

Application feature improvements in support of human health assessments: optimisations for epidemiology data extraction

Sean Watford, Krista Christensen*, Elizabeth Radke-Farabaugh, Andy Shapiro

> Center for Public Health and Environmental Assessment Office of Research and Development U.S. Environmental Protection Agency 10 November 2023



The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.

⇒EPA

EPA's Integrated Risk Information System (IRIS)

- Database of health effects information on hundreds of environmental pollutants
- IRIS assessments contribute to decisions across EPA and other health agencies
- Focus is on toxicity due to lifetime exposure
- Provides toxicity values for cancer and noncancer effects
- Have no direct regulatory impact until combined with extent of exposure, cost of cleanup, available technology, and other regulatory options that are the purview of other EPA programs



EPA's mission is to protect human health and the environment. EPA's IRIS Program supports this mission by identifying and characterizing the health hazards of chemicals found in the environment. Each IRIS assessment can cover a chemical, a group of related chemicals, or a complex mixture.

Basic Information

IRIS Assessments

- Browse A to Z List of Chemicals
 Browse by Organ/System
- Guidance & Tools

Learn About IRIS

IRIS Process

History of IRIS

- Assessments in Development



By Chemical, CASRN, or Keyword

Search the IRIS database of final assessr

Search

EPA's Integrated Risk Information System (IRIS)

IRIS assessments contribute to EPA decisions such as:

- Health-based national standards
- Health-based clean-up levels at local sites
- Health-based advisory levels
- Ranking across chemicals
- Information for the general public
- Cost-benefit analyses

SEPA

- **Clean Air Act (CAA)** Safe Drinking Water Act (SDWA) **Food Quality Protection Act (FQPA)** Supports **Comprehensive Environmental Response,** Compensation, and Liability Act (CERCLA) **Toxic Substances Control Act (TSCA)** RIS **Resource Conservation and Recovery Act (RCRA) Agency Strategic Goals Regions and States** Broad **Children's Health** Input to
 - Environmental Justice





IRIS Assessment Development Process

- These systematic review (SR) methods are resource- and time-intensive, yet increasingly a foundational part of the chemical assessment process
- Software applications and tools can help to:
 - Improve user interfaces and interactions (UI/UX)
 - Standardize data exchange formats
 - Utilize artificial intelligence for (semi-)automation.
- Focus for this presentation is the IRIS program's use of the Health Assessment Workspace Collaborative (HAWC)
 - HAWC has data extraction features for both animal toxicology and epidemiology studies.
 - We will describe recent updates made by the HAWC team in coordination with EPA epidemiologists to update data extraction features including:
 - Updates to UI/UX
 - Increased flexibility to accommodate partial extractions





The Health Assessment Workspace Collaborative (HAWC)

SEPA What is HAWC?

- Web-based content management system for human health data
 - Used by EPA, NTP, WHO/IARC, CalEPA, and others
- Open-source Python application
 - Source code is available at <u>https://github.com/usepa/hawc</u>
 - EPA deployment available for EPA projects and collaborators: <u>hawc.epa.gov</u>
 - Public deployment is available at <u>https://hawcproject.org</u>*
 - Custom, private deployments are possible (MIT licensed)
- Assessments with tiered access
 - Managed read/write access
 - Assessments can be public



Environmental Topics Laws & Regulations Report a Viol

Report a Violation About EPA

Health Assessment Workspace Collaborative (HAWC) Contact Us

Search EPA.gov

Q



Researchers and regulatory agencies around the world conduct assessments to determine the potential for chemicals and other pollutants to pose a risk to human health and the environment. These assessments typically consist of a critical review of available studies, identification of health and environmental effects, and characterization of exposure-response relationships and uncertainties in the data. HAWC aims to facilitate team collaboration by scientists who develop these assessments and enhance transparency of the assessment process by providing online access to the data and expert decisions used to evaluate the potential human health and environmental hazard and risk of chemical exposures.

Background Information

- Learn about HAWC
- History of HAWC
- Frequent Questions

HAWC Assessments

- Public Assessments
- Decent Du
 - Recent Public Assessments
 - ORD Assessment PFPrA (2023)

Resources

- Publications
- User Login
- Technical Support

* Mention of or referral links to non-EPA sites does not imply official EPA endorsement of or responsibility for the opinions, ideas, data, or products presented at those locations, or guarantee the validity of the information provided.

SEPA What is HAWC?

Endpoint

Clinical Observation

Feed Consumption

- Key modules:
 - Literature screening
 - Risk of bias/study evaluation
 - Animal bioassay data extraction
 Epidemiology data extraction
 - Interactive summary tables and visualizations
- Interactive "supplemental materials" for reports and data re-use





Epidemiology Data Extraction

- Multiple data fields are extracted from epidemiology studies during assessment development
 - Study design features (e.g., design, population, sample size, time period)

- Chemicals (chemical name and other identifying information like CAS registry number)
- Exposures and exposure levels (e.g., route of exposure, exposure source, distribution of exposure)
- Adjustment factors (e.g., characteristics used for matching or statistical adjustment in multivariate models)
- Outcomes (e.g., assay or instrument used, timing of measurement)
- Quantitative results (e.g., relative risks point and variance estimates)
- The extraction features are flexible and enable the development and storage of information for multiple purposes: evidence maps, toxicological reviews, visualizations, and interoperability with other tools

SEPA Evidence Maps

- High-level summary of the available literature used to inform prioritization, and Scoping and Problem Formulation activities
- Uses tailored criteria to identify and categorize references as potentially relevant to human health risk assessment
- Used to identify key data gap and characterize level of effort and scientific issues to be considered
- Can be used as a tool for identifying needed expertise

Epidemiology studies examining exposure to PFAS

Epidemiological Studies Examining Exposure to PFAS

CAS-RN Outcome Overview of Epidemiological Evidence Base (AII) (AII) Expand Health Effect Category to Outcome by clicking the small [+] icor Population Chemical References Grand Pregnant 8:2-FTOH Health Effect Category 3M (2000)-5412700 Adult Tota Aimuzi et al. (2019)-5387.. 🛛 NEtFOSE 1 Cancer 12 PFDDA Aimuzi et al. (2020)-6512.. Cardiovascular 30 PFDeS Ait Bamai et al. (2020)-6.. DEU-A Developmental 26 Endocrine 30 Exposure Measure Study Design Hepatic Expand Exposure Measure by clicking the Case-control small [+] icon 22 Immune Cohort DIOMONITO Metabolic 37 Cross-sectional drinking wa 1 Nervous 21 occupationa Ecological 1 2 Other 3 5 Grand Total 193 Grand Total 193 Reproductive, female 27 Reproductive, male 1 14 **Overall Study Confidence** Respiratory 5 Some references have more than one overall confidence rating 11 Urinary 75 Grand Tota 64 Low 61 31 193 Uninformative 12 Epidemiological Study Details Sub Measured Effect Estimate Ci Lcl population Outcome Effect/Endpoints Comparison Cilld Chemical SMR for bladder cancer in the high exposure 3M (2000) Bladder cancer Bladder cancer 2083 16.1 33 47.1 ഒ group compared to no workplace exposure

Toxicological Reviews

Public Assessments	ODD IDIS Asses	smont DFH	$[v\Lambda(2022)]$			Actio	ons 🔻						
ORD IRIS Assessment	UND INIS ASSES	Sillent FFI	IAA (2023)										
PFHXA (2023)	Assessment name	ORD IRIS Assessment	PFHxA										
Literature review	DSSTox substance identifiers	Он	Common name	Perfluorohexanoic	acid								
Management dashboard	(DTXSID)	FF	DTXSID	DTXSID3031862	DTXSID3031862								
Study list	-	FF	CASRN	307-24-4	Assessi	ment en	dpoints (1	082 found)					
Study evaluation		FF	SMILES	OC(=O)C(F)(F)C(F)	F) Filter by endpoint name (ex: heart weight)			♦ ↑ study	ţ	25 per	r page 🗢	×	۹
Endpoint list	-	F	Molecular weight	314.054	Study	Experiment	Animal group	Endpoint	Units	NEL	LEL	BMD	BMDL
Summary tables			Chemical information prov	ided by USEPA Chemicals [Chengelis, 2009, 2850404	90-Day Oral	Male Crl:CD(SD) Rats	Peroxisomal Beta Oxidation	mg/kg- day	-	200.0		-
Visualizations	Year	2023			Chengelis,	90-Day Oral	Male Crl:CD(SD)	Liver Weight, Relative to Brain	mg/kg-	50.0	200.0	-	
Downloads	Version	Final			2009, 2850404 Chengelis	90-Day Oral	Male Crl+CD(SD)	Kidney Weight Relative to Brain	mg/kg.	200.0			
About HAWC	Objective	Final IRIS Toxicologica	al Review of Perfluorohexano	ic Acid (PFHxA, CASRI	2009, 2850404	So-bay orac	Rats	Runey weight, Relative to Dram	day	200.0			Ĩ
HAWC Resources		information, visit EPA	ormation, visit EPA's IRIS website for PFHxA.										
	Authors	U.S. Environmental Pr Environmental Assess	S. Environmental Protection Agency, Office of Research and Develo nvironmental Assessment				n and						
	Vocabulary	EPA Environmental he	ealth vocabulary										

• The HAWC page for the IRIS program's toxicological review of PFHxA

 Includes extensive information such as individual study evaluations and visualizations for risk of bias and extracted data from relevant studies

SEPA Interoperability with Other Tools

• Data can be moved from other systematic review tools into HAWC. Data can be accessed through API to support visualizations or data analysis



Epidemiology Data Extraction Module

		DTYCID					
ame		DIXSID		Edit			
anadium		DTXSID2040282: Van	adium (CASRN 7440-62-2)				
				Special char			
Name*		DSST	DSSTox substance identifier				
This field is commonly used in visualizations,	so consider using a common acronym, e.g., BPA instead	d of Bisphenol A	x substance identifier (recommended). When using an identifie	r, chemical name and CASRN are standardi			
				Save			
posure Measurements							
xposure Measurements Name	Measurement Type		Measurement Timing	Edit			
xposure Measurements ame laternal serum	Measurement Type Biomonitoring		Measurement Timing 15-19 wks, 24-28 wks, 31-36 wks gestation	Edit			
xposure Measurements Iame Iaternal serum	Measurement Type Biomonitoring		Measurement Timing 15-19 wks, 24-28 wks, 31-36 wks gestation	Edit C C C C C C C C C C C C C C C C C C C			
xposure Measurements ame laternal serum Name*	Measurement Type Biomonitoring	Εχρο	Measurement Timing 15-19 wks, 24-28 wks, 31-36 wks gestation	Edit Carlos Constantes Constante			
xposure Measurements ame aternal serum Name*	Measurement Type Biomonitoring	Expo	Measurement Timing 15-19 wks, 24-28 wks, 31-36 wks gestation sure measurement types*	Edit Constant of the second s			
xposure Measurements lame laternal serum Name*	• Biomonitoring	Expo Bio Air Fo	Measurement Timing 15-19 wks, 24-28 wks, 31-36 wks gestation sure measurement types* pmonitoring	Edit Carlos Constant Edit Special char			

\$EPA

Epidemiology Data Extraction Module

Chemicals						+ Add Row	2		
Name		DTXSID				Edit			
Vanadium	Chemicals								+ Add Row
Name*	Name			DTXSID				Edit	
	Vanadium			DTXSID2040282: Vanadium (CASRN 7440-62-2)				0	
This field is common									
	Exposure Measurements								+ Add Row
Exposure Mea	Name		Measurement Type	Measurement Timing			Edit		
Name	Maternal serum		Biomonitoring	ing 15-19 wks, 24-28 wks, 31-36 wks gestation					
Maternal serum								1	
Name*	Exposure Levels								+ Add Row
A unique name for th	Name	Chemical	Exposure		Central Tend	dency	Comments	Edit	
	Vanadium Maternal Serum	Vanadium	Maternal serum		2.3 ng/L		average across pregnancy	6	



Epidemiology Data Extraction Module

Chemicals	_	Create new study-po	pulation								
Name		Population Summary*		Study name (if a	pplicable)	Study design*		Source*			
Vanadium	Chemicals					instanty	¢	3	¢		
Name*	Name Vanadium	Briefly describe the study population. Try t outside a typical general population sampl a specific health condition, specific enviror assisted living facility, farmers, etc.], expos This field may be used in visualizations as	o capture anything les (e.g.,people with nments [e.g., ure scenario, etc. a summary of the	Study name assigne cohorts.	ed by authors. Typically available for	Select the most appropriate one study design applies (e analyses of baseline measu ("cohort") and clarify differe or b) select "other" and pro	e design from the list. If more than .g., a cohort with cross-sectional ires), can either a) select one design ent timing in remaining extraction wide details in comments.				
	Exposure N	Istudy, so it is important to be consistent wi	thin the assessment.								
	Name	Population age category*	ulation age category*		ulation age category*		etails	Sex*		Population race/ethnicity	
	Maternal serun	Adults Children and adolescents <18 yrs				Male and Female	÷				
Exposure Mea	Exposure L	Pregnant women Other Select all that apply. Note: do not select "P pregnant women are only included as part	regnant women" if								
Maternal serum	Vanadium Mate	population sample	28 (2009) - 20 (2009)								
Name*		Overall study population N*	Year(s) of enrollme	ent	Year(s) or length of follow-up	Countries	Other geographic	c information			
A unique name for t	his exposure that t	Enter the total number of participants enrolled in the study (after exclusions). Note: Sample size for specific result can be extracted in qualitative data extraction Inclusion/Exclusion Criteria			Susceptibility		Comments				
		Normal + BIUS X2	x: <u>A</u> 🕷 🗞	1≣ ≡ "	Normal ≑ B I U S <u>I</u> x	x₂ x² <u>A</u> ∦ � ⊟	≡ 99 Normai ≎ <i>I</i> x	BIUS X ₂ X ² A	ø ≡ ≡ n		

\$EPA

Epidemiology Data Extraction Module

Exposure Levels

Chemicals		Create ne	Name	Chemic	al	Exposure		Level		Comr	ments	
Name		Pagulation Sumu	Vanadium Maternal Serun	n Vanadiu	ım	Maternal serum		2.3 ng/L		avera	age across pi	regnancy
Vanadium	Chemicals	Population Sumi										
	Name	Briefly describe the										
Name*	Vanadium	outside a typical ge	Outcomes									
		a specific nearth co assisted living facili										
This field is common		This field may be us study, so it is împor	System		Effect	Effect Effect Detail			Endpoi		ome	
	Exposure Measurem	Population age c	Developmental		Clinical Observation					Body Weight		
	Name	Adults										
	Maternal serum	Children and										
		Other	Adjustment Factor	rs								
Exposure Mea		Select all that apply pregnant women a										
Exposure mea	Exposure Levels	population sample	Name		Description				Comments			
Name		Overall study po	Set A		sex gestational age gestat	ional diabetes mellitus 🛽 hyperten:	ion physical activity incom	ne	-			
Maternal serum	Name Vanadium Maternal Serum				mode of delivery education	mode of delivery education parity pre-pregnancy BMI maternal age						
		Enter the total num										
Name*		enrolled in the stud Note: Sample size fi-										
		be extracted in qua	Data Extractions									
A unique name for th	his exposure that will help you id	Inclusion/Exclus								1		
		Normal :	-		Exposure		Estimate			C C I		
		<u>I</u> ×	Group	Outcome	Level	Timing	уре	N	Value	Confidence	C	omments
			Continuous	Body Weight	Vanadium Maternal Serum	Birth	Jnit increase	227	-45.48 [-86.06, -4.9]	-	-	
			Continuous	Body Weight	Vanadium Maternal Serum	Birth	Jnit increase	111	-64.73 [-125.17, -4.29]	-	-	



Epidemiology Data Extraction Module: Visualization Integration

Data visualizations

Visual	Data Custor	nize						Dash	board:	Outcom	e System
				Outcome S	ystem by S	tudy Design				Sh	ort Cita
	Cancer	8	13		5	1	1	28		1	Akins J
	Cardiovascular		13				129	142		1	Albouv
	Dermal				1			1		1	Alson F
	Developmental	1	8					9		1	Alsopr
	Endocrine			3	1		5	9		1	Amaral
	Gastrointestinal	1			2		5	8		1	Arnold
	Hematologic			3	1		12	16		1	Ashor A
	Hepatic		1				5	6		1	Babate
	Immune		1		1		11	13		1	Bahad
L L	Metabolic		5	1			28	34			Dahad
)ste	Multi-System						1	1		1	Bahado
ഗ	Musculoskeletal				1		31	32		1	Dabad
	Nervous		3		1		16	20		De	esign So
	Ocular		1	1		1		3		215	Genera
	Other		1				6	7		3	Occupa
	Reproductive		2					2			D C1
	Respiratory		1		1		48	50		Aş	ge Prom
	Urinary	1	2		1		7	11		200	Adults
	Whole Body		3	2	1		43	49		22	Childre
	Grand Total	11	39	5	7	2	154	218		3	Pregna
		ontrol	ohort	tional	gical	ested	nized d trial	Total		Cl	nemical
		6 0	Ŭ	sect		ž °	olled	. pu		13	Beetro
		Cas		Cross-	ш Ш	cas	Rar contro	Gra		100	Beetro
			I		l Study Desia	l n			⊕ ± .	1	Chard a

Out	tcome	System by Study Design	~	5
	She	ort Citation		
	1	Akins JD et al. 2021	Ø	^
	1	Albouy-Llaty M et al. 2016	ľ	
	1	Alsop P and Hauton D 2016	C	l.
	1	Amaral AL et al. 2021	C	I.
	1	Arnold JT et al. 2021	Z	I.
	1	Ashor AW et al. 2016	Z	1
	1	Babateen AM et al. 2022	ľ	i -
	1	Bahadoran Z et al.	ď	I.
	1	Bahadoran Z et al. 2016	C	e
	1	Pabadaran 7 at al. 2017	~7	•
	De	sign Source		
	215	General population		
	3	Occupational		
	Ag	e Profile		
	200	Adults		
	22	Children and adolescents <	18 yı	rs
	3	Pregnant women		
	Ch	emical Name		
	13	Beetroot		A
	100	Beetroot juice		

- 1 Chard and rhubarb gel
- 1 Green leafy vegetable juice
- 1 Guava fruit juice
- 1 Lettuce
- 2 Lettuce juice
- 70 Nitrate

3 Nitrate-rich vegetables

1 Nitroto /Nitrito

Data querying

 Filter by outcome name 	ne (ex: B vitamins and risk of cano	er)	↑ Study 🗘	25 per page	÷ × Q
Study reference	Study Desigr	ı	S		
× Akins JD et al. 2021 ex: Smith et al. 2010	Case-contro Nested case Case report Case series Randomized	or -control d controlled tri	al 🗸	r 🔶	
	Apply Filt	ters Car	ncel		
Study	Apply Filt Study Population	ters Car Design	ncel	Outco	ome
Study Akins JD et al. 2021	Apply Filt Study Population Young BL men	ters Car Design Randomi	ized controlled	Outco d trial peak t	ome plood flow
Study Akins JD et al. 2021 Akins JD et al. 2021	Apply Filt Study Population Young BL men Young BL men	ters Car Design Randomi Randomi	ized controlled	Outco d trial peak t d trial total t	o me blood flow
Study Akins JD et al. 2021 Akins JD et al. 2021 Akins JD et al. 2021	Apply Filt Study Population Young BL men Young BL men Young BL men Young BL men	ters Car Design Randomi Randomi	ized controlled	Outco d trial peak t d trial total t d trial chang	ome olood flow olood flow ge in blood pressure
Study Akins JD et al. 2021 Akins JD et al. 2021 Akins JD et al. 2021 Akins JD et al. 2021	Study Population Young BL men	ters Can Design Randomi Randomi Randomi	ized controlled ized controlled ized controlled ized controlled	Outco d trial peak I d trial total E d trial chang d trial peak f condu condu	one plood flow plood flow ge in blood pressure forearm vascular ictance
Study Akins JD et al. 2021 Akins JD et al. 2021 Akins JD et al. 2021 Akins JD et al. 2021 Akins JD et al. 2021	Apply Filt Study Population Young BL men Young BL men	ters Can Design Randomi Randomi Randomi Randomi	ized controlled ized controlled ized controlled ized controlled ized controlled	Outco d trial peak b d trial total b d trial total b d trial chang d trial peak f condu condu d trial peak f	ome olood flow olood flow ge in blood pressure forearm vascular icctance velocity



Epidemiology Data Extraction Module: Visualization Integration

Study	Population	Overall Study Confidence	Design	Exposure Window	Regression Coefficient	Exposure Comparison	Regression coefficient β [change in mean BWT (g)] β [change in mean BWT (g)]
Buck Louis, 2018, 5016992	NICHD Fetal Growth Studies (2009-2013), United States, 2106 mother-infant pairs	High	Cohort (Prospective)	Trimester 1	0.13	1 SD increase	► 95% confidence interval
Manzano-Salgado, 2017a, 4238465	INMA cohort (2003-2008) 1202 mother-infant pairs	High	Cohort (Prospective)	Trimester 1	-4.75	Quartile 2	⊢
					3.56	Quartile 3	⊢ ● (
					-5.34	Quartile 4	⊢ − ●−−−−1
					-14.82	In-unit (ng/mL) increase	⊢_ ● ¦(
Bach, 2016, 3981534	Aarhus Birth Cohort (2008-2013), Denmark, 1507 mother-infant pairs	High	Cohort (Prospective)	Trimester 1-2	-26	Quartile 2	
					-72	Quartile 3	
					10	Quartile 4	↓ → → ↓
Lenlers, 2016, 5617416	INUENDO (2002-2004), Greenland/Poland/Ukraine, 1,321 mother-infant pairs	Medium	Cohort (Prospective)	Trimester 2-3	-43.45	In-unit (ng/mL) increase	
Cao, 2018, 5080197	Zhoukou City Longitudinal Birth Cohort (2013-2015), China, 282 mother-infant pairs	Low	Cohort (Prospective)	At birth	138.1	Tertile 2	· · · · · · · · · · · · · · · · · · ·
					80.4	Tertile 3	• • • • • • • • • • • • • • • • • • •
Li, 2017, 3981358	GBCS (2013), China, 321 mother-infant pairs	Low	Cross-sectional	At birth	-45.6	In-unit (ng/mL) increase	F€
Shi, 2017, 3827535	Haidan Hospital (2012) 170 mother-infant pairs	Low	Cross-sectional	At birth	52.68	log10-unit (ng/mL) increase	•
Xu, 2019, 5381338	Cross-sectional study (2016-2017), China, 98 mother-infant pairs	Low	Cross-sectional	At birth	-133.4	In-unit (ng/mL) increase	• • • • • • • • • • • • • • • • • • •
Callan, 2016, 3858524	AMETS (2008-2011), Australia, 98 mother-infant pairs	Low	Cross-sectional	Trimester 3	14	In-unit (ng/mL) increase	• • • • • • • • • • • • • • • • • • •
							-300 -250 -200 -150 -100 -50 0 50 100 150 200 250 300

Data visualization: Example of a forest plot showing results from multiple studies evaluating changes in birth weight



Epidemiology Data Extraction Module: API and advanced cleanup

- Added an application programming interface (API) layer to be able to create all objects via the API
- Added bulkcleanup features to cleanup manually extracted data and standardize



Bulk cleanup





Population	Outcome	Commen	Expo	osure Contrast adjOl (EWPM)	र						
USA (TX 1996-zoc., a case-control, 오♂ (60,613 case-mothers; 244,927 control-mothers)	anencephaly 3,985 (1.6%)	controls and 10 (1.7%) case	es Q group	0 vs >0 1.09		4	nine biff.de				
							spina binda				
		-				Colorit.	Population description		p-tr 027; 9-14 (0.7-1.1%	exposed cases per group	
	Selection	Exposure	Accortainment	Confounding	Analysis	Selective	Metric Description		adjOR		
	Selection	Measurement	Ascertainment			Reporting	Adjustment factors		birth year		
	+	-	++	+	-	+	1		geographic maternal age		
									race/ethnicity		
	Click on any cell ab	ove to view details.					Dose response		not-applicable		
	+ Show all details						Statistical power		not reported or calculated		
	Data type(s)		Epidemiol	ogy			Prevalence incidence Number of cases: 1276 (97.2%) for referent, 14 (1.1%) for 0.001-42.27, 9 (0.7%) for 42.7 far::100.00 far::100.00			.7%) for 42.28-1490.26, 14 (1.1%)	
	Full citation		Maternal re study. Brer	esidential proximity nder JD. Shinde MU.	to chlorinated solve Zhan FB. Gong X. La	ent emissions a anglois PH. Env	Results by group		101 - 1490.20		
	All all and the		PACKODO	, , , , , , , , , , , , , , , , , , ,	, , , ,	· · · · · · · · · · · · · ·	Group ^a	N	Estimate ()	95% confidence intervals	p-value
	Abstract		BACKGRO	UND: Some studies	have noted an assoc	ciation between	0	239716	1	-	
			air emissio	u solvents and birth	on birth defects	, but data are ta	0.01-42.27	1345	1.74	1.02 - 2.99	
			an emissio	ins of these solvents	on birth delects.		42.28-1490.26	1340	1.23	0.63 - 2.4	
			METHODS	: With data from the	Texas Birth Defects	Registry for bi	>1490.26	1337	7 1.66 0.94 - 2.91		
			the relation	n between maternal	residential proximit	ty to industrial	^e Trend-test result: 0.027.		·		
			birth defec	ts in offspring of 60,	613 case-mothers a	nd 244,927 cor	Forest plot				
			exposures	to solvent emission	s were estimated wit	th metrics that	-				
			industrial s	sources and annual	amounts of chemica	als released. Lo				spina bifida	
			ratios and	95% confidence inte	ervals for the associa	ations between	0-			9	
			chlorinated	d solvents and selec	ted birth defects, ind	cluding neural	neural 0.01-42-27-				
			congenital heart defects. All risk estimates were adjusted for ver		liusted for vear	>1490.26 -					
							100m				10



- IRIS assessments are developed using systematic review methods, which can be time and resource intensive
- Software tools such as HAWC can aid in streamlining multiple steps of assessment development, including the extraction of data from epidemiology studies. These data extractions are integrated with visualization capabilities and can be produced with minimal data processing.
- The flexibility provided with the updated data extraction in HAWC enables the development and storage of information for evidence maps, toxicological reviews, visualizations, and interoperability with other tools.



Acknowledgements

Code & Product Contributors

Casey Hansen (EPA/ORAU) Matt Munns (EPA/ORAU) Danny Peterson (EPA/ORAU) Danny Rabstejnek (EPA/ORAU) Carrie Smithing (EPA/ORAU) Andy Shapiro (EPA) Sean Watford (EPA)

Infrastructure & Support

Christine Alvarez (EPA) Jaspreet Gill (GDIT) Maureen Johnson (EPA) Dahnish Shams (EPA) Emily Snyder (EPA) Michele Taylor (EPA) Shane Thacker (EPA) Kris Thayer (EPA)







So what is HAWC again? Parlez-vous code?

- A Python web-application
 - A web-application data entry in/out
 - APIs for automation of data in/out
 - Data science compute environment
- A relational database
 - Mostly relational data
 - Also binary/nosql data
- An interactive frontend
 - Dynamic visualizations
- An open-source application
 - Can collaborate with anyone



Research and Development

Interactive Displays: Data Extraction

Chemical	Endpoint	Study	Animal Description	Route	Exposure Duration	
6:2 Fluorotelomer alcohol	Liver Weight, Absolute	Mukerji et al. 2015	P0 Mouse, Crl:CD-1(ICR)BR (♀)	oral gavage	14d pre-mating, 14d mating, gestation, lactation	
			P0 Mouse, Crl:CD-1(ICR)BR (ੋ)	oral gavage	109 d (premating-sacrifice)	••••
		Serex T et al. 2014	Rat, Crl:CD(SD) (♀)	oral gavage	90 d	
			Rat, Crl:CD(SD) (ೆ)	oral gavage	90 d	•• <u> </u>
		Unnamed report (2005a) (ECHA summary)	Rat, Crl:CD(SD) (ೆ⊋)	oral gavage	28 d	+++++++
	Liver Weight, Relative	Mukerji et al. 2015	P0 Mouse, Crl:CD-1(ICR)BR (♀)	oral gavage	14d pre-mating, 14d mating, gestation, lactation	••
			P0 Mouse, Crl:CD-1(ICR)BR (ੋ)	oral gavage	109 d (premating-sacrifice)	•• -
		ECHA, 2007, 5701160	Rat, Crl:CD(SD) (♀)	oral gavage	28d (1dose/d)	+ +
		Serex T et al. 2014	Rat, Crl:CD(SD) (♀)	oral gavage	90 d	•• <u> </u>
		ECHA, 2007, 5701160	Rat, Crl:CD(SD) (்)	oral gavage	28d (1dose/d)	•••
		Serex T et al. 2014	Rat, Crl:CD(SD) (ೆ)	oral gavage	90 d	
6:2 Fluorotelomer methacrylate	Liver Weight, Absolute	ECHA, 2007, 6299223	Rat, Crl:CD(SD) (♀)	oral gavage	28d (1dose/d)	••
			Rat, Crl:CD(SD) (ீ)	oral gavage	28d (1dose/d)	••
	Liver Weight, Absolute, Recovery	ECHA, 2007, 6299223	Rat, Crl:CD(SD) (♀)	oral gavage	28d (1dose/d)	••
			Rat, Crl:CD(SD) (ீ)	oral gavage	28d (1dose/d)	••
	Liver Weight, Relative	ECHA, 2007, 6299223	Rat, Crl:CD(SD) (♀)	oral gavage	28d (1dose/d)	••
			Rat, Crl:CD(SD) (්)	oral gavage	28d (1dose/d)	••
	Liver Weight, Relative, Recovery	ECHA, 2007, 6299223	Rat, Crl:CD(SD) (♀)	oral gavage	28d (1dose/d)	••
			Rat, Crl:CD(SD) (♂)	oral gavage	28d (1dose/d)	••
Trifluoroacetic acid	Liver Weight, Absolute	Unnamed Report (2010a) (ECHA Summary)	P0 Rat, Crl:CD(SD)IGS BR (♀)	oral gavage	GD 6-19	+++- <u>A</u>
		Unnamed Report (2012b) (ECHA Summary)	P0 Rat, Crl:CD(SD)IGS BR (♀)	oral gavage	up to 57 d (premating-lactation)	•••
			P0 Rat, Crl:CD(SD)IGS BR (ੇ)	oral gavage	38 d (premating-termination)	+• 🔺
		Saillenfait et al. 1997	P0 Rat, Sprague-Dawley (♀)	oral gavage	GD 10-20	
			F1 Rat, Sprague–Dawley (ở⊋)	oral gavage	GD 10-20	→→→
		Unnamed Report (2016a) (ECHA Summary)	Rat, Wistar Rj:Wi (lops Han) (♀)	oral diet	90 d	•
			Rat, Wistar Rj:Wi (lops Han) (්)	oral diet	90 d	•
	Liver Weight, Relative	Unnamed Report (2012b) (ECHA Summary)	P0 Rat, Crl:CD(SD)IGS BR (유)	oral gavage	up to 57 d (premating-lactation)	•••
			P0 Rat, Crl:CD(SD)IGS BR (♂)	oral gavage	38 d (premating-termination)	•• <u>^</u>
		Saillenfait et al. 1997	P0 Rat, Sprague-Dawley (♀)	oral gavage	GD 10-20	← <u>∧</u> <u>∧</u>
			F1 Rat, Sprague–Dawley (∛⊋)	oral gavage	GD 10-20	• • • • •
		Unnamed Report (2016a) (ECHA Summary)	Rat, Wistar Rj:Wi (lops Han) (우)	oral diet	90 d	•
			Rat, Wistar Rj:Wi (lops Han) (♂)	oral diet	90 d	0 100 200 30

	Liver We	ight, Absolute				
	Endpoint [Details				
	ndpoint name	Liver Weight, Absolute				
treatment-related	System	Hepatic				
v treatment-related	Organ	Liver				
	Effect	Clinical Observation				
	Effect subtype	Organ Weight				
	Diagnostic description	Liver, Weight				
	Observation time	90 d				
	Data reported?	*				
	Data extracted?	√				
	Values estimated?	-				
	Location in literature	Table 5				
	Expected response adversity direction					
•	NEL	25 mg/kg-day				
	LEL	125 mg/kg-day				
	Monotonicity	-				
	Trend result	not reported				
•	Results notes	"Following 90 days of dosing, effects on organ weights were present in the testes, liver and kidney of males (Table 5) and in livers and kidneys				
_	A					

300 400 500 600 700 800 900 1,0001,100

Dose (mg/kg-day)



Dataset

Dose (mg/kg- day)	Number of Animals	Response (g)	Standard Deviation
0	10	15.94	1.9
5	10	16.09	1.9
25ª	10	16.62	2.02
125 ^{b,c}	10	19.09	1.89
250 ^b	8	22.84	2.39

a NEL (No effect level)

^b Significantly different from control (p < 0.01) ^c LEL (Lowest effect level)