

ExpoCastDB: A Publicly Accessible Database for Observational Exposure Data

Presented to: CompTox Communities of Practice

Sumit Gangwal, Ph.D.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

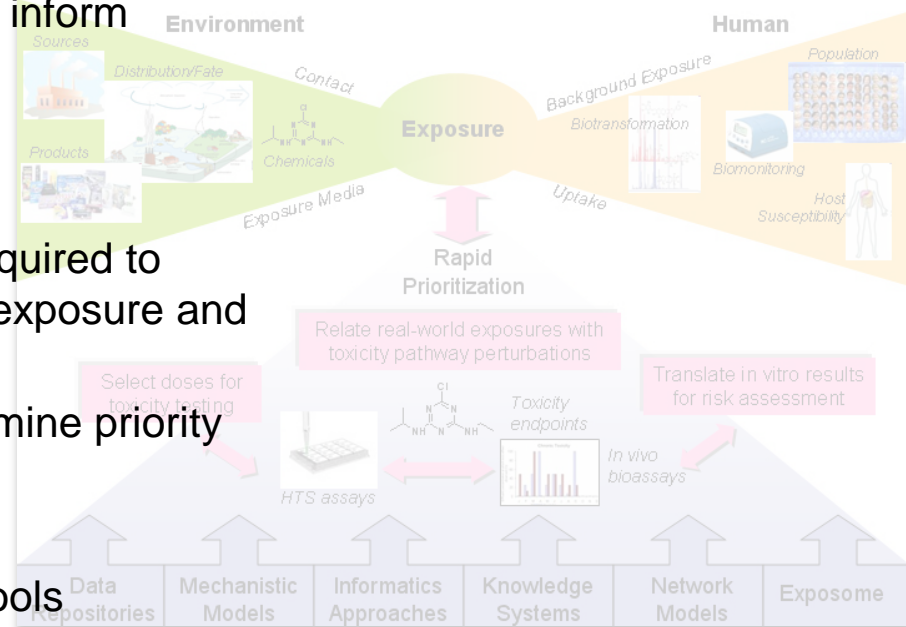


COMPUTATIONAL
TOXICOLOGY

- ExpoCast project
 - ExpoCast DB motivation and goals
- ACToR – Aggregated Computational Toxicology Resource
 - Exposure taxonomy
- ExpoCastDB initial release
 - Studies Included and QC of data
 - Demo
 - Applications
 - Future phases

ExpoCast™: Exposure science for prioritization and toxicity testing

- Recognizes critical need for exposure info to inform
 - Chemical design and evaluation
 - Health risk management
- Goal
 - Advance characterization of exposure required to **translate** findings in comptox to support exposure and risk assessment
 - Together with ToxCast™ help EPA determine priority chemicals
- Approach
 - Mine and apply scientific advances and tools
 - Develop novel approaches for evaluating chemicals based on **biologically-relevant** human exposure



Cohen Hubal et al., J Toxicol Environ Health, Part B, 2010

Priority Exposure Research for Computational Toxicology

- Accessible and linkable exposure databases
- Exposure-based screening tools for accelerated chemical prioritization

Motivation for ExpoCastDB

- Increase access to human exposure data
 - Support exposure modeling
 - Advance public health through improved management of chemical risks

Sexton et al., Arch Environ Health, 1992;
Wagener et al., Annu Rev Public Health, 1995
- Linkages between exposure and toxicity data must be implemented

Sheldon and Cohen Hubal, EHP, 2009

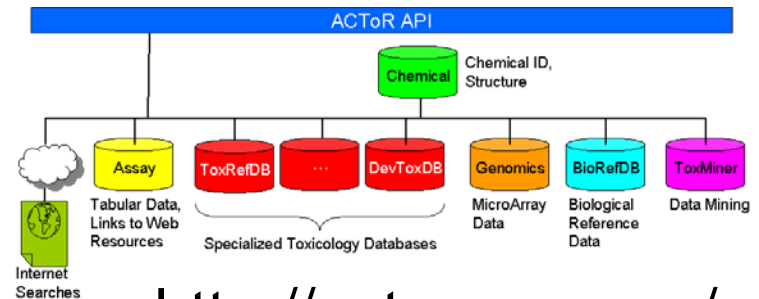
 - Achieve vision of better integration of exposure considerations into the risk assessment process

NAS, 2007, “Toxicity Testing in the 21st Century: A Vision and a Strategy”
 - Consolidation facilitates bioinformatic interrogation
 - Estimated proportion of human disease attributable to environmental pollutants is roughly the same as that attributable to genetic factors

Rappaport, JESEE, 2011

ExpoCastDB goals

- Consolidate observational human exposure data, improve access and provide links to health related data
 - House measurements from human exposure studies
 - Encourage standardized reporting of observational exposure information
- Provide separate interface with inner workings of ACToR
 - Facilitate linkages with toxicity data, environmental fate data, chemical manufacture information



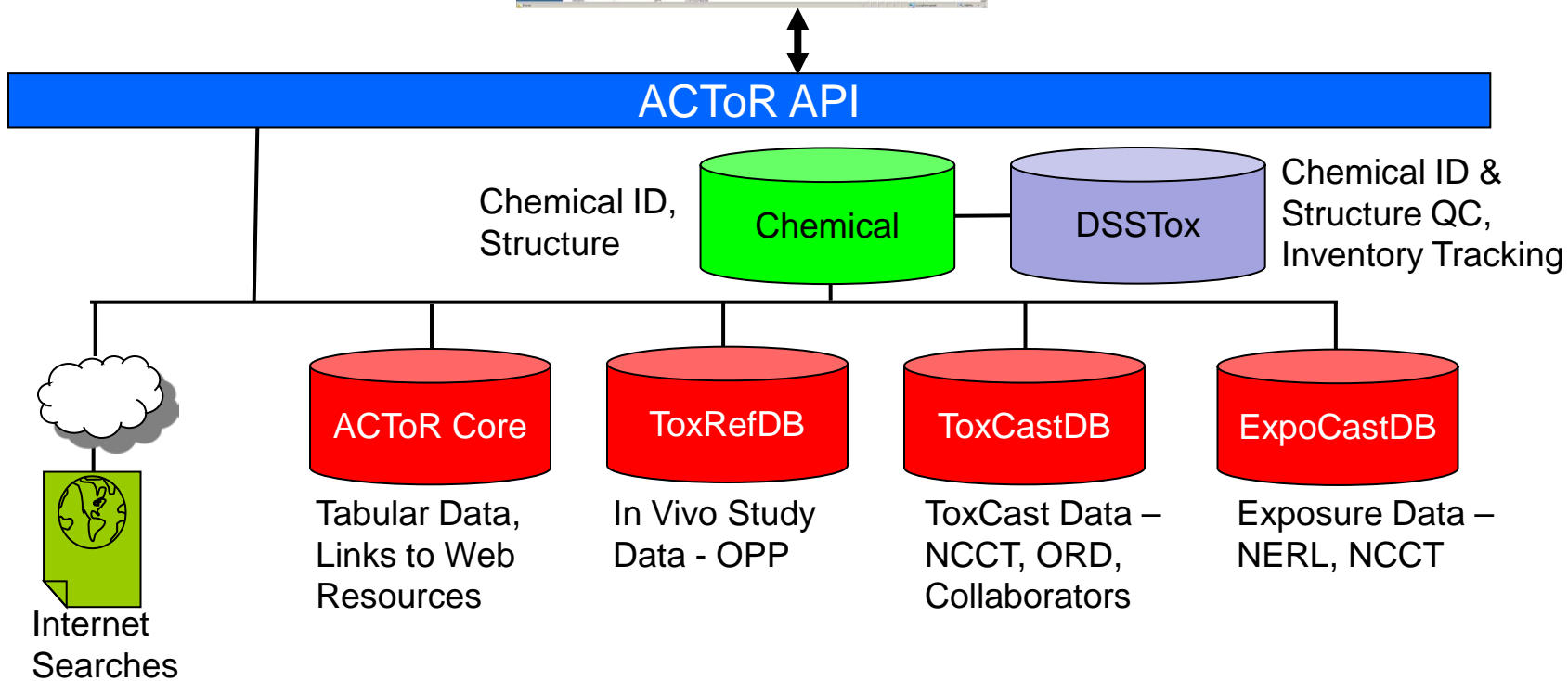
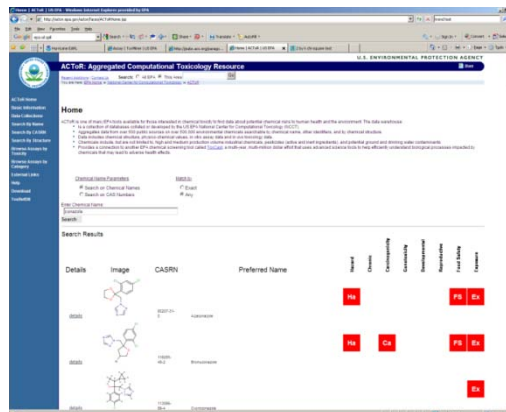
<http://actor.epa.gov/>

- Provide basic user functions
 - Visualization (e.g., scatterplots, probability plots, goodness-of-fit)
 - Obtain summary statistics and estimate distributional parameters
 - Download customized datasets

ACToR

Aggregated Computational Toxicology Resource

<http://actor.epa.gov/>



ACToR goals and data sources

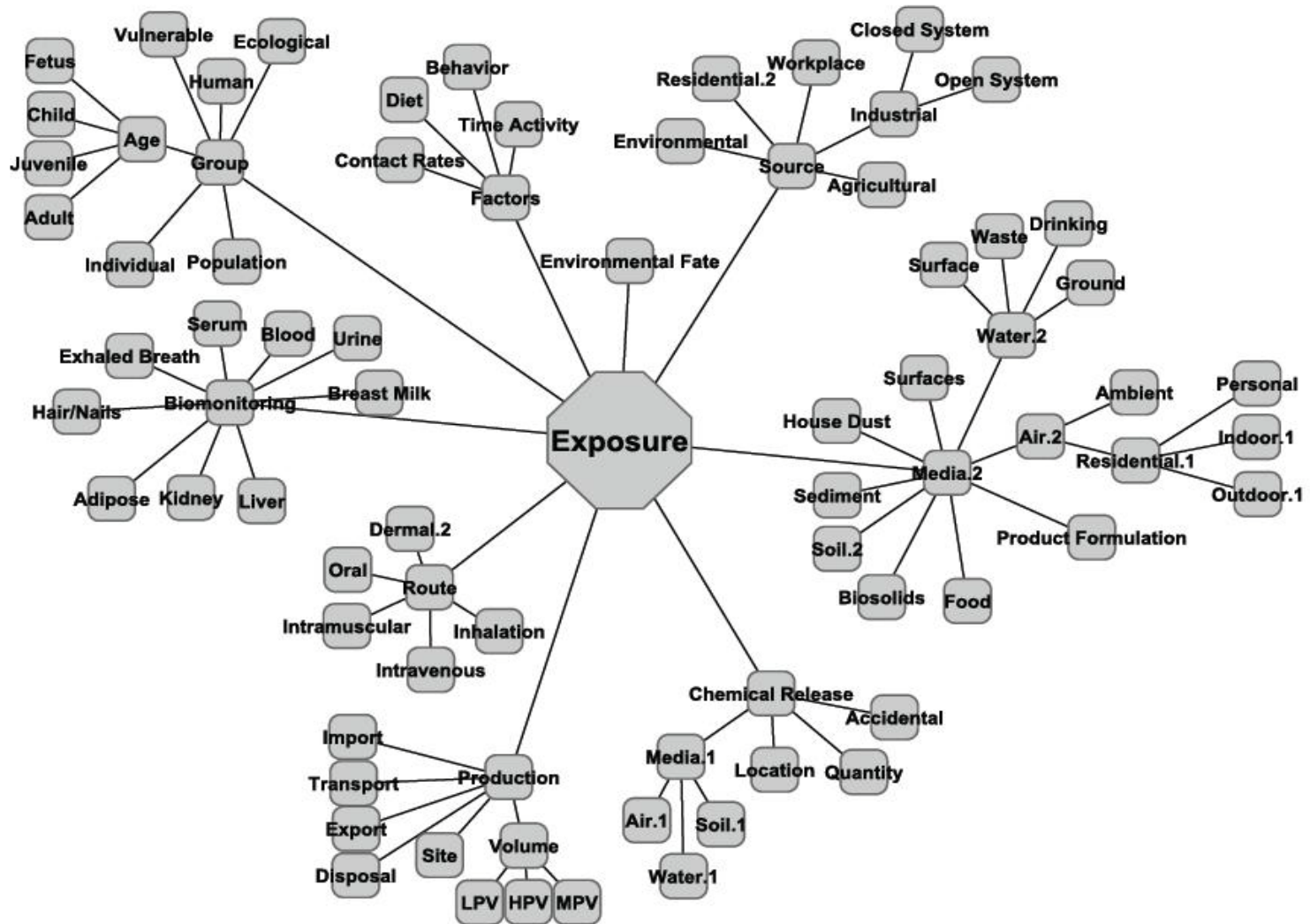
- Compile all publicly available information on environmental chemicals

Category	Count
Data Sets	580
Chemicals	546,956
Assays	3,213
Assay Components	7,221
Data Points	6,662,296

- EPA (OPP, OPPT, NCEA, NERL)
- FDA, NIH, CDC, OSHA, USDA
- States and other countries
- Universities
- NGOs
- Companies

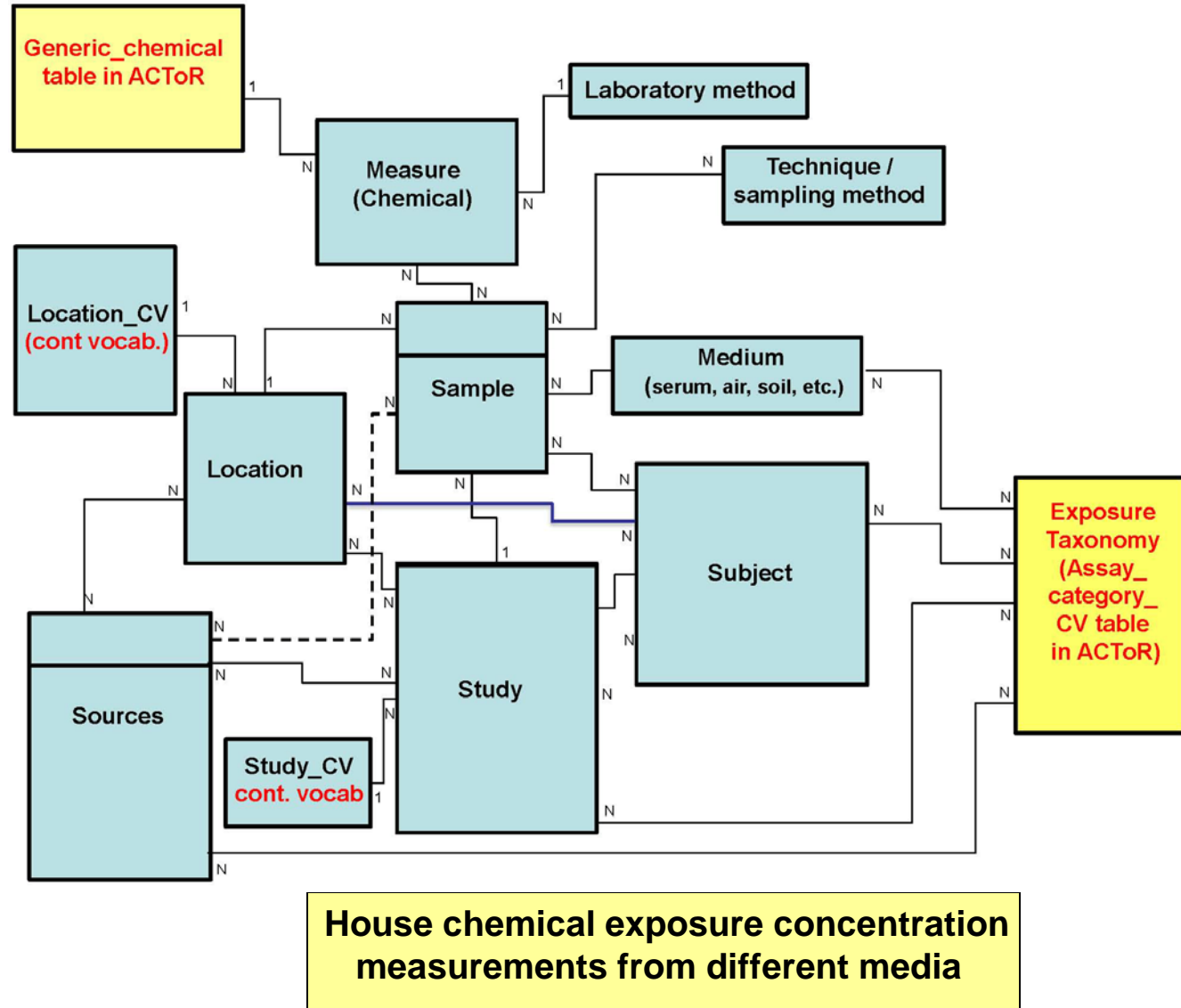
- Make data available for downloading, data mining
 - Available through data.gov
 - Entire DB can be downloaded and installed locally
- Make it easy to see data gaps
 - Provides resource for EPA testing programs
- Make it widely used
 - over 2000 regular users

ACToR exposure taxonomy



ExpoCastDB initial release

- Data from NERL studies
 - 1) American Health Homes Survey
 - 2) HUD Child Care Center Survey (“CCC”)
 - 3) CTEPP – NC
 - 4) CTEPP – OH
- Full raw data sets available to download
- Browse data capability
 - By study name, chemical list, media list
- Descriptive statistics capabilities:
 - Detection Freq., N, max, MDL (if result concentration is < MDL, then result = 2/3 * MDL), Units
 - Mean, Std. dev., geometric mean, geometric std. dev. (GSD), 25th and 75th percentile



- American Health Home Survey (AHHS):
 - 500 homes underwent pesticide sampling across US.
 - Isopropranol wetted wipe media used to sample floor surfaces.
 - 24 residential-use insecticides represented by 5 chemical classes:
 - Organochlorines, Organophosphates, Neonicotinoid, Synergist, Pyrethrins/Pyrethroids

- HUD Child Care Center Survey (“CCC”)
 - 168 child care centers in 30 primary US sampling units.
 - Wipe samples from indoor surfaces (floors, tabletops, desks) and soil samples collected.
 - Samples for pesticides (organophosphates and pyrethroids), lead, and allergens collected at multiple locations in each center.



ExpoCastDB: Initial studies (con't.)

- Children's Total Exposure to Persistent Pesticides and Other Persistent Organic Pollutants (CTEPP) – NC, OH
 - Observation measurement study designed to determine what commonly used chemicals are found in home and/or day care environments in six NC and six OH counties.
 - Samples of food, drinking water, air, urine, dust, soil, transferable residues on floors, and surface wipes collected.
 - Measured total exposure of 257 preschool children (ages 2 to 5 years) and their primary adult caregivers to more than 50 different pesticides and other chemicals.



Transferable residues sample



Outdoor air sample

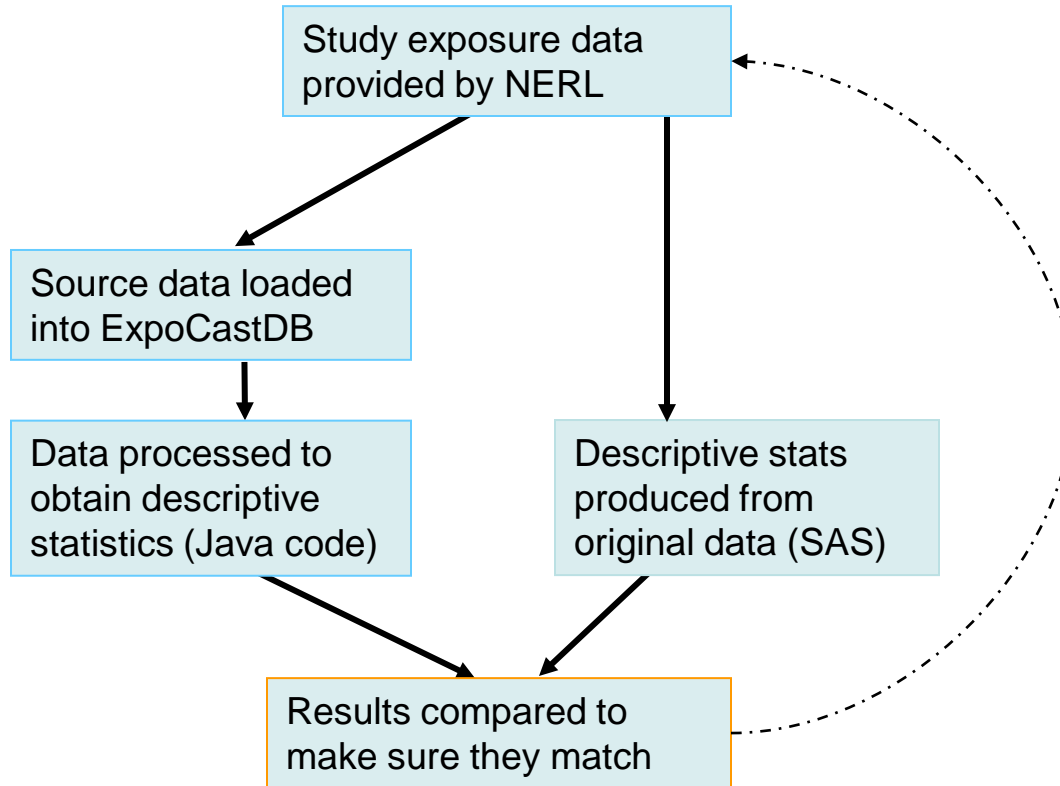


Hand wipe sample



Soil sample

Quality control of ExpoCastDB data



ExpoCastDB demo

ExpoCastDB:

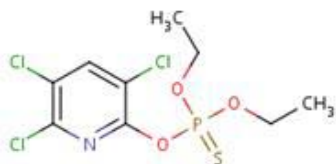
<http://actorpreprod.epa.gov/actor/faces/ExpoCastDB/Home.jsp> (beta)

<http://134.67.216.45:9192/actor/faces/ExpoCastDB/Home.jsp> (beta)

ExpoCastDB: Applications

- Investigate occurrence and co-occurrence of chemicals
 - How frequently was a particular pesticide (e.g., esfenvalerate) detected in nationwide AHHS study?
- Comparisons across studies
 - How does avg. permethrin loading differ between homes in AHHS and daycares in CCC?
- Download individual sample-level structured data
 - Derive input distributions for probabilistic exposure models (such as SHEDS)
 - Use with readily available algorithms (U.S. EPA, 1992) and exposure factors (U.S. EPA, 1997) to produce quick deterministic estimates of uptake resulting from exposure to specific chemicals

Chemical: CHLORPYRIFOS



CASRN
ACToR

2921-88-2
[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 %
Exposure; Media; Surfaces - Total Residue	479	78.0	8.61E-4	ng/cm2	136.0	0.493	6.28	0.0138	10.7	0.00266	0.0627

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	117	13.0	5.0	ng/g	1150.0						
Exposure; Media; Surfaces - Total Residue	248	38.0	0.084	ng/cm2	27.6	0.34	1.89	0.0968	2.82	0.056	0.132

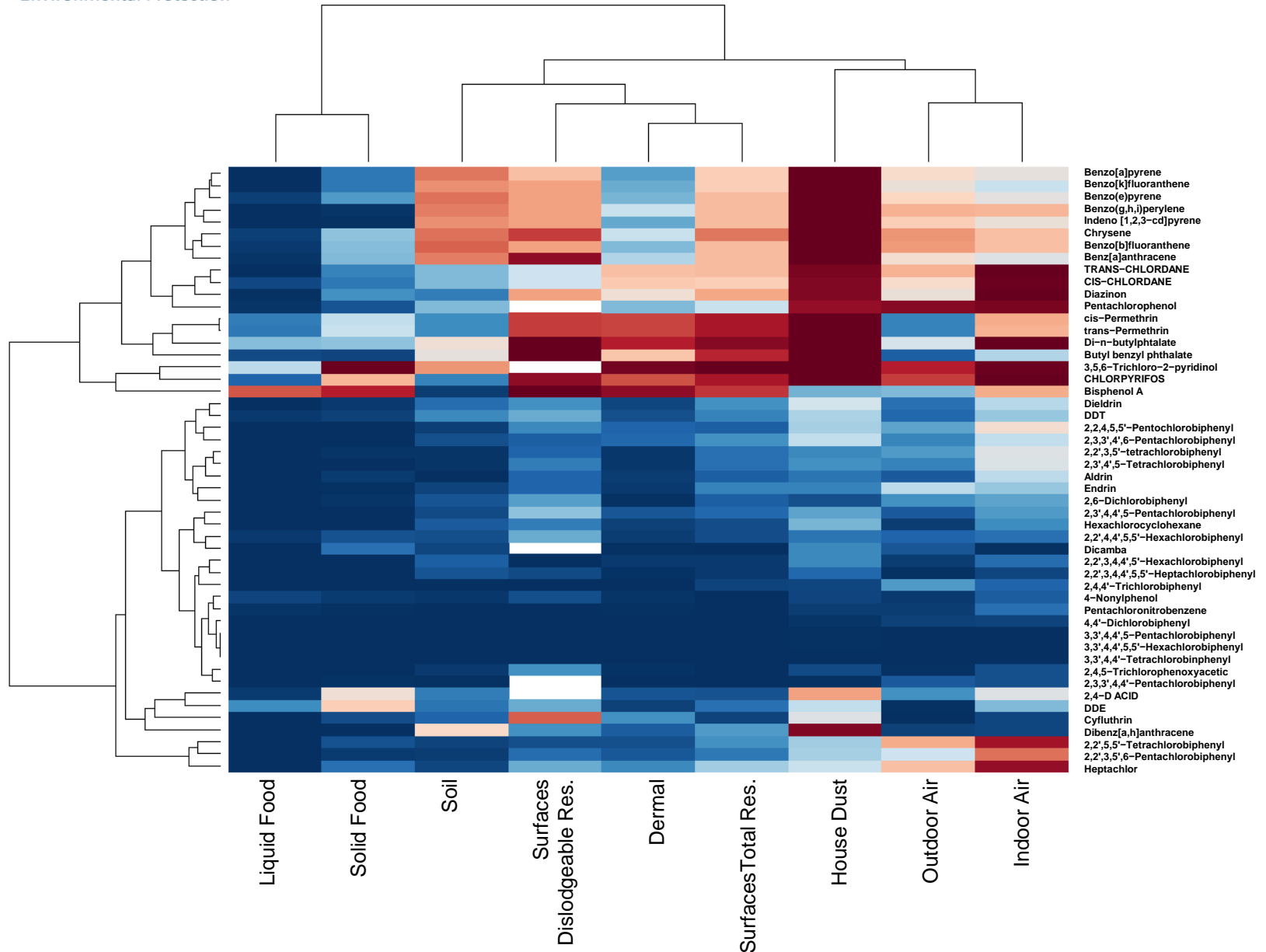
CTEPP NC

Media/Sample Type	N	Detection Freq.(%)	Median MDL	Units	Max	Mean	SD	Geom. Mean	Geom. SD	25th %	75th %
Exposure; Biomonitoring; Dermal	223	80.0	0.00317	ng/cm2	0.744	0.04	0.0799	0.0152	4.07	0.00495	0.0365
Exposure; Media; Air	288	92.0	0.09	ng/m3	391.0	9.48	29.4	1.45	7.65	0.28	6.7
Exposure; Media; Air; Residential	288	92.0	0.09	ng/m3	391.0	9.48	29.4	1.45	7.65	0.28	6.7
Exposure; Media; Air; Residential; Indoor	148	100.0	0.09	ng/m3	391.0	17.5	39.3	6.45	3.96	2.25	17.4
Exposure; Media; Air; Residential; Outdoor	140	83.0	0.09	ng/m3	45.9	0.998	4.02	0.299	3.64	0.11	0.638
Exposure; Media; Food	153	63.0	0.08	ng/g	19.7	0.513	1.68	0.188	3.4	0.0533	0.39
Exposure; Media; Food	148	11.0	0.03	ng/mL	1.71						
Exposure; Media; House Dust	141	100.0	2.0	ng/g	15100.0	389.0	1330.0	137.0	3.66	51.1	293.0
Exposure; Media; House Dust	141	100.0	3.47	ng/m2	51600.0	1540.0	5990.0	174.0	7.6	38.2	786.0
Exposure; Media; Soil	142	18.0	0.49	ng/g	1170.0						
Exposure; Media; Surfaces - Dislodgable Residue	18	94.0	4.39	ng/m2	719.0	142.0	230.0	47.4	4.52	15.2	140.0
Exposure; Media; Surfaces - Total Residue	50	90.0	6.9	ng/m2	2080.0	212.0	414.0	67.3	4.54	29.1	135.0

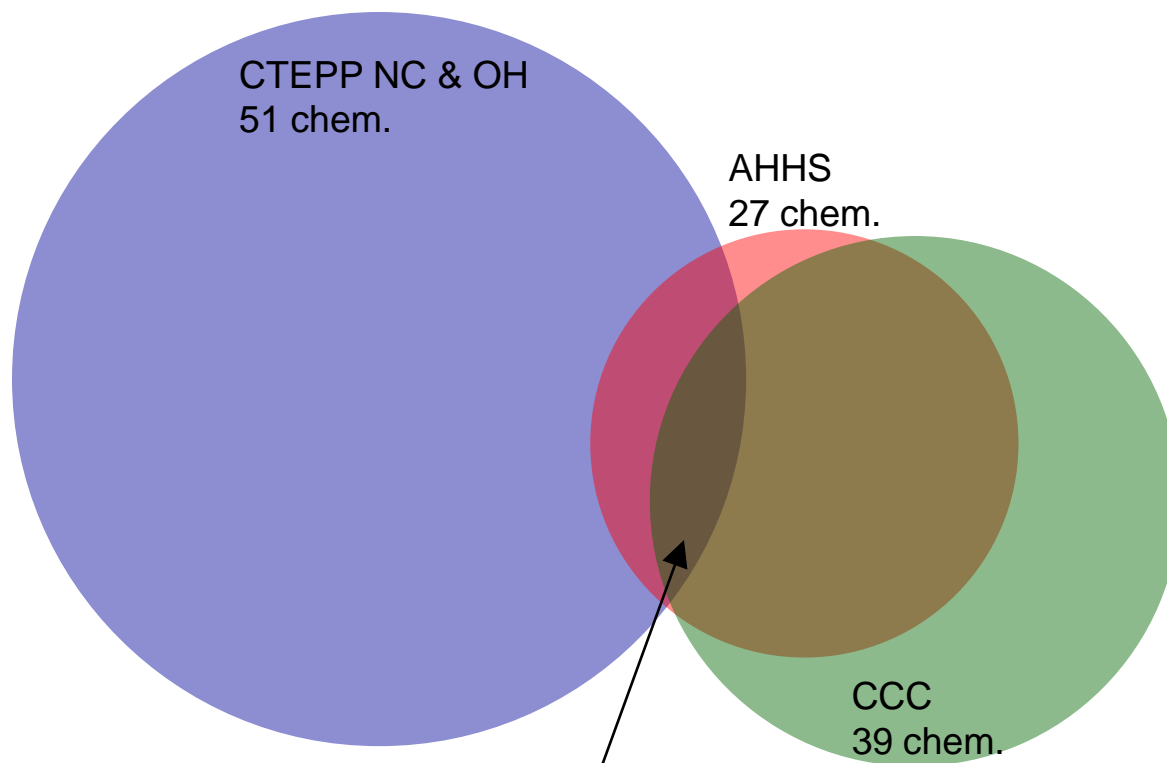
CTEPP OH

Media/Sample Type	N	Detection Freq.(%)	Median MDL	Units	Max	Mean	SD	Geom. Mean	Geom. SD	25th %	75th %
Exposure; Biomonitoring; Dermal	221	60.0	0.00303	ng/cm2	14.8	0.104	1.01	0.00785	4.9	0.00242	0.0211
Exposure; Media; Air	273	88.0	0.09	ng/m3	98.0	3.54	10.6	0.74	5.49	0.21	2.08
Exposure; Media; Air; Residential	273	88.0	0.09	ng/m3	98.0	3.54	10.6	0.74	5.49	0.21	2.08
Exposure; Media; Air; Residential; Indoor	147	99.0	0.09	ng/m3	98.0	6.24	13.8	2.26	3.75	0.93	5.82
Exposure; Media; Air; Residential; Outdoor	126	75.0	0.09	ng/m3	6.5	0.389	0.746	0.201	2.85	0.06	0.393
Exposure; Media; Food	153	66.0	0.08	ng/g	3.51	0.347	0.56	0.179	2.95	0.0533	0.385
Exposure; Media; Food	153	7.2	0.03	ng/mL	0.65						
Exposure; Media; House Dust	143	100.0	2.0	ng/g	49600.0	775.0	4610.0	81.0	4.83	25.6	178.0
Exposure; Media; House Dust	143	100.0	3.38	ng/m2	54100.0	1200.0	5120.0	102.0	7.58	24.6	400.0
Exposure; Media; Soil	143	39.0	0.5	ng/g	2930.0	23.7	245.0	0.946	4.78	0.333	3.47
Exposure; Media; Surfaces - Dislodgable Residue	13	85.0	4.39	ng/m2	32200.0	2500.0	8920.0	25.5	11.5	4.87	41.0
Exposure; Media; Surfaces - Total Residue	39	69.0	6.9	ng/m2	38600.0	1270.0	6260.0	27.7	8.63	4.6	65.9

Heatmap of CTEPP NC chemical detection frequencies in different media



Chemical overlap of CTEPP (NC,OH), AHHS, CCC studies



Only 5 chemicals overlap:

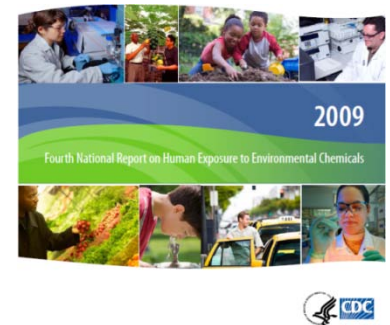
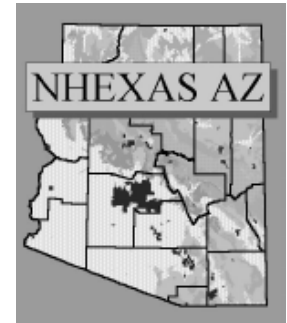
Name	CASRN
CHLORPYRIFOS	2921-88-2
cis-Permethrin	61949-76-6
Cyfluthrin	68359-37-5
Diazinon	333-41-5
trans-Permethrin	61949-77-7

ExpoCastDB prototype application: Future phases

Next phase study data sets that can be added:

- National Human Exposure Assessment Survey (NHEXAS) – Arizona, “MD” (longitudinal), Region 5
- Minnesota Children’s Pesticide Exposure survey (MNCPEs ~ 90 kids)
- Non-occupational pesticide exposure study (NOPES)
- Total Exposure Assessment Methodology Study (TEAM)
- Particle Total Exposure Assessment Methodology (PTEAM)
- NHANES – National Health and Nutrition Examination Survey
- Other exposure data curated from literature . . .
 - Nanomaterials exposure data (measured particle counts in air)

Gangwal et al. 2011, Environ Health Perspect
“Informing Selection of Nanomaterial Concentrations for ToxCast In Vitro Testing based on Occupational Exposure Potential”



ExpoCastDB prototype application: Future phases

Next phase website application capabilities:

- Combine chemical concentration measurements from multiple studies and allow filters/search criteria:
 - Chemical concentrations must be in same media.
 - Allow for user input of entering chemicals, media, or study.
 - When combining data sets, calculate detection frequency criteria ($\geq 25\%$) based on entire data set for all studies selected.
- Data visualization including log probability plots.
- Additional meta data browsing by ancillary information (location, subject, etc.).

Acknowledgments

EPA National Center for Computational Toxicology (NCCT), ORD

Elaine Cohen Hubal
Shad Mosher
Richard Judson, Ann Richard, David Reif
David Dix
Robert Kavlock

EPA ORD

Peter Egeghy (National Exposure
Research Laboratory—NERL)
Marsha Morgan – NERL
Nicolle Tulve – NERL

Academia, Industry and others

Tommy Cathey – Lockheed Martin
Ruthann Rudel, Robin Dodson – Silent Spring Institute, Newton MA



This work was reviewed by EPA and approved for presentation but does not necessarily reflect Agency policy