ExpoCastDB – Observational human exposure database

Release v2011Q3a (August 2011)
Produced by the U.S. Environmental Protection Agency,
National Center for Computational Toxicology

Table of Contents

Introduction

ExpoCastDB was developed by the Environmental Protection Agency's (EPA) National Center for Computational Toxicology (NCCT) in partnership with EPA's National Exposure Research Laboratory (NERL) to store data from observational studies of potential human exposure to environmental chemicals. ExpoCastDB is one of four main repositories that compose EPA's Aggregated Computational Toxicology Resource (ACToR), an online datawarehouse which provides linkages among publicly accessible hazard, exposure, use and risk resources. Currently, ExpoCastDB includes measurements of chemical concentrations, mainly pesticides, in environmental and biological media from the American Healthy Homes Survey, the First National Environmental Health Survey of Child Care Centers and the Children's Total Exposure to Persistent Pesticides and Other Persistent Organic Pollutants study. The data from each study were categorized into the entities shown below in the conceptual data model. Through this web interface the data is provided in an accessible manner. Descriptive statistics on chemical concentrations from an individual medium are presented from each study. The entire set of structured measurement data (extracted from the database) is available for download from each individual study.
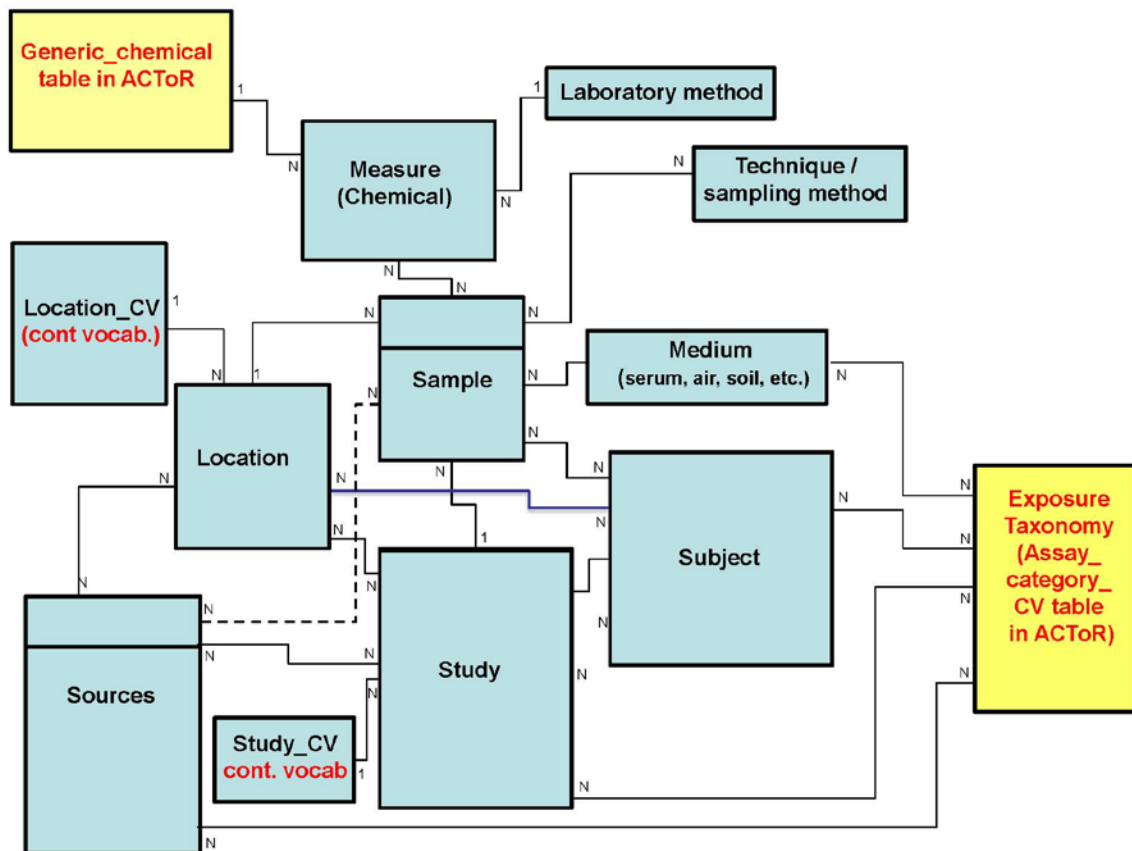
Questionnaire and laboratory metadata are not included in ExpoCastDB. Additional background information and examples of the utility of ExpoCastDB are being prepared in a manuscript.

ExpoCastDB is anticipated to encourage standardized reporting of observational exposure information and facilitate cross study analyses. Future versions of ExpoCastDB will provide additional user functionality including access to additional data sets, enhanced filtering capability to facilitate downloading of customized data sets by combining data from multiple studies, capability to estimate distributional parameters for exposure modeling, and visualization tools.

TOP

Overview: How to Use ExpoCastDB

In ExpoCastDB chemicals are organized by study, media type, and chemical. Summary statistics are available and the data are also downloadable for customized analysis.

Browsing

One can find information on ExpoCastDB chemicals by either browsing or by searching for a particular chemical. There are three main views:
- [Study List](#): Displays all studies in the database.
- [Chemical List](#): Displays all chemicals in the database including visual structure.
- [Media List](#): Displays all media in the database.

Searching

One can search for chemicals either by entering the name or Chemical Abstract Services (CAS) registration number.
- [Search by Name](#)- allows one to search for chemicals by their common and their many alternate names.
- [Search by CAS](#)- enables one to search by CAS registration number

After conducting a search, it is possible to go back and edit what has been entered. To do this, do not click the back button as your search results will be lost. Instead re-click on the "Search by (name/CAS/structure)" link.
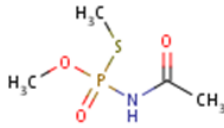
Search Results

The search results are composed of a table with the following columns:

Structure – structural drawing of the chemical
CASRN – CAS Registration Number
Name – the chemicals preferred name
Summary Statistics – displays a "data" link to a page listing descriptive statistics for the chemical for each study and media or sample type

Summary Statistics

ExpoCastDB

You are here: EPA Home » National Center for Computational Toxicology ExpoCastDB » Chemical

ACToR | ToxRefDB | ToxCastDB | ExpoCastDB | DSSTox

Home | Basic Info | Study List | Chemical List | Media List | Help

Chemical: Acephate

| CASRN | 30560-19-1 |
| ACToR | Find in ACToR DB |

American Healthy Homes Survey (AHHS)

Media/Sample Type	N	Detection Freq.(%)	Median MDL	Units	Max	Mean	SD	Geom. Mean	Geom. SD	25th %	75th %

Child Care Center Survey (CCC)

Media/Sample Type	N	Detection Freq.(%)	Median MDL	Units	Max	Mean	SD	Geom. Mean	Geom. SD	25th %	75th %
Exposure; Media; Soil	116	51.0	25.0	ng/g	1900.0	56.7	184.0	28.6	2.24	16.7	33.2
Exposure; Media; Surfaces - Total Residue	247	1.6	0.0269	ng/cm2	1.62						

Clicking on the "data" link will take you to the "Summary Statistics" page (above), where the chemical's statistics for each study and media/sample type are presented. There is also a link-out to the ACToR core database page for that chemical ("Find in ACToR DB"). The statistics are listed in a separate table for each study and includes:

Media/Sample – displays the media from which samples were collected.

N – displays the total number of samples for that chemical.

Detection Freq. (%) –displays the number of samples with values above the detection limit divided by the total number of samples (N), shown as a percentage. If the DF is <25%, then only the maximum measured value 'Max' is listed, otherwise if DF ≥25% all statistics are displayed.

Median MDL – displays the median, or middle value, of the method detection limit (MDL) for the particular chemical in the particular medium.

Units – displays the units in which the chemical level was reported, for example mg/L or ng/g.

Max – displays the largest recorded value of the chemical for all the samples in that medium.

Mean – displays the arithmetic average value of the chemical for all the samples in that medium.

SD – displays the Standard Deviation (SD) of the measured values for the chemical for all the samples in that medium, and represents the amount of variation among the data set from the mean.

Geom. Mean – displays the Geometric Mean of all the samples for the chemical in that medium. The Geometric Mean is the most appropriate descriptor of central tendency for a lognormal distribution and is defined as the nth root of the result of all values multiplied together with n = the number of values.

Geom. SD – displays the Geometric SD of all the samples for the chemical in that medium, and represents the amount of variations among a data set whose preferred average is the Geometric Mean.

25th % – displays the calculated 25th Percentile (value that 25% of the measurements fall below) for the data set in that medium.

75th % – displays the calculated 75th Percentile (value that 75% of the measurements fall below) for the data set in that medium.

TOP

Study List

All information in ExpoCastDB is obtained from various study results. The entire list of Study collections in ExpoCastDB can be seen by selecting Study List in the top left hand navigation bar. For each Study, the following data are presented:

Name – displays the title of the individual study performed
Description – an abstract of the study performed
Samples – the total number of samples collected in the study
Chemicals – the total number of unique chemicals assessed in the study
Measurements – the total number of analytical results from all samples within the study
Summary Statistics – displays a link to the page with summary statistics tables for each media type used in that study, including the chemical Name and CASRN, followed by the same statistics in the Summary Statistics page above
All Structured Data – displays a link to download all raw measurement data from the study as a zip file
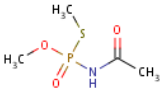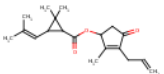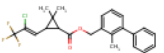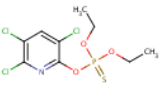
TOP

Chemical List

The Chemical List page shows all unique chemicals represented in all studies in ExpoCastDB. This page can be reached by clicking on Chemical List from the home screen.

The Chemical List table is composed similar to the Search Results table with the following columns:

Structure – structural drawing of the chemical
CASRN – CAS Registration Number
Name – the chemicals preferred name
Summary Statistics – displays a "data" link to a page listing descriptive statistics for the chemical for each study and medium or sample type in the same order as the Summary Statistics page above

TOP

Media List

This page is accessed by clicking on "Media List" and provides a complete list of all media and sample types used from all Studies in the ExpoCast database. Data is listed as a table with four columns:

Media/Sample Type – the media or sample types analyzed for chemicals

Num Chemicals – the number of unique chemicals assessed in that medium/sample for all studies

Num Measurements – the total number of measurements of all chemicals in all samples in each study in that medium or sample type

Summary Statistics – displays a link to a Summary Statistics page similar to the one from Study List but with tables for each Study and only for the medium or sample selected

TOP

Glossary

ACToR

(Aggregated Computational Toxicology Resource) is a collection of databases collated or developed by the EPA National Center for Computational Toxicology

Assay

An assay is a collection of data for substances from one data collection. Currently, an assay can be thought of a simple table. An assay falls into one data type category but may have multiple phenotypes. An assay can have more than one row or entry for the same substance, and elements in the data matrix can be empty.

Assay category

Assays are organized into a number of categories that describe the broad type of data presented. Several of these categories describe the level of biological organization being probed, while others describe the class of information being presented. The current sets of categories are:

> **PhysicoChemical**- physical and chemical properties (in vitro and/or in silico)
> **Biochemical**- chemical processes in living organisms that are non-cell based
> **Genomics**- gene expression values or signatures
> **Cellular**- cell-based assay
> **Tissue**- tissue slice assay
> **Organ**- focus on organs
> **Organism**- focus on organisms (animal testing)
> **In vivo toxicology (tabular primary) -** tabulated results from primary animal-based studies of chemical effects
> **In vivo toxicology (study listing primary**) - primary studies are available but have not been tabulated
> **In vivo toxicology (tabular secondary**)-tabulated data from secondary sources for in vivo toxicology studies
> **In vivo toxicology (summary calls**) - derived summary determinations of risk
> **In vivo toxicology (summary report via URL**) - links to text reports on the web for which specific data values are not directly accessible in tabular form
> **General Descriptive information**- a brief description of the chemical
> **Regulation**- listings of chemicals that fall under specific environmental laws, government mandates, or standards

Assay component

An assay component defines one column or element of an assay. A component has a unique ID, a name, a description, a data type, and optionally units.

Assay phenotype

Some assays are characterized by toxicology phenotypes. This allows one to organize the data in ACToR into broad toxicity areas. The current set of phenotypes is:

> **Hazard** – Information on basic harm that can be caused by a chemical. Information under this category includes workplaces safety and first aid in case of exposure.
> **Acute Toxicity** – Information on health effects due to short term exposure.
> **Subchronic Toxicity** - Information on health effects due to intermediate term exposure.
> **Chronic Toxicity** - Information on health effects due to long term exposure.
> **Carcinogenicity** – Information from studies of cancer-causing ability of chemicals.
> **Genetic Toxicity**- Information on the ability of chemicals to cause DNA damage.
> **Reproductive Toxicity**- Information on the ability of a chemical to damage an organism's reproductive ability.
> **Neurotoxicity** - Information on the ability of a chemical to damage nerve cells or tissues.
> **Developmental Neurotoxicity** – Information on chemicals that cause neurological deficits during development
> **Immunotoxicity** - determines how a chemical affects the immune system
> **Dermal Toxicity**- includes studies about chemicals and what level of toxicity results from the substances being applied through the skin. May or may not have the skin as a target organ.
> **Respiratory Toxicity**- contains data on how the chemicals that affect the respiratory system
> **Nephrotoxicity**- measures to what degree that the chemical that affect the kidneys
> **Endocrine**- contains data about if and how the chemical affect hormone signaling and downstream processes
> **Cardiotoxicity**- contains data about how the chemical affects the heart
> **Ecotoxicity**- includes data about how chemicals affect non-human species such as fish and amphibians
> **Food Safety**- includes data that determines if the food can be safely used as an ingredient, additive or food wrapper
> **Toxicity other**- contains other information about a chemical's toxicity
> **PK/metabolism**- Information on pharmacokinetics and metabolism of xenobiotic chemicals

Assay result

An assay result is one data point for a single substance and a single assay component.

Assay types

There are two main types of assays: phenotypes and categories.

CAS

CAS (Chemical Abstract Services) Registry Number ([for more information](#))

Some examples of number in CAS format are:
7439-92-1
7440-50-8
79-34-5
59325
39001-02-0
59001050

Chemical

A chemical is defined by a unique chemical ID in the database and can be either a substance or a compound.

Chemical structure

Diagram of a chemical- can be used to search for information about chemicals.

Compound

A compound is an entity with a chemical ID and chemical structure information, which may be a 2 or 3 dimensional molfile or a string representation. This can be SMILES or InCHII.

Data collection

A data collection is at minimum a set of substances with corresponding CAS registry numbers and names. Additional information may include chemical structures and assays. As mentioned above, a generic chemical links together data from many data collections on all substances that share a common CAS registry number. All data is initially compiled as part of a set of Data Collections.

Detection Frequency

This category displays the number of samples with values above the detection limit divided by the total number of samples (N), shown as a percentage.

Exposure data

Exposure studies measure the amount of a substance that people and animals may potentially contact and/or take into their bodies. These data may not directly explain how the person was exposed nor do these data provide information on the potential for health effects related to the exposures.

Generic chemical

A generic chemical aggregates all data from all data collections for substances with a single given CAS number. It will have links to one or more substances and all of their related assay data, as well as all synonyms derived from the substances.

HPV and MPV

HPV stands for High Production Volume industrial chemicals and MPV for Medium Production Volume industrial chemicals

Includes the chemical structure and calculated physical chemical properties of compounds produced or imported into the United States

Inert ingredient

An inert ingredient means any substance other than an active ingredient. Inert ingredients tend to be carriers for the active ingredients. (for more information)

In vitro

An experiment that is performed outside of a living organisms (for examples test tubes)

In vivo

Experimentation done on or inside of living organisms- other wise known as animal testing

Media/Medium

Soil, water, air, dust, biota (plants and animals), or any other parts of the environment that can contain contaminants

MESH

The U.S. National Library of Medicine's controlled vocabulary used for indexing articles for MEDLINE/ PubMed. MESH terminology provides a consistent way to retrieve information that may use different terminology for the same concepts.

Method Detection Limit

The minimum concentration of a substance that can be measured and reported with 99% confidence that the analyte concentration is greater than zero

SMILES

SMILES (Simplified Molecular Input Line Entry System) is a *line notation* (a typographical method using printable characters) for entering and representing molecules and reactions. (for more information)

Study

A study is a collection of a specific data collection of chemical exposure.

Substance

A substance is an entity with a chemical ID, one or more names (including a CAS number) and potentially a URL pointing to primary data. One special name for the substance is the "source name sid" which is a unique alphanumeric label from the source, which allows a unique link back to the source.

TOP