour (mg/hour); and
- Detect flag.

The following data manipulations were performed on the raw TEAM breath data:

1. The number of measured samples was added to the aggregated data table and calculated for each record using the following equation:

   Number measured = Total number of samples x (% measurable x 0.01)

2. The detect flag was added to the aggregated data table and updated for each record having a *number measured samples* calculated to be greater than zero (see Step 1 above).

HPV/pesticide inert chemicals that were measured in monitoring samples from this data source were considered for priority setting purposes.

4.1.5 Summary of Human Biological Monitoring Pathway Chemicals

Table 4-1 presents the list of HPV/pesticide inert chemicals that were detected in monitoring samples from the Human Biological Monitoring Pathway data sources, along with an indication of the data source in which they appeared.

Table 4-1
HPV/Pesticide Inert Chemicals Present in Human Biological Monitoring Pathway Data Sources

Chemical Name	CAS Number	Total Number of Data Sources	NHANES VOC	NHANES CDC	NHATS	TEAM-Breath
Butyl benzyl phthalate	85687	2	No	Yes	Yes	No
Dibutyl phthalate	84742	2	No	Yes	Yes	No
Diethyl phthalate	84662	2	No	Yes	Yes	No
Acetone	67641	1	Yes	No	No	No
Bicyclo[3.1.1]hept-2-ene, 2,6,6-trimethyl-	80568	1	No	No	No	Yes
Cyclotetrasiloxane, octamethyl-	556672	1	No	No	Yes	No
di-sec-octyl phthalate	117817	1	No	Yes	No	No
Dimethyl phthalate	131113	1	No	Yes	No	No
Hexanedioic acid, bis(2-ethylhexyl) ester	103231	1	No	No	Yes	No
Isophorone	78591	1	No	No	Yes	No
Methyl ethyl ketone	78933	1	Yes	No	No	No
Phenol, 2,6-bis(1,1-dimethylethyl)-4-methyl-	128370	1	No	No	Yes	No
Phosphoric acid, triphenyl ester	115866	1	No	No	Yes	No
Toluene	108883	1	Yes	No	No	No

4.2 Ecological Biological Monitoring Exposure Pathway

Relevant data were extracted from the following data sources to determine the presence of HPV/pesticide inert chemicals in ecological tissues:

- National Sediment Quality Database: 1980 to 1999 [or National Sediment Inventory (NSI)] Tissue Data;
- National Fish Tissue Study (NFTS) Data; and
- National Water Quality Assessment (NAWQA) Program Aquatic Animal Tissue Data.

4.2.1 National Sediment Inventory (NSI) Tissue Data

EPA's Office of Science and Technology initiated the National Sediment Inventory (NSI), in order to compile sediment quality information from available electronic databases into one centralized, easily accessible location. The NSI database includes approximately 4.6 million records for more than 50,000 monitoring stations across the country for the following:

- Sediment chemistry - measures the chemical concentration of sediment-associated contaminants;
- Tissue residue - measures chemical contaminants in the tissue of organisms, and
- Toxicity - measures the lethal and sublethal effects of contaminants in environmental media on various test organisms.

To efficiently collect usable information for inclusion in the NSI, EPA sought data that were available in electronic format, represented broad geographic coverage, and represented specific sampling locations identified by latitude and longitude coordinates. At a minimum, EPA required that electronically available data include the monitoring program, sampling date, latitude and longitude coordinates, and measured units for inclusion in the Inventory. The NSI includes data from the following data storage systems and monitoring programs:

- Selected data sets from EPA's Storage and Retrieval System (STORET);
- National Oceanic and Atmospheric Administration's (NOAA) Query Manager Data System;
- State of Washington Department of Ecology's Sediment Quality Information System (SEDQUAL);
- Selected data sets from USGS's WATer Data STOrage and REtrieval System (WATSTORE);

- EPA's Environmental Monitoring and Assessment Program (EMAP);
- Data compiled for the previous 1998 Report to Congress - entitled "The Incidence and Severity of Sediment Contamination in Surface Waters of the United States;"
- Chesapeake Bay Program;
- Upper Mississippi River System data compilation prepared by the USGS;
- Indiana Department of Environmental Management Sediment Sampling Program;
- Oklahoma Reservoir Fish Tissue Monitoring Program, 1990-1998; and
- Houston Ship Channel Toxicity Study.

4.2.1.1 Data Source Location

Information on NSI data may be accessed at: http://www.epa.gov/waterscience/cs/nsidbase.html.

4.2.1.2 Data Source Contact

Bob Shippen
Standards & Health Protection Division
Office of Science and Technology, Mail Code 4305T
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460
Phone: (202) 566-0391
E-mail: shippen.robert@epa.gov

4.2.1.3 Data Format

NSI data are contained in an Access database: National Sediment Quality Survey Database - 1980 to 1999. Limited copies of the database are available on CD.

4.2.1.4 Frequency of Updates to the Data

Additional updates to this data set are not expected.

4.2.1.5 Data Quality

This database contains environmental monitoring data from a variety of sources. Most of the monitoring data are from state and federal monitoring programs that require documented quality assurance and quality control procedures. To be included in this database, adequate information on the data source, sampling date, analyte measured, units, and data qualifiers (e.g., "not-detected") were required. EPA compiled the data using a number of different sampling strategies. Component sources contain data from different spatial sampling plans, sampling methods, and analytical methods.

4.2.1.6 Data Manipulation

The National Sediment Quality Survey Database was obtained. Supporting database documentation was reviewed to identify relevant data tables for NSI tissue data. The following tables were identified:

- (1)SITE - Site;
- (2)STUDY - Study;
- (3a)STATION - Station;
- (4c)SMPTISS - Tissue sample;
- (5c)CHEMTISS - Tissue residue;
- (8a)CHEMDICT - Dictionary of chemical;
- (8b)QUALIFY - Concentration qualifier;
- (8c)SPECIES - Dictionary of species; and
- (8e)TISSTYPE - Type of tissue.

The following data manipulations were performed on the raw data tables summarized above:

1. The raw data files described above were incorporated into two databases, one for sediment and one for sub-sediment. The specific source files for tissue data are (4c)SMPTISS and (5c)CHEMTISS.

2. The raw data files were combined into a comprehensive data set that included the following data elements:

    - Site identifier;
    - Study identifier code;
    - Station identifier;
    - Sample identifier;
    - Date sample collected;
    - Species code;
    - Tissue code;
    - Percent lipid;
    - Tissue chemistry flag;
    - Abbreviated chemical name;

- Measured concentration;
- Assigned qualifier for concentration;
- Units of concentration for chemical;
- Data missing, Y or N?;
- Full chemical name;
- CAS number;
- Descriptive name of site;
- U.S. EPA Region of site location: 11 for Canada;
- County; and
- State.

3. Additional data fields that were added to the aggregated data table include:

   - Detect flag; and
   - Fish flag.

4. A final data table was created that contained a subset of data where the sample date is >= 1990. EPA determined that only samples taken from 1990 to 1999 should be included. Note that sample date information was not available for a small fraction of samples (0.17 percent); therefore, these samples were excluded from the analysis.

5. Chemical name and CAS numbers were reviewed for completeness.

6. The *(8c)SPECIES* file was reviewed to identify[6] and flag fish species. This file was then used to update the data table's fish flag to "Yes".

7. The detect flag was updated to "Yes" where the assigned qualifier for concentration was one of the following:

   - # - Use data without qualification;
   - D - Diluted analysis; and
   - J - Estimate.

   The following qualifiers were considered non-detects:

   - < - Detected, but less then calculated method detection limit;
   - <J - Estimate, less than;
   - ND - Analyte not detected; and
   - NJ - Estimated, non-detect.

[6]Species were initially identified based on common names (e.g., fish, darters, shiners, sculpin, chub, etc.). FishBase, a web site with a search engine powered by a relational database of fish information at: http://www.fishbase.org/home.htm, was used to identify any additional fish species (Froese, R. and D. Pauly. Editors. 2005. FishBase).

HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

4.2.2 National Fish Tissue Study (NFTS) Data

EPA's Office of Water's Office of Science and Technology conducted a screening-level study to estimate the national distribution of selected persistent, bioaccumulative and toxic chemical residues in fish tissue from lakes and reservoirs of the continental United States. EPA worked with partner agencies over a four-year period (2000 - 2003) to collect fish from 500 randomly selected lakes and reservoirs of the estimated 260,000 lakes and reservoirs in the continental United States. The lakes were divided into six size categories, ranging from 2.5 to over 900,000 surface acres, with a similar number of lakes in each category.

The National Fish Tissue four-year study defines the national background levels for 265 chemicals in fish, establishes a baseline to track the progress of pollution control activities, and identifies areas where contaminant levels are high enough to warrant further investigation. Analyses were conducted in fish tissues for the following chemicals:

- Mercury;
- Arsenic;
- 17 dioxins and furans;
- 159 PCB congeners;
- 43 pesticides; and
- 40 other organics (e.g., phenols).

The study involved the collection of predator and bottom-dwelling fish from 500 randomly selected lakes and reservoirs of the continental United States (excluding the Great Lakes) over a period of 4 years (~125 lakes per year). Sampling teams applied consistent methods nationwide to collect composites of one predator species and one bottom-dwelling species at each lake. Composites consisted of 5 adult fish of similar size that were large enough to provide 560 grams (20 ounces) of tissue for analysis of fillets for predators and whole bodies for bottom dwellers.

Although the study was to begin during the summer of 1999, full implementation did not commence until 2000. For this reason, samples collected during the 1999 mobilization and 2000 implementation periods cumulatively represent Year 1 of the Study. The study design resulted from a comprehensive planning effort that included a national workshop involving more than 50 scientists from state, federal, and tribal agencies to obtain technical input on sampling design, target analytes, sampling methods and data management.

4.2.2.1 Data Source Location

National Fish Tissue data information can be accessed at: http://www.epa.gov/waterscience/fishstudy/overview.htm. Data can be requested from the contact listed below.

4.2.2.2 Data Source Contact

Leanne Stahl
U.S. EPA, Office of Science and Technology (4305T)
1200 Pennsylvania Avenue, NW
Washington, DC 20460
Phone: (202) 566-0404
E-mail: stahl.leanne@epa.gov

4.2.2.3 Data Format

NFTS data are available in an electronic format. Data are contained on two CDs in Microsoft Excel spreadsheets.

4.2.2.4 Frequency of Updates to the Data

Data for Years 1, 2, 3, and 4 of the Study are currently available.

4.2.2.5 Data Quality

Both field and laboratory quality control procedures were implemented. These are described below. In addition, data received were subject to three-levels of review.

Field QC

- Development and implementation of the Sample Collection Activities QAPP, field sampling plan and SOPs.
- Collection of replicate fish samples on 10% of the lakes.
- Use of experienced fisheries biologists to ensure use of proper procedures.
- Distribution of standardized sampling kits to control contamination and ensure proper documentation.
- Daily tracking and coordination of sample shipments through a centralized source.
- Regional orientation/training workshops to ensure all field personnel understood objectives and design of study and to ensure consistent application of required sample collection, handling, and shipping procedures.

Laboratory QC

- Development and implementation of the Analytical Control and Assessment Activities QAPP.

- Use of centralized Sample Prep Laboratory to minimize variability during sample grinding, homogenizing, and compositing.

- Identification of quantifiable measurement quality objectives (MQOs).

- Implementation of standardized sample tracking, lab analysis, data reporting, and data review procedures

- Use of pure and traceable reference standards.

- Demonstration of instrument calibration and system performance.

- Periodic calibration verification.

- Verification that each laboratory could achieve the required detection and quantitation levels.

- Analysis of initial and ongoing QC samples to demonstrate each laboratory's ability to achieve precise and accurate results with the method.

- Analysis of blanks to demonstrate the absence of contamination.

All analytical data generated during Year 1 through Year 4 of the study were subjected to the three levels of review described below:

- A pre-qualification review was performed on data submitted by each laboratory to demonstrate that the labs were qualified to prepare and/or analyze tissue samples collected during the study.

- Each submission of sample results was carefully scrutinized to verify that the samples were analyzed as directed and that supporting QC results demonstrated the quality of results generated. In evaluating these submissions, data reviewers employed a suite of standardized data qualifiers and abbreviated qualifier codes to consistently and accurately document the quality of all data generated so that both the primary data users (statisticians) at EPA Headquarters and secondary data users within states, tribes, and other organizations could make informed decisions regarding their use.

- A third level of data review was performed at the conclusion of the data review process to determine if overall data quality supported study objectives. These end-of-year evaluations indicated that all MQOs were met for every year of the study.

4.2.2.6 Data Manipulation

1. Data for Years 1 and 2 of the Study were initially obtained and data for Years 3 and 4 of the Study were subsequently obtained when they were made available. Supporting data documentation was reviewed to identify relevant data. Data for Years 1 and 2 were contained in MS Excel spreadsheets, organized by state and year. Data for Years 3 and 4 were combined into one MS Excel Spreadsheet.

2. Each spreadsheet for Years 1 and 2 contained the following four tabs (where "XX" represents the state abbreviation):

    - XX_RESULTS - worksheet contains the analytical results for all of the composite samples;

    - XX_SAMPLE_DESCRIPTION - worksheet describes the individual samples used to create the composite samples;

    - XX_SAMPLER_INFO - worksheet describes the information related to the field sampler responsible for the collection of samples at a particular water body; and

    - XX_TRIP_INFO - worksheet describes the information related to the sampling trip made to a particular water body.

3. The spreadsheet for Years 3 and 4 contained the following six tabs:

    - YEAR3_RESULTS(A-M) - worksheet contains the analytical results from Year 3 for all of the composite samples in states beginning with letters A through M;

    - YEAR3_RESULTS(N-Z) - worksheet contains the analytical results from Year 3 for all of the composite samples in states beginning with letters N through Z;

    - YEAR3_SAMPLE_DESCRIPTION - worksheet describes the individual samples from Year 3 used to create the composite samples;

    - YEAR4_RESULTS(A-M) - worksheet contains the analytical results for all of the composite samples in states beginning with letters A through M;

    - YEAR4_RESULTS(N-Z) - worksheet contains the analytical results for all of the composite samples in states beginning with letters N through Z;

    - YEAR4_SAMPLE_DESCRIPTION - worksheet describes the individual samples from Year 4 used to create the composite samples.

4. Supporting data documentation was reviewed to identify relevant data. Raw spreadsheet data files were imported into a database. "Results" and "Sample Description" worksheets were imported for all years.

5. The individual worksheets for Years 1 through 4 were combined into two data tables, one containing Results data and one containing Sample Description data.

6. An aggregated data table was created based on the Results data that includes the following relevant data elements:

   - Analyte;
   - CAS number;
   - Sample number;
   - Composite sample ID;
   - Fish species;
   - Site name;
   - State;
   - Sample date;
   - Sample type;
   - Concentration amount;
   - Detection limit;
   - Units; and
   - Data qualifiers (SCC Code).

7. The species, sample date, and sample type information was updated from the Sample Description data by linking on sample number and composite sample number.

8. A listing of all data qualifier flags used in the data set was reviewed. Records with data qualifiers indicating an estimated value were considered valid samples. Only standard samples were included in this analysis (i.e., quality assurance samples were excluded). NFTS concentration data are only provided for relevant samples.

HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

4.2.3 National Water Quality Assessment (NAWQA) Program Aquatic Animal Tissue Data

The U.S. Geological Survey (USGS) began its NAWQA program in 1991, systematically collecting chemical, biological, and physical water quality data from 42 study units (basins) across the nation. A study unit is a major hydrologic system in which NAWQA studies are focused, is geographically defined by a combination of groundwater and surface water features, and usually encompasses more than 10,000 square miles. USGS collects data from all study units, and uses them, in addition to information from other programs, agencies,

and researchers, to produce a national assessment or "national synthesis." The data warehouse currently contains the following data:

- Chemical concentrations in water, bed sediment, and aquatic organism tissues for approximately 500 chemical constituents;
- Site, basin, well, and network characteristics with descriptive variables;
- Daily stream flow information for fixed sampling sites;
- Ground water levels for sampled wells;
- 6,400 surface water sites and 7,000 wells;
- 44,000 nutrient samples, 26,000 pesticide samples, and 8,000 VOC samples; and
- 2,650 samples of bed sediment and aquatic organism tissues.

4.2.3.1 Data Source Location

NAWQA data may be queried and downloaded at: http://infotrek.er.usgs.gov/servlet/page?_pageid=543&_dad=portal30&_schema=PORTAL30.

4.2.3.2 Data Source Contact

Sandy Williamson, National Database Team Leader
USGS Wisconsin District
Office: (253) 428-3600 ext. 2683

4.2.3.3 Data Format

NAWQA data can be retrieved from the NAWQA website in tab- or comma-delimited format.

4.2.3.4 Frequency of Updates to the Data

Data are continuously updated.

4.2.3.5 Data Quality

The similar design of each investigation and use of standard methods allows for comparisons among the results measured at the various study units. All ground water samples

are analyzed at the USGS's National Water Quality Laboratory. Quality control check samples are run daily.

4.2.3.6 Data Manipulation

National Water Quality Assessment (NAWQA) data available since the program's inception were obtained from the USGS. Supporting database documentation was reviewed to identify relevant data tables. The following fields were identified, and the raw data files were imported into a database and were aggregated into a data table.

- State Postal Code;
- County;
- Study Unit Id;
- Station Id;
- Land Use Group;
- Land Use Code;
- Scientific Name;
- Common Name;
- Biological Part;
- Sample Medium Description;
- Result Datetime;
- Parameter Code;
- Parameter Name;
- Report Units;
- Value Remark; and
- Value.

This analysis was conducted for raw data where the sample medium description = "ANIMAL TISSUE". The following data manipulations were performed on the aggregated data table:

1. The following fields were added to the data table:

   - Chemical name;
   - CAS number; and
   - Detect?.

2. Parameter chemical names and CAS numbers were downloaded from the following USGS website: http://infotrek.er.usgs.gov/traverse/f?p=NAWQA:23:2887845841475415. Entering "%" in Search Parameter Codes will provide a list of all parameters. A unique list of chemical names and updated CAS numbers from various sources was created.

3. The detect flag was updated to "Yes" where the concentration value remark was not " < " (indicating a detection below the detection limit). Concentration value

remarks of "E" were considered detects. "E" represents estimated values for all detections that were below the maximum detection limit, above the highest calibration standard, or otherwise less reliable than average because of sample- or compound-specific considerations.

HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

4.2.4 Summary of Ecological Biological Monitoring Exposure Pathway Chemicals

Table 4-2 presents the list of HPV/pesticide inert chemicals that were detected in monitoring samples from the Ecological Biological Monitoring Exposure Pathway data sources, along with an indication of the data source in which they appeared.

Table 4-2
HPV/Pesticide Inert Chemicals Present in Ecological Biological Monitoring Exposure Pathway Data Sources

Chemical Name	CAS Number	Total Number of Data Sources	NSI Tissue	NAWQA Tissue	NFT Study
Butyl benzyl phthalate	85687	2	Yes	No	Yes
di-sec-octyl phthalate	117817	2	Yes	No	Yes
Dibutyl phthalate	84742	2	Yes	No	Yes
Diethyl phthalate	84662	2	Yes	No	Yes
Dimethyl phthalate	131113	2	Yes	No	Yes
Isophorone	78591	2	Yes	No	Yes
Phenol	108952	2	Yes	No	Yes
2-Pentanone, 4-methyl-	108101	1	Yes	No	No
3-Cyclohexene-1-methanol, .alpha.,.alpha.,4-trimethyl-	98555	1	No	No	Yes
Acetone	67641	1	Yes	No	No
Benzene, 1-methyl-4-(1-methylethyl)-	99876	1	No	No	Yes
Benzenemethanol	100516	1	Yes	No	No
Benzoic acid	65850	1	Yes	No	No
Methyl ethyl ketone	78933	1	Yes	No	No
Phenol, 4-nitro-	100027	1	No	No	Yes
Toluene	108883	1	Yes	No	No
Water	7732185	1	No	Yes	No

4.3 Chemicals in Drinking Water Exposure Pathway

Relevant data were extracted from the following data sources to determine the presence of HPV/pesticide inert chemicals in drinking water:

- National Contaminant Occurrence Database (NCOD);
- National Human Exposure Assessment Survey (NHEXAS) Drinking and Tap Water;
- TEAM Drinking Water Data;
- National Stream Quality Accounting Network (NASQAN) Surface Water and Sediment Data; and
- National Water Quality Assessment (NAWQA) Ground Water, Surface Water, and Sediment Data.

4.3.1 National Contaminant Occurrence Database (NCOD)

NCOD is a repository of drinking water quality data, mandated by Congress in the 1996 Safe Drinking Water Act (SDWA) Amendments. The database contains physical, chemical, microbial, and radiological information for monitored contaminants. The primary goals of NCOD are to:

- Identify contaminants for the Candidate Contaminant List;
- Select contaminants for future regulation;
- Develop new National Primary Drinking Water Regulations (NPDWRs);
- Revise existing NPDWRs; and
- Provide drinking water information to the general public.

Drinking water sample data were collected at public water systems for both regulated and unregulated contaminants. Unregulated contaminants are those for which health based standards have not been established under SDWA at the time of monitoring, while regulated contaminants are those for which health based standards have been established.

Data for unregulated contaminants were collected in two separate efforts, Round 1 and Round 2. Round 1 data include public water system monitoring sample results for 62 (then) unregulated contaminants collected from 40 states and primacy entities between 1988 and 1992. Round 1 data were stored in the Unregulated Contaminant Monitoring Information System (URCIS) database. Round 2 data include public water system monitoring sample results for 48 (then) unregulated contaminants collected between 1993 and 1997. These data were collected from 35 states and primacy entities and were stored in the EPA Safe Drinking Water Information System (SDWIS/FED).

SDWA requires EPA to review the NPDWRs for regulated contaminants at least once every six years and revise them, if appropriate, to maintain or increase public health protection. From 1993 to 1997, EPA conducted detailed contaminant occurrence analyses for 61 regulated contaminants, using drinking water compliance monitoring data provided by a cross section of 16 states. These data were compiled into a data set for the NPDWR Six Year Review and are limited to those with confirmed drinking water source and population served information.

Note that twenty chemicals appear in both the unregulated and regulated data sets, indicating that these chemicals were regulated after Round 1 data collection efforts took place.

4.3.1.1 Data Source Location

NCOD data may be obtained from the NCOD web site at: http://www.epa.gov/safewater/data/ncod/index.html.

4.3.1.2 Data Source Contact

Additional information regarding NCOD data can be obtained by contacting EPA's Office of Ground Water and Drinking Water (OGWDW).

4.3.1.3 Data Format

Data are available for download from the EPA OGWDW website as Excel Pivot Tables or tab delimited text files. Excel Pivot Tables were used in this analysis.

4.3.1.4 Frequency of Updates to the Data

NCOD drinking water data sets (Rounds 1 and 2 and Six Year Review) are static and are not updated. Round 1 data were collected between approximately 1988 and 1992, Round 2 data between 1993 and 1997, and Six Year Review data between 1993 and 1997. Data are also available from the 1999 Unregulated Contaminant Monitoring Rule (UCMR), and this data set is updated periodically as new data are received. However, monitoring for UCMR is scheduled from 2001 through 2003, and EPA did not recommend using these data for analytical purposes until the data collection effort was complete. The final data sets were not available for this analysis.

4.3.1.5 Data Quality

The NCOD drinking water data have been extensively checked for data quality and analyzed for national representativeness. Both data sets (Rounds 1 and 2 and Six-Year Review) are accompanied by a description of the data collection methods used and the quality assurance procedures followed.

Since 1992, EPA has made an effort to assure the quality of Round 1 (URCIS) data; data quality assessments and preliminary analyses of the URCIS data are presented in the occurrence and data report prepared for EPA Office of Ground Water and Drinking Water's Chemical Monitoring Revisions (CMR) project (EPA 816-R-99-006). The URCIS database was reviewed and edited to remove problematic data and ensure data quality, and some records were excluded from the data set. After editing, a second review of the 3.5 million records was undertaken, including the compilation of various descriptive statistics to enable a more detailed review for data bias and representativeness. Some incomplete state data were excluded to reduce the potential for introducing bias into the data analysis.

Significant data review, formatting, and data quality checking and editing were required of Round 2 (SDWIS/FED) data. Data records were reviewed extensively for data quality considerations including reporting consistencies, uniform and valid coding, data completeness, correct and consistent use of analytical units, and any inherent bias in the raw records. Some records were either deleted (such as when water source or system type codes were invalid) or converted (when data units conversion appeared straightforward). After this initial data editing and filtering, a basic analysis of the 4.21 million records was undertaken, similar to the Round 1 data review, and inappropriate data were excluded.

Only standard SDWA compliance samples were used in the Six-Year review; samples identified as "special," "duplicative," "investigation," or samples of unknown type were not included in the data set. Because raw data from the states were submitted in a variety of formats, each data set was examined before and after data were received and reviewed to ensure it contained the basic data elements necessary to conduct a consistent analysis. Where ambiguities or errors in certain data elements could not be resolved, those particular data elements were not included in the data set.

4.3.1.6 Data Manipulation

The following steps were taken to create an NCOD data set:

1. NCOD data for Rounds 1 and 2 were combined with data from the 6-Year Review to create a comprehensive data set that contained the following data:

    - Chemical name;
    - CAS number;
    - Number of analyses; and
    - Number of detects.

2. A flag was added to the aggregated data table to indicate whether the chemical was considered an HPV/pesticide inert.

3. A detect flag was added and updated to "Yes", as all chemicals in NCOD have detects.

HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

4.3.2 National Human Exposure Assessment Survey (NHEXAS) Drinking and Tap Water

EPA designed the NHEXAS program to evaluate comprehensive human exposure to multiple chemicals from multiple routes on both a community and regional scale, as well as the association between exposure, environmental concentrations, and personal activities. Samples were collected from 1995 to 1998. EPA completed Phase I field sample collection and laboratory analyses of NHEXAS data in 1998. Preliminary results were reported in 15 journal articles published in the September-October 1999 issue of the *Journal of Exposure Analysis and Environmental Epidemiology*; however, only five of the of these articles provide information potentially relevant to EDSP priority setting activities. Table 4-3 summarizes the relevant NHEXAS journal articles used in EDSP priority setting.

Four of the five EDSP-relevant articles provide drinking and tap water data included in this pathway: Clayton et al., O'Rourke et al., Robertson, et al., and Thomas. Specifically, the drinking water collected in NHEXAS pertains to flushed tap water, standing tap water, tap water, and drinking water data. The fifth EDSP-relevant article does not provide drinking or tap water data.

Chemicals were considered detected unless it was clear that the chemical was not detected in any of the samples. Where appropriate, the total number of detects was calculated from the number of samples and percent measurable. HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

Table 4-3
NHEXAS Journal Articles Relevant to EDSP Priority Setting Activities

Reference	Data Fields Available	Description of Data available
Clayton, C.A.; Pellizzari, E.D.; Whitmore, R.W.; Perritt, R.L.; and J.J. Quackenboss. National Human Exposure Assessment Survey (NHEXAS): distributions and associations of lead, arsenic and volatile organic compounds in EPA Region 5. *Journal of Exposure Analysis and Environmental Epidemiology*. (1999) 9:381-392.	Compound Medium Number of Samples Percentage Measurable Mean Median 90th Percentile	Provides extensive exposure data to arsenic, lead, benzene, chloroform, tetrachloroethene, and trichloroethene from air, dust, water, food, beverages, and urine for ~2500 participants.
Gordon, Sidney M.; Callahan, Patrick J.; Nishioka, Marcia G.; Brinkman, Marielle C.; O'Rourke, Mary Kay; Lebowitz, Michael D.; and Moschandreas, Demetrios M. Residential Environmental Measurements in the National Human Exposure Assessment Survey (NHEXAS) Pilot Study in Arizona: Preliminary Results for Pesticides and VOCs. *Journal of Exposure Analysis and Environmental Epidemiology*. (1999) 9: 456-470.	Compound Media Number of Samples Number of Detects Percentage of Detects Minimum Concentration Maximum Concentration Concentration Units Location Season Median Concentration 75th Percentile 90th Percentile	Reports personal (e.g., blood, urine, dermal wipes, 24 h duplicate diet) and microenvironmental (e.g., indoor and outdoor air, house dust, foundation soil) data obtained for selected pesticides (chloropyrifos and diazinon), VOCs (benzene, toluene, trichloroethene, formaldehyde, and 1,3-butadiene), and dermal wipe data obtained for selected pesticides (chloropyrifos and diazinon). Note: The Gordon report does not provide drinking or tap water data.

Table 4-3 (Continued)

Reference	Data Fields Available	Description of Data available
O'Rourke, Mary Kay; Van de Water, Peter K.; Jin, Shan; Rogan, Seumas P.; Weiss, Aaron D.; Gordon, Sydney M.; Moschandreas, Demetrios M.; and Lebowitz, Michael D. Evaluations of primary metals from NHEXAS Arizona: distributions and preliminary exposures. *Journal of Exposure Analysis and Environmental Epidemiology*. (1999) 9: 435-445.	Media Metal Detection Method Number of Samples Percentage of Samples Below the Detection Limit Maximum Concentration Median Concentration 75^{th} Percentile 90^{th} Percentile Concentration Units	Reports results for lead, arsenic, cadmium, chromium, and nickel in various media (air, soil, house dust, food, beverage, and water).
Robertson, Gary L.; Lebowitz, Michael D.; O'Rourke, Mary Kay; Gordon, Sydney; and Moschandreas, Demetrios. National Human Exposure Assessment Survey (NHEXAS) study in Arizona – introduction and preliminary results. *Journal of Exposure Analysis and Environmental Epidemiology*. (1999) 9: 427-434.	Compound Number of Samples Media Method Median Concentration 90^{th} Percentile Concentration Units	Provides demographic data. Also reports data for lead (representing metals) and chloropyrifos (representing pesticides), and benzene (representing VOCs) in various media (yard soil, foundation soil, house dust, indoor air, outdoor air, drinking water, food, and beverage).

Table 4-3 (Continued)

Reference	Data Fields Available	Description of Data available
Thomas, Kent W.; Pellizzari, Edo D.; and Berry, Maurice R. Population-based dietary intakes and tap water concentrations for selected elements in the EPA Region V National Human Exposure Assessment Survey (NHEXAS). *Journal of Exposure Analysis and Environmental Epidemiology*. (1999) 9: 402-413.	Compound Percent measurable Mean Concentration Standard Deviation 10^{th} Percentile 25^{th} Percentile 50^{th} Percentile 75^{th} Percentile 90^{th} Percentile Medium Concentration Units Number of Samples	Reports results for dietary intake and tap water concentrations of lead, arsenic, cadmium, chromium, barium, copper, manganese, nickel, selenium, vanadium, and zinc.

4.3.2.1 Data Source Location

The NHEXAS drinking water (DW) data were obtained from four journal articles published in the September-October 1999 issue of the *Journal of Exposure Analysis and Environmental Epidemiology*, as discussed above. The articles by Clayton et al., O'Rourke et al., Robertson, and Thomas are summarized in Table 4-3.

4.3.2.2 Data Source Contact

The NHEXAS DW data are contained in the four articles previously mentioned and are published in the September-October 1999 issue of the *Journal of Exposure Analysis and Environmental Epidemiology*.

4.3.2.3 Data Format

The raw NHEXAS DW data are available in hard copy format in four articles published in the September-October 1999 issue of the *Journal of Exposure Analysis and Environmental Epidemiology*, as previously discussed. These data were manually extracted from these hard copies and compiled into an Excel spreadsheet.

4.3.2.4 Frequency of Updates to the Data

The 1998 NHEXAS Phase I field sample collection results were published in the September-October 1999 issue of the *Journal of Exposure Analysis and Environmental Epidemiology*. No additional phases of the program were completed due to budget limitations and other Agency priorities.

4.3.2.5 Data Quality

All of the NHEXAS data appear to be of high quality. Established methods were used for collecting and analyzing data. Quality control and quality assurance samples were collected and analyzed, including reagent blanks, field blanks, duplicate samples, and spiked samples. Samples were split and analyzed in multiple laboratories. Audit samples were also analyzed when appropriate audit samples were available.

4.3.2.6 Data Manipulation

Drinking water data collected from 1995-1998 that were presented in four hard-copy journal articles (see Table 4-3: Clayton et al., O'Rourke et al., Robertson et al.., and Thomas et al.) were manually extracted and saved to an Excel spreadsheet.

1. Each of the measured concentrations were converted into geometric means to provide a consistent measure for comparison.

2. The drinking water data were extracted. The following data elements are included in the converted raw drinking water data:

    - Compound;
    - CAS number;
    - Central tendency (i.e., mean, median, 90^{th} percentile, or maximum of the measured concentrations converted into geometric means);
    - Units (ug/L);
    - Central tendency measurement label (e.g., 'GM estimated from AM', 'Median', 'GM calculated from 90^{th} percentile', 'GM calculated from Maximum', or 'Detection Limit');
    - Number of samples;
    - Percent measurable;
    - Population;
    - Medium (e.g., 'Drinking Water', 'Flushed Tap Water', 'Standing Tap Water', and 'Tap Water');
    - Data type (Food and Drinking Water);
    - Location (e.g., 'EPA Region V' or 'Arizona'); and
    - Source (e.g., 'Clayton', 'O'Rourke...', 'Robertson', or 'Thomas').

The following data manipulations were performed on the NHEXAS DW aggregated data table containing the raw data:

1. The NHEXAS DW data table contains all of the converted raw data elements for the NHEXAS drinking water records. The following data elements were added to the data table after the converted raw data were imported:

    - Number of measured samples;
    - Study date (i.e., month and year range for the sample); and
    - Detect flag.

2. The number of measured samples was added to the data table and calculated for each record using the following equation:

Number measured = Total number of samples x (% measurable x 0.01)

3. The detect flag was added to the data table and updated for each record having either a *central tendency measurement* label, indicating that the measurement was NOT a detection limit, or having a calculated *number of measured samples*.

HPV/pesticide inert chemicals that were measured in monitoring samples from this data source were considered for priority setting purposes.

4.3.3 TEAM-Drinking Water Data

The TEAM data source is described in Section 4.1.4. Data manipulations performed on TEAM drinking water data are described in the following sections.

4.3.3.1 Data Manipulation

Drinking water data collected from 1981 to 1984 that were contained in two hard-copy volumes of *The TEAM Study* were manually extracted and saved to a spreadsheet.

The raw TEAM drinking water (DW) data were extracted, and the following data elements are included in the TEAM DW aggregated data table:

- Compound;
- CAS number;
- Central tendency (i.e., geometric mean of the measured concentrations or a detection limit) (ug/L);
- Units (ug/L);
- Central tendency measurement label (e.g., 'Geometric Mean', 'Minimum Quantifiable Limit', or 'Limit of Detection');
- Total number of samples;
- Percent measurable;
- Population;
- Medium (e.g., 'Water');
- Location (e.g., 'Contra Costa' County, CA; 'Los Angeles', CA; 'New Jersey', 'North Carolina', or 'North Dakota'); and

- Season (e.g., 'Fall', 'Spring', 'Winter', or a particular month).
- Year; and
- Detect flag.

The following data manipulations were performed on the raw TEAM DW data:

1. The number of measured samples was added to the aggregated data table and calculated for each record using the following equation:

   Number measured = Total number of samples x (% measurable x 0.01)

2. The detect flag was added to the aggregated data table and updated for each record having a *number measured samples* calculated to be greater than zero (see Step 1 above).

HPV/pesticide inert chemicals that were measured in monitoring samples from this data source were considered for priority setting purposes.

4.3.4 National Stream Quality Accounting Network (NASQAN) Surface Water and Sediment Data

The U.S. Geological Survey's (USGS) National Stream Quality Accounting Network (NASQAN) focuses on monitoring the water quality of the nation's largest river systems, including those monitoring sites listed in Table 4-4.

Table 4-4
NASQAN Sites

NASQAN Site (Site Number)
Alabama River at Claiborne, Alabama (02429500)
Arkansas River at David Terry L&D, below Little Rock, Arkansas (07263620)
Arroyo Colorado at Harlingen, Texas (08470400)
Colorado River above Diamond Creek (09404200)
Colorado River above Imperial Dam, Arizona (09429490)
Colorado River at Lees Ferry, Arizona (09380000)
Colorado River at NIB above Morelos Dam, near Andrade, California (09522000)
Colorado River below Hoover Dam, Arizona-Nevada (09421500)
Colorado River near Cisco, Utah (09180500)
Columbia River at Beaver Army Terminal, near Quincy, Oregon (14246900)

Table 4-4 (Continued)

NASQAN Site (Site Number)
Columbia River at Northport, Washington (12400520)
Columbia River at Vernita Bridge, near Priest Rapids Dam, Washington (12472900)
Columbia River at Warrendale, Oregon (14128910)
Cumberland River at Smithland, Kentucky (03438500)
Green River at Green River, Utah (09315000)
Lower Atchafalaya River at Melville, Louisiana (07381495)
Minnesota River near Jordan, Minnesota (05330000)
Mississippi River at Clinton, Iowa (05420500)
Mississippi River at Grafton, Illinois (05587455)
Mississippi River at St. Francisville, Louisiana (07373420)
Mississippi River at Thebes, Illinois (07022000)
Mississippi River below Lock and Dam 2, at Hastings, Minnesota (05331580)
Missouri River at Garrison Dam, North Dakota (06338490)
Missouri River at Hermann, Missouri (06934500)
Missouri River at Omaha, Nebraska (06610000)
Missouri River at Pierre, South Dakota (06440000)
Missouri River near Culbertson, Montana (06185500)
Ohio River at Cannelton Dam, Kentucky (03303280)
Ohio River at Dam 53 near Grand Chain, Illinois (03612500)
Ohio River at Greenup Dam, Kentucky (03216600)
Pecos River near Langtry, Texas (08447410)
Platte River at Louisville, Nebraska (06805500)
Porcupine River near Fort Yukon, Alaska (15389000)
Rio Grande at El Paso, Texas (08364000)
Rio Grande at Falcon Dam, Texas (08461300)
Rio Grande at Foster Ranch, near Langtry, Texas (08377200)
Rio Grande at Laredo, Texas (08459200)
Rio Grande below Amistad Reservoir, Texas (08450900)
Rio Grande below Rio Conchos near Presidio, Texas (08374200)
Rio Grande near Brownsville, Texas (08475000)
San Juan River near Bluff, Utah (09379500)
Snake River at Burbank, Washington (13353200)
St. Lawrence River at Cornwall, Ontario, near Massena, New York (04264331)

Table 4-4 (Continued)

NASQAN Site (Site Number)
Susquehanna River at Conowingo Dam, Maryland (01578310)
Tanana River at Nenana, Alaska (15515500)
Tennessee River near Paducah, Kentucky (03609750)
Tombigbee River below Coffeeville Lock and Dam, Alabama (02469762)
Wabash River at New Harmony, Indiana (03378500)
Willamette River at Portland, Oregon (14211720)
Yellowstone River near Sidney, Montana (06329500)
Yukon River at Eagle, Alaska (15356000)
Yukon River at Pilot Station, Alaska (15565447)
Yukon River near Stevens Village, Alaska (15453500)

The primary goals of NASQAN are to characterize these sites, to determine regional source areas for the monitored substances, and to assess the effects of human influences on observed concentrations and amounts of the monitored substances. USGS uses NASQAN data to develop and compare constituent mass fluxes (i.e., the amount of material that passes through a station expressed as tons per day) among stations and across spatial miles. Data collected from the NASQAN monitoring sites are stored in the USGS National Water Information System. Monitoring data available from 1990 to present can be downloaded from NWISWeb.

4.3.4.1 Data Source Location

NASQAN data may be downloaded at: http://nwis.waterdata.usgs.gov/usa/nwis/qwdata. Water quality data for the specific NASQAN sites can be retrieved. Note that NWIS holds other data in addition to NASQAN; therefore the query results should be limited to the monitoring sites listed in Table 4-4.

4.3.4.2 Data Source Contact

Jennifer Morace
U.S. Geological Survey
Telephone: (503) 251-3229
Email: jlmorace@usgs.gov

4.3.4.3 Data Format

NASQAN data are available in an electronic format.

4.3.4.4 Frequency of Updates to the Data

Data are continuously updated.

4.3.4.5 Data Quality

The NASQAN Quality Assurance (QA) program was developed to support the data quality objective of annual flux estimation. NASQAN has a highly dispersed team of field personnel because of the national scale of the network. Therefore, ensuring consistency across the network is a critical element of the QA program. Major QA elements include well-defined protocols for sample collection, sample processing, chemical analysis, and data review. Extensive training is provided for field personnel to ensure that uniform procedures are used in sample collection and processing, and that data review is conducted according to established data-quality criteria. A field audit of all sampling crews was performed early in the program by each basin coordinator.

An additional component of the QA program consists of the collection and evaluation of quality control (QC) samples, including blanks, replicates, and field-matrix spikes for pesticides. Data from these samples document the performance of the overall process of sample collection, processing, and analysis. Quality-control samples typically comprise 10-20 percent of the total number of samples submitted for analysis.

For the purpose of consistent data review, data quality criteria have been defined by the NASQAN program, and provide the basis for national data review procedures. The criteria are based on various ranges of statistical variation, including the distribution of previous data for the site, where available. NASQAN maintains a web-based interactive review process wherein questionable data are flagged for special review and input by District personnel. Data-quality indicator (DQI) codes of "Q" for all constituents are available where exceedances occur for NASQAN criteria. Note that this DQI code is generated from the simple comparison of sample data to the statistical criteria, and does not necessarily imply that the data are questionable. In many cases, data are flagged simply because they represent extreme conditions, which are targeted by the NASQAN sampling strategy.

4.3.4.6 Data Manipulation

National Stream Quality Accounting Network (NASQAN) data were obtained from the National Water Information System (NWIS). Data for both surface water and sediment were available. Two data tables were created, one for each data set, and were treated as separate data sets for priority setting.

Data files were available in standard American Standard Code for Information Interchange (ASCII) format. Each data file provides documentation for the file in the header of each file. The following raw data files were downloaded:

- Parameter code definitions and units of measure;

- Suspended sediment data;
- Major inorganics;
- Minor inorganics;
- Organics; and
- Other.

The following data manipulations were performed on the raw data tables summarized above:

1. Supporting database documentation was reviewed to identify relevant raw data files which were imported into a database.

2. Each raw data table was appended to create one aggregated data table. Description fields were added to the data table and updated to the corresponding description of the codes (based on the documentation).

3. CAS numbers were reviewed and updated by the parameter table and various other sources.

4. The number of detects were calculated for each chemical by counting samples where the remark associated with the concentration value is not "<". Only "regular" samples (i.e., non-QA samples) collected from 1990 to present were included in the analysis.

HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

4.3.5 National Water Quality Assessment (NAWQA) Ground Water, Surface Water, and Sediment Data

The U.S. Geological Survey (USGS) began its NAWQA program in 1991, systematically collecting chemical, biological, and physical water quality data from 42 study units (basins) across the nation. A study unit is a major hydrologic system in which NAWQA studies are focused, is geographically defined by a combination of groundwater and surface water features, and usually encompasses more than 10,000 square miles. USGS collects data from all study units, and uses them, in addition to information from other programs, agencies, and researchers, to produce a national assessment or "national synthesis." The data warehouse currently contains the following data:

- Chemical concentrations in water, bed sediment, and aquatic organism tissues for approximately 500 chemical constituents;

- Site, basin, well, and network characteristics with descriptive variables;

- Daily stream flow information for fixed sampling sites;

- Ground water levels for sampled wells;
- 6,400 surface water sites and 7,000 wells;
- 44,000 nutrient samples, 26,000 pesticide samples, and 8,000 VOC samples; and
- 2,650 samples of bed sediment and aquatic organism tissues.

4.3.5.1 Data Source Location

NAWQA data may be queried and downloaded at: http://infotrek.er.usgs.gov/servlet/page?_pageid=543&_dad=portal30&_schema=PORTAL30.

4.3.5.2 Data Source Contact

Sandy Williamson, National Database Team Leader
USGS Wisconsin District
Office: (253) 428-3600 ext. 2683

4.3.5.3 Data Format

NAWQA data can be retrieved from the NAWQA website in tab- or comma-delimited format.

4.3.5.4 Frequency of Updates to the Data

Data are continuously updated.

4.3.5.5 Data Quality

The similar design of each investigation and use of standard methods allows for comparisons among the results measured at the various study units. All ground water samples are analyzed at the USGS's National Water Quality Laboratory. Quality control check samples are run daily.

4.3.5.6 Data Manipulation

National Water Quality Assessment (NAWQA) data were obtained from USGS. Supporting database documentation was reviewed to identify relevant data tables. The raw data file was imported into a database and a comprehensive data set was created with the following relevant fields:

- State Postal Code;
- County;
- Study Unit ID;
- Station Id;
- Primary Water Use;
- Land Use Group;
- Land Use Code;
- Sample Medium Desc;
- Result Datetime;
- Parameter Code;
- Parameter Name;
- Report Units;
- Value Remark; and
- Value.

This analysis was conducted for three data sets: groundwater, surface water, and sediment. Each data set was segregated based on the sample medium description (i.e., GROUNDWATER, SURFACE WATER, BOTTOM MATERIAL). Three summary tables were created, one for each data set, and were treated as separate data sets for priority setting.

The following data manipulations were performed on the raw data tables summarized above:

1. Additional fields were added to the aggregated data table, including:

    - Chemical Name;
    - Parameter Code - Numeric;
    - CAS Number; and
    - Detect?.

2. Parameter chemical names and CAS numbers were downloaded from the following USGS website: http://infotrek.er.usgs.gov/docs/nawqa_www/nwq_paramx.htm. Entering "%" in Search Parameter Codes will provide a list of all parameters.

3. For the Ground Water data set, unique groundwater use codes were reviewed to identify water uses primarily associated with drinking water. The complete list of groundwater uses is described below:

    - Aquaculture;
    - Commercial;
    - Dewater;
    - Domestic;
    - Fire;
    - Industrial;
    - Industrial (Cooling);
    - Institutional;

- Irrigation;
- Other;
- Power;
- Public Supply;
- Recreation;
- Stock; and
- Unused.

EPA identified water that will be primarily used as drinking water, including:

- Domestic;
- Recreational;
- Stock; and
- Public supply.

5. A new data table was created that contains only a subset of groundwater records, all associated with the four relevant uses identified in Step 4.

6. The detect flag was updated to "Yes" where the concentration value remark was not " < " (indicating a detection below the detection limit). Concentration value remarks of "E" were considered detects. "E" represents estimated values for all detections that were below the maximum detection limit, above the highest calibration standard, or otherwise less reliable than average because of sample- or compound-specific considerations.

HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

4.3.6 Summary of Drinking Water Exposure Pathway Chemicals

Table 4-5 presents the list of HPV/pesticide inert chemicals that were detected in monitoring samples from the Chemicals Drinking Water Exposure Pathway data sources, along with an indication of the data source in which they appeared.

Table 4-5
HPV/Pesticide Inert Chemicals Present in Drinking Water
Exposure Pathway Data Sources

Chemical Name	CAS Number	Total Number of Data Sources	NCOD	NHEXAS-DW	TEAM-DW	NASQAN-SW	NASQAN-Sed	NAWQA-SW	NAWQA-GW	NAWQA-Sed
Toluene	108883	5	Yes	No	Yes	Yes	No	Yes	Yes	No
Benzene, dimethyl-	1330207	4	Yes	No	No	Yes	No	Yes	Yes	No
di-sec-octyl phthalate	117817	4	Yes	No	No	Yes	Yes	No	No	Yes
1H-Purine-2,6-dione, 3,7-dihydro-1,3,7-trimethyl-	58082	3	No	No	No	Yes	No	Yes	Yes	No
Benzene, 1-methyl-4-(1-methylethyl)-	99876	3	Yes	No	No	No	No	Yes	Yes	No
Sulfur	7704349	3	No	No	No	Yes	Yes	No	No	Yes
2-Pentanone, 4-methyl-	108101	2	No	No	No	No	No	Yes	Yes	No
Acetone	67641	2	No	No	No	No	No	Yes	Yes	No
Butyl benzyl phthalate	85687	2	No	No	No	No	Yes	No	No	Yes
Dibutyl phthalate	84742	2	No	No	No	No	Yes	No	No	Yes
Diethyl phthalate	84662	2	No	No	No	No	Yes	No	No	Yes
Dimethyl phthalate	131113	2	No	No	No	No	Yes	No	No	Yes
Ethane, 1,1'-oxybis-	60297	2	No	No	No	No	No	Yes	Yes	No
Furan, tetrahydro-	109999	2	No	No	No	No	No	Yes	Yes	No
Methyl ethyl ketone	78933	2	No	No	No	No	No	Yes	Yes	No
Phenol	108952	2	No	No	No	No	Yes	No	No	Yes
Ethanol, 2-butoxy-, phosphate (3:1)	78513	1	No	No	No	Yes	No	No	No	No

Table 4-5 (Continued)

Chemical Name	CAS Number	Total Number of Data Sources	NCOD	NHEXAS-DW	TEAM-DW	NASQAN-SW	NASQAN-Sed	NAWQA-SW	NAWQA-GW	NAWQA-Sed
Hexanedioic acid, bis(2-ethylhexyl) ester	103231	1	Yes	No	No	No	No	No	No	No
Isophorone	78591	1	No	No	No	No	No	No	No	Yes

4.4 Indoor Air Exposure Pathway

Relevant data were extracted from the following data sources to determine the presence of HPV/pesticide inert chemicals in indoor air:

- EPA/Office of Research and Development (ORD) Journal Articles;
- NHEXAS - Indoor and Personal Air Data; and
- TEAM Air Data.

4.4.1 EPA/Office of Research and Development (ORD) Journal Articles

The Office of Research and Development (ORD) journal articles consist of eight journal articles and reports cited in the September 2005 FR Notice, seven of which contain indoor and/or personal air data. These documents are summarized as follows:

- *Brown et al.* – A comprehensive compilation and analysis of U.S. and European literature, with data on residences, office buildings, schools, and other buildings. Includes indoor concentrations and indoor/outdoor (I/O) ratios for samples collected between 1978 and 1990;

- *Daisey et al.* – A field study of indoor and outdoor concentrations and I/O ratios from 12 office buildings in northern California with 3 different types of ventilation systems;

- *Immerman et al.* – Reports data from the *Nonoccupational Pesticide Exposure Study (NOPES)*, an EPA field study with indoor and outdoor concentrations and I/O ratios from 350 samples taken in homes in Jacksonville, Florida and Chicopee-Springfield, Massachusetts during several seasons and years;

- *Kelly et al.* – Provides summary outdoor air data for 189 Hazardous Air Pollutants (HAPs);

- *Samfield* – A literature survey of U.S. and foreign indoor air concentrations through the late 1980's rom residences, office buildings, schools, and other commercial buildings;

- *Shah et al.* – A literature survey based on the *National Ambient Volatile Organic Compounds (VOCs) Database*, which provides measurements of indoor and outdoor air concentrations in the U.S. through 1986;

- *Sheldon et al.* – A field study of indoor and outdoor concentrations and I/O ratios from 128 homes in Woodland, California; and

- *Shields et al.* – A field study of indoor and outdoor concentrations from 70 commercial buildings with different occupant densities collected between March 18 and April 29, 1991.

All ORD data (both indoor and outdoor) were previously compiled while populating EDPSD v.2, and data overlaps were eliminated to ensure that "double-counting" did not occur. The indoor and personal air data were incorporated into a database in two separate data tables. Indoor and/or personal air data are available for approximately 230 compounds. Note that EPA excluded the Kelly article from this analysis, as this article only provides outdoor air data.

The ORD articles include all of the indoor air data that were collected in the TEAM study; therefore, TEAM data were analyzed along with the ORD data rather than in a separate TEAM indoor air database.

4.4.1.1 Data Source Location

The ORD indoor and personal air data were obtained from seven documents/journal articles, as discussed above. The articles by Brown et al., Daisey et al., Immerman et al., Samfield, Shah et al., Sheldon et al., and Shields et al. are summarized in Table 4-7 in Section 4.4.1.6.

4.4.1.2 Data Source Contact

The ORD indoor/personal air data is contained in the seven articles previously mentioned and summarized in Table 4-7 in Section 4.4.1.6.

4.4.1.3 Data Format

The raw ORD indoor/personal air data are available in hard copy format in seven documents. These data were manually extracted from these hard copies and compiled into spreadsheets.

4.4.1.4 Frequency of Updates to the Data

The ORD data sources are hard copy and final; they will not be updated in their current form; however, the databases and sources they reported on may be updated in the future, and additional studies may result. For instance, the *National Ambient Volatile Organic Compounds (VOCs) Data Base* was intended to be updated every three years; however, the 1988 report was never finalized and no updated reports exist. Updates of the *Indoor Air Quality Data Base for Organic Compounds* were planned; however, no updated publications on this database appear to be available. Most of the other data consist of reported field measurements that are not likely to be updated.

4.4.1.5 Data Quality

All of the data, except for the data in the *Indoor Air Quality Data Base for Organic Compounds* (used by Samfield), appear to be of high quality. Note that the data input into the *Indoor Air Quality Data Base for Organic Compounds* were not screened for quality before being entered. Also, a cursory review of the database indicates some minor discrepancies such as citation of a reference that does not exist and listing incorrect compound synonyms. Although these errors are minor, they could be indicative of more significant errors in the entered data. See Table 4-6 for a discussion of data quality for each data source.

Table 4-6
Summary of Data Quality of ORD Sources

Reference	Quality Indicators and Considerations
Brown	No information provided on screening procedures; published in peer-reviewed journal.
Daisey	Published in a peer-reviewed journal.
Immerman (NOPES)	Published EPA report. Study used duplicate samples, matrix spike samples, independent analysis of samples by a second laboratory, analytical and field blanks, and system audits.
Samfield	Data were not screened for quality before being entered into the database. Determination of quality must be made by referring to the original reference. A cursory review of the database indicates some minor discrepancies including citation of a reference that does not exist and listing incorrect synonyms for compounds.
Shah	Summary of EPA database update work. Data were screened for inconsistencies, duplication, numerical errors, unsupported validation procedures, and unpublished methods before being included in the database.
Sheldon	Published California EPA Air Resources Board (CARB) report. Study used standard operating procedures, systems audits, field blanks, spiked control samples, method blanks and controls, and duplicate samples.
Shields	Published in a peer-reviewed journal; used field and laboratory blanks. Data compared well to other studies.

In addition, some of the ORD data that were manually extracted from the hard copy documents were verified against the original documents to ensure the quality of the electronic data collection used for priority setting purposes.

4.4.1.6 Data Manipulation

Indoor and personal air data collected from 1978-1991 and "pre-1990" data that were contained in seven hard-copy documents pertaining to the ORD indoor and personal air data were manually extracted into spreadsheets. Table 4-7 provides full citations and additional

descriptions for each of these eight documents (Brown et al., Daisey et al., Immerman et al., Samfield, Shah et al., Sheldon et al., and Shields et al.). These data were originally summarized and used for EDSP ranking efforts in 2000.

Table 4-7
Citations and Descriptions of ORD Data Sources

Reference	Data Fields Available	Description of Data Available	Comments
Brown, S.K.; Sim, M.R.; Abramson, M.J.; and Gray, C.N. "Concentrations of Volatile Organic Compounds in Indoor Air – A Review". *Indoor Air*. (1994) 4: 123-134.	Compound No. Buildings Measured Type of Building No. of Measurements Weighted Average Geometric Mean 90^{th} Percentile 98^{th} Percentile Average Maximum Concentration	Data for approximately 37 volatile organic compounds collected in indoor air of buildings of different classifications (dwellings, offices, schools, hospitals) and categories (established, new, and complaint buildings).	Data overlaps with Shah and Samfield articles. All studies performed between 1978 and 1990.
Daisey, J.M.; Hodgson, A.T.; Fisk, W.J.; Mendell, M.J.; and Brinke, J. "Ten Volatile Organic Compounds in Twelve California Office Buildings: Classes, Concentrations, and Sources". *Atmospheric Environment*. (1994) 28: 3557-3562.	Compound Class Geometric Mean Geometric Standard Deviation Minimum Concentration Maximum Concentration Minimum I/O Ratio Maximum I/O Ratio	A field study of indoor and outdoor concentrations and I/O ratios of 39 compounds from 12 office buildings in northern CA with 3 different types of ventilation systems.	Information limited to one type of building and one geographic area. Sample dates not readily available.

Table 4-7 (Continued)

Reference	Data Fields Available	Description of Data Available	Comments
Immerman, Frederick W. and Schaum, John L. *Final Report of the Nonoccupational Pesticide Exposure Study (NOPES)*. EPA/600/3-90/003 (NTIS PB90-152224). January 1990.	Analyte Pesticide Use Level Season Type of Sample Estimated Percent of Population Exposed Weighted Arithmetic Mean Concentration Standard Error Median Concentration Maximum Concentration Percent Detectable Standard Error in Percent Detectable Percent Undetectable Standard Error in Percent Undetectable Number of Samples	Reports data from the Nonoccupational Pesticide Exposure Study (NOPES), an EPA field study with indoor and outdoor concentrations and I/O ratios from 350 samples taken in homes in Jacksonville, FL and Chicopee-Springfield, MA. Provides data on 32 household pesticides from indoor, outdoor, and personal microenvironments.	NOPES data were collected in three phases: – Phase I: Summer 1986, Jacksonville, FL; – Phase II: Spring 1987, Florida and Massachusetts; and – Phase III: Winter 1988, Florida and Massachusetts.
Samfield, Max M. *Indoor Air Quality Data Base for Organic Compounds*. EPA-600-R-92-025 (NTIS PB92-158468).	Number of Carbon Atoms Compound Empirical Formula Molecular Weight Odor Threshold (micrograms/cubic meter) Minimum Measured Concentration Maximum Measured Concentration Median Measured Concentration Mean Measured Concentration Standard Deviation Reference No. Sampling Time (hours) Building Type	Data on over 220 organic compounds ranging in molecular weight from 30 to 446 and containing one to 31 carbon atoms. Data were obtained from apartments, automobiles, commercial and office buildings, hospitals, mobile homes, nursing homes, residences, and schools and were collected between 1979 and 1990.	Computerized version of the data existed at EPA but was not made publicly available. Updates were planned but there is no indication that any ever occurred. Includes data from Shah and NOPES.

Table 4-7 (Continued)

Reference	Data Fields Available	Description of Data Available	Comments
Shah, Jitendra J. and Singh, Hanwant B. "Distribution of Volatile Organic Chemicals in Outdoor and Indoor Air. A National VOCs Data Base". *Environmental Science and Technology*. (1988) 22: 1381-1388.	Database No. CAS No. Chemical Name Number of Data Points Average Daily Concentration Median Daily Concentration Lower Quartile Upper Quartile Indoor/Ambient Indicator Site Type	Summary data for the National Ambient Volatile Organic Compounds (VOCs) Data Base. Database contains information for 320 chemicals, but the paper summarizes data for less than 100 chemicals.	Database first published by EPA in 1980's and was updated in 1986. Individual sample dates not available.
Sheldon, L.; Clayton, A.; Jones, B.; Keever, J.; Perritt, R.; Smith, D.; Whitaker, D.; and Whitmore, R. *Indoor Pollutant Concentrations and Exposures: Final Report*. California Air Resources Board, Contract A833-156. January 1992.	Compound Class Detection Percentages Sample Type Median Air Concentration Maximum Concentration 90^{th} Percentile 75^{th} Percentile 50^{th} Percentile 25^{th} Percentile 10^{th} Percentile Minimum Concentration Geometric Mean Geometric Standard Error Arithmetic Mean Arithmetic Standard Error I/O Geometric Mean I/O Standard Error Study Name Number of Samples	A field study of indoor and outdoor concentrations and I/O ratios for approximately 28 compounds from 128 homes in Woodland CA. Also contains comparison data from residential studies in two other urban centers.	Data for the main study were collected during June 1990. Data from Los Angeles comparison study were collected during January and May of 1984 and February and July of 1987. Data from Contra Costa comparison study were collected during June 1984.

Table 4-7 (Continued)

Reference	Data Fields Available	Description of Data Available	Comments
Shields, Helen C.; Fleisher, Daniel M.; and Weschler, Charles J. "Comparisons among VOCs Measured in Three Types of U.S. Commercial Buildings with Different Occupant Densities". *Indoor Air*. (1996) 6: 2-17.	Indoor/Outdoor Indicator Occupants/1000 ft^2 Ventilation Supply Air Distribution Median No. of Floors Median Age (years) Carpeting Photocopiers Smoking (in sampled area) Smoking (in breakroom) Compound Geometric Mean Geometric Standard Deviation Detection Percentages Maximum Concentration Second Highest Concentration Geometric Mean for I/O Ratios Geometric Standard Deviation for I/O Ratios	A field study of indoor and outdoor concentrations from 70 commercial buildings with different occupant densities. Includes data for 31 compounds collected in 70 buildings in 27 states and the District of Columbia.	Most samples collected between March 18 and April 29, 1991.

The following steps were taken to manipulate and summarize the ORD data.

1. Each of the measured concentrations were converted into geometric means.

2. Indoor and personal air data were extracted. The following data elements are included in the converted raw air data:

    - Compound;
    - CAS number;
    - Central tendency (e.g., geometric mean, arithmetic mean, median, or maximum of the measured concentrations or a detection limit - converted into a universal geometric mean, as described above);
    - Central tendency measurement label (e.g., 'Geometric Mean', 'GM calculated from AM', 'GM calculated from Average', 'GM estimated from AM', 'Median', 'GM estimated from Max', 'GM estimated from Min' or 'Detection Limit');
    - Number of samples;
    - Percent measurable;
    - Building type (e.g., 'Automobile', 'Commercial', 'Dwelling' (some indicated 'New' or 'Mobile'), 'Hospital' (some indicated 'New'), 'Nursing Home', 'Office' (some indicated 'New'), 'Residential', 'School' (some indicated 'New'), some types were unknown);
    - Data type (e.g., 'Indoor', 'Personal', 'Workplace'); and
    - Source (e.g., 'Brown', 'Daisey', 'Immerman'/'NOPES', 'Samfield', 'Shah', 'Sheldon', 'Sheldon-CC' (Contra Costa County, CA data), 'Sheldon-LA' (Los Angeles, CA data), 'Shields' or 'Robertson').

The following data manipulations were performed on the ORD air data contained in the imported table:

1. The ORD indoor and personal air records were extracted from the combined indoor air data in the imported table and saved to a new "ORD air" data table.

2. The ORD indoor and personal air data were divided and saved into two new data tables, one containing only ORD indoor air data and the other containing only ORD personal (including "workplace") air data.

3. The number of detected samples was added to the data tables and calculated for each record using the following equation:

Number of detects = Number of samples x (Percent detects x 0.01)

4. The detect flag was added to the ORD indoor and personal data tables and updated for each record having either a calculated *number of detects* or a valid median or geometric mean. Chemicals were considered detected unless it was clear that the chemical was not detected in the sample (e.g., if number of detects was reported as 0 or the central tendency measurement was reported as the detection limit).

HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

4.4.2 NHEXAS - Indoor and Personal Air Data

EPA designed the NHEXAS program to evaluate comprehensive human exposure to multiple chemicals from multiple routes on both a community and regional scale, as well as the association between exposure, environmental concentrations, and personal activities. Samples were collected from 1995 to 1998. EPA completed Phase I field sample collection and laboratory analyses of NHEXAS data in 1998. Preliminary results were reported in 15 journal articles published in the September-October 1999 issue of the *Journal of Exposure Analysis and Environmental Epidemiology*; however only five of the of these articles provide information potentially relevant to EDSP priority setting activities.

Table 4-8 summarizes the relevant NHEXAS journal articles used in EDSP priority setting. Four of the five EDSP-relevant articles provide the indoor and personal air data included in this pathway: Clayton et al., O'Rourke et al., Gordon et al., and Robertson, et al. The fifth EDSP-relevant article does not provide indoor or personal air data.

Chemicals were considered detected unless it was clear that the chemical was not detected in any of the samples. Where appropriate, the total number of detects was calculated from the number of samples and percent measurable. HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

Table 4-8
NHEXAS Journal Articles Relevant to EDSP Priority Setting Activities

Reference	Data Fields Available	Description of Data available
Clayton, C.A.; Pellizzari, E.D.; Whitmore, R.W.; Perritt, R.L.; and J.J. Quackenboss. National Human Exposure Assessment Survey (NHEXAS): distributions and associations of lead, arsenic and volatile organic compounds in EPA Region 5. *Journal of Exposure Analysis and Environmental Epidemiology*. (1999) 9:381-392.	Compound Medium Number of Samples Percentage Measurable Mean Median 90th Percentile	Provides extensive exposure data to arsenic, lead, benzene, chloroform, tetrachloroethene, and trichloroethene from air, dust, water, food, beverages, and urine for ~2500 participants.
Gordon, Sidney M.; Callahan, Patrick J.; Nishioka, Marcia G.; Brinkman, Marielle C.; O'Rourke, Mary Kay; Lebowitz, Michael D.; and Moschandreas, Demetrios M. Residential Environmental Measurements in the National Human Exposure Assessment Survey (NHEXAS) Pilot Study in Arizona: Preliminary Results for Pesticides and VOCs. *Journal of Exposure Analysis and Environmental Epidemiology*. (1999) 9: 456-470.	Compound Media Number of Samples Number of Detects Percentage of Detects Minimum Concentration Maximum Concentration Concentration Units Location Season Median Concentration 75th Percentile 90th Percentile	Reports personal (e.g., blood, urine, dermal wipes, 24 h duplicate diet) and microenvironmental (e.g., indoor and outdoor air, house dust, foundation soil) data obtained for selected pesticides (chloropyrifos and diazinon), VOCs (benzene, toluene, trichloroethene, formaldehyde, and 1,3-butadiene), and dermal wipe data obtained for selected pesticides (chloropyrifos and diazinon).

Table 4-8 (Continued)

Reference	Data Fields Available	Description of Data available
O'Rourke, Mary Kay; Van de Water, Peter K.; Jin, Shan; Rogan, Seumas P.; Weiss, Aaron D.; Gordon, Sydney M.; Moschandreas, Demetrios M.; and Lebowitz, Michael D. Evaluations of primary metals from NHEXAS Arizona: distributions and preliminary exposures. *Journal of Exposure Analysis and Environmental Epidemiology*. (1999) 9: 435-445.	Media Metal Detection Method Number of Samples Percentage of Samples Below the Detection Limit Maximum Concentration Median Concentration 75^{th} Percentile 90^{th} Percentile Concentration Units	Reports results for lead, arsenic, cadmium, chromium, and nickel in various media (air, soil, house dust, food, beverage, and water).
Robertson, Gary L.; Lebowitz, Michael D.; O'Rourke, Mary Kay; Gordon, Sydney; and Moschandreas, Demetrios. National Human Exposure Assessment Survey (NHEXAS) study in Arizona – introduction and preliminary results. *Journal of Exposure Analysis and Environmental Epidemiology*. (1999) 9: 427-434.	Compound Number of Samples Media Method Median Concentration 90^{th} Percentile Concentration Units	Provides demographic data. Also reports data for lead (representing metals) and chloropyrifos (representing pesticides), and benzene (representing VOCs) in various media (yard soil, foundation soil, house dust, indoor air, outdoor air, drinking water, food, and beverage).

Table 4-8 (Continued)

Reference	Data Fields Available	Description of Data available
Thomas, Kent W.; Pellizzari, Edo D.; and Berry, Maurice R. Population-based dietary intakes and tap water concentrations for selected elements in the EPA Region V National Human Exposure Assessment Survey (NHEXAS). *Journal of Exposure Analysis and Environmental Epidemiology*. (1999) 9: 402-413.	Compound Percent measurable Mean Concentration Standard Deviation 10^{th} Percentile 25^{th} Percentile 50^{th} Percentile 75^{th} Percentile 90^{th} Percentile Medium Concentration Units Number of Samples	Reports results for dietary intake and tap water concentrations of lead, arsenic, cadmium, chromium, barium, copper, manganese, nickel, selenium, vanadium, and zinc. Note: The Thomas report does not provide indoor or personal air data.

4.4.2.1 Data Source Location

The NHEXAS air data were obtained from four journal articles published in the September-October 1999 issue of the *Journal of Exposure Analysis and Environmental Epidemiology*, as discussed above. The articles by Clayton et al., O'Rourke et al., Gordon et al., and Robertson et al. are summarized in Table 4-8.

4.4.2.2 Data Source Contact

Not applicable.

4.4.2.3 Data Format

The raw NHEXAS air data are available in hard copy format in four articles published in the September-October 1999 issue of the *Journal of Exposure Analysis and Environmental Epidemiology*, as previously discussed. These data were manually extracted from the hard copies and compiled into a spreadsheet.

4.4.2.4 Frequency of Updates to the Data

The 1998 NHEXAS Phase I field sample collection results were published in the September-October 1999 issue of the *Journal of Exposure Analysis and Environmental Epidemiology*. No additional phases of the program were completed due to budget limitations and other Agency priorities.

4.4.2.5 Data Quality

All of the NHEXAS data appear to be of high quality. Established methods were used for collecting and analyzing data. Quality control and quality assurance samples were collected and analyzed, including reagent blanks, field blanks, duplicate samples, and spiked samples. Samples were split and analyzed in multiple laboratories. Audit samples were also analyzed when appropriate audit samples were available.

In addition, the NHEXAS data that were manually extracted from the hard copy documents were spot-checked against the original documents to ensure the quality of the electronic data collection used for EDSP.

4.4.2.6 Data Manipulation

Indoor and personal air data collected from 1995-1998 that were contained in four hard-copy journal articles pertaining to the NHEXAS air data (see Table 4-8: Clayton et al., Gordon et al., O'Rourke et al., and Robertson et al.) were manually extracted into a spreadsheet.

1. Each of the measured concentrations were converted into geometric means to provide a consistent measure for comparison.

2. The air data were extracted. The following data elements are included in the converted raw air data:

    - Compound;
    - CAS number;
    - Central tendency (i.e., mean, median, 90^{th} percentile, or maximum of the measured concentrations converted into geometric means);
    - Central tendency measurement label (e.g., 'GM estimated from AM', 'Median', 'GM calculated from 90^{th} percentile', 'GM calculated from Maximum', or 'Detection Limit');
    - Number of samples;
    - Percent measurable;
    - Building type (e.g., 'Dwelling');
    - Data type (e.g., 'Indoor Air' or 'Personal Air');
    - Source (e.g., 'Clayton', 'Gordon, 'O'Rourke', or 'Robertson').

The following data manipulations were performed on the NHEXAS air data:

1. The NHEXAS indoor and personal air records were extracted from the combined indoor air data in the imported table and saved to a new data table.

2. A number of detected samples field was added to the data table and calculated for each record using the following equation:

    Number of detects = Number of samples x (Percent detects x 0.01)

3. A detect flag field was added to the NHEXAS air table and updated for each record having either a *central tendency measurement label*, indicating that the measurement was either NOT a detection limit, or having a calculated *number of detects*.

4. The NHEXAS indoor and personal air data were divided and saved into two new data tables: one containing only NHEXAS indoor air data and the other containing only NHEXAS personal air data.

HPV/pesticide inert chemicals that were detected in monitoring samples from this data source were considered for priority setting purposes.

4.4.3 TEAM Air Data

As previously discussed, the ORD articles include all of the indoor air data collected in the TEAM study; therefore, TEAM data were analyzed along with the ORD data rather than in a separate TEAM indoor air database.

4.4.4 Summary of Indoor Air Exposure Pathway Chemicals

Table 4-9 presents the list of HPV/pesticide inert chemicals that were detected in monitoring samples from the Indoor Air Exposure Pathway data sources, along with an indication of the data source in which they appeared.

Table 4-9
HPV/Pesticide Inert Chemicals Present in Indoor Air
Exposure Pathway Data Sources

Chemical Name	CAS Number	Total Number of Data Sources	ORD-Indoor	ORD-Personal	NHEXAS-Indoor	NHEXAS-Personal
Benzene, dimethyl-	1330207	2	Yes	Yes	No	No
Bicyclo[3.1.1]hept-2-ene, 2,6,6-trimethyl-	80568	2	Yes	Yes	No	No
di-sec-octyl phthalate	117817	2	Yes	Yes	No	No
Toluene	108883	2	Yes	No	Yes	No
1-Butanol	71363	1	Yes	No	No	No
1-Hexanol, 2-ethyl-	104767	1	Yes	No	No	No
1-Propanol, 2-methyl-	78831	1	Yes	No	No	No
2-Pentanone, 4-methyl-	108101	1	Yes	No	No	No
2-Propanol	67630	1	Yes	No	No	No
2-Propenoic acid, butyl ester	141322	1	Yes	No	No	No
9-Octadecenoic acid (9Z)-	112801	1	Yes	No	No	No
Acetic acid ethyl ester	141786	1	Yes	No	No	No
Acetic acid, 2-methylpropyl ester	110190	1	Yes	No	No	No
Acetic acid, butyl ester	123864	1	Yes	No	No	No
Acetic acid	64197	1	Yes	No	No	No
Acetone	67641	1	Yes	No	No	No
Benzaldehyde	100527	1	Yes	No	No	No
Benzene, 1-methyl-4-(1-methylethyl)-	99876	1	Yes	No	No	No
Butane	106978	1	Yes	No	No	No

Table 4-9 (Continued)

Chemical Name	CAS Number	Total Number of Data Sources	ORD-Indoor	ORD-Personal	NHEXAS-Indoor	NHEXAS-Personal
Butane, 2-methyl-	78784	1	Yes	No	No	No
Butanoic acid	107926	1	Yes	No	No	No
Butyl benzyl phthalate	85687	1	Yes	No	No	No
Cyclohexane	110827	1	Yes	No	No	No
Cyclohexanone	108941	1	Yes	No	No	No
Cyclohexene, 1-methyl-4-(1-methylethenyl)-, (4R)-	5989275	1	Yes	No	No	No
Cyclopentasiloxane, decamethyl-	541026	1	Yes	No	No	No
Cyclotetrasiloxane, octamethyl-	556672	1	Yes	No	No	No
Dibutyl phthalate	84742	1	Yes	No	No	No
Diethyl phthalate	84662	1	Yes	No	No	No
Dimethyl phthalate	131113	1	Yes	No	No	No
Dodecanoic acid	143077	1	Yes	No	No	No
Ethanamine, N-ethyl-	109897	1	Yes	No	No	No
Ethanol, 2-butoxy-	111762	1	Yes	No	No	No
Ethanol	64175	1	Yes	No	No	No
Ethanol, 2-ethoxy-	110805	1	Yes	No	No	No
Ethanone, 1-phenyl-	98862	1	Yes	No	No	No
Heptane	142825	1	Yes	No	No	No
Hexadecanoic acid	57103	1	Yes	No	No	No
Hexane	110543	1	Yes	No	No	No
Methanol	67561	1	Yes	No	No	No
Methyl ethyl ketone	78933	1	Yes	No	No	No
Octadecanoic acid	57114	1	Yes	No	No	No
Phenol	108952	1	Yes	No	No	No
Propane, 2-methyl-	75285	1	Yes	No	No	No
Propane	74986	1	Yes	No	No	No
Propanoic acid, 2-methyl-, monoester with 2,2,4-trimethyl-1,3-pentanediol	25265774	1	Yes	No	No	No
Propanoic acid, 2-methyl-, 2,2-dimethyl-1-(1-methylethyl)-1,3-propanediyl ester	6846500	1	Yes	No	No	No
Tetradecanoic acid	544638	1	Yes	No	No	No

5.0 INTEGRATION OF PATHWAY DATA

A unique list of HPV/pesticide inert chemicals was generated for each exposure pathway evaluated, and each list contains only those chemicals that are both and HPV and pesticide inert chemical. Each of the exposure pathway lists were then combined to generate a unique list of HPV/pesticide inert chemicals represented in one or more exposure pathways. Table 5-1 presents the number of HPV/pesticide inert chemicals according to the number and types of pathways in which they were observed.

Table 5-1
Number of HPV/Pesticide Inert Chemicals
According to the Number of Pathways in which they were Observed

Number (Type) of Pathways	Number of HPV/ Pesticide Inert Chemicals
4 (Human, Eco, Water, Air)	8
3 (Human, Eco, Water)	1
3 (Human, Eco, Air)	0
3 (Human, Water, Air)	0
3 (Eco, Water, Air)	3
2 (Human, Eco)	0
2 (Human, Water)	1
2 (Human, Air)	2
2 (Eco, Water)	0
2 (Eco, Air)	0
2 (Water, Air)	1
1 (Human)	2
1 (Eco)	5
1 (Water)	5
1 (Air)	34

Table 5-2 presents the complete list of 62 HPV/pesticide inert chemicals, along with an indication of the pathways in which they were observed.

Table 5-2
HPV/Pesticide Inert Chemicals
According to the Pathways in which they were Observed

Chemical Name	CAS Number	Total Pathways	Human Biological Monitoring Pathway	Ecological Biological Monitoring Pathway	Drinking Water Pathway	Indoor Air Pathway
Acetone	67641	4	Yes	Yes	Yes	Yes
Butyl benzyl phthalate	85687	4	Yes	Yes	Yes	Yes
di-sec-octyl phthalate	117817	4	Yes	Yes	Yes	Yes
Dibutyl phthalate	84742	4	Yes	Yes	Yes	Yes
Diethyl phthalate	84662	4	Yes	Yes	Yes	Yes
Dimethyl phthalate	131113	4	Yes	Yes	Yes	Yes
Methyl ethyl ketone	78933	4	Yes	Yes	Yes	Yes
Toluene	108883	4	Yes	Yes	Yes	Yes
2-Pentanone, 4-methyl-	108101	3	No	Yes	Yes	Yes
Benzene, 1-methyl-4-(1-methylethyl)-	99876	3	No	Yes	Yes	Yes
Isophorone	78591	3	Yes	Yes	Yes	No
Phenol	108952	3	No	Yes	Yes	Yes
Benzene, dimethyl-	1330207	2	No	No	Yes	Yes
Bicyclo[3.1.1]hept-2-ene, 2,6,6-trimethyl-	80568	2	Yes	No	No	Yes
Cyclotetrasiloxane, octamethyl-	556672	2	Yes	No	No	Yes
Hexanedioic acid, bis(2-ethylhexyl) ester	103231	2	Yes	No	Yes	No
1-Butanol	71363	1	No	No	No	Yes
1-Hexanol, 2-ethyl-	104767	1	No	No	No	Yes
1-Propanol, 2-methyl-	78831	1	No	No	No	Yes
1H-Purine-2,6-dione, 3,7-dihydro-1,3,7-trimethyl-	58082	1	No	No	Yes	No
2-Propanol	67630	1	No	No	No	Yes
2-Propenoic acid, butyl ester	141322	1	No	No	No	Yes
3-Cyclohexene-1-methanol, .alpha.,.alpha.,4-trimethyl-	98555	1	No	Yes	No	No
9-Octadecenoic acid (9Z)-	112801	1	No	No	No	Yes
Acetic acid, butyl ester	123864	1	No	No	No	Yes
Acetic acid ethyl ester	141786	1	No	No	No	Yes
Acetic acid, 2-methylpropyl ester	110190	1	No	No	No	Yes
Acetic acid	64197	1	No	No	No	Yes
Benzaldehyde	100527	1	No	No	No	Yes

Table 5-2 (Continued)

Chemical Name	CAS Number	Total Pathways	Human Biological Monitoring Pathway	Ecological Biological Monitoring Pathway	Drinking Water Pathway	Indoor Air Pathway
Benzenemethanol	100516	1	No	Yes	No	No
Benzoic acid	65850	1	No	Yes	No	No
Butane	106978	1	No	No	No	Yes
Butane, 2-methyl-	78784	1	No	No	No	Yes
Butanoic acid	107926	1	No	No	No	Yes
Cyclohexane	110827	1	No	No	No	Yes
Cyclohexanone	108941	1	No	No	No	Yes
Cyclohexene, 1-methyl-4-(1-methylethenyl)-, (4R)-	5989275	1	No	No	No	Yes
Cyclopentasiloxane, decamethyl-	541026	1	No	No	No	Yes
Dodecanoic acid	143077	1	No	No	No	Yes
Ethanamine, N-ethyl-	109897	1	No	No	No	Yes
Ethane, 1,1'-oxybis-	60297	1	No	No	Yes	No
Ethanol	64175	1	No	No	No	Yes
Ethanol, 2-butoxy-	111762	1	No	No	No	Yes
Ethanol, 2-ethoxy-	110805	1	No	No	No	Yes
Ethanol, 2-butoxy-, phosphate (3:1)	78513	1	No	No	Yes	No
Ethanone, 1-phenyl-	98862	1	No	No	No	Yes
Furan, tetrahydro-	109999	1	No	No	Yes	No
Heptane	142825	1	No	No	No	Yes
Hexadecanoic acid	57103	1	No	No	No	Yes
Hexane	110543	1	No	No	No	Yes
Methanol	67561	1	No	No	No	Yes
Octadecanoic acid	57114	1	No	No	No	Yes
Phenol, 4-nitro-	100027	1	No	Yes	No	No
Phenol, 2,6-bis(1,1-dimethylethyl)-4-methyl-	128370	1	Yes	No	No	No
Phosphoric acid, triphenyl ester	115866	1	Yes	No	No	No
Propane, 2-methyl-	75285	1	No	No	No	Yes
Propane	74986	1	No	No	No	Yes
Propanoic acid, 2-methyl-, 2,2-dimethyl-1-(1-methylethyl)-1,3-propanediyl ester	6846500	1	No	No	No	Yes
Propanoic acid, 2-methyl-, monoester with 2,2,4-trimethyl-1,3-pentanediol	25265774	1	No	No	No	Yes

Table 5-2 (Continued)

Chemical Name	CAS Number	Total Pathways	Human Biological Monitoring Pathway	Ecological Biological Monitoring Pathway	Drinking Water Pathway	Indoor Air Pathway
Sulfur	7704349	1	No	No	Yes	No
Tetradecanoic acid	544638	1	No	No	No	Yes
Water	7732185	1	No	Yes	No	No

Because there were a large number of chemicals from which to select, it was necessary to establish priorities within the pathways. In choosing which HPV/pesticide inert chemicals to select for the initial screening list, EPA gave highest priority to chemicals that appeared in four exposure pathways, followed by chemicals that appeared in three pathways. For those chemicals that appeared in three pathways, EPA gave highest priority to those chemicals appearing in human biological monitoring exposure data.

This resulted in a total nine chemicals on the draft list of HPV/pesticide inert chemicals (i.e., 10 in four pathwaysand one in three pathways where the human biological monitoring exposure pathway was represented). Table 5-3 presents the draft initial list of nine HPV/pesticide inert chemicals to undergo screening in the EDSP, along with an indication of the pathways in which they appeared. Because this list of HPV/pesticide inert chemicals was selected on the basis of exposure potential only, it should not be construed as a list of known or likely endocrine disruptors.

Table 5-3
High Production Volume Pesticide Inerts

Chemical Name	CAS Number	Total Pathways	Human	Eco	Water	Air
Acetone	67641	4	•	•	•	•
Butyl benzyl phthalate	85687	4	•	•	•	•
Dibutyl phthalate	84742	4	•	•	•	•
Diethyl phthalate	84662	4	•	•	•	•
Dimethyl phthalate	131113	4	•	•	•	•
Di-sec-octyl phthalate	117817	4	•	•	•	•
Methyl ethyl ketone	78933	4	•	•	•	•
Toluene	108883	4	•	•	•	•
Isophorone	78591	3	•	•	•	
Total = 9 HPV/pesticide inert chemicals						