http://www.orcid.org/0000-0002-2668-4821

EPA United States Environmental Protection Agency

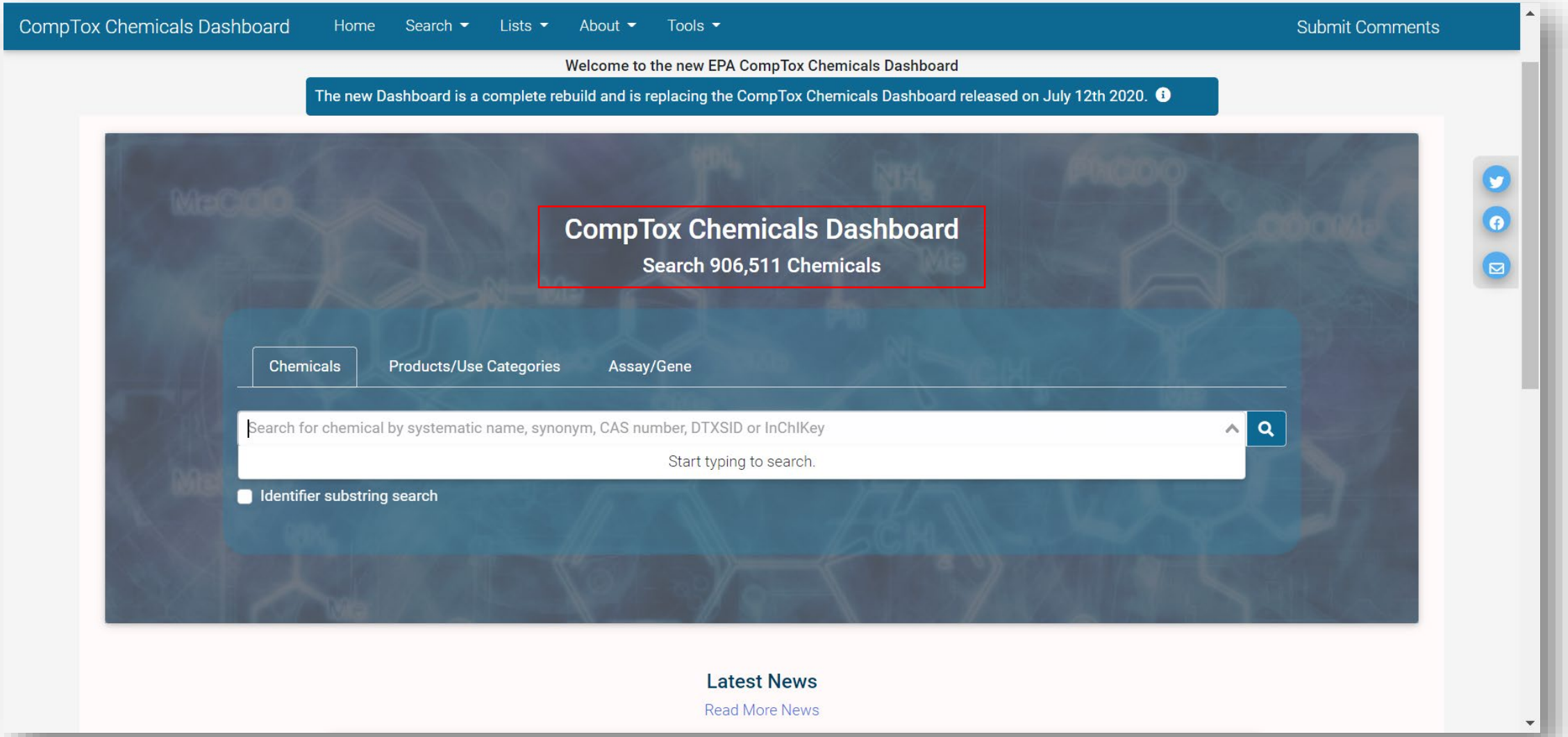
Accessing Environmental Chemistry Data via Data Dashboards

Antony J. Williams

Center for Computational Toxicology and Exposure, US-EPA, RTP, NC

Federal Environmental Symposium – March 28th 2022

CompTox Chemicals Dashboard >906k chemicals

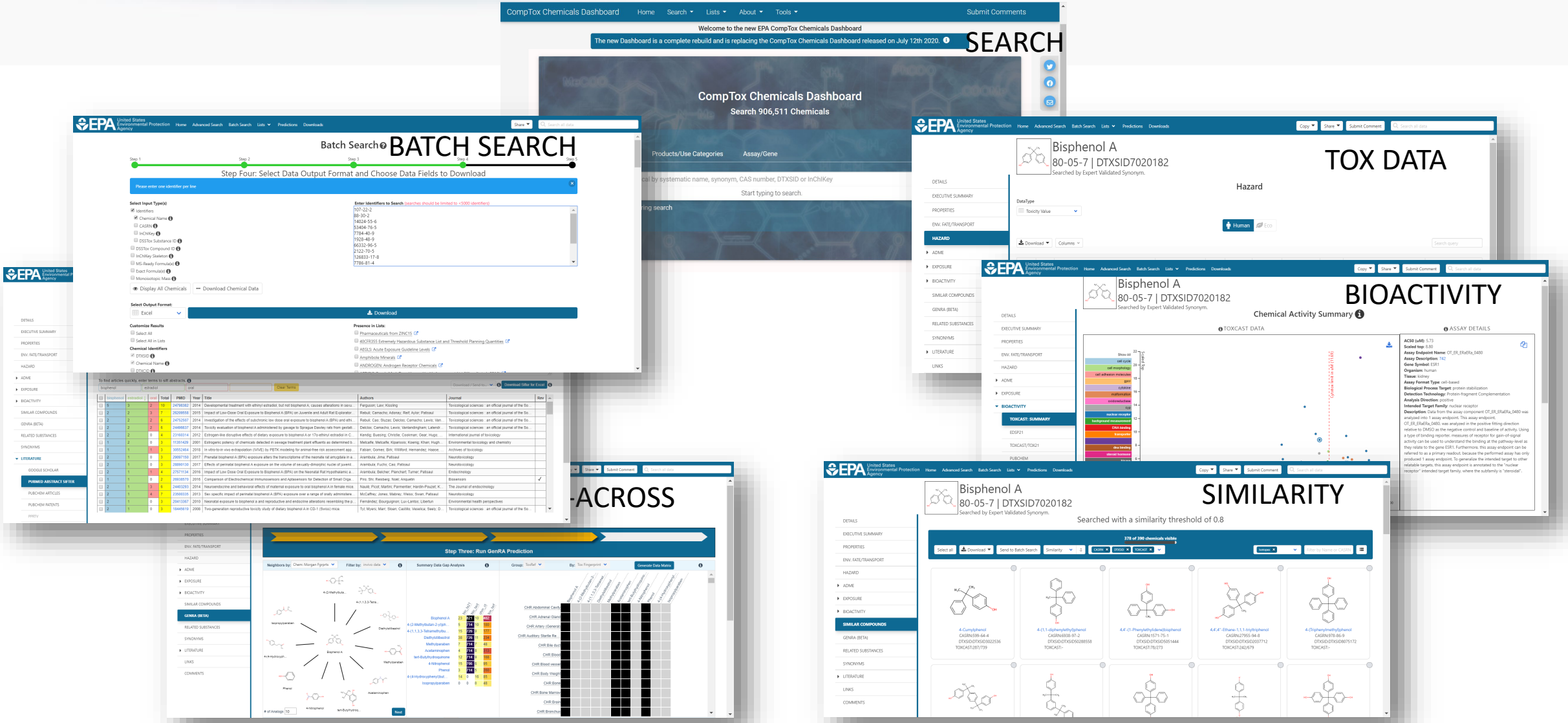


The Charge for the Dashboard

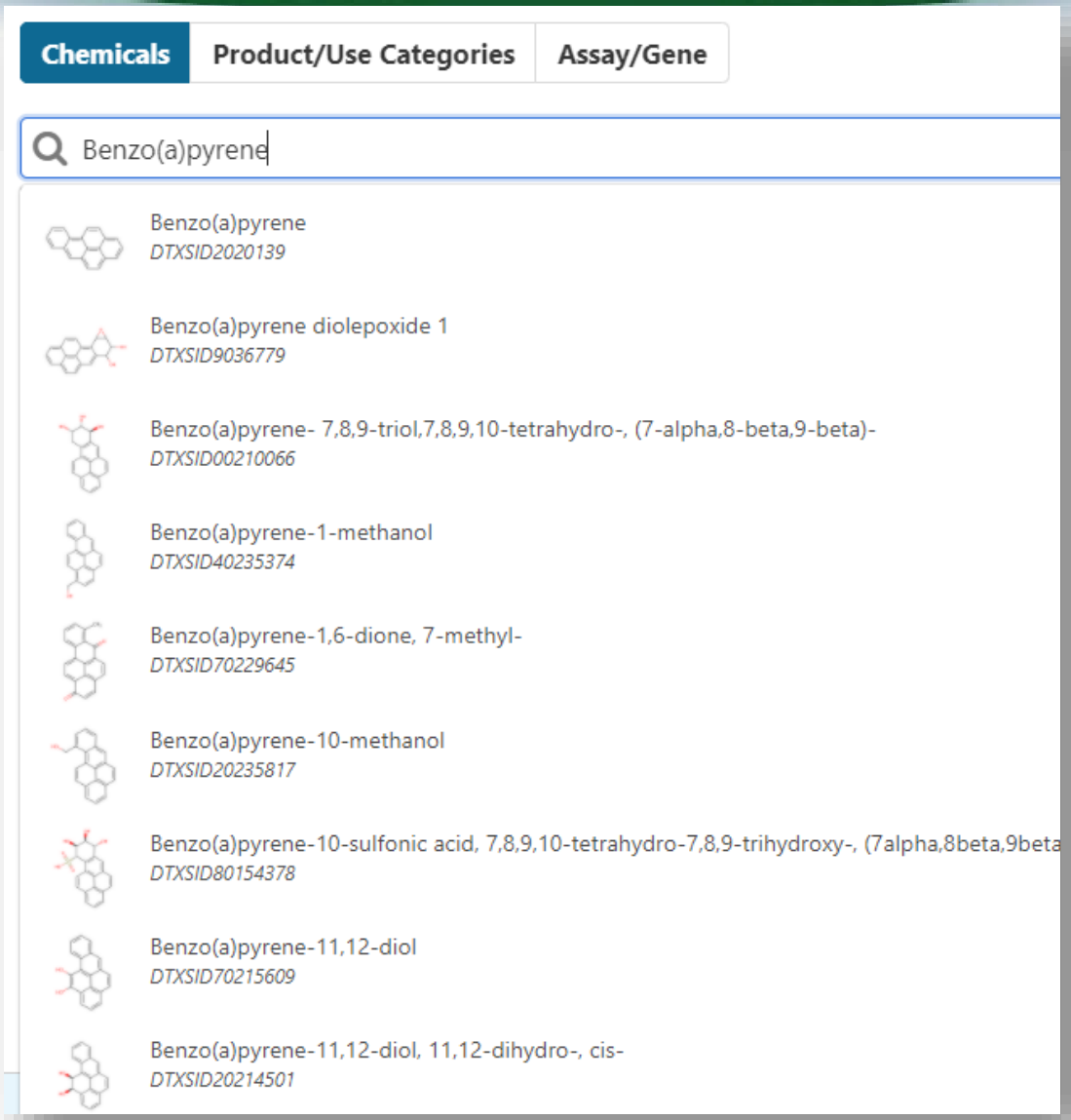
- Develop a “first-stop-shop” for environmental chemical data to support EPA and partner decision making:
  - **Centralized location** for relevant chemical data
  - Chemistry, exposure, hazard and dosimetry
  - Combination of existing data and predictive models
  - Publicly accessible, periodically updated, curated
- Easy access to data improves efficiency and ultimately accelerates chemical risk assessment

CompTox Chemicals Dashboard
https://comptox.epa.gov/dashboard

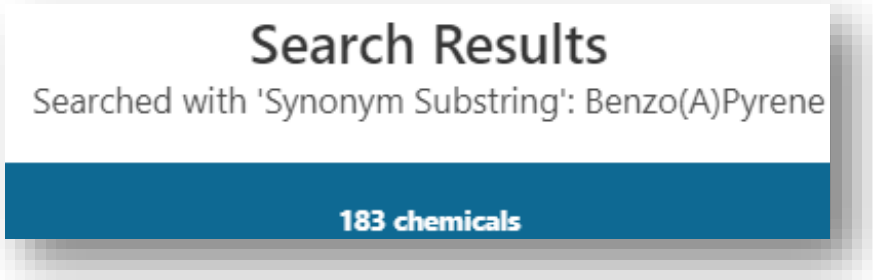
EPA United States Environmental Protection Agency



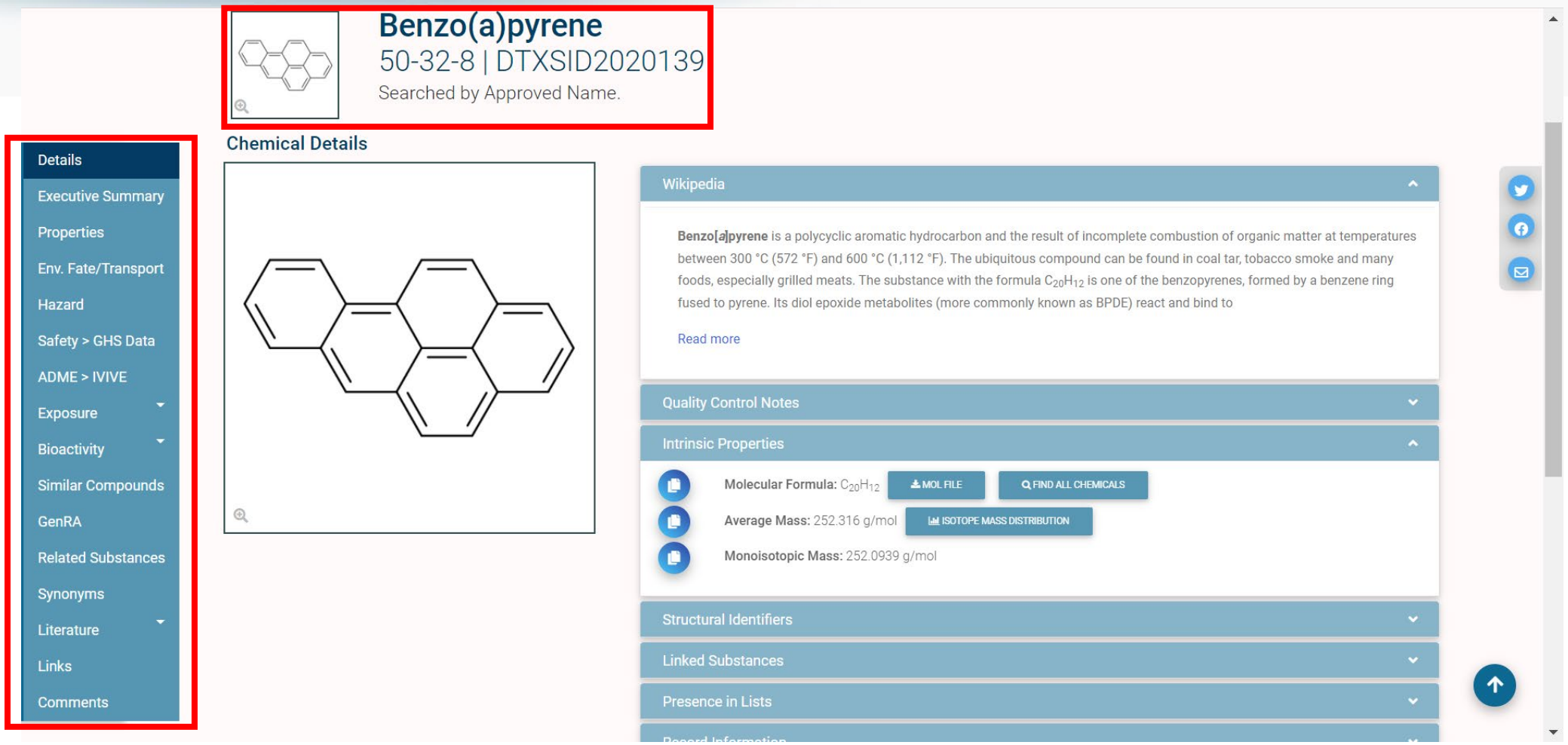
BASIC Search



- Type ahead search using Names, synonyms and CASRNs
- Millions of identifiers
- Substring search

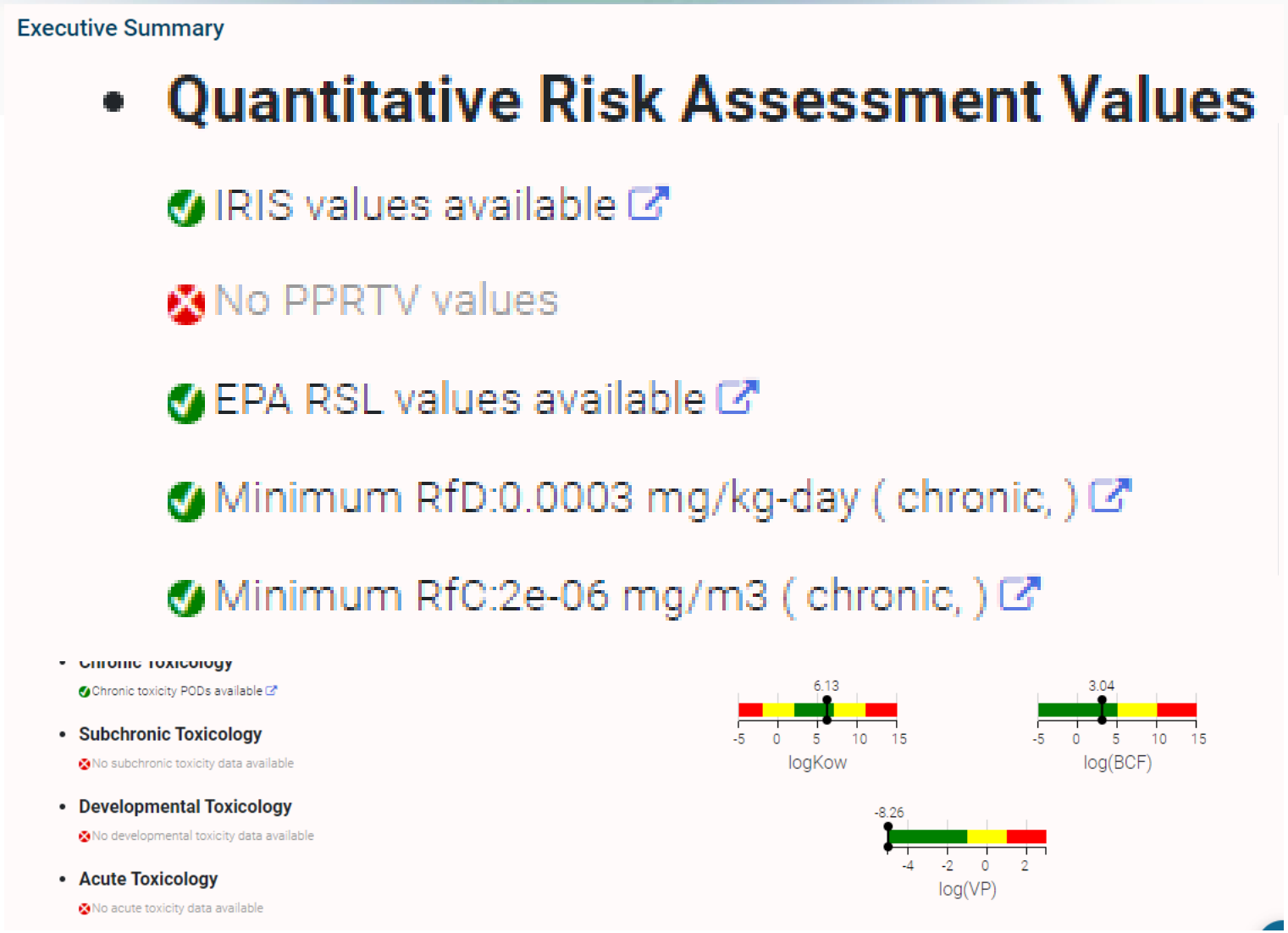


Detailed Chemical Pages



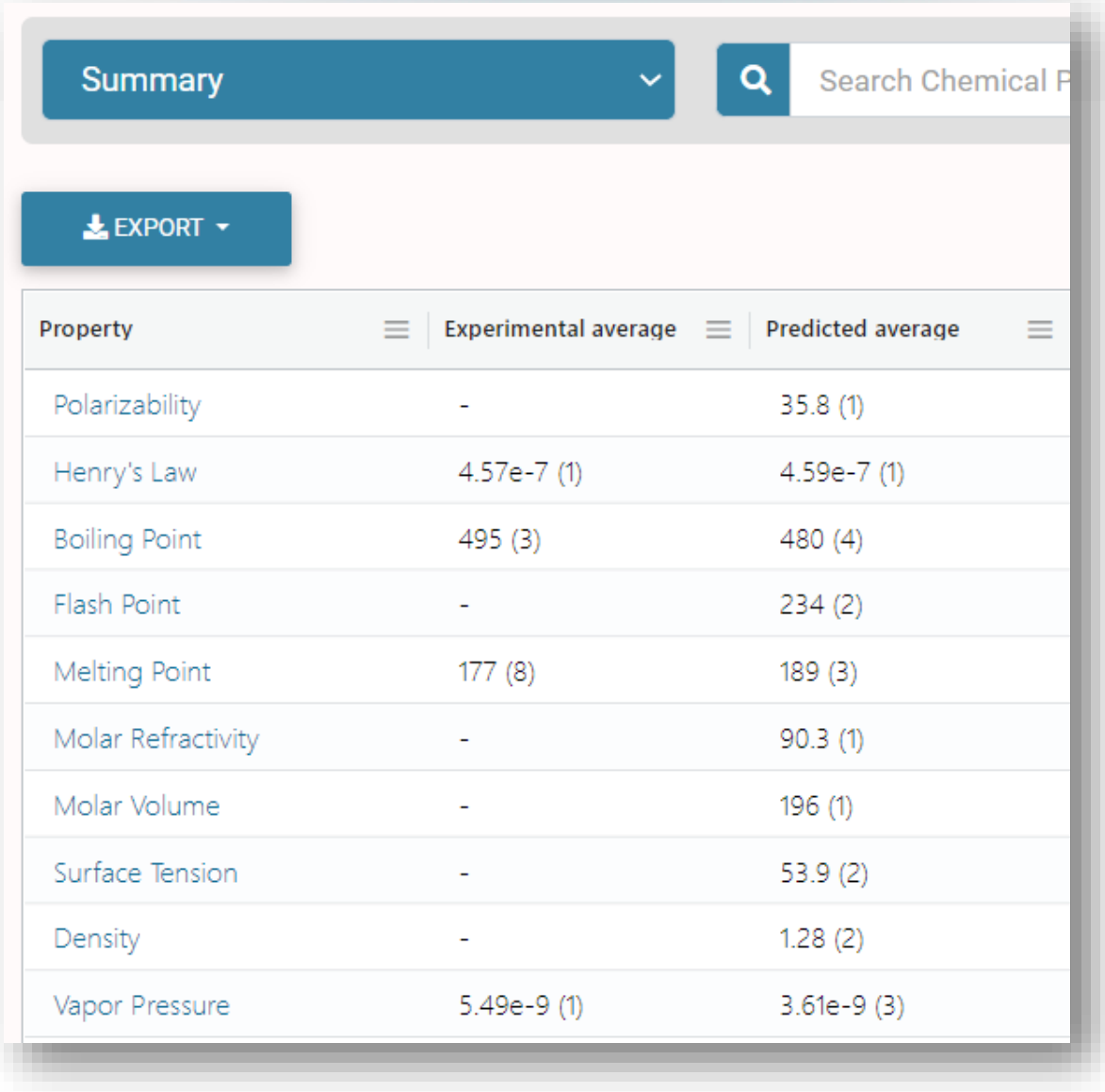
- Chemical page: Wikipedia snippet when available, intrinsic properties, structural identifiers, linked substances

“Executive Summary”



- Overview of toxicity-related info
  - Quantitative values
  - Physchem. and Fate & Transport
  - Adverse Outcome Pathway links
  - *In vitro* bioactivity summary plot

Experimental and Predicted Data



Property	Experimental average	Predicted average
Polarizability	-	35.8 (1)
Henry's Law	4.57e-7 (1)	4.59e-7 (1)
Boiling Point	495 (3)	480 (4)
Flash Point	-	234 (2)
Melting Point	177 (8)	189 (3)
Molar Refractivity	-	90.3 (1)
Molar Volume	-	196 (1)
Surface Tension	-	53.9 (2)
Density	-	1.28 (2)
Vapor Pressure	5.49e-9 (1)	3.61e-9 (3)

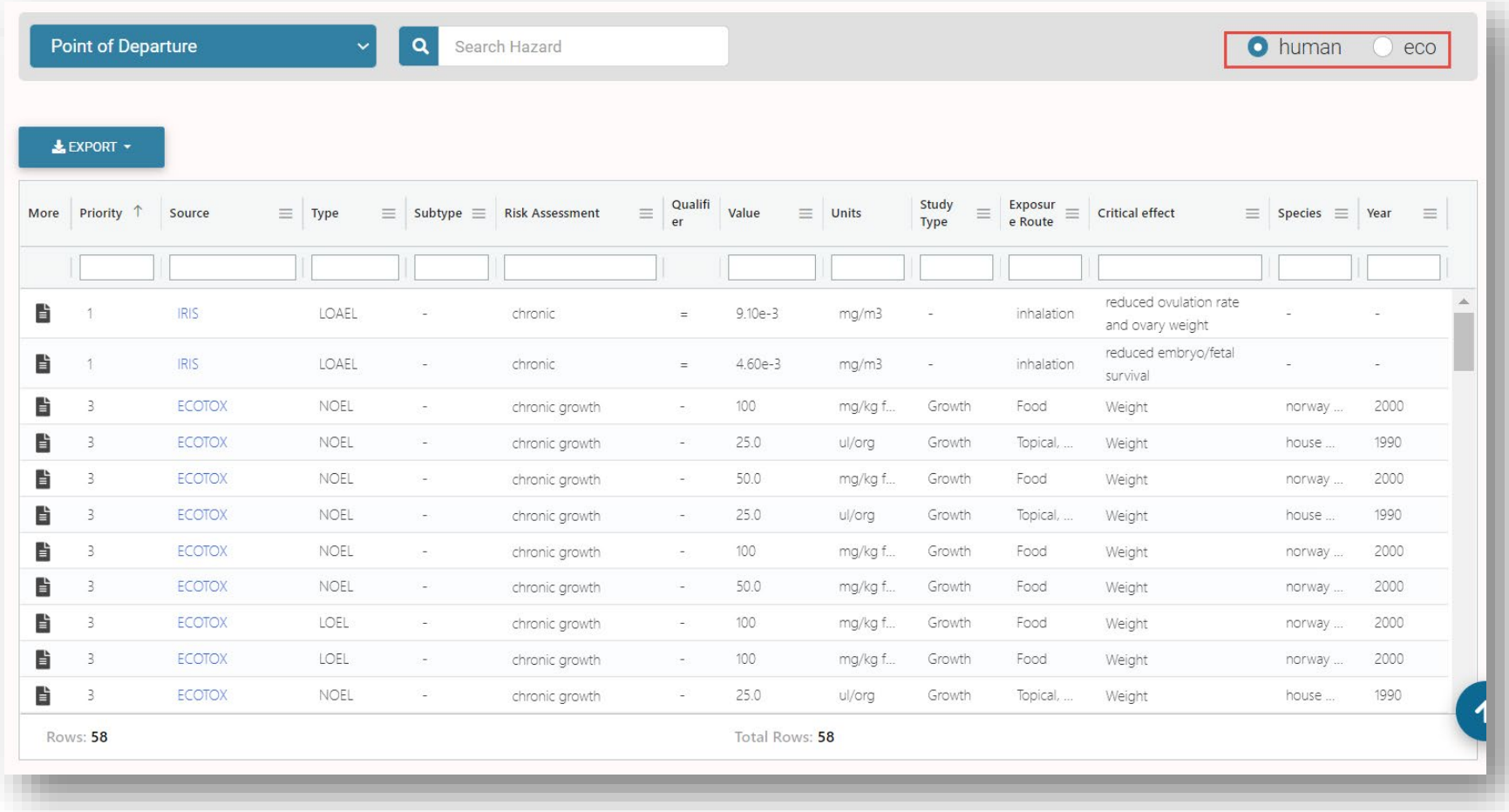
- Physchem and Fate & Transport experimental and predicted data
- Data can be downloaded as Excel, TSV and CSV files

Chemical Hazard Data

United States Environmental Protection Agency

ToxVal Database

- >50k chemicals
- >770k tox. values
- >30 sources of data
- ~5k journals cited
- ~70k citations



More	Priority	Source	Type	Subtype	Risk Assessment	Qualifier	Value	Units	Study Type	Exposure Route	Critical effect	Species	Year
	1	IRIS	LOAEL	-	chronic	=	9.10e-3	mg/m3	-	inhalation	reduced ovulation rate and ovary weight	-	-
	1	IRIS	LOAEL	-	chronic	=	4.60e-3	mg/m3	-	inhalation	reduced embryo/fetal survival	-	-
	3	ECOTOX	NOEL	-	chronic growth	-	100	mg/kg f...	Growth	Food	Weight	norway ...	2000
	3	ECOTOX	NOEL	-	chronic growth	-	25.0	ul/org	Growth	Topical, ...	Weight	house ...	1990
	3	ECOTOX	NOEL	-	chronic growth	-	50.0	mg/kg f...	Growth	Food	Weight	norway ...	2000
	3	ECOTOX	NOEL	-	chronic growth	-	25.0	ul/org	Growth	Topical, ...	Weight	house ...	1990
	3	ECOTOX	NOEL	-	chronic growth	-	100	mg/kg f...	Growth	Food	Weight	norway ...	2000
	3	ECOTOX	NOEL	-	chronic growth	-	50.0	mg/kg f...	Growth	Food	Weight	norway ...	2000
	3	ECOTOX	LOEL	-	chronic growth	-	100	mg/kg f...	Growth	Food	Weight	norway ...	2000
	3	ECOTOX	LOEL	-	chronic growth	-	100	mg/kg f...	Growth	Food	Weight	norway ...	2000
	3	ECOTOX	NOEL	-	chronic growth	-	25.0	ul/org	Growth	Topical, ...	Weight	house ...	1990

Rows: 58 Total Rows: 58

Safety Data

EPA United States Environmental Protection Agency

GHS Data

Print Page

PUBCHEM > BENZO[A]PYRENE > LABORATORY CHEMICAL SAFETY SUMMARY (LCSS) > GHS CLASSIFICATION

CID 2336

Benzo[a]pyrene

GHS Classification

Showing 6 of 6

Pictogram(s)	 Irritant, Health Hazard, Environmental Hazard
Signal	Danger
GHS Hazard Statements	H317: May cause an allergic skin reaction [**Warning** Sensitization, Skin] H340: May cause genetic defects [**Danger** Germ cell mutagenicity] H350: May cause cancer [**Danger** Carcinogenicity] H360FD: May damage fertility; May damage the unborn child [**Danger** Reproductive toxicity] H400: Very toxic to aquatic life [**Warning** Hazardous to the aquatic environment, acute hazard] H410: Very toxic to aquatic life with long lasting effects [**Warning** Hazardous to the aquatic environment, long-term hazard]
Precautionary Statement Codes	P201, P202, P261, P272, P273, P280, P281, P302+P352, P308+P313, P321, P333+P313, P363, P391, P405, and P501 (The corresponding statement to each P-code can be found at the GHS Classification page.)

Sources of Exposure to Chemicals

EPA United States Environmental Protection Agency

Chemical Weight Fractions (CWF)

Search Chemical Weight Fractions

EXPORT

Product Name	Product Use Category	Categorization Subtype	Minimum Weight Fraction	Maximum Weight Fraction	Data Type	Source	Product Count
48743 pah mixture	Not yet Categorized		-	-	reported	SIRI	1
asphalt cement penetration 60-70	Not yet Categorized		-	-	reported	SIRI	1
base-neutral 4 1ml methylene chl...	Not yet Categorized		-	-	reported	SIRI	1
base neutral calibration checkco...	Not yet Categorized		-	-	reported	SIRI	1
benzo (a) pyrene_ 98%_ b1008-0	Not yet Categorized		-	-	reported	SIRI	1
benzo (a) pyrene_ md-1956	Not yet Categorized		0.990	1.00	reported	SIRI	1
blasocut 2000 cf art no 875	Not yet Categorized		0.00	1.00e-3	reported	SIRI	1
blasocut 2000 universal_ 870	Not yet Categorized		-	-	reported	SIRI	1
blasocut 2000 universal art_ 870	Not yet Categorized		0.00	1.00e-3	reported	SIRI	1
blasocut 4000 strong_ 872	Not yet Categorized		-	-	reported	SIRI	1
blasocut 4000 universal art_ 872	Not yet Categorized		0.00	1.00e-3	reported	SIRI	1
clp-011a clp base/neutrals check ...	Not yet Categorized		-	-	reported	SIRI	1

Rows: **29**

What about PFAS?

EPA United States Environmental Protection Agency

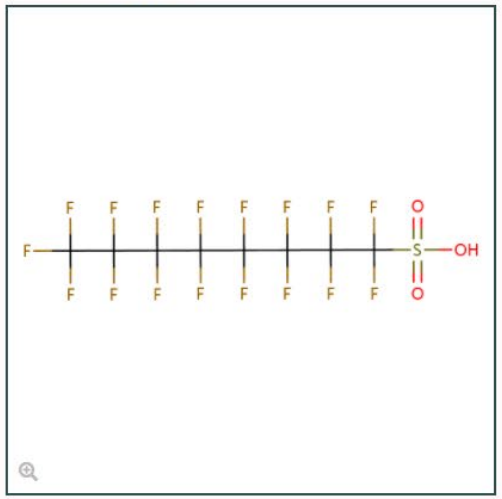
Perfluorooctanesulfonic acid

1763-23-1 | DTXSID3031864

Searched by DTXSID3031864.

- Details
- Executive Summary
- Properties
- Env. Fate/Transport
- Hazard
- Safety > GHS Data
- ADME > IVIVE
- Exposure
- Bioactivity
- Similar Compounds
- GenRA
- Related Substances
- Synonyms
- Literature
- Links
- Comments

Chemical Details



Wikipedia

Perfluorooctanesulfonic acid (**PFOS**) (conjugate base **perfluorooctanesulfonate**) is an anthropogenic (human-made) fluorosurfactant, now regarded as a global pollutant. PFOS was the key ingredient in Scotchgard, a fabric protector made by 3M, and related stain repellents. In many contexts, PFOS refers to the parent sulfonic acid and its various salts of perfluorooctanesulfonate. These are all colorless or white, water soluble solids. Although of low acute toxicity, PFOS has

Read more

Quality Control Notes

Intrinsic Properties

Molecular Formula: $C_8HF_{17}O_3S$ MOL FILE FIND ALL CHEMICALS

Average Mass: 500.13 g/mol ISOTOPE MASS DISTRIBUTION

Monoisotopic Mass: 499.937494 g/mol

Structural Identifiers

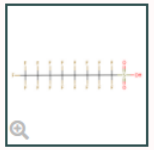
Linked Substances

Presence in Lists

Record Information

Are there Similar Compounds?

EPA United States Environmental Protection Agency



Perfluorooctanesulfonic acid

1763-23-1 | DTXSID3031864

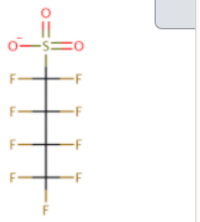
Searched by DTXSID3031864.

Searched with a similarity threshold of 0.8

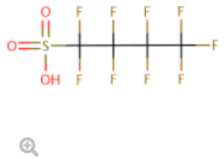
- Details
- Executive Summary
- Properties
- Env. Fate/Transport
- Hazard
- Safety > GHS Data
- ADME > IVIVE
- Exposure
- Bioactivity
- Similar Compounds
- GenRA
- Related Substances
- Synonyms
- Literature
- Links
- Comments

Search Results | Similarity | SEND 127 TO BATCH SEARCH | TILE INFO | FILTER | EXPORT

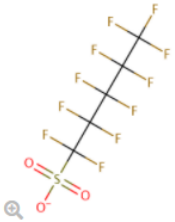
Showing 127 of 127 chemicals



Perfluorobutanesulfonate
DTXSID : DTXSID60873015
CASRN : 45187-15-3
TOXCAST :
Similarity : 1.00

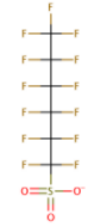


Perfluorobutanesulfonic a...
DTXSID : DTXSID5030030
CASRN : 375-73-5
TOXCAST :
Similarity : 1.00

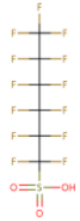


Perfluoropentanesulfonate
DTXSID : DTXSID70892479
CASRN : 175905-36-9
TOXCAST :
Similarity : 1.00

Perfluoropentanesulfonic ...
DTXSID : DTXSID8062600
CASRN : 2706-91-4
TOXCAST :
Similarity : 1.00

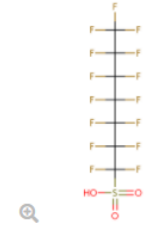


Perfluorohexanesulfonate
DTXSID : DTXSID80873012
CASRN : 108427-53-8
TOXCAST :
Similarity : 1.00

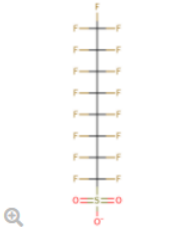


Perfluorohexanesulfonic a...
DTXSID : DTXSID7040150
CASRN : 355-46-4
TOXCAST :
Similarity : 1.00

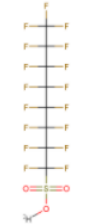
Perfluoroheptanesulfonate
DTXSID : DTXSID20892505
CASRN : 146689-46-5
TOXCAST :
Similarity : 1.00



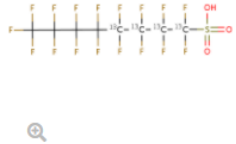
Perfluoroheptanesulfonic ...
DTXSID : DTXSID8059920
CASRN : 375-92-8
TOXCAST :
Similarity : 1.00



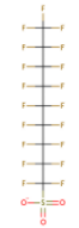
Perfluorooctanesulfonate
DTXSID : DTXSID80108992
CASRN : 45298-90-6
TOXCAST :
Similarity : 1.00



Heptadecafluorooctane-1-...
DTXSID : DTXSID00892720
CASRN : NOCAS_892720
TOXCAST :
Similarity : 1.00

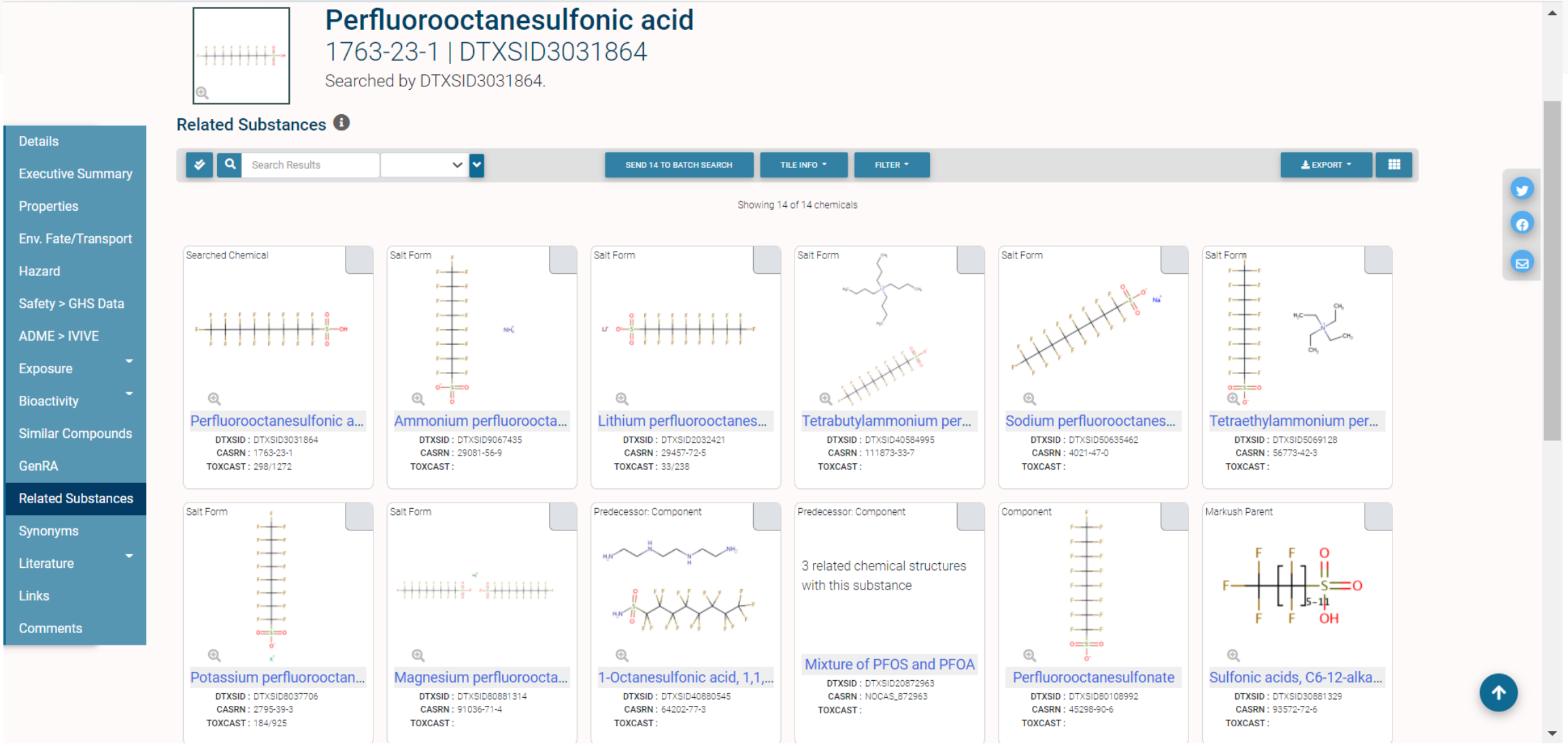


heptadecafluoro(1,2,3,4-1...
DTXSID : DTXSID80894101
CASRN : 960354-19-2
TOXCAST :
Similarity : 1.00



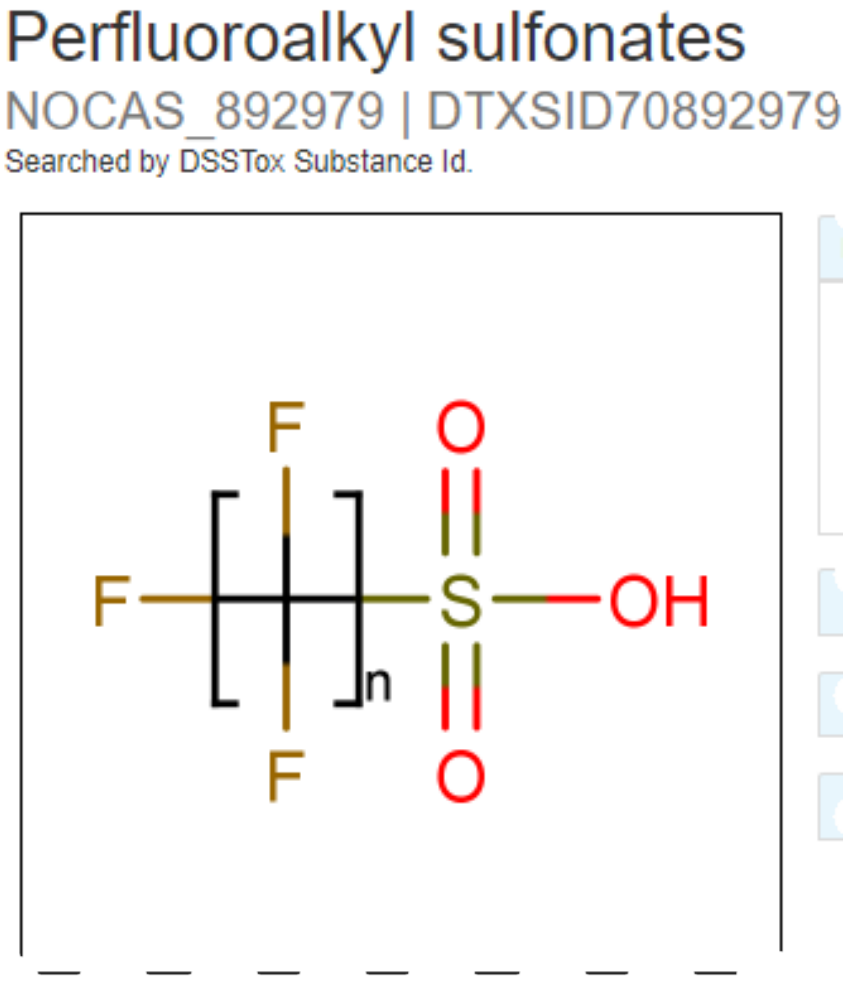
Perfluorononanesulfonate
DTXSID : DTXSID60873010
CASRN : 474511-07-4
TOXCAST :
Similarity : 1.00

Relationships in the data



Markush Chemicals

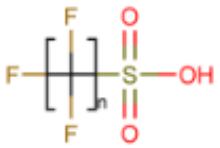
- PFOS is a member of linear perfluoroalkyl sulfonates



…and their Markush Children…

- Linear perfluoroalkyl sulfonates has children…

Searched Chemical



Perfluoroalkyl sulfonates
DTXSID: DTXSID70892979
CASRN: NOCAS_892979

Markush Child

Perfluorobutanesulfonic acid
DTXSID: DTXSID5030030
CASRN: 375-73-5

Markush Child

Perfluorooctanesulfonic acid
DTXSID: DTXSID3031864
CASRN: 1763-23-1

Markush Child

Perfluorodecanesulfonic acid
DTXSID: DTXSID3040148
CASRN: 335-77-3

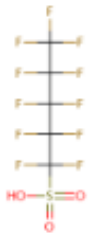
Markush Child

Perfluorohexanesulfonic acid
DTXSID: DTXSID7040150
CASRN: 355-46-4

Markush Child

Perfluoroheptanesulfonic acid
DTXSID: DTXSID8059920
CASRN: 375-92-8

Markush Child



Perfluoropentanesulfonic acid
DTXSID: DTXSID8062600
CASRN: 2706-91-4

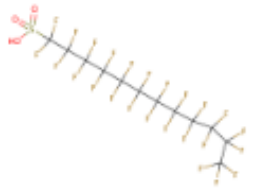
Markush Child

Perfluorononanesulfonic acid
DTXSID: DTXSID8071356
CASRN: 68259-12-1

Markush Child

Perfluoropropanesulfonic acid
DTXSID: DTXSID30870531
CASRN: 423-41-6

Markush Child



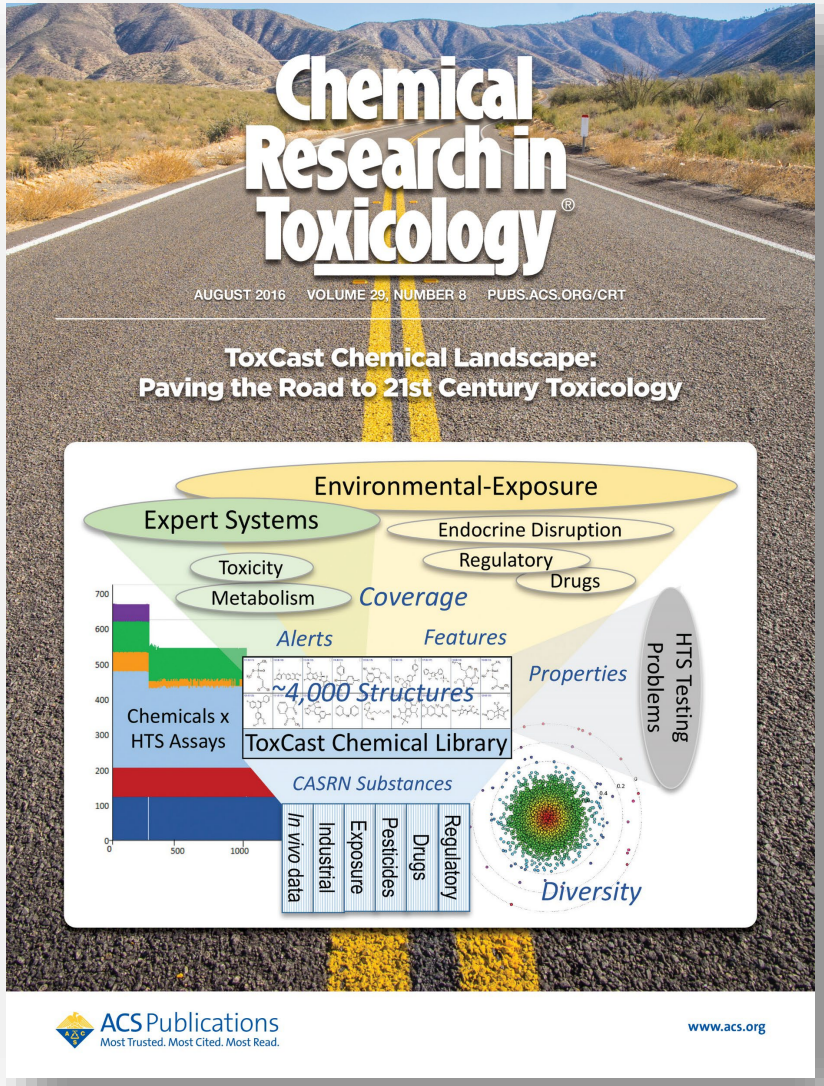
Perfluorododecanesulfonic acid (PFDOS)
DTXSID: DTXSID20873011
CASRN: 79780-39-5

EPA
United States
Environmental Protection
Agency

Bioactivity Data

ToxCast

EPA United States Environmental Protection Agency



ToxCast Chemical Landscape: Paving the Road to 21st Century Toxicology

Ann M. Richard[*†], Richard S. Judson[†], Keith A. Houck[†], Christopher M. Grulke[†], Patra Volarath[‡], Inthirany Thillainadarajah[§], Chihae Yang[∥⊥], James Rathman[⊥#], Matthew T. Martin[†], John F. Wambaugh[†], Thomas B. Knudsen[†], Jayaram Kancherla[▽], Kamel Mansouri[▽], Grace Patlewicz[†], Antony J. Williams[†], Stephen B. Little[†], Kevin M. Crofton[†], and Russell S. Thomas[†]

View Author Information

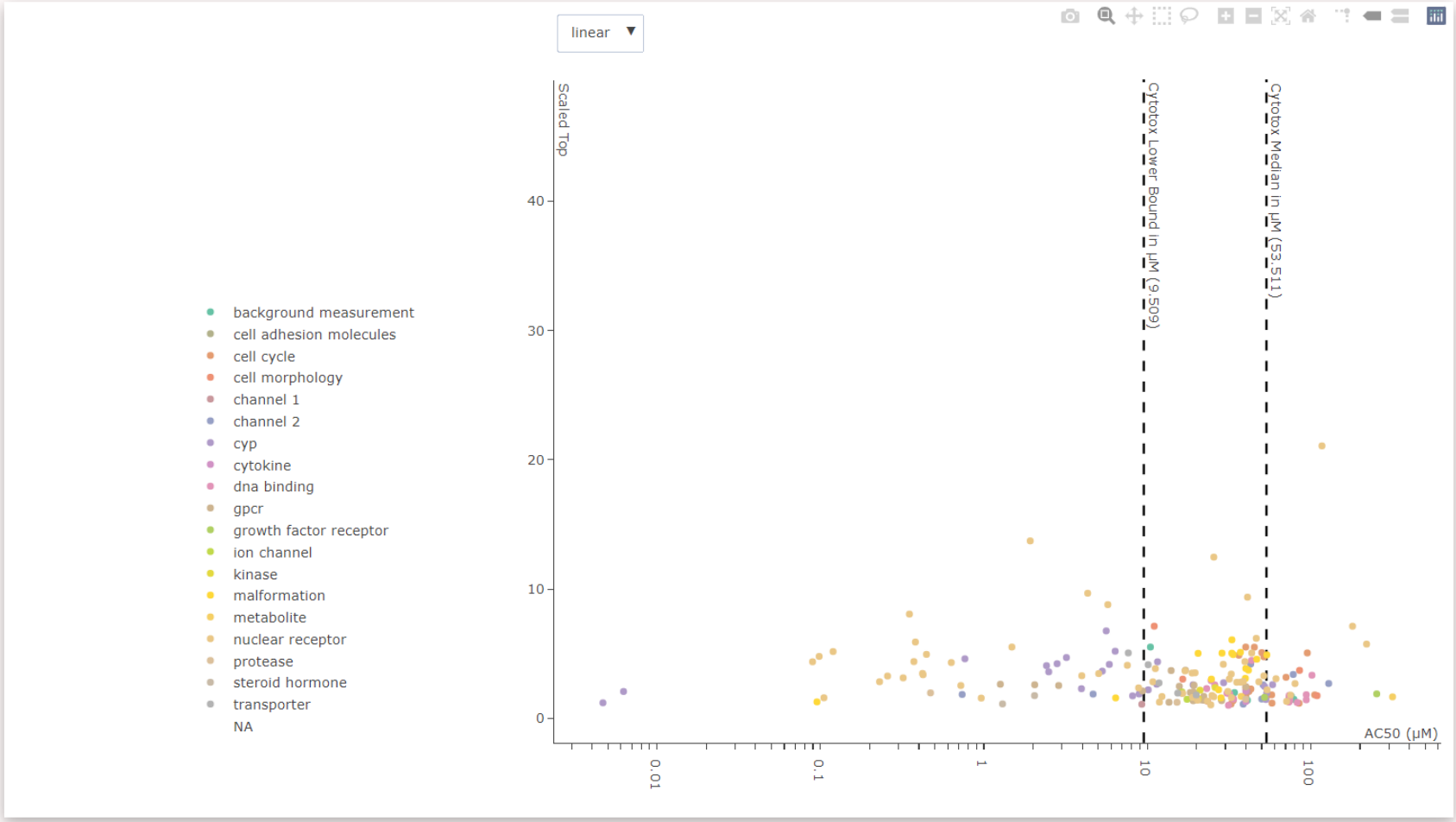
Cite this: *Chem. Res. Toxicol.* 2016, 29, 8, 1225–1251
Publication Date: July 1, 2016
https://doi.org/10.1021/acs.chemrestox.6b00135

ToxCast/Tox21 Data

United States Environmental Protection Agency

- Details
- Executive Summary
- Properties
- Env. Fate/Transport
- Hazard
- Safety > GHS Data
- ADME > IVIVE
- Exposure
- Bioactivity
- Similar Compounds
- GenRA
- Related Substances
- Synonyms
- Literature
- Links
- Comments

Bioactivity - TOXCAST Summary



Full transparency in terms of data

- Concentration Response data

Concentration Response Data

Analytical Data on Tox21 Browser

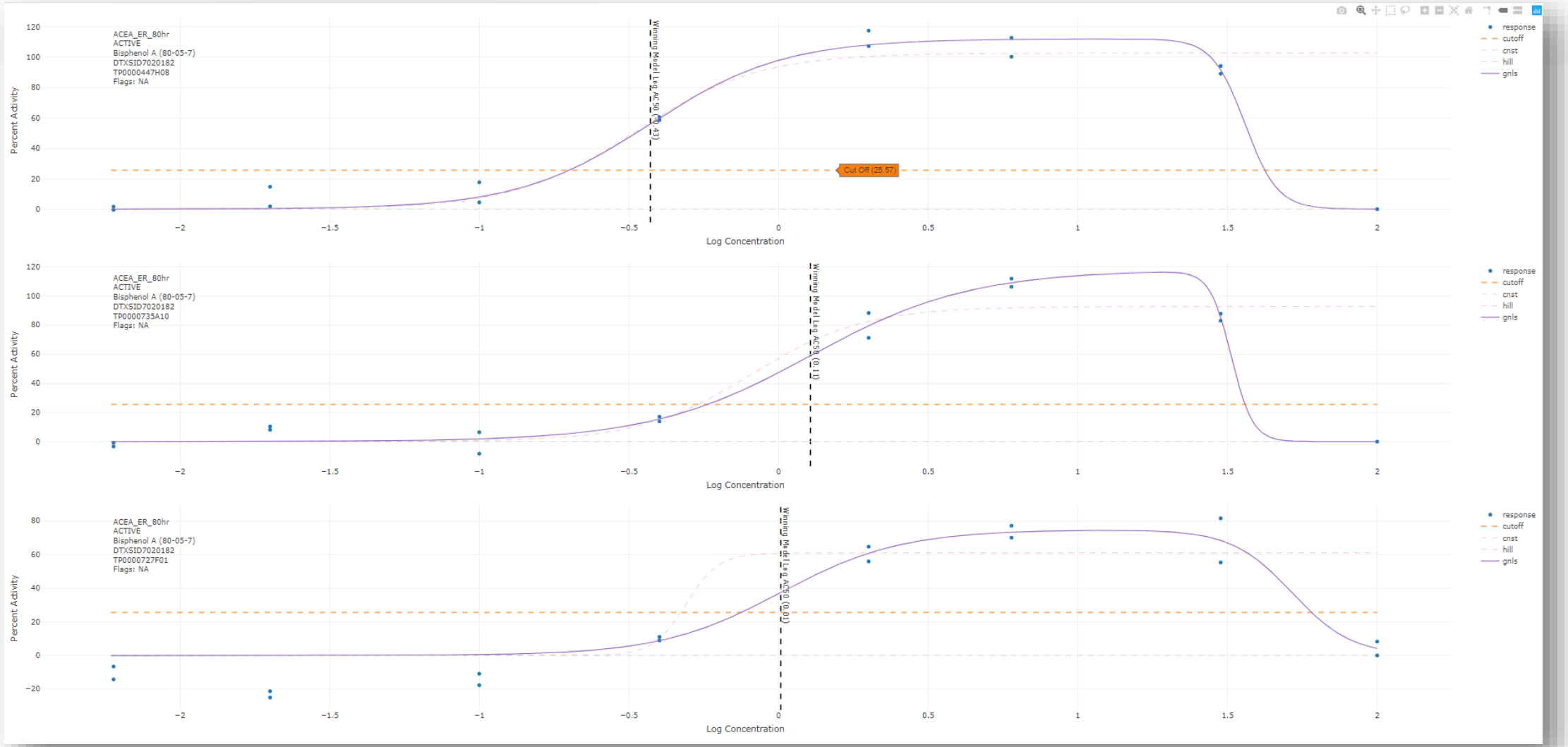
EXPORT

Name	Description	Endpoint Name	Active	Deta...	Rep....	All P...	Gene	Intended Target	Cell Line	Cell For...
ASSAY SOURCE: ACEA	ACEA Biosciences	ACEA_AR_agonist_80hr	Inactive				AR	steroidal	prostate	cell line
ASSAY SOURCE: ACEA	ACEA Biosciences	ACEA_AR_agonist_AUC_viability	Active				null	cytotoxicity	prostate	cell line
ASSAY SOURCE: ACEA	ACEA Biosciences	ACEA_AR_antagonist_80hr	Active				AR	steroidal	prostate	cell line
ASSAY SOURCE: ACEA	ACEA Biosciences	ACEA_AR_antagonist_AUC_viability	Active				null	cytotoxicity	prostate	cell line
ASSAY SOURCE: ACEA	ACEA Biosciences	ACEA_ER_80hr	Active				ESR1	steroidal	breast	cell line
ASSAY SOURCE: ACEA	ACEA Biosciences	ACEA_ER_AUC_viability	Inactive				null	cytotoxicity	breast	cell line
ASSAY SOURCE: APR	Apredica	APR_HepG2_CellCycleArrest_1h_dn	Inactive				null	proliferation	liver	cell line
ASSAY SOURCE: APR	Apredica	APR_HepG2_CellCycleArrest_1h_up	Inactive				null	arrest	liver	cell line
ASSAY SOURCE: APR	Apredica	APR_HepG2_CellCycleArrest_24h_dn	Inactive				null	proliferation	liver	cell line
ASSAY SOURCE: APR	Apredica	APR_HepG2_CellCycleArrest_24h_up	Inactive				null	arrest	liver	cell line
ASSAY SOURCE: APR	Apredica	APR_HepG2_CellCycleArrest_72h_dn	Inactive				null	proliferation	liver	cell line
ASSAY SOURCE: APR	Apredica	APR_HepG2_CellCycleArrest_72h_up	Inactive				null	arrest	liver	cell line
ASSAY SOURCE: APR	Apredica	APR_HepG2_CellLoss_1h_dn	Inactive				null	cytotoxicity	liver	cell line
ASSAY SOURCE: APR	Apredica	APR_HepG2_CellLoss_1h_up	Inactive				null	proliferation	liver	cell line

Rows: 1,398 Total Rows: 1,398

Full access to concentration-response curves

EPA United States Environmental Protection Agency



Use Models Derived from the Data

EPA United States Environmental Protection Agency

Screening Chemicals for Estrogen Receptor Bioactivity Using a Computational Model

Patience Browne[*†], Richard S. Judson[‡], Warren M. Casey[§], Nicole C. Kleinstreuer[∥], and Russell S. Thomas[‡]

View Author Information

Cite this: *Environ. Sci. Technol.* 2015, 49, 14, 8804–8814
Publication Date: June 12, 2015
https://doi.org/10.1021/acs.est.5b02641

Vol. 124, No. 7 | Research

CERAPP: Collaborative Estrogen Receptor Activity Prediction Project

Kamel Mansouri, Ahmed Abdelaziz, Aleksandra Rybacka, Alessandra Roncaglioni, Alexander Tropsha, Alexandre Varnek, Alexey Zakharov, Andrew Worth, Ann M. Richard, Christopher M. Grulke, Daniela Trisciuzzi, Denis Fourches, Dragos Horvath, Emilio Benfenati, Eugene Muratov, Eva Bay Wedebye, Francesca Grisoni, Giuseppe F. Mangiatordi, ... See all authors

Published: 1 July 2016 | https://doi.org/10.1289/ehp.1510267 | Cited by: 76

Development and Validation of a Computational Model for Androgen Receptor Activity

Nicole C. Kleinstreuer[*†], Patricia Ceger[‡], Eric D. Watt[§], Matthew Martin[§], Keith Houck[§], Patience Browne[∥], Russell S. Thomas[§], Warren M. Casey[†], David J. Dix[⊥], David Allen[‡], Srilatha Sakamuru[#], Menghang Xia[#], Ruili Huang[#], and Richard Judson[§]

View Author Information

Cite this: *Chem. Res. Toxicol.* 2017, 30, 4, 946–964
Publication Date: November 18, 2016
https://doi.org/10.1021/acs.chemrestox.6b00347

Vol. 128, No. 2 | Research

CoMPARA: Collaborative Modeling Project for Androgen Receptor Activity

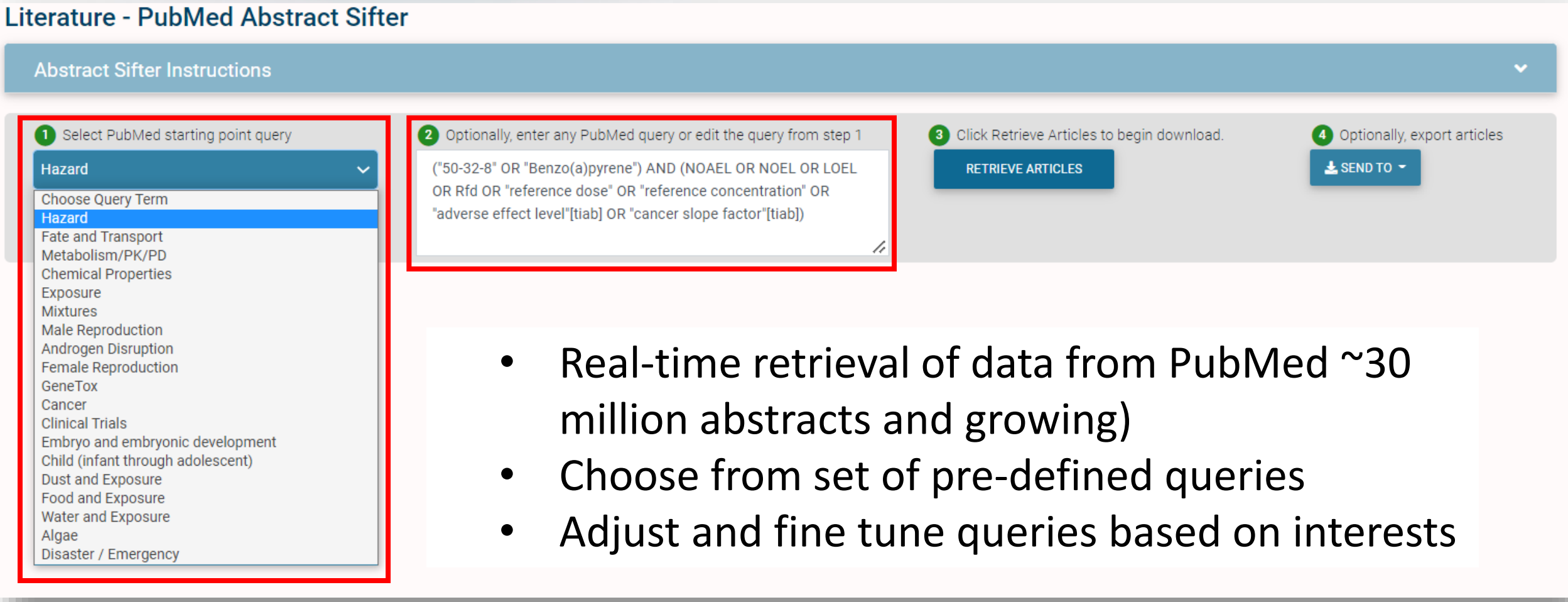
Kamel Mansouri, Nicole Kleinstreuer, Ahmed M. Abdelaziz, Domenico Alberga, Vinicius M. Alves, Patrik L. Andersson, Carolina H. Andrade, Fang Bai, Ilya Balabin, Davide Ballabio, Emilio Benfenati, Barun Bhhatarai, Scott Boyer, Jingwen Chen, Viviana Consonni, Sherif Farag, Denis Fourches, Alfonso T. García-Sosa, Paola Gramatica, Francesca Grisoni, ... See all authors

Published: 7 February 2020 | CID: 027002 | https://doi.org/10.1289/EHP5580 | Cited by: 2

EPA
United States
Environmental Protection
Agency

Searching Literature and the Internet

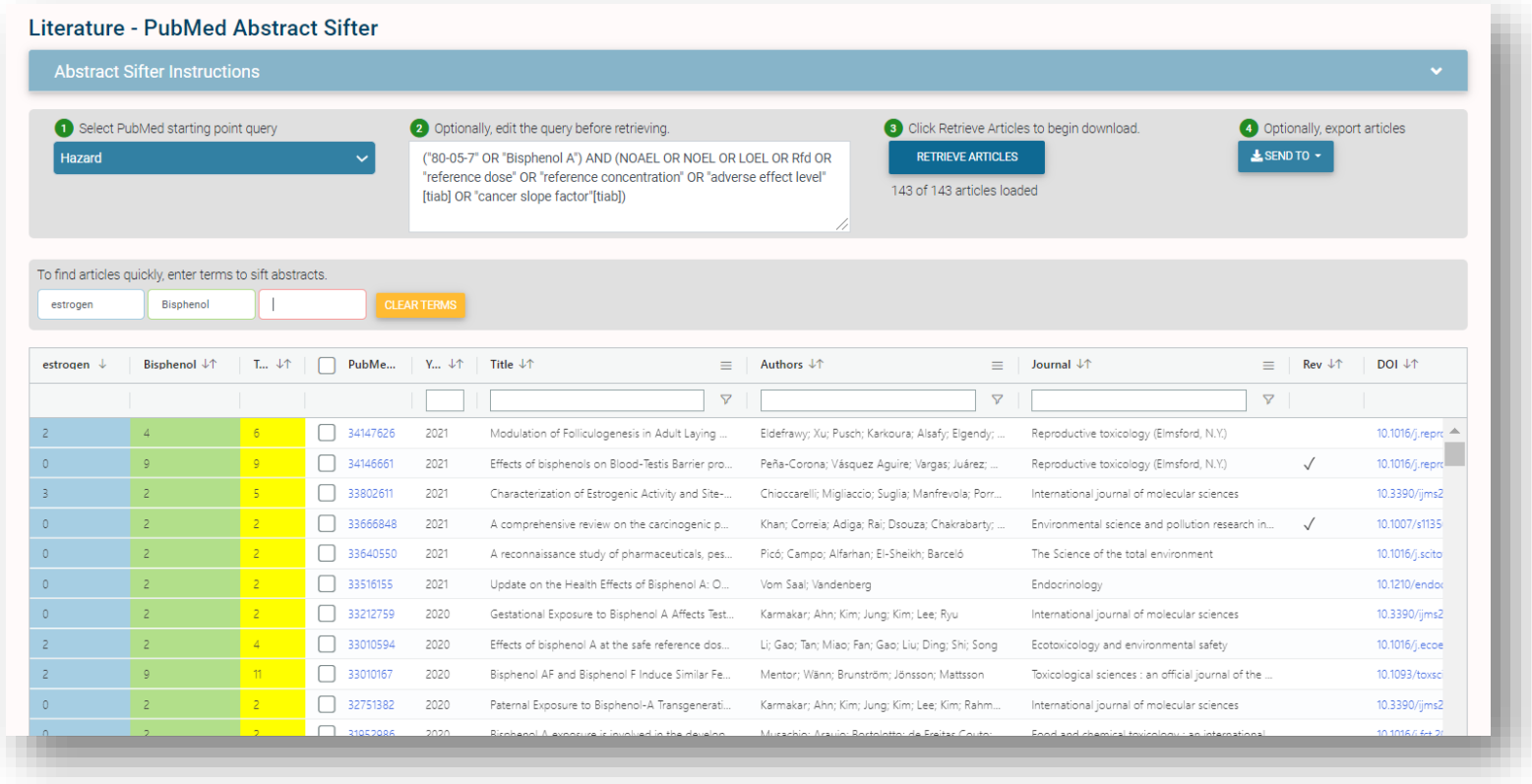
Literature Searching



- Real-time retrieval of data from PubMed ~30 million abstracts and growing)
- Choose from set of pre-defined queries
- Adjust and fine tune queries based on interests

Literature Searching

- “Sifting” of results using multiple terms
- Frequency counting terms
- Color highlighting of terms
- Download list to Excel
- Send list to PubMed for downloading ref. file
- Direct link via PubMed ID



What's the best way to search the internet for chemical data?

- We know how complex chemicals identifiers are…
  - CASRN(s)
  - Hundreds of names (maybe)
  - SMILES
  - InChIs
  - EINECS, EC numbers

- What can WE do to help you navigate the internet?

External Links – Also use Identifiers
Names, CASRN, PubChem IDs, InChIs…

Benzo(a)pyrene

50-32-8 | DTXSID2020139

Searched by DSSTox Substance Id.

General

- EPA Substance Registry Service
- PubChem
- Chemspider
- CPCat
- DrugBank
- Wikipedia
- MSDS Lookup
- ChEMBL
- ToxPlanet
- ACS Reagent Chemicals
- Wolfram Alpha
- ECHA Infocard
- ChemAgora
- Consumer Product Information Database
- ChEBI
- NIST Chemistry Webbook
- WEBWISER
- PubChem Safety Sheet
- Consumer Product Information Database
- PubChem: Chemical Vendors

Toxicology

- ACToR
- DrugPortal
- CCRIS
- ChemView
- CTD
- eChemPortal
- Gene-Tox
- HSDB
- ACToR PDF Report
- CREST
- National Air Toxics Assessment
- ECOTOX
- ChemView
- Chemical Checker
- BindingDB
- CalEPA OEHHA
- NIOSH IDLH Values
- LactMed
- ECOTOX

Publications

- Toxline
- PPRTVWEB
- PubMed
- IRIS Assessments
- EPA HERO
- NIOSH Skin Notation Profiles
- NIOSH Pocket Guide
- RSC Publications
- BioCaddie DataMed
- Springer Materials
- Bielefeld Academic Search Engine
- CORE Literature Search
- Google Books (Text Search)
- Google Patents (Text search)
- Google Scholar (Text search)
- Google Patents (Structure search)
- Google Books (Structure Search)
- Google Scholar (Structure search)
- Federal Register

Analytical

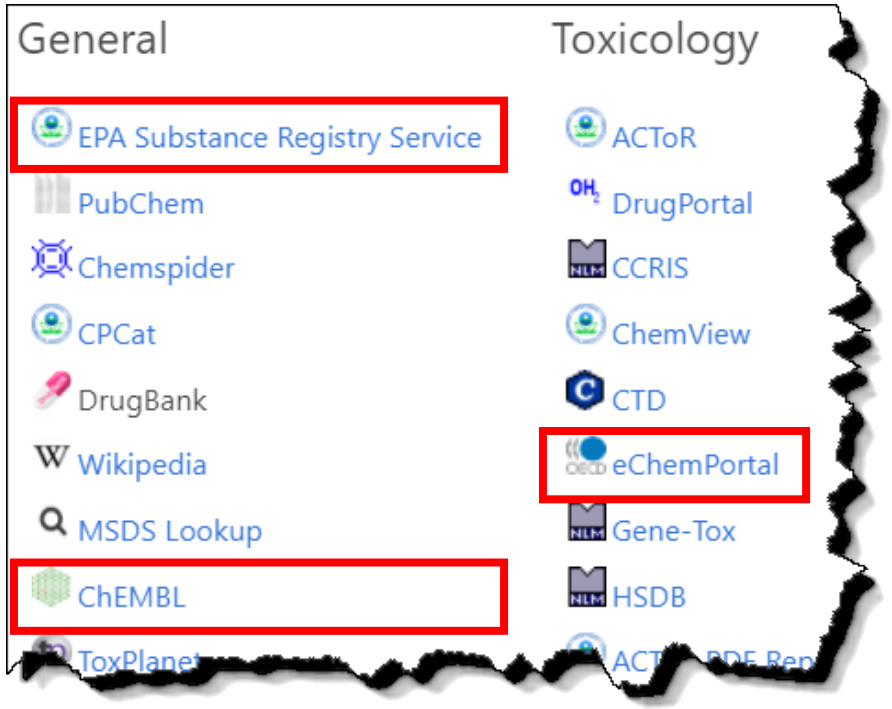
- RSC Analytical Abstracts
- Tox21 Analytical Data
- MONA: MassBank North America
- mzCloud
- NIST IR Spectrum
- NIST MS Spectrum
- MassBank
- NIST Antoine Constants
- IR Spectra on PubChem
- NIST Kovats Index values
- Protein DataBank
- National Environmental Methods Index

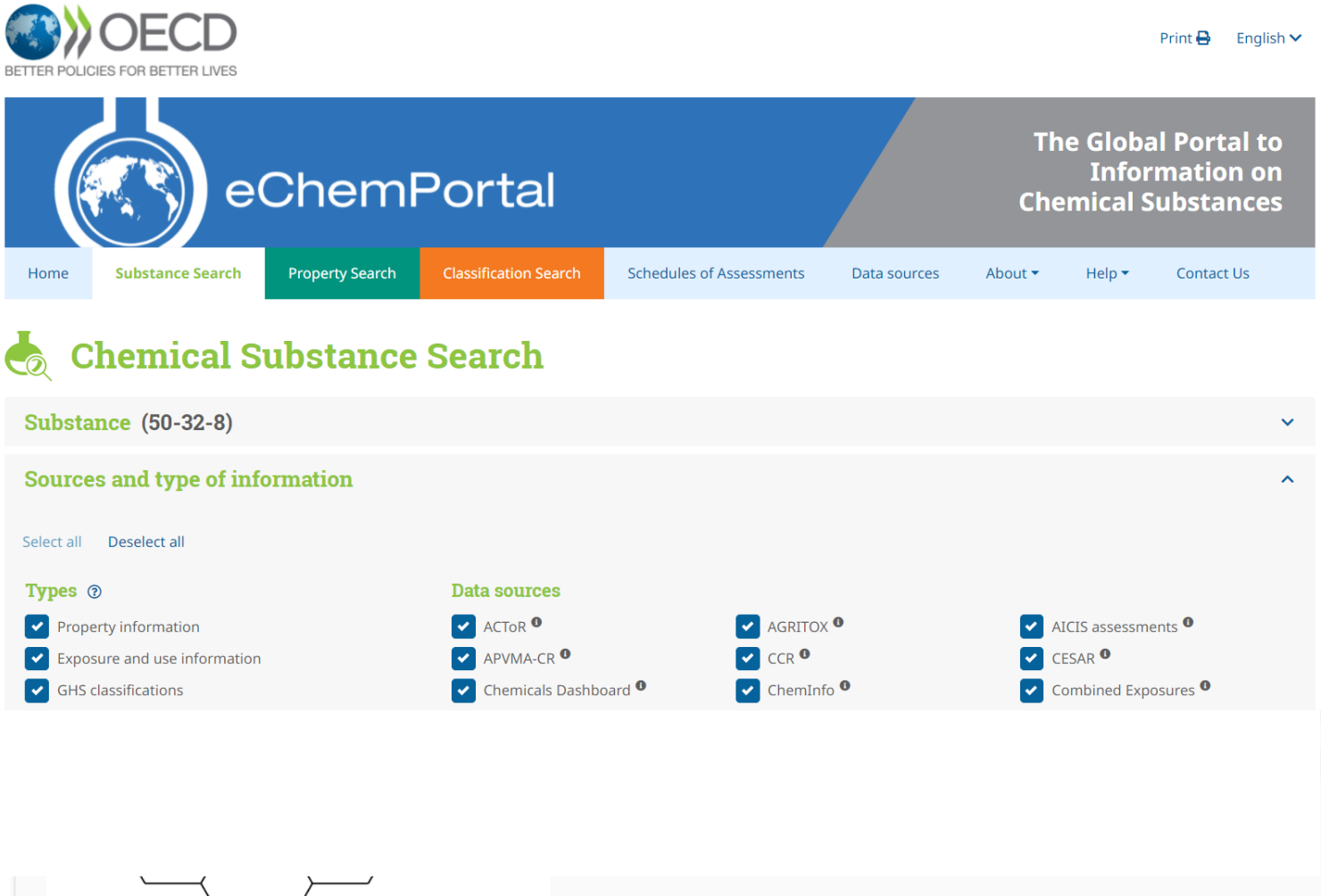
Prediction

- 2D NMR HSQC/HMBC Prediction
- Carbon-13 NMR Prediction
- Proton NMR Prediction
- ChemRTP Predictor
- LSERD

External Links

- Links to ~90 websites providing access to additional data on the chemical of interest





EPA
United States
Environmental Protection
Agency

Chemical Lists and Categories

A List of Lists of Chemicals

https://comptox.epa.gov/dashboard/chemical_lists

EPA United States Environmental Protection Agency

Chemical Lists

Search Chemical Lists

EXPORT | COPY URL

Showing 36 of 319 Records

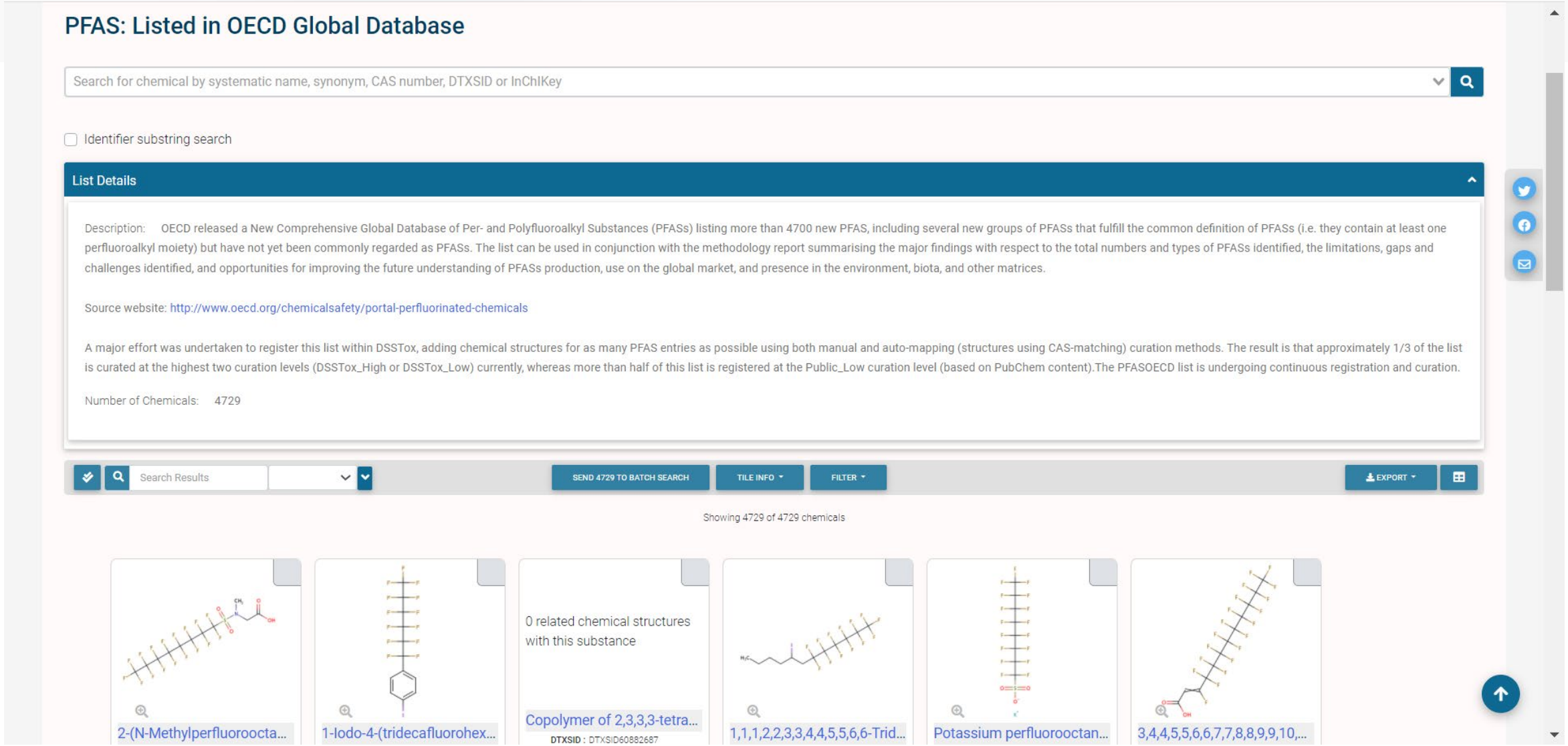
List Acronym	List Name	# Chemicals	Updated	List Description
PFAS				
EPAPFASDW537	PFAS|EPA|WATER: Existing EPA DW Method ...	19	2019-11-16	EPA has recently revised method 537.1 for the PFAS on this list to detect them in drinking water.
EPAPFASDWTREAT	PFAS|EPA|WATER: Drinking Water Treatment...	9	2019-11-16	EPA is gathering and evaluating treatment effectiveness and cost data for removing these PFAS from drinking water systems.
EPAPFASINSOL	PFAS|EPA: Chemical Inventory Insoluble in ...	43	2021-11-21	Per- and Polyfluoroalkyl Substances (PFASs) in EPA's expanded ToxCast chemical inventory that were determined to be insoluble in DMSO above 5mM concentration. These PFAS chemicals were successfully procured from commercial suppliers (with a small number provided by National Toxicology Program partners) but deemed unsuitable for testing due to limited DMSO solubility. For a complete list of solubilized PFAS in EPA's inventory, see https://comptox.epa.gov/dashboard/chemical-lists/EPAPFASINV
EPAPFASINV	PFAS|EPA: ToxCast Chemical Inventory	430	2021-11-21	Per- and Polyfluoroalkyl Substances (PFAS) included in EPA's expanded ToxCast chemical inventory and available for testing. These PFAS chemicals were successfully procured from commercial suppliers (with a small number provided by National Toxicology Program partners) and were deemed suitable for testing (i.e., solubilized in DMSO above 5mM, and not gaseous or highly reactive). All or portions of this inventory are being made available to EPA researchers and collaborators to be analyzed and tested in various high-throughput screening (HTS) and high-throughput toxicity (HTT) assays. The https://comptox.epa.gov/dashboard/chemical-lists/EPAPFAS75S1 list is a prioritized subset of this larger chemical inventory.

Rows: **36 of 319**

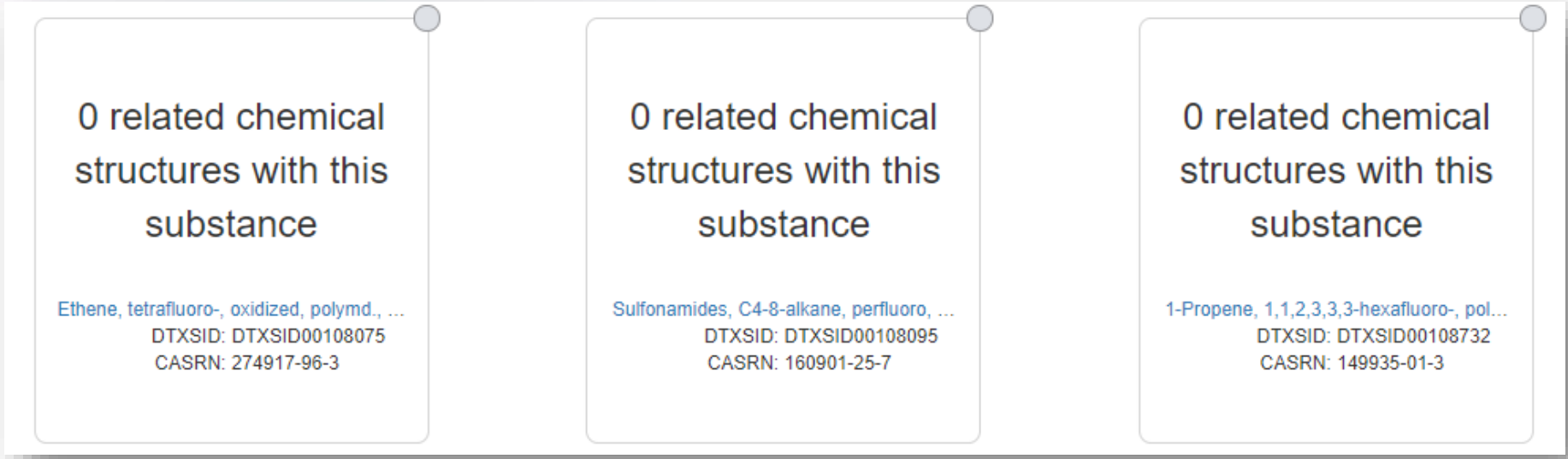
The OECD List of PFAS

http://www.oecd.org/chemicalsafety/portal-perfluorinated-chemicals/

EPA United States Environmental Protection Agency



Example PFAS-UVCBs



Ethene, tetrafluoro-, oxidized, polymd., reduced, decarboxylated, C6 fraction
274917-96-3 | DTXSID00108075

1-Propene, 1,1,2,3,3,3-hexafluoro-, polymer with 1,1-difluoroethene, ethene, 1,1,2,2-tetrafluoroethene and 1,1,2-trifluoro-2-(trifluoromethoxy)ethene
149935-01-3 | DTXSID00108732

PFAS List Paper Just Published

EPA United States Environmental Protection Agency

frontiers in Environmental Science | Toxicology, Pollution and the Environment

ORIGINAL RESEARCH article

Assembly and curation of lists of per- and polyfluoroalkyl substances (PFAS) to support environmental science research

Provisionally accepted
The final version of the article will be published here soon pending final quality checks

Notify me

Antony J. Williams[1*], Linda G. Gaines[2], Christopher Grulke[1], Charles N. Lowe[1], Gabriel Sinclair[3], Vicente Samano[4], Inthirany Thillainadarajah[4], Bryan Meyer[4], Grace Patlewicz[1] and Ann M. Richard[1]

Batch Searching

Batch Searching

- Singleton searches are great but…
- …we generally want data on LOTS of chemicals!

- Typical questions
  - What are the structures for a set of chemical names? Set of CASRNs?
  - Can I get chemical lists in Excel files? As a list of SMILES strings? Can I get an SDF file?
  - Can I include predicted properties? OPERA? TEST?
  - Are “these chemicals” screened in Toxcast?
  - I need masses and formulae for a list of chemicals

Batch Search

EPA United States Environmental Protection Agency

Batch Search

1 Select Input Type(s)

- Substance Identifiers
  - Chemical Name
  - CASRN
  - InChIKey
  - DSSTox Substance ID
- DSSTox Compound ID
- InChIKey Skeleton
- MS-Ready Formula(e)
- Exact Formula(e)
- Monoisotopic Mass

2 Enter Identifiers to Search

(Please enter one identifier per line. Processing time increases with number of inputs.)

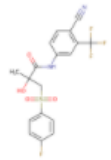
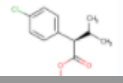
```
DTXSID9020374
DTXSID9020827
DTXSID2022678
DTXSID4023381
DTXSID9044164
DTXSID7032004
DTXSID4022361
DTXSID8021771
```

3 DISPLAY ALL CHEMICALS or CHOOSE EXPORT OPTIONS

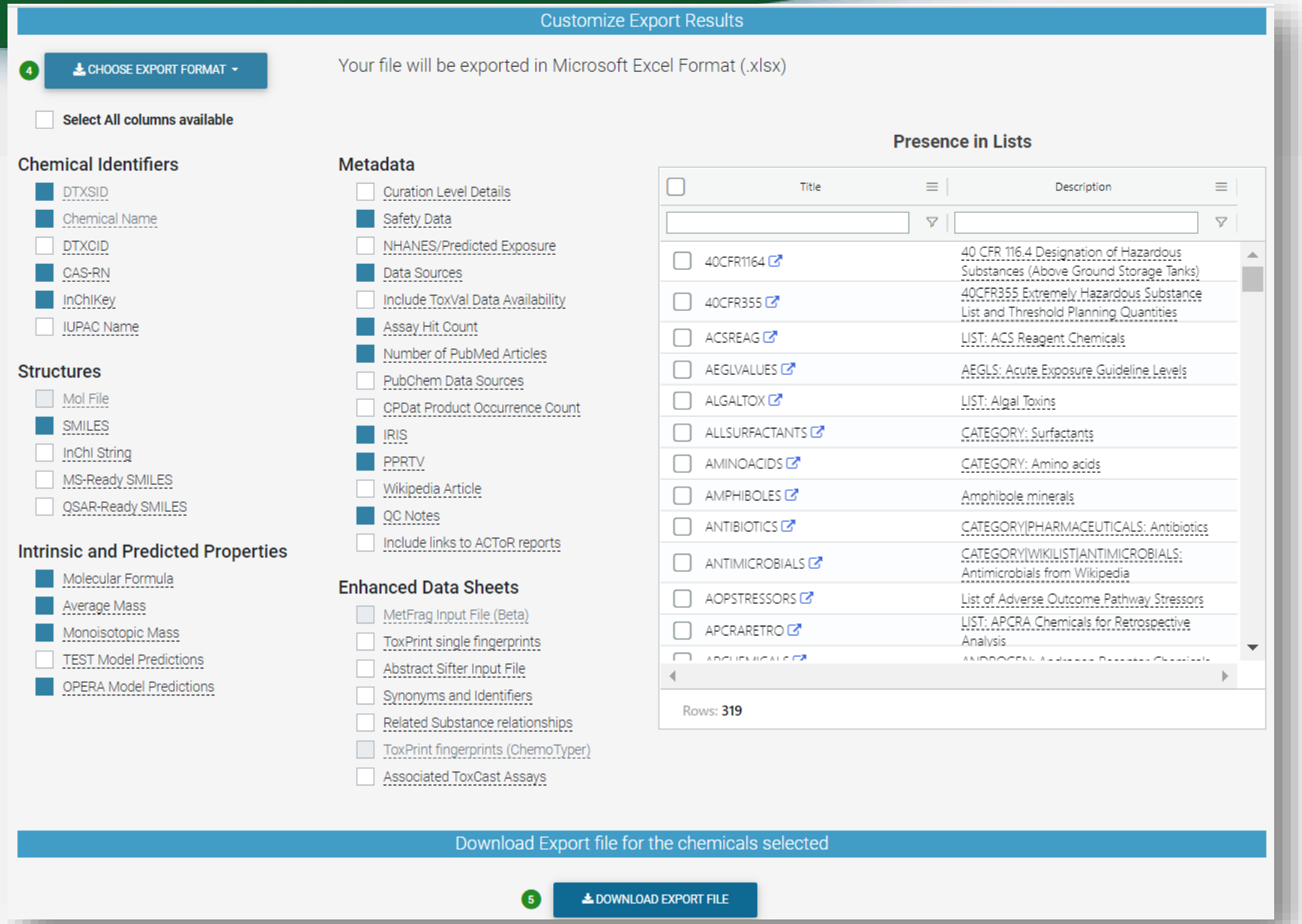
45% loaded

45 Chemicals Found from 110 Input(s)

REPLACE IDENTIFIERS WITH SELECTED CHEMICALS

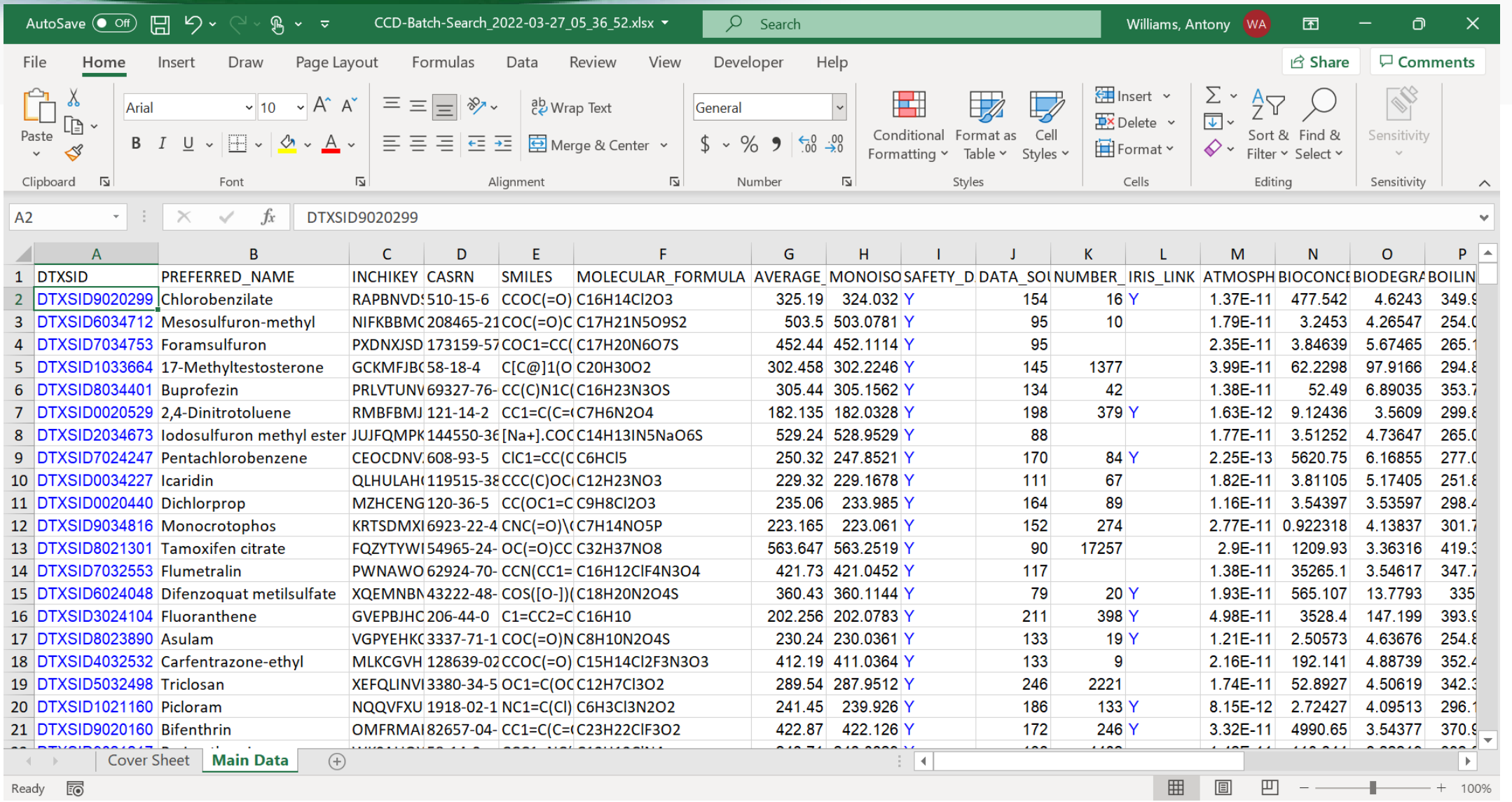
Structure	DTXSID	Preferred Name	CASRN	Mono. Mass	Mol. Formula
	DTXSID2022678	Bicalutamide	90357-06-5	430.061041	C18H14F4N2O4S
	DTXSID3020621	(R,R)-Fenvalerate	67614-33-9	419.128821	C25H22ClNO3

Batch Search – Excel, CSV, SDF file



Batch Search

United States Environmental Protection Agency



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	DTXSID	PREFERRED_NAME	INCHIKEY	CASRN	SMILES	MOLECULAR_FORMULA	AVERAGE_	MONOISO	SAFETY_D	DATA_SOU	NUMBER_	IRIS_LINK	ATMOSPH	BIOCONCE	BIODEGRA	BOILIN
2	DTXSID9020299	Chlorobenzilate	RAPBNVDS	510-15-6	CCOC(=O)	C16H14Cl2O3	325.19	324.032	Y	154	16	Y	1.37E-11	477.542	4.6243	349.9
3	DTXSID6034712	Mesosulfuron-methyl	NIFKBBMC	208465-21	COC(=O)C	C17H21N5O9S2	503.5	503.0781	Y	95	10		1.79E-11	3.2453	4.26547	254.0
4	DTXSID7034753	Foramsulfuron	PXDNXJSD	173159-57	COC1=CC(	C17H20N6O7S	452.44	452.1114	Y	95			2.35E-11	3.84639	5.67465	265.1
5	DTXSID1033664	17-Methyltestosterone	GCKMFJBC	58-18-4	C[C@]1(O	C20H30O2	302.458	302.2246	Y	145	1377		3.99E-11	62.2298	97.9166	294.8
6	DTXSID8034401	Buprofezin	PRLVTUNV	69327-76-	CC(C)N1C(	C16H23N3OS	305.44	305.1562	Y	134	42		1.38E-11	52.49	6.89035	353.7
7	DTXSID0020529	2,4-Dinitrotoluene	RMBFBMJ	121-14-2	CC1=C(C=(	C7H6N2O4	182.135	182.0328	Y	198	379	Y	1.63E-12	9.12436	3.5609	299.8
8	DTXSID2034673	Iodosulfuron methyl ester	JUJFQMPK	144550-36	[Na+].COC	C14H13IN5NaO6S	529.24	528.9529	Y	88			1.77E-11	3.51252	4.73647	265.0
9	DTXSID7024247	Pentachlorobenzene	CEOCDNV	608-93-5	ClC1=CC(C	C6HCl5	250.32	247.8521	Y	170	84	Y	2.25E-13	5620.75	6.16855	277.0
10	DTXSID0034227	Icaridin	QLHULAH(	119515-38	CCC(C)OC(	C12H23NO3	229.32	229.1678	Y	111	67		1.82E-11	3.81105	5.17405	251.8
11	DTXSID0020440	Dichlorprop	MZHCENG	120-36-5	CC(OC1=C	C9H8Cl2O3	235.06	233.985	Y	164	89		1.16E-11	3.54397	3.53597	298.4
12	DTXSID9034816	Monocrotophos	KRTSDMXI	6923-22-4	CNC(=O)\(	C7H14NO5P	223.165	223.061	Y	152	274		2.77E-11	0.922318	4.13837	301.7
13	DTXSID8021301	Tamoxifen citrate	FQZYTYWI	54965-24-	OC(=O)CC	C32H37NO8	563.647	563.2519	Y	90	17257		2.9E-11	1209.93	3.36316	419.3
14	DTXSID7032553	Flumetralin	PWNAWO	62924-70-	CCN(CC1=	C16H12ClF4N3O4	421.73	421.0452	Y	117			1.38E-11	35265.1	3.54617	347.7
15	DTXSID6024048	Difenzoquat metilsulfate	XQEMNBN	43222-48-	COS([O-])(	C18H20N2O4S	360.43	360.1144	Y	79	20	Y	1.93E-11	565.107	13.7793	335
16	DTXSID3024104	Fluoranthene	GVEPBJHC	206-44-0	C1=CC2=C	C16H10	202.256	202.0783	Y	211	398	Y	4.98E-11	3528.4	147.199	393.9
17	DTXSID8023890	Asulam	VGPYEHKC	3337-71-1	COC(=O)N	C8H10N2O4S	230.24	230.0361	Y	133	19	Y	1.21E-11	2.50573	4.63676	254.8
18	DTXSID4032532	Carfentrazone-ethyl	MLKCGVH	128639-02	CCOC(=O)	C15H14Cl2F3N3O3	412.19	411.0364	Y	133	9		2.16E-11	192.141	4.88739	352.4
19	DTXSID5032498	Triclosan	XEFQLINVI	3380-34-5	OC1=C(OC	C12H7Cl3O2	289.54	287.9512	Y	246	2221		1.74E-11	52.8927	4.50619	342.3
20	DTXSID1021160	Picloram	NQQVFXU	1918-02-1	NC1=C(Cl)	C6H3Cl3N2O2	241.45	239.926	Y	186	133	Y	8.15E-12	2.72427	4.09513	296.1
21	DTXSID9020160	Bifenthrin	OMFRMAI	82657-04-	CC1=C(C=(	C23H22ClF3O2	422.87	422.126	Y	172	246	Y	3.32E-11	4990.65	3.54377	370.9

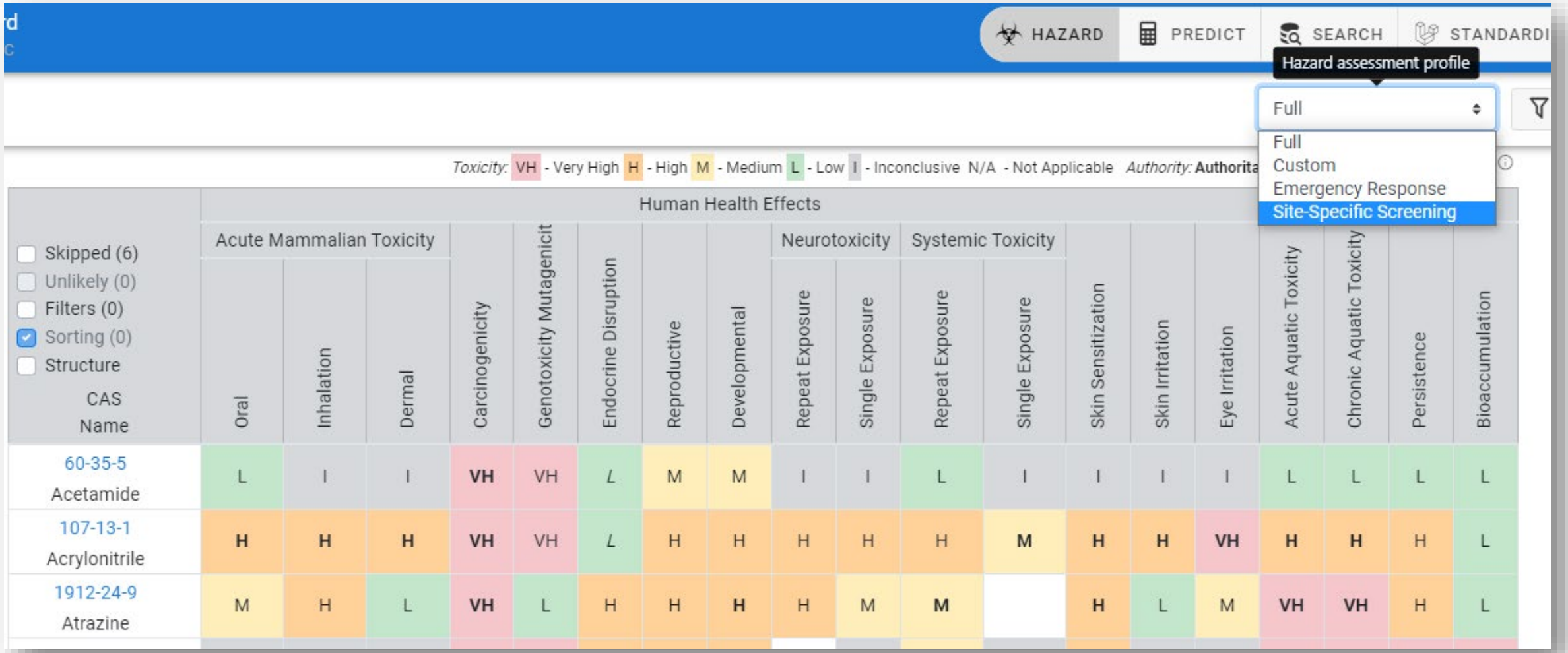
EPA
United States
Environmental Protection
Agency

Cheminformatics “PoC Modules”

Six Modules

- Hazard Comparison Profiling – profile chemicals based on hazard
- Alerts – structure, substructure, SMARTS based alerts and flags
- Predict – batch prediction using WebTEST (100s of structures)
- Search – structure/substructure/similarity searches
- Standardize – convert structures into QSAR/MS-Ready forms
- ToxPrints – generate ToxPrint substructural fragments and profile

Module 1: Hazard Module



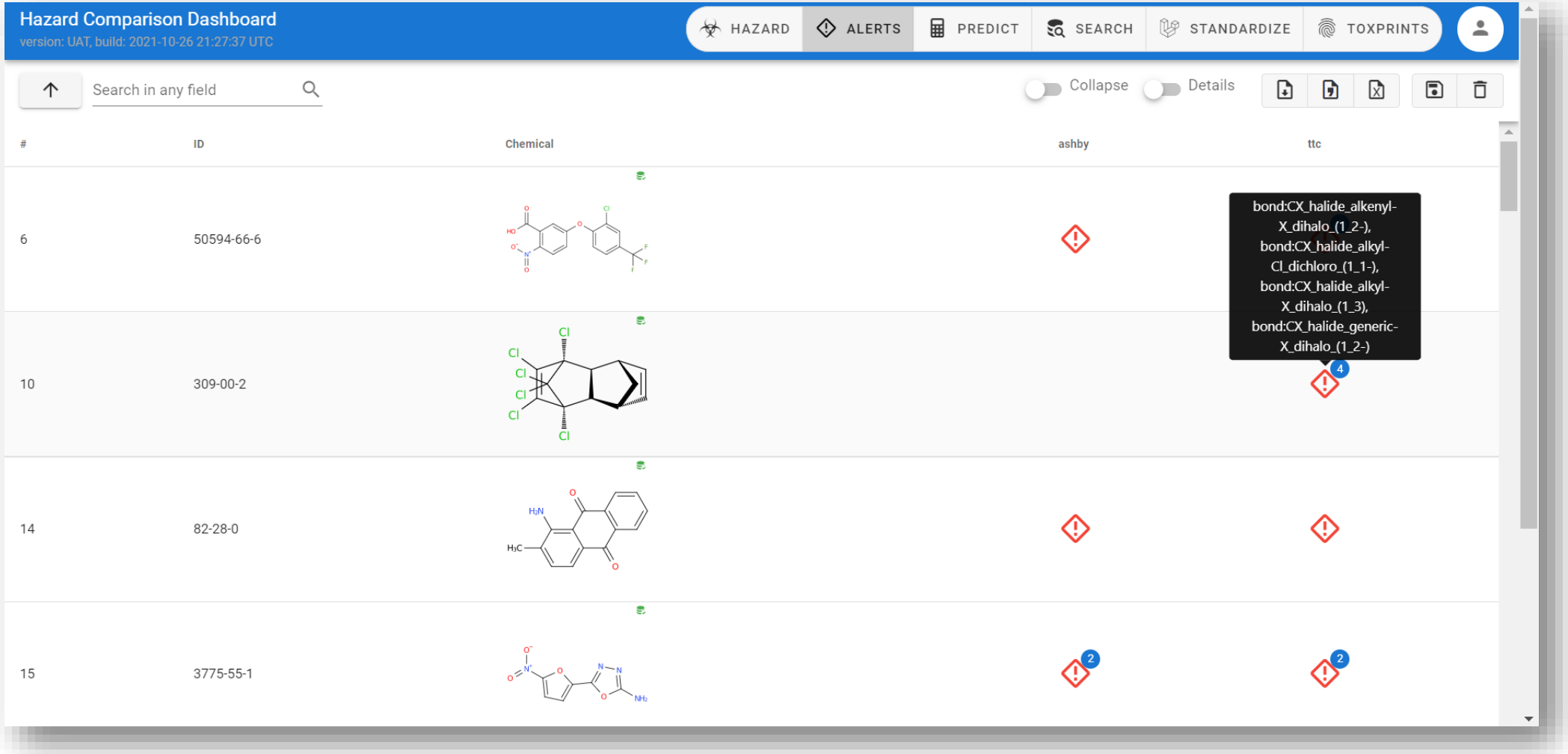
Export SDF or Excel File

- Most useful form for your needs..Excel

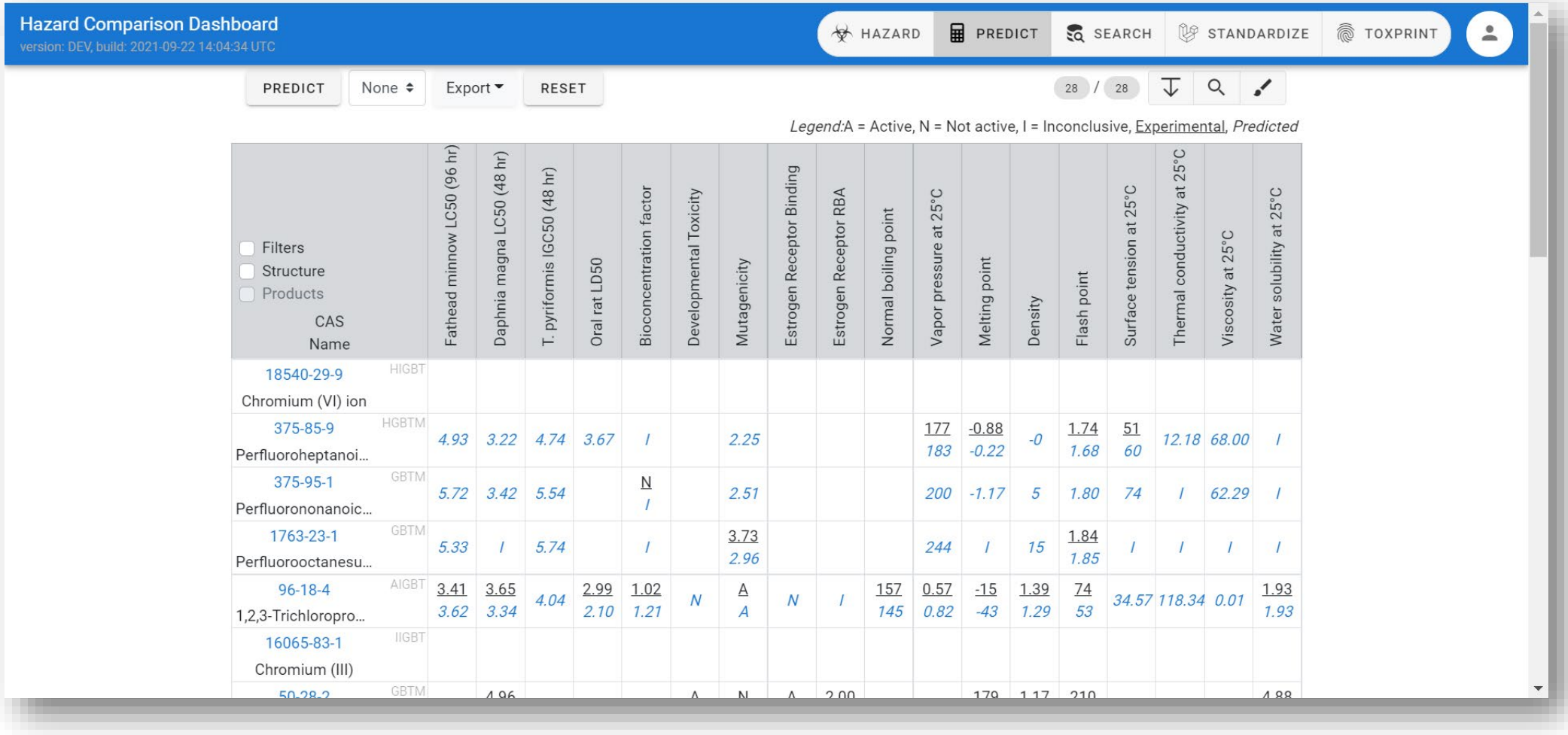
The Hazard Comparison Dashboard is a prototype tool and a compilation of information sourced from many sites, databases and sources including U.S. Federal and state sources and international bodies that saves the user time by providing information in
The data are not reviewed by USEPA – the user must apply judgment in use of the information. The results do not indicate EPA's position on the use or regulation of these chemicals.

			VH - Very High	H - High	M - Medium	L - Low	I - Inconclusive	No Data		Authoritative	Screening	QSAR Model	
			Human Health Effects								Ecotoxicity		Fa
			Acute Mammalian Toxicity					Neurotoxicity		Systemic Toxicity									
DTXSID	CAS	Name	Oral	Inhalation	Dermal	Carcinogenicity	Genotoxicity Mutagenicity	Endocrine Disruption	Reproductive	Developmental	Repeat Exposure	Single Exposure	Repeat Exposure	Single Exposure	Skin Sensitization	Skin Irritation	Eye Irritation	Acute Aquatic Toxicity	Chronic Aquatic Toxicity	Persistence	
DTXSID6020438	107-06-2	1,2-Dichloroethane	M	H	L	VH	VH	L	H	L	H	H	M	M	M	H	H	L	L	H	
DTXSID7024031	156-60-5	(E)-1,2-Dichloroethylene	M	L	L	I	L	I	I	L	I		M	H	I	H	H	L		H	
DTXSID5021380	79-00-5	1,1,2-Trichloroethane	M	M	M	VH	L	L	I	L	H		H	H	I	H	M	M	M	H	
DTXSID0020448	78-87-5	1,2-Dichloropropane	M	H	L	VH	VH	L	M	M			L	H	H	H	H	M	M	H	
DTXSID1020437	75-34-3	1,1-Dichloroethane	M	M	I	VH	VH	L	I	L	I	H	L	M	I	H	H	L	M	H	
DTXSID3020203	106-99-0	1,3-Butadiene	L	L	I	VH	VH	L	H	H			H	M	I	I	H	M	L	L	
DTXSID6020430	95-50-1	1,2-Dichlorobenzene	M	H	L	I	VH	H	I	L	H		H	M	I	H	H	VH	VH	H	
DTXSID0021383	79-01-6	Trichloroethylene	L	M	L	VH	VH	I	H	H	H	H	H	M	H	H	H	H	VH	H	
DTXSID8020250	56-23-5	Carbon tetrachloride	H	H	H	VH	VH	H	M	L		H	H	H	H	H	H	M	M	H	
DTXSID4020533	123-91-1	1,4-Dioxane	M	M	L	VH	VH	L	I	L	H	H	H	M	I	H	H	H	L	H	
DTXSID6020856	872-50-4	N-Methyl-2-pyrrolidone	L	L	L	I	L	L	H	H	M		M	M	I	H	H	H		L	

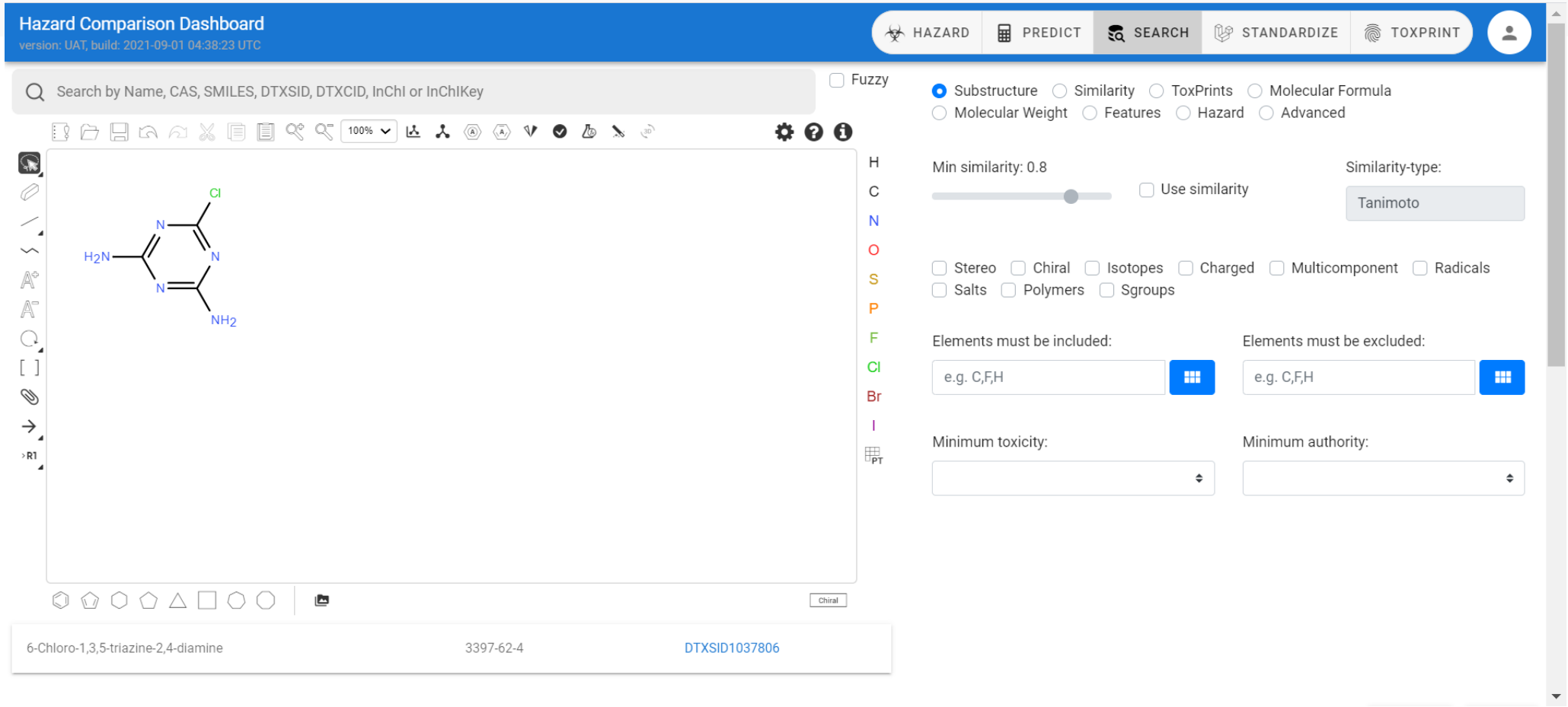
Module 2: Alerts



Module 3: WebTEST Batch Prediction



Module 4: Structure/Substructure/Similarity

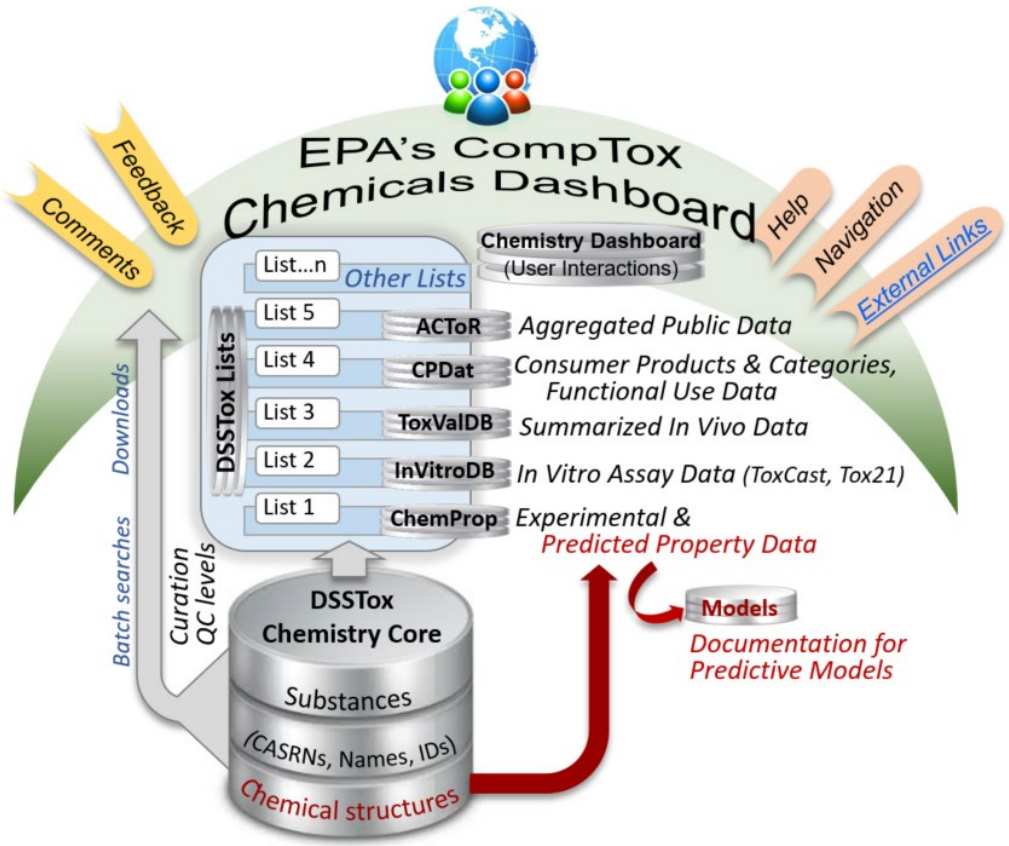


These Modules will be public shortly

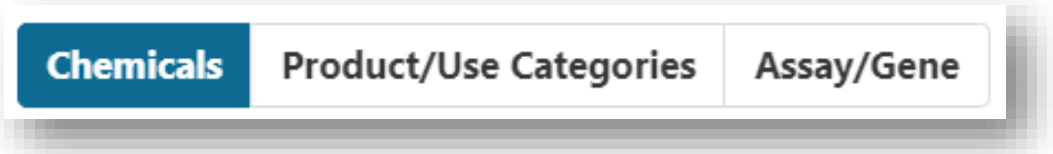
- These modules, and their future updates, are presently being deployed to the public

- Intended URL is : https://cheminformatics.epa.gov

Summary and Conclusion



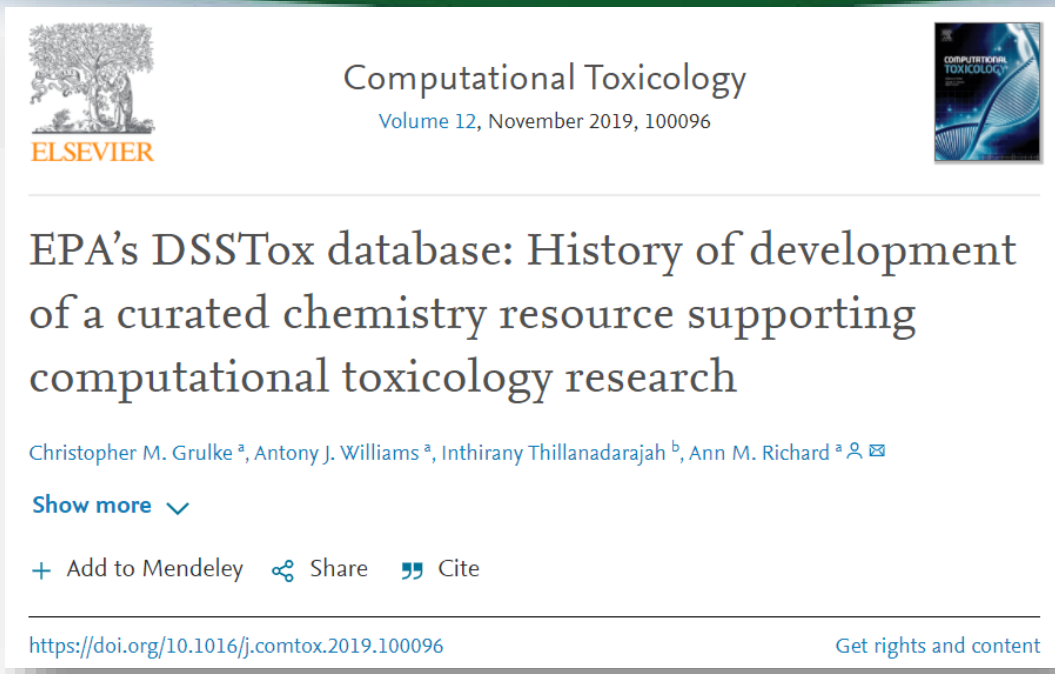
- CompTox Chemicals Dashboard - a central hub for environmental data
  - ~900k chemical substances
  - Integrating property data, hazard data, exposure data, *in vitro* bioactivity data
  - Interrogation of bioactivity data -
  - Multiple types of searches



  - Batch search for thousands of chemicals
  - Real-time property and toxicity predictions
  - Downloadable files – CSV, TSV and Excel

Some Related Publications of Interest

EPA United States Environmental Protection Agency



ELSEVIER

Computational Toxicology

Volume 12, November 2019, 100096

EPA's DSSTox database: History of development of a curated chemistry resource supporting computational toxicology research

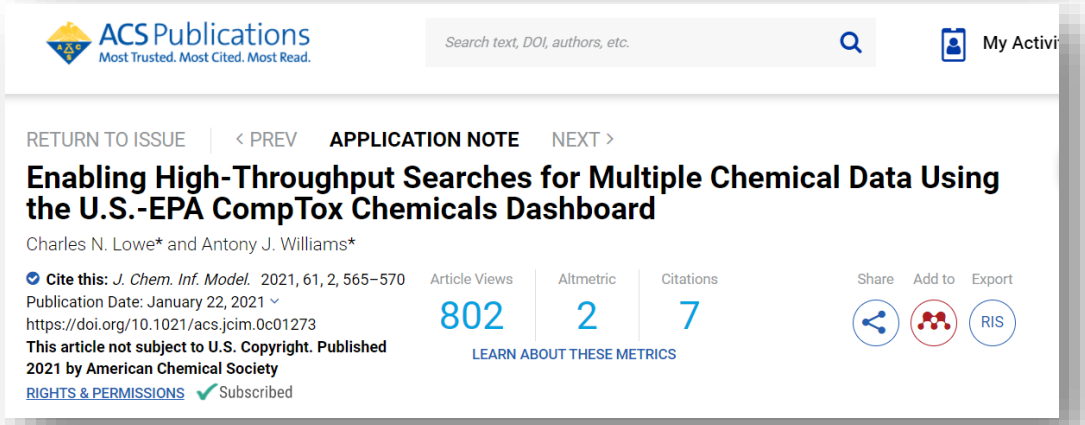
Christopher M. Grulke [a], Antony J. Williams [a], Inthirany Thillanadarajah [b], Ann M. Richard [a]

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.comtox.2019.100096

Get rights and content



ACS Publications
Most Trusted. Most Cited. Most Read.

Search text, DOI, authors, etc.

My Activi

RETURN TO ISSUE | < PREV | APPLICATION NOTE | NEXT >

Enabling High-Throughput Searches for Multiple Chemical Data Using the U.S.-EPA CompTox Chemicals Dashboard

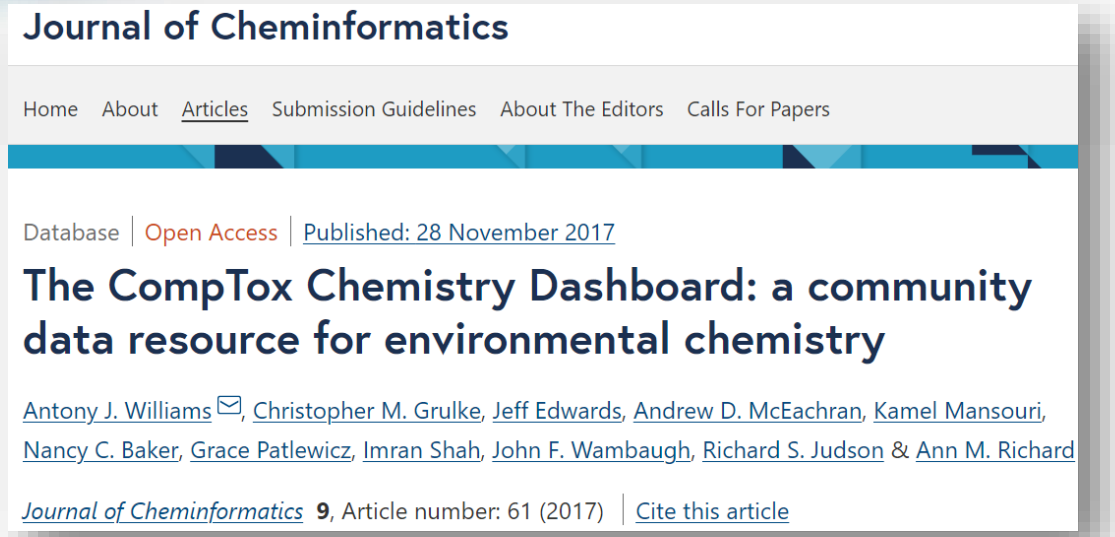
Charles N. Lowe* and Antony J. Williams*

Cite this: *J. Chem. Inf. Model.* 2021, 61, 2, 565−570
Publication Date: January 22, 2021
https://doi.org/10.1021/acs.jcim.0c01273

Article Views 802 | Altmetric 2 | Citations 7

LEARN ABOUT THESE METRICS

Share Add to Export RIS



Journal of Cheminformatics

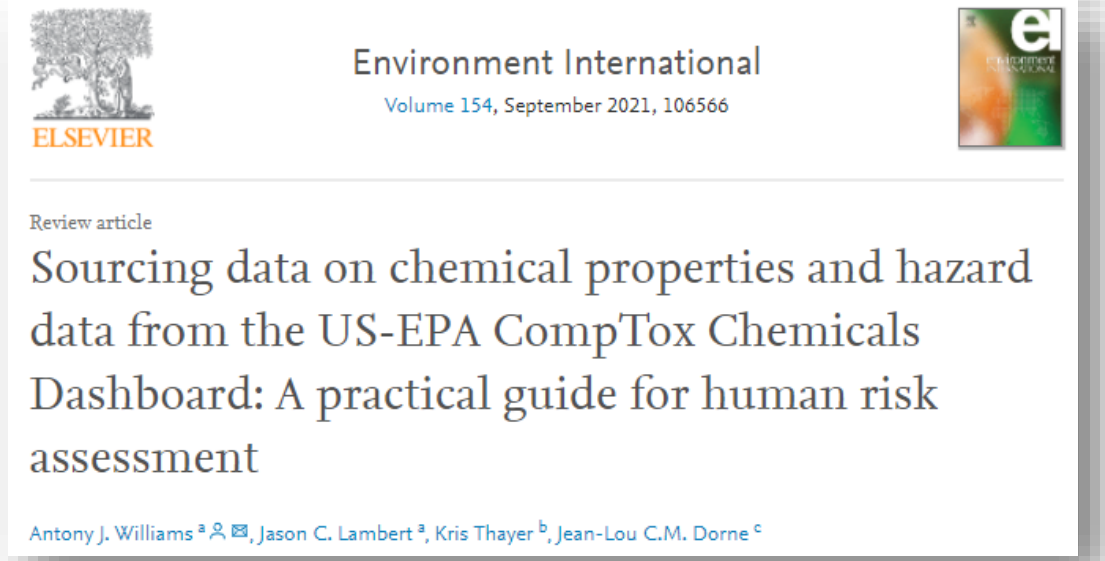
Home About Articles Submission Guidelines About The Editors Calls For Papers

Database | Open Access | Published: 28 November 2017

The CompTox Chemistry Dashboard: a community data resource for environmental chemistry

Antony J. Williams, Christopher M. Grulke, Jeff Edwards, Andrew D. McEachran, Kamel Mansouri, Nancy C. Baker, Grace Patlewicz, Imran Shah, John F. Wambaugh, Richard S. Judson & Ann M. Richard

Journal of Cheminformatics **9**, Article number: 61 (2017) | Cite this article



ELSEVIER

Environment International

Volume 154, September 2021, 106566

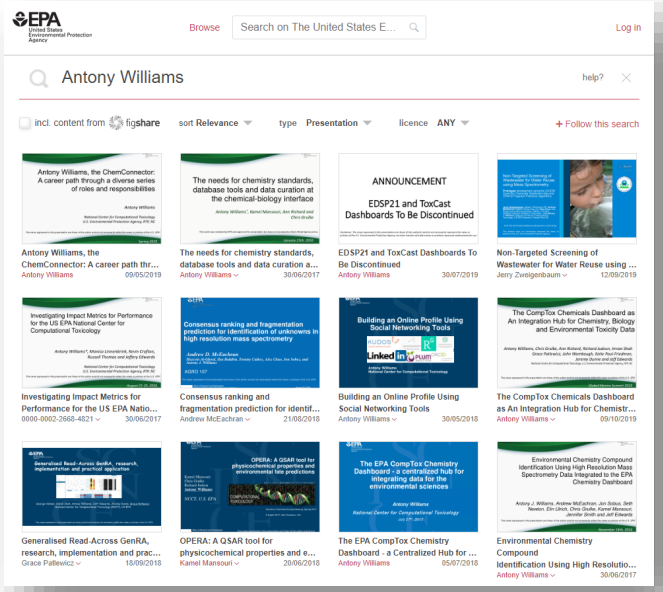
Review article

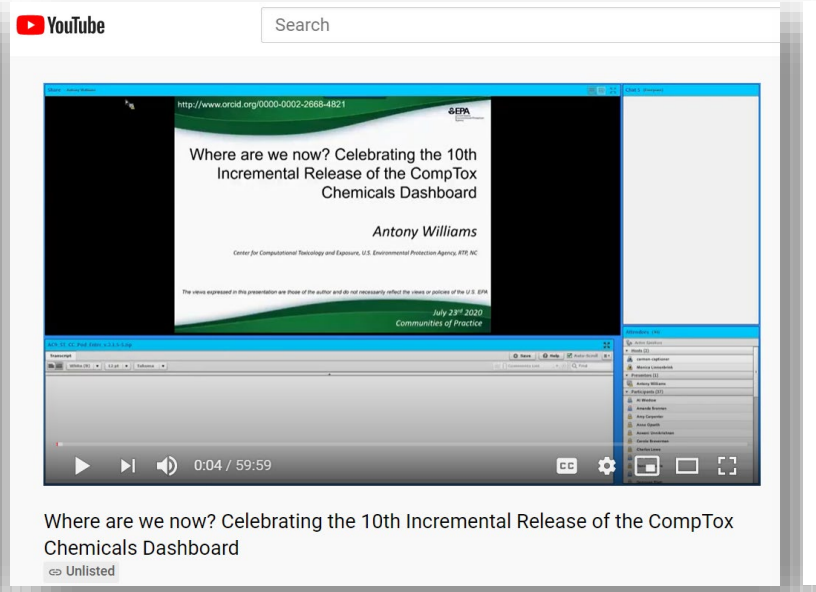
Sourcing data on chemical properties and hazard data from the US-EPA CompTox Chemicals Dashboard: A practical guide for human risk assessment

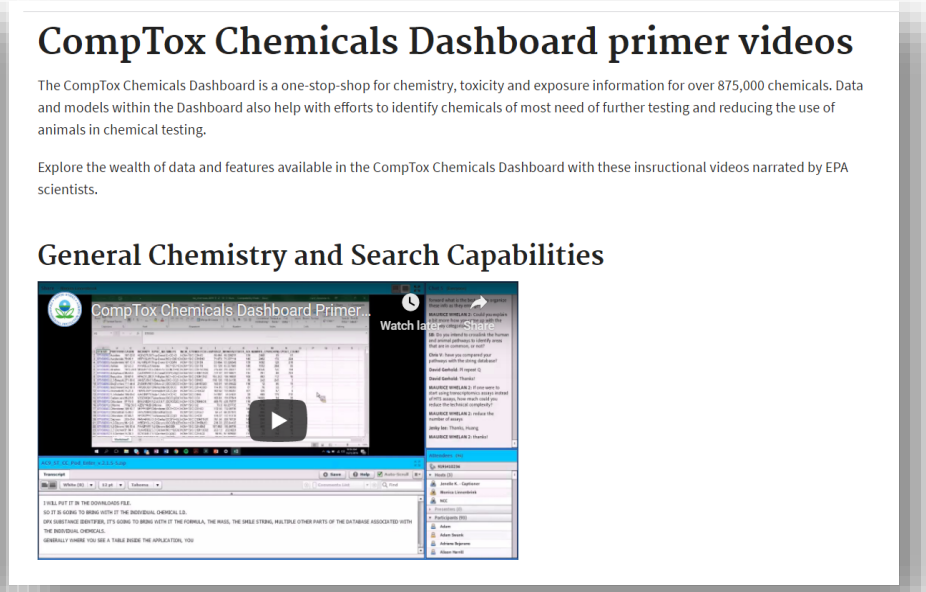
Antony J. Williams [a], Jason C. Lambert [a], Kris Thayer [b], Jean-Lou C.M. Dorne [c]

You want to know more…

- Lots of resources available
  - Presentations: https://tinyurl.com/w5hqs55
  - Communities of Practice Videos: https://rb.gy/qsbno1
  - Manual: https://rb.gy/4fgydc
  - Latest News: https://comptox.epa.gov/dashboard/news_info







Acknowledgments

- Contact:Williams.Antony@epa.gov

- Feedback and follow-up is welcomed! Your questions help
- The dashboard is based on the efforts of many more team members than us. Many collaborators provide data also.

EPA's Center for Computational Toxicology and Exposure