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

EPA CompTox chemicals dashboard: An online resource for environmental chemists

Antony Williams, *Chris Grulke, Jeremy Dunne and Jeff Edwards*

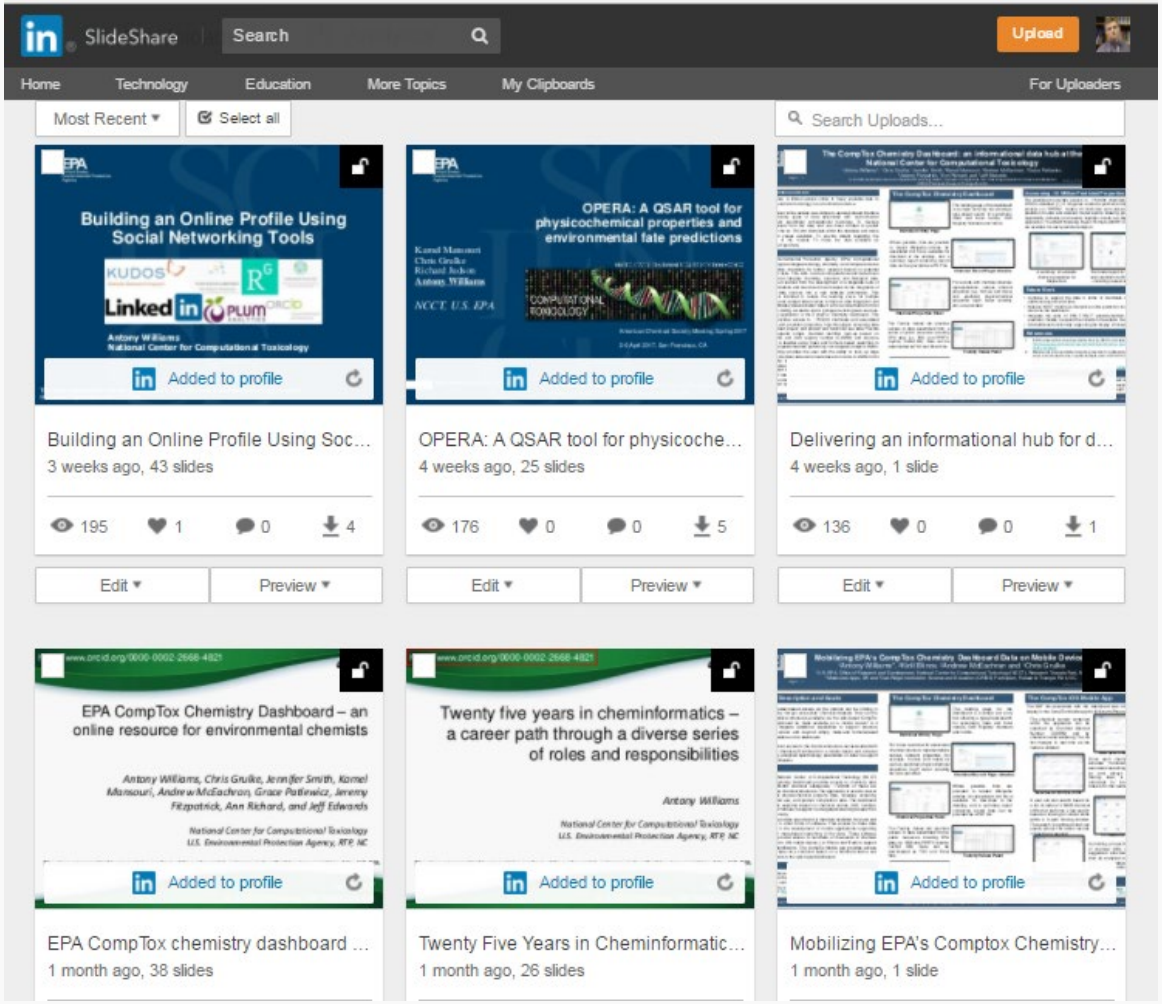
National Center for Computational Toxicology, U.S. Environmental Protection Agency, RTP, NC

Spring 2019
ACS Spring Meeting, Orlando

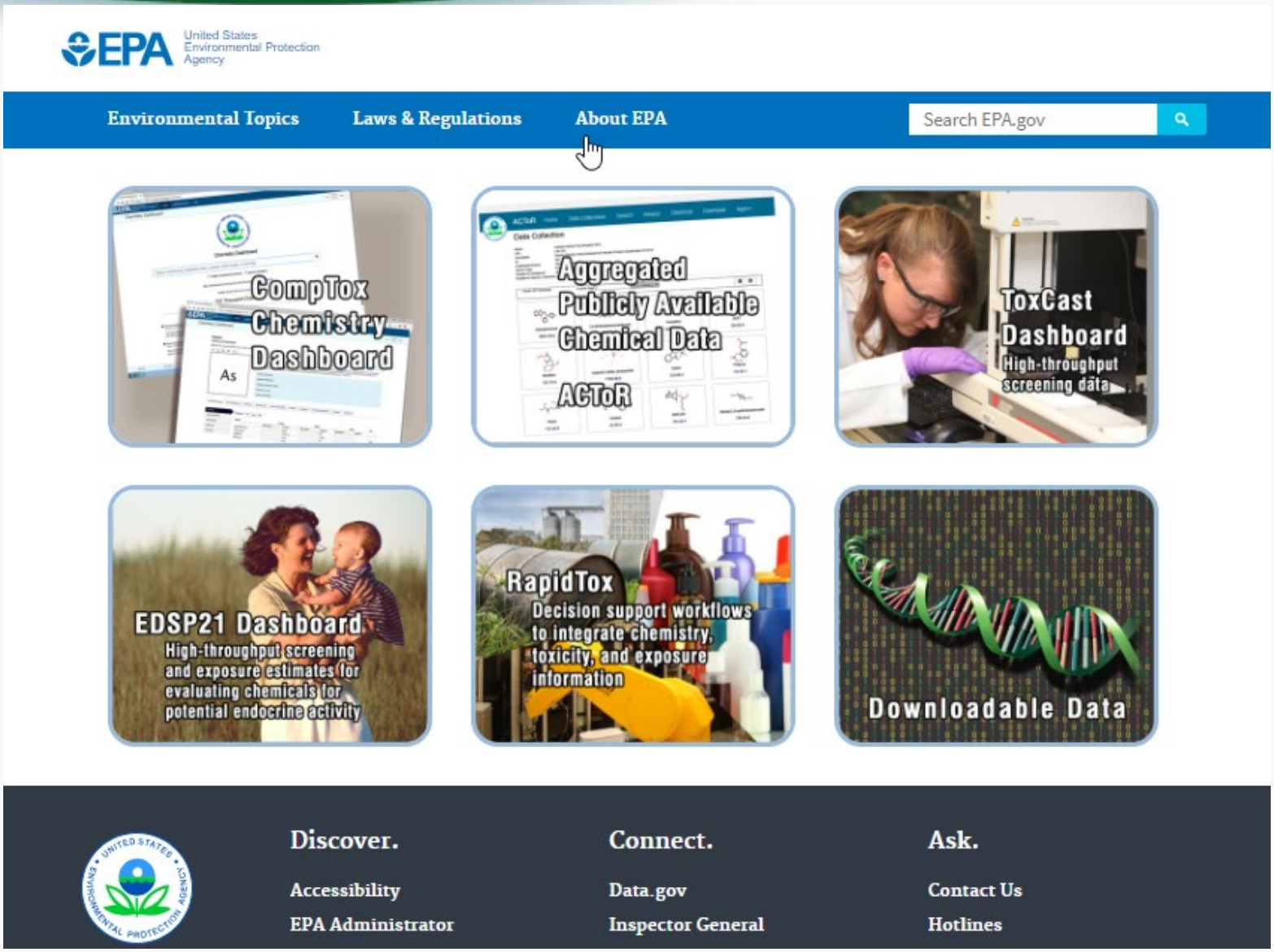
Limit notetaking if you wish

EPA United States Environmental Protection Agency

www.slideshare.net/AntonyWilliams



CompTox Portal



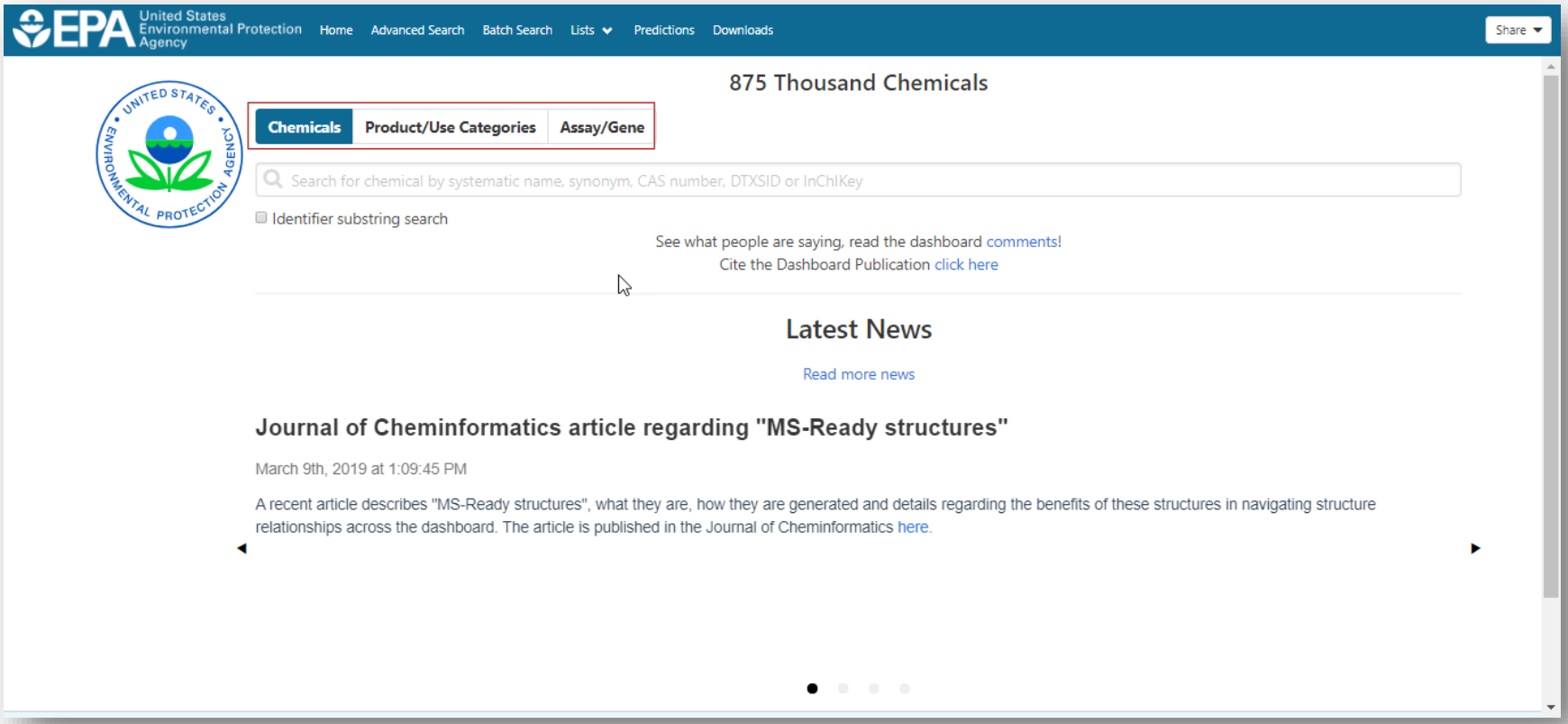
CompTox Chemicals Dashboard

- A **publicly accessible website** delivering access:
  - ~875,000 chemicals with related property data
  - Experimental and predicted physicochemical property data
  - Experimental Human and Ecological hazard data
  - Integration to “biological assay data” for 1000s of chemicals
  - Information regarding consumer products containing chemicals
  - Links to other agency websites and public data resources
  - “Literature” searches for chemicals using public resources
  - “Batch searching” for thousands of chemicals
  - Real time prediction of physchem and toxicity endpoints

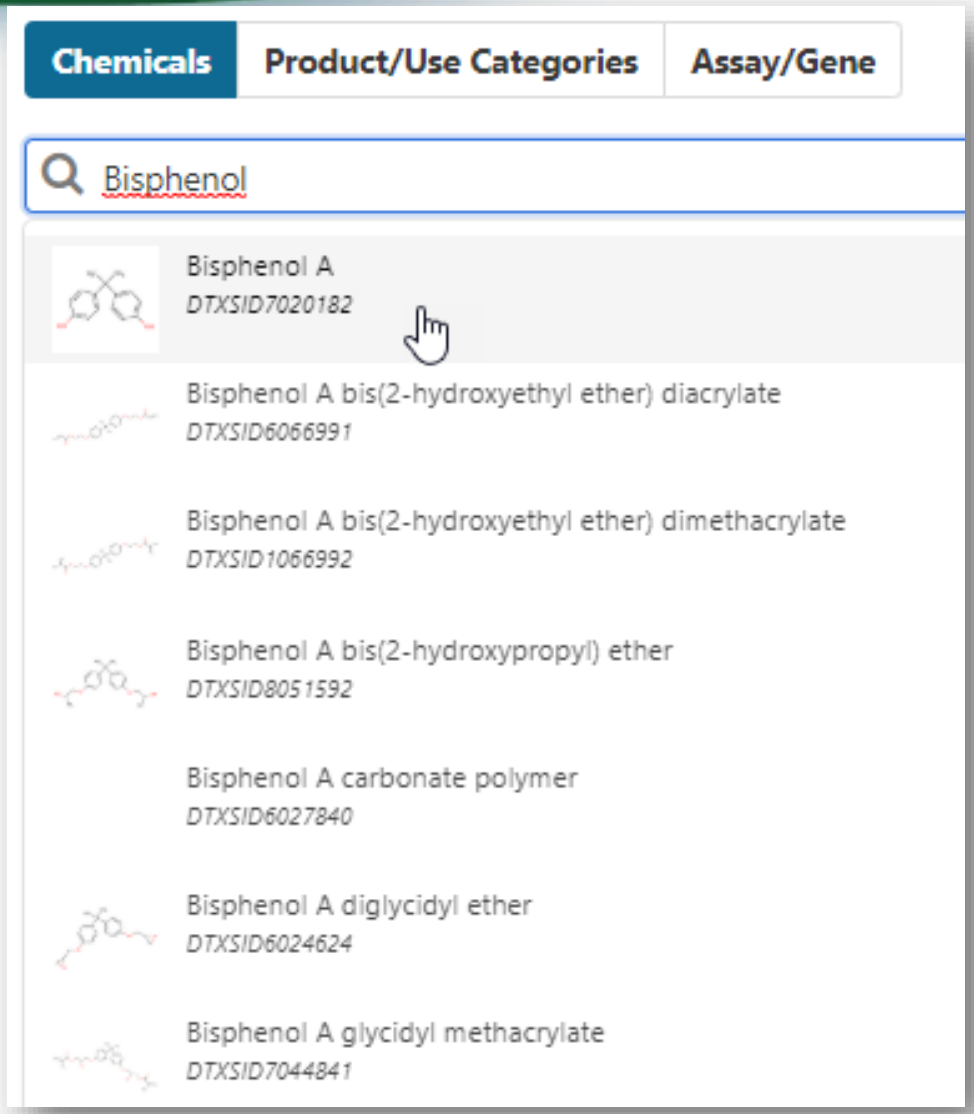
CompTox Chemicals Dashboard

https://comptox.epa.gov/dashboard

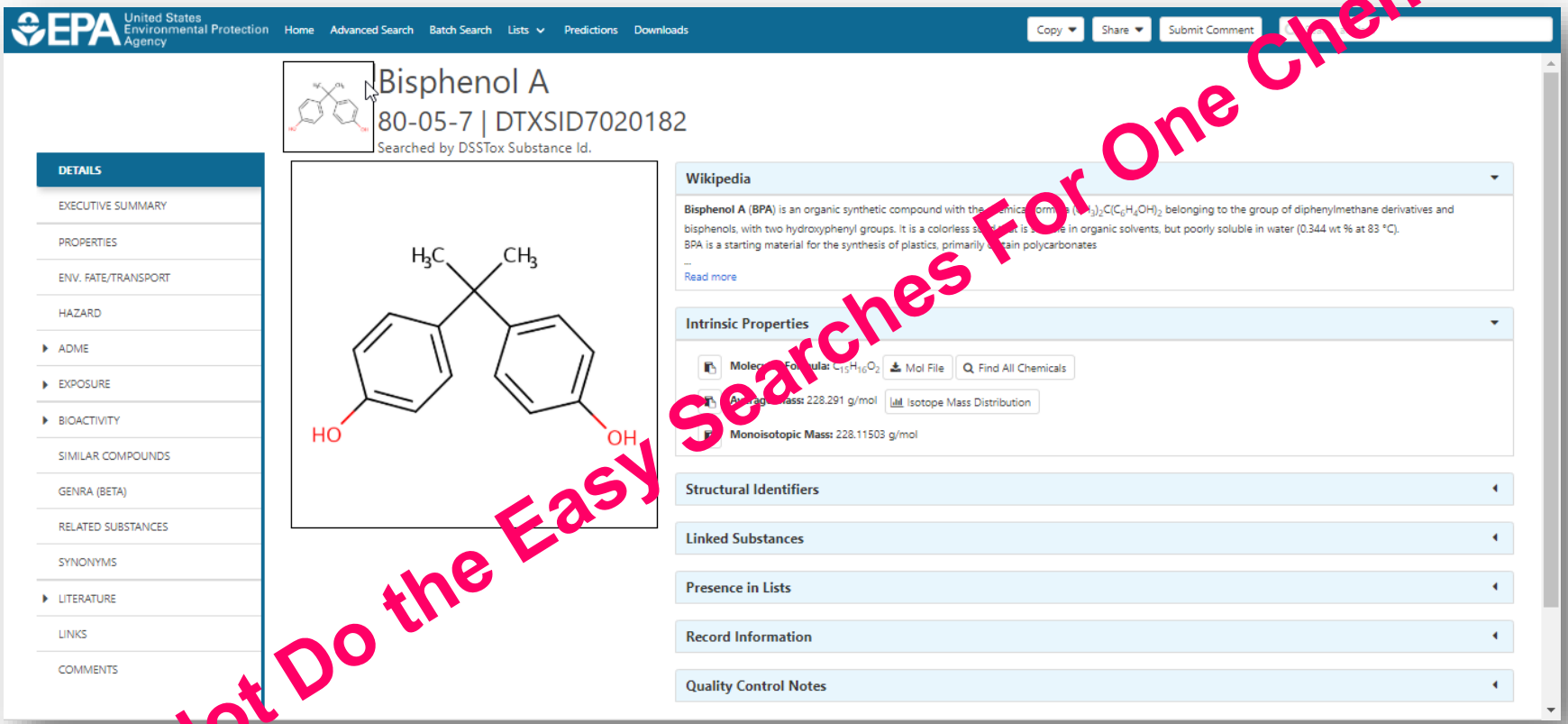
875k Chemical Substances



BASIC Search



Detailed Chemical Pages

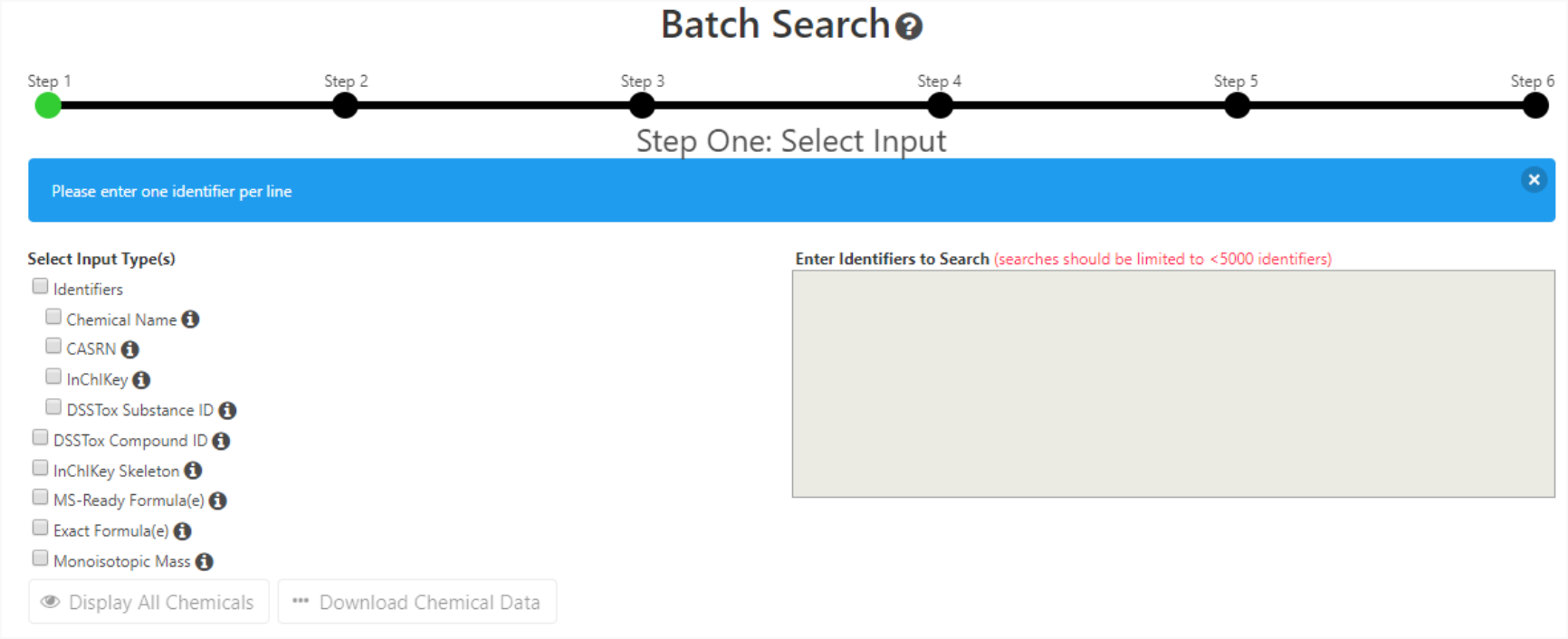


Let's Not Do the Easy Searches For One Chemical

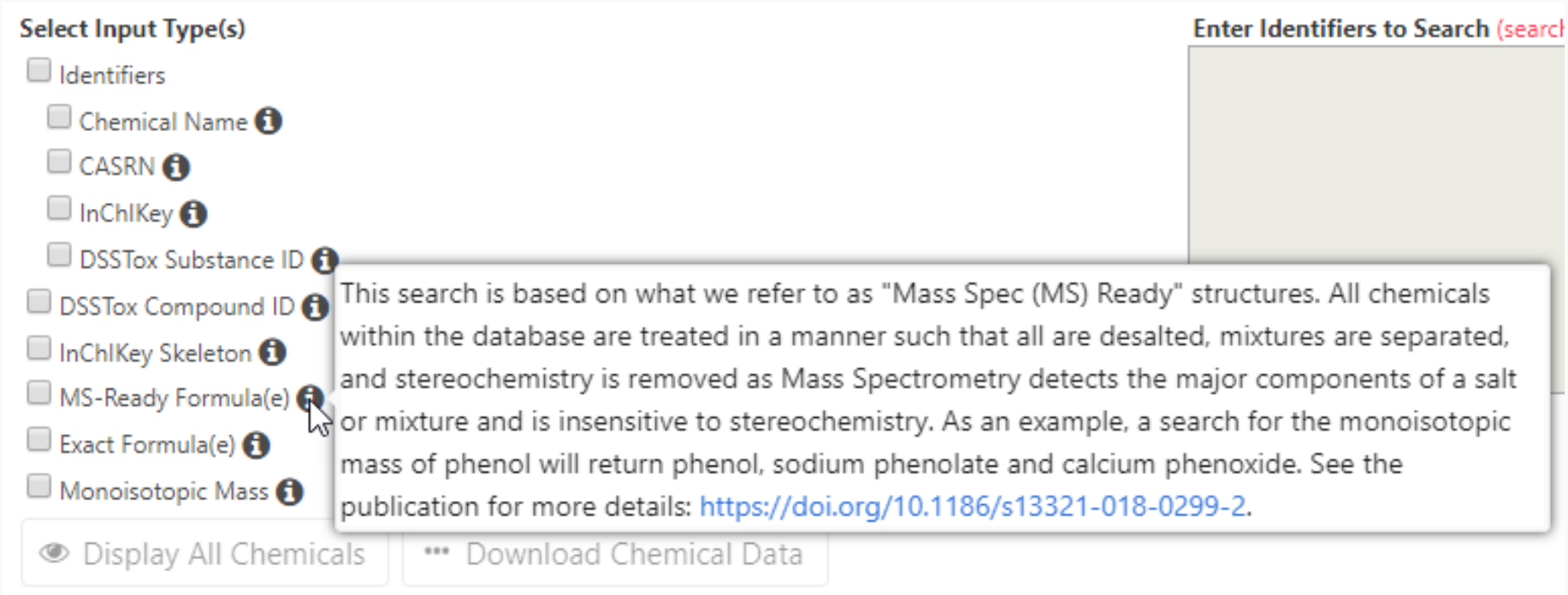
THE POWER OF BATCH SEARCHING

Batch Access to Underlying Data

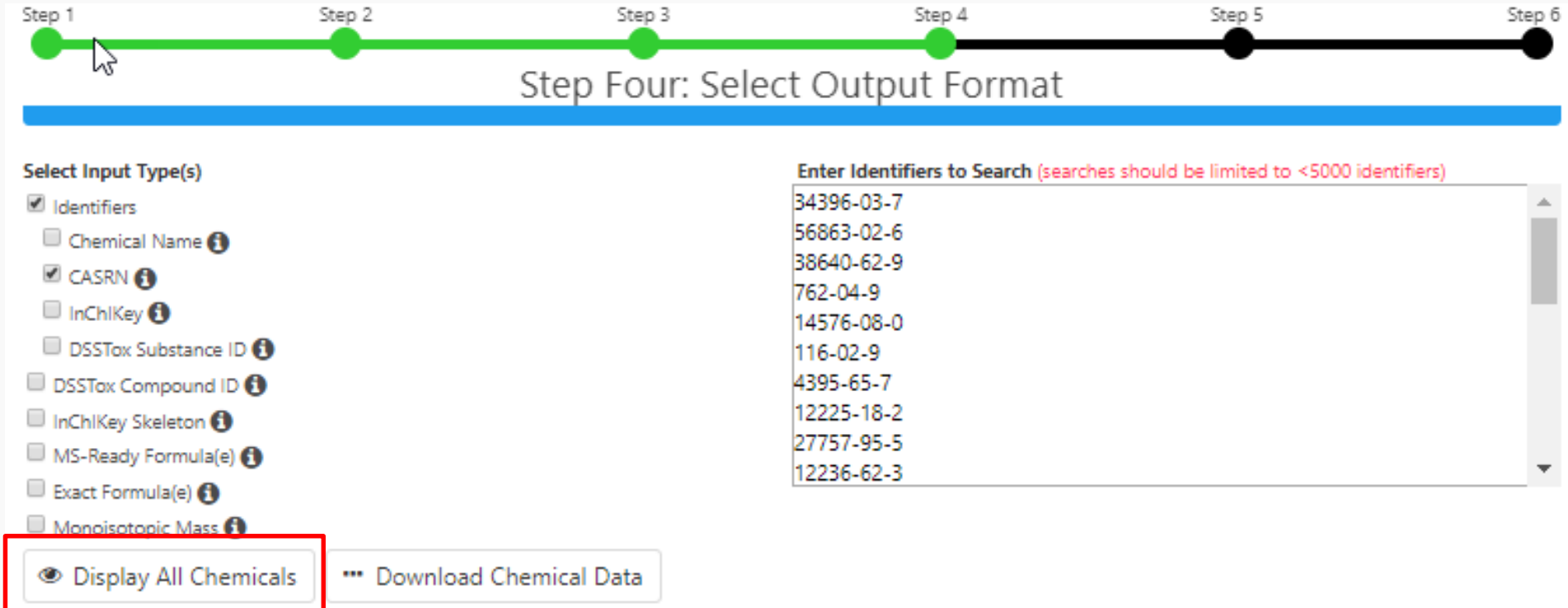
https://comptox.epa.gov/dashboard/dsstoxdb/batch_search



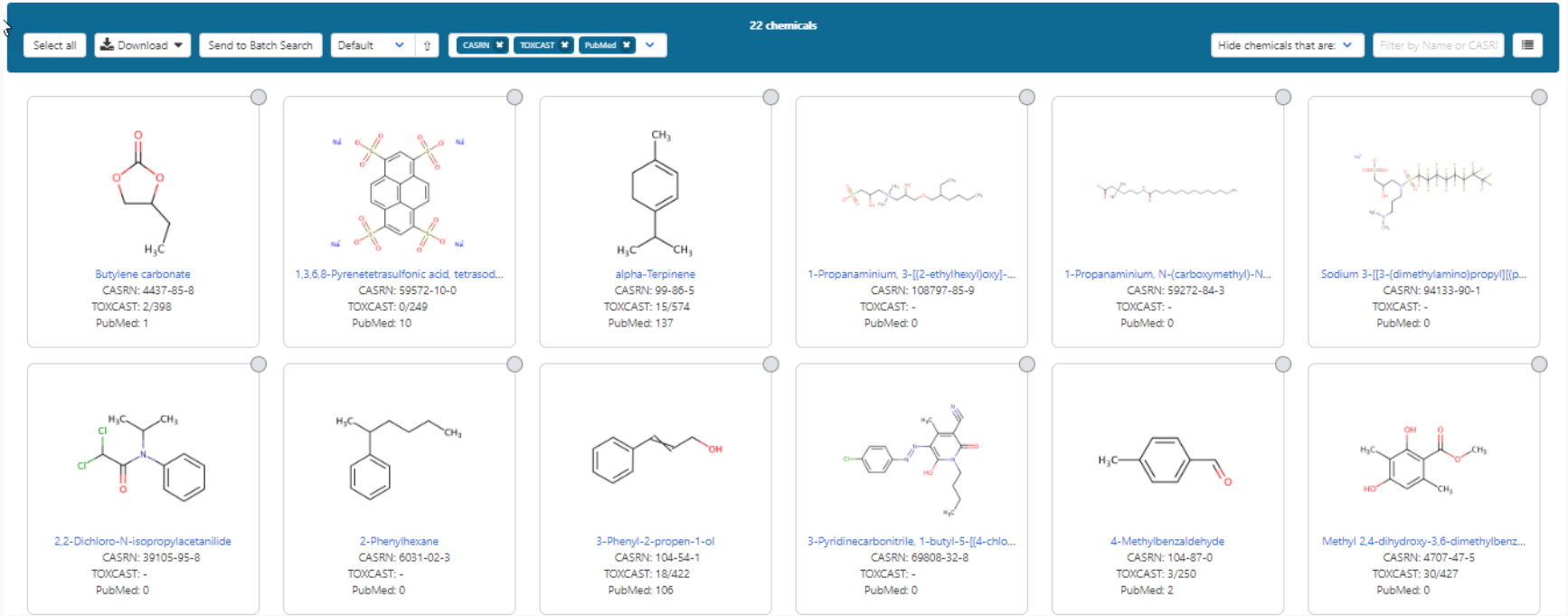
Inputs


Select Input Type(s)
Enter Identifiers to Search (search
Identifiers
Chemical Name
CASRN
InChIKey
DSSTox Substance ID
DSSTox Compound ID
InChIKey Skeleton
MS-Ready Formula(e)
Exact Formula(e)
Monoisotopic Mass
This search is based on what we refer to as "Mass Spec (MS) Ready" structures. All chemicals within the database are treated in a manner such that all are desalted, mixtures are separated, and stereochemistry is removed as Mass Spectrometry detects the major components of a salt or mixture and is insensitive to stereochemistry. As an example, a search for the monoisotopic mass of phenol will return phenol, sodium phenolate and calcium phenoxide. See the publication for more details: https://doi.org/10.1186/s13321-018-0299-2.
Display All Chemicals
Download Chemical Data

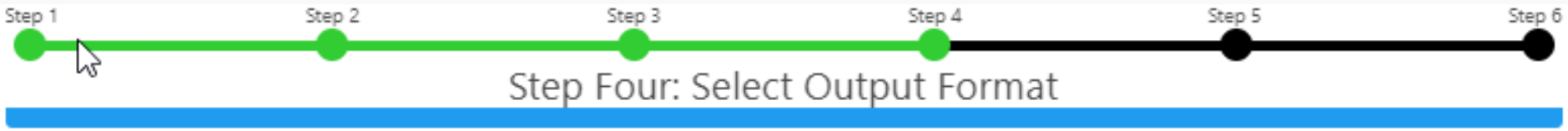
CAS Numbers…


Step 1
Step 2
Step 3
Step 4
Step 5
Step 6
Step Four: Select Output Format
Select Input Type(s)
Identifiers
Chemical Name
CASRN
InChIKey
DSSTox Substance ID
DSSTox Compound ID
InChIKey Skeleton
MS-Ready Formula(e)
Exact Formula(e)
Monoisotopic Mass
Enter Identifiers to Search (searches should be limited to <5000 identifiers)
34396-03-7
56863-02-6
38640-62-9
762-04-9
14576-08-0
116-02-9
4395-65-7
12225-18-2
27757-95-5
12236-62-3
Display All Chemicals
Download Chemical Data

Display Chemicals



CAS Numbers…



Select Input Type(s)

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

Display All Chemicals | Download Chemical Data

Enter Identifiers to Search (searches should be limited to <5000 identifiers)

34396-03-7
56863-02-6
38640-62-9
762-04-9
14576-08-0
116-02-9
4395-65-7
12225-18-2
27757-95-5
12236-62-3

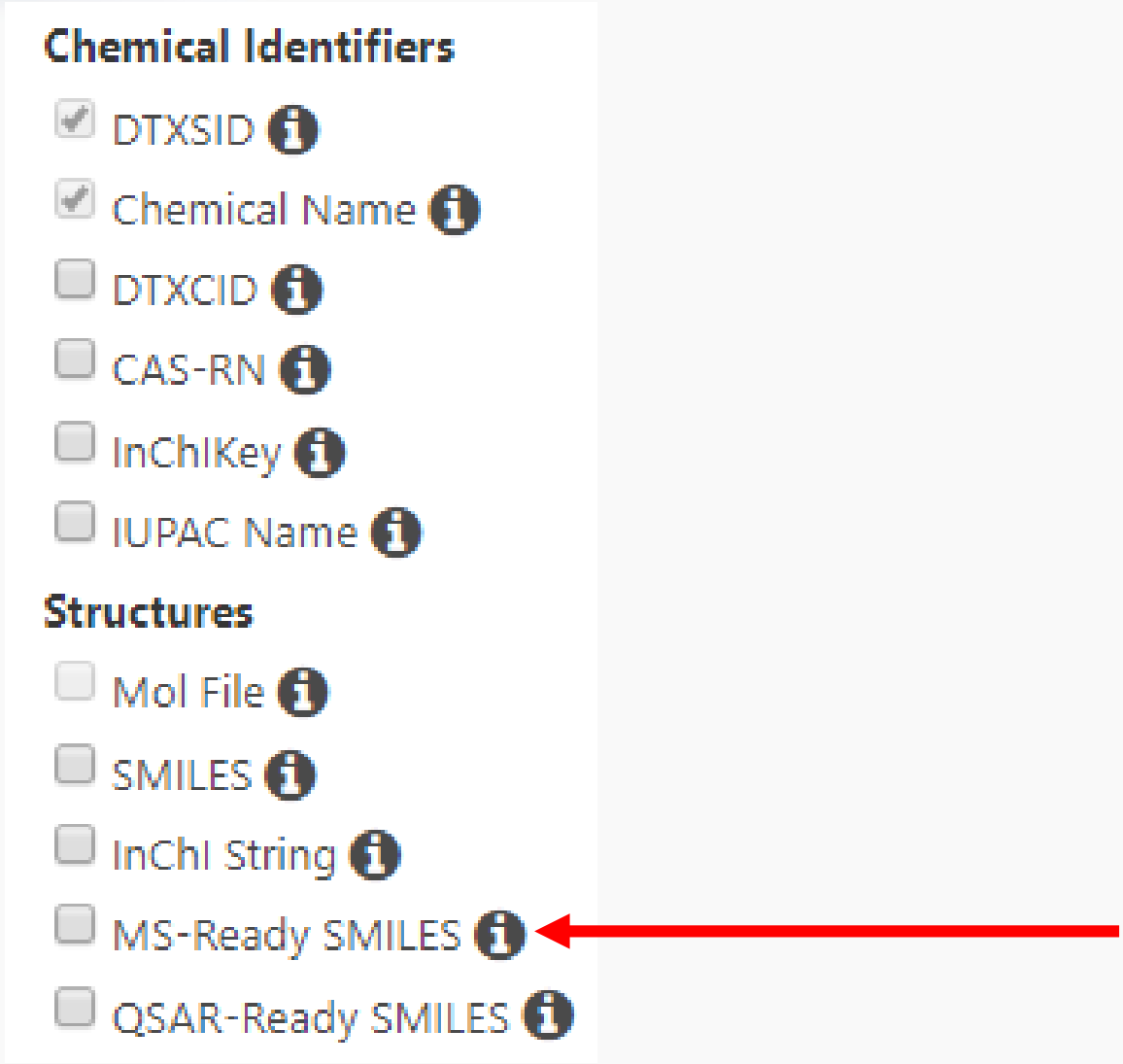
Select Output Format:

Download as...

- TSV
- CSV
- Excel
- SDF

Chemical Identifiers

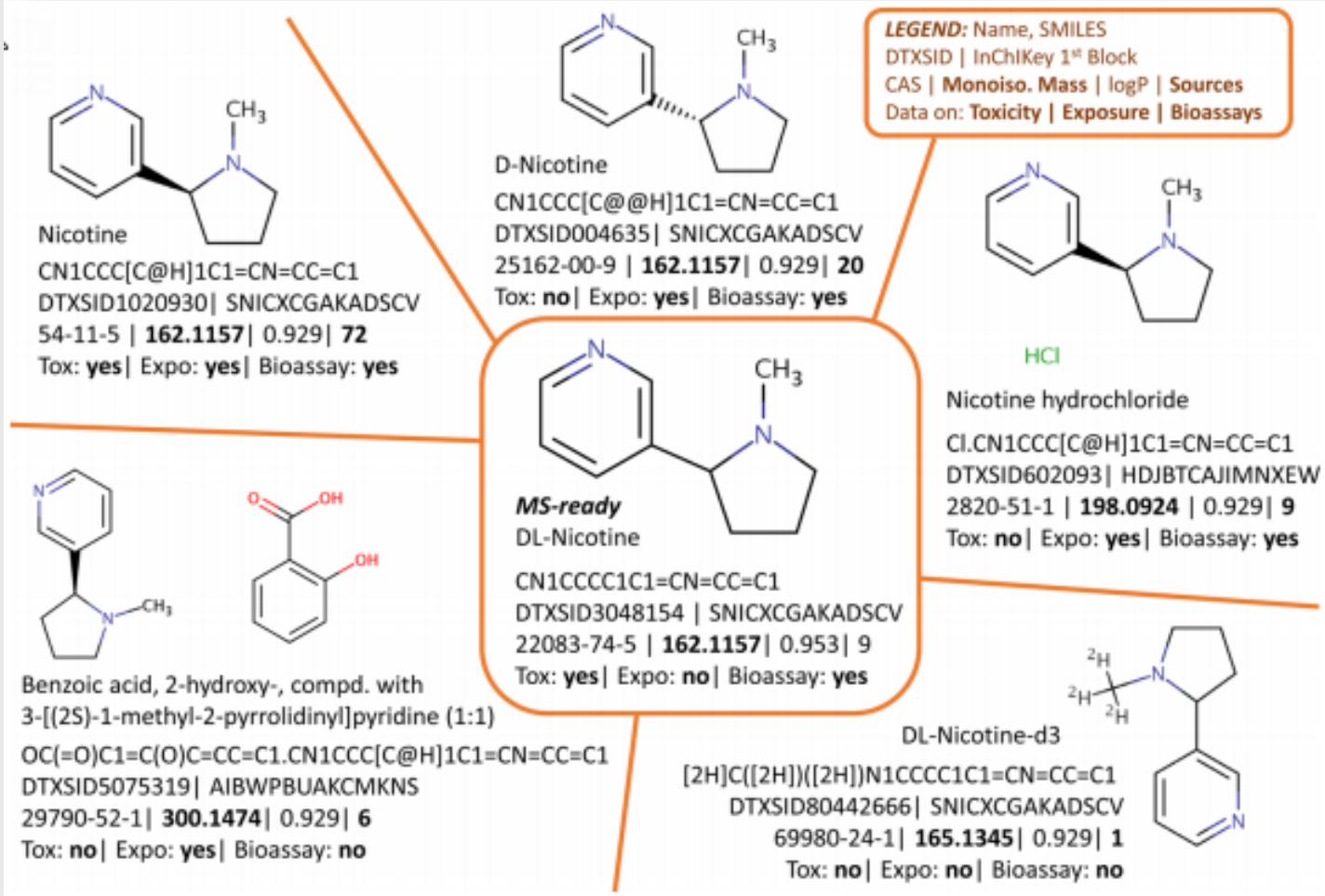
Possible Outputs…



Environmental Science & Technology
Viewpoint
pubs.acs.org/est

Open Science for Identifying "Known Unknown" Chemicals

Emma L. Schymanski*,† and Antony J. Williams*,‡



“MS-Ready Structures”

https://doi.org/10.1186/s13321-018-0299-2

McEachran *et al. J Cheminform (2018) 10:45*
https://doi.org/10.1186/s13321-018-0299-2

Journal of Cheminformatics

METHODOLOGY Open Access

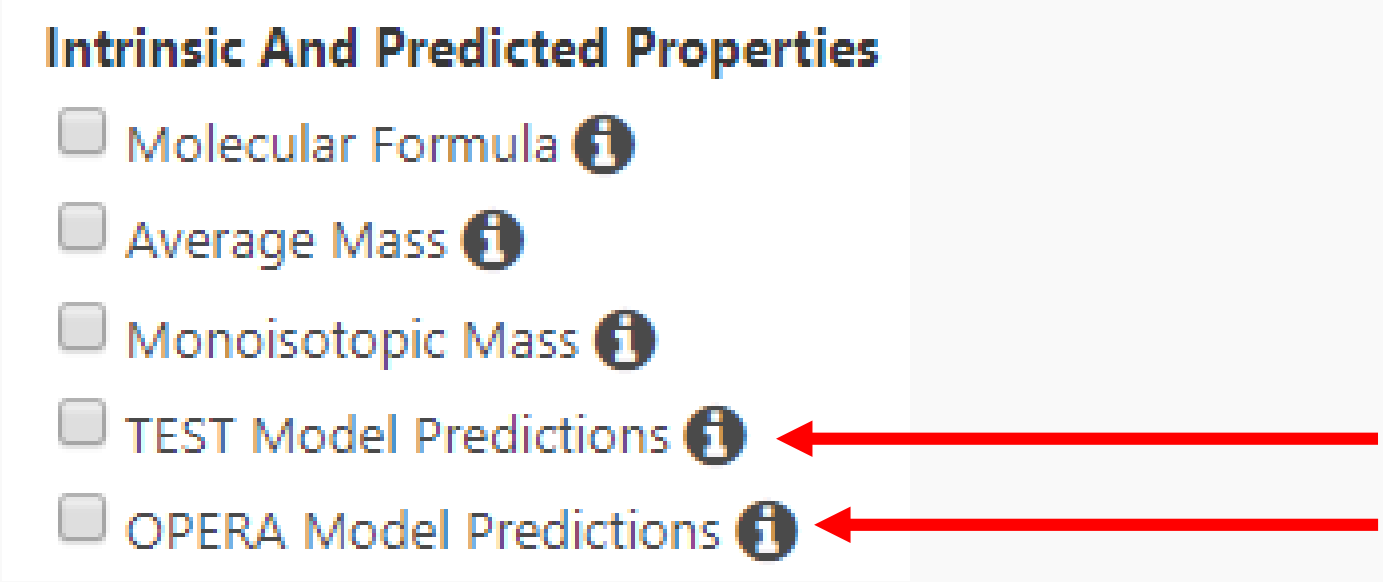
“MS-Ready” structures for non-targeted high-resolution mass spectrometry screening studies

Andrew D. McEachran[1,2*], Kamel Mansouri[1,2,3], Chris Grulke[2], Emma L. Schymanski[4], Christoph Ruttkies[5] and Antony J. Williams[2*]

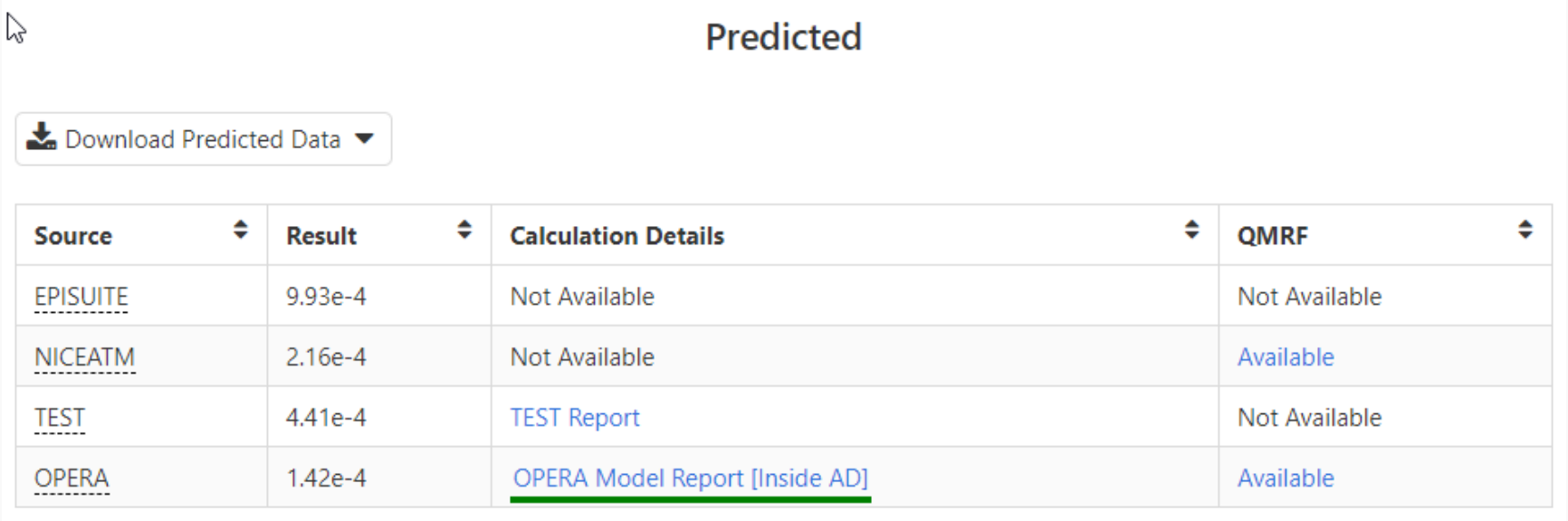
MS-Ready SMILES

INPUT	FOUND_BY	DTXSID	PREFERRED_NAME	MS_READY_SMILES
59572-10-0	CAS-RN	DTXSID90	1,3,6,8-Pyrenetetrasulfonic acid, tetrasodium s	OS(=O)(=O)C1=CC(=C2C=CC3=C(C=C
104-54-1	CAS-RN	DTXSID90	3-Phenyl-2-propen-1-ol	OCC=CC1=CC=CC=C1
4395-65-7	CAS-RN	DTXSID90	C.I. Solvent Blue 68	NC1=C2C(=O)C3=C(C=CC=C3)C(=O)C
12225-18-2	CAS-RN	DTXSID90	C.I. Pigment Yellow 97	COC1=CC(Cl)=C(OC)C=C1NC(=O)C(N=
14576-08-0	CAS-RN	DTXSID90	Cyclohexene, 4-(1-methoxy-1-methylethyl)-1-n	COC(C)(C)C1CCC(C)=CC1
4707-47-5	CAS-RN	DTXSID90	Methyl 2,4-dihydroxy-3,6-dimethylbenzoate	COC(=O)C1=C(O)C(C)=C(O)C=C1C
34396-03-7	CAS-RN	DTXSID90	Silane, trimethoxy(2,4,4-trimethylpentyl)-	CO[Si](CC(C)CC(C)(C)C)(OC)OC
2475-46-9	CAS-RN	DTXSID90	1-(Methylamino)-4-(2-hydroxyethylamino)anthr	CNC1=CC=C(NCCO)C2=C1C(=O)C1=C
94133-90-1	CAS-RN	DTXSID90	Sodium 3-[[3-(dimethylamino)propyl][(perfluoro	CN(C)CCCN(CC(O)CS(O)(=O)=O)S(=O
762-04-9	CAS-RN	DTXSID90	Diethyl phosphite	CCOP(O)OCC

Possible Outputs…



TEST and OPERA Predictions



Predicted

Download Predicted Data

Source	Result	Calculation Details	QMRF
EPISUITE	9.93e-4	Not Available	Not Available
NICEATM	2.16e-4	Not Available	Available
TEST	4.41e-4	TEST Report	Not Available
OPERA	1.42e-4	OPERA Model Report [Inside AD]	Available

Transparency for prediction models

EPA United States Environmental Protection Agency

Predicted

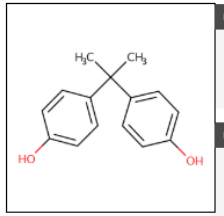
Download Predicted Data

Source	Result	Calculation Details	QMRF
EPISUITE	3.64	Not Available	Not Available
NICEATM	2.40	Not Available	Available
ACD/Labs Consensus	3.63	Not Available	Not Available
ACD/Labs	3.43	Not Available	Not Available
OPERA	3.35	OPERA Model Report [Inside AD]	Available

OPERA Models: LogP: Octanol-Water

Bisphenol A

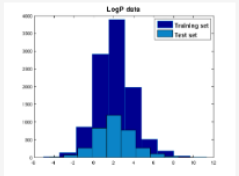
80-05-7 | DTXSID7020182

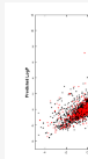


Model Results

Predicted value: 3.35
Global applicability domain: Inside
Local applicability domain index: 0.877
Confidence level: 0.813

Model Performance



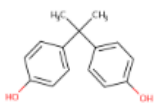


QMRF

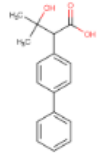
Weighted KNN model

5-fold CV (75%)		Training (75%)	
Q2	RMSE	R2	RMSE
0.850	0.690	0.860	0.670

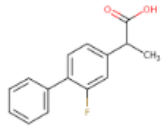
Nearest Neighbors from the Training Set



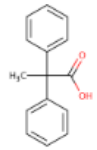
Bisphenol A
Measured:3.32
Predicted: 3.35076



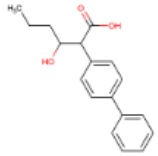
BUTANOIC ACID,2-(4-BIPHENYLYL)-3-HYDROXY-3-METHY
Measured:3.25
Predicted: 3.39062



Flurbiprofen
Measured:4.16
Predicted: 3.94445

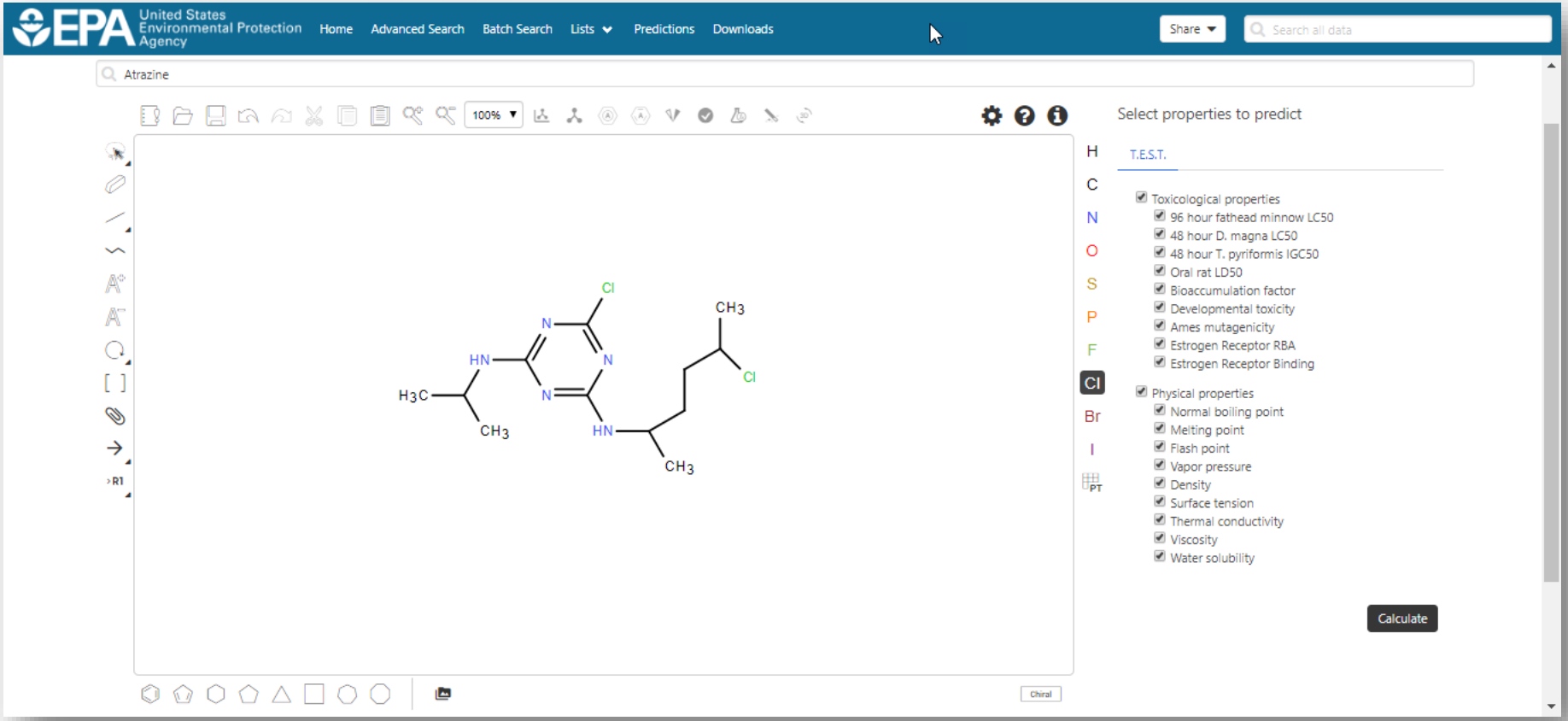


2,2-Diphenylpropionic acid
Measured:2.69
Predicted: 2.84603



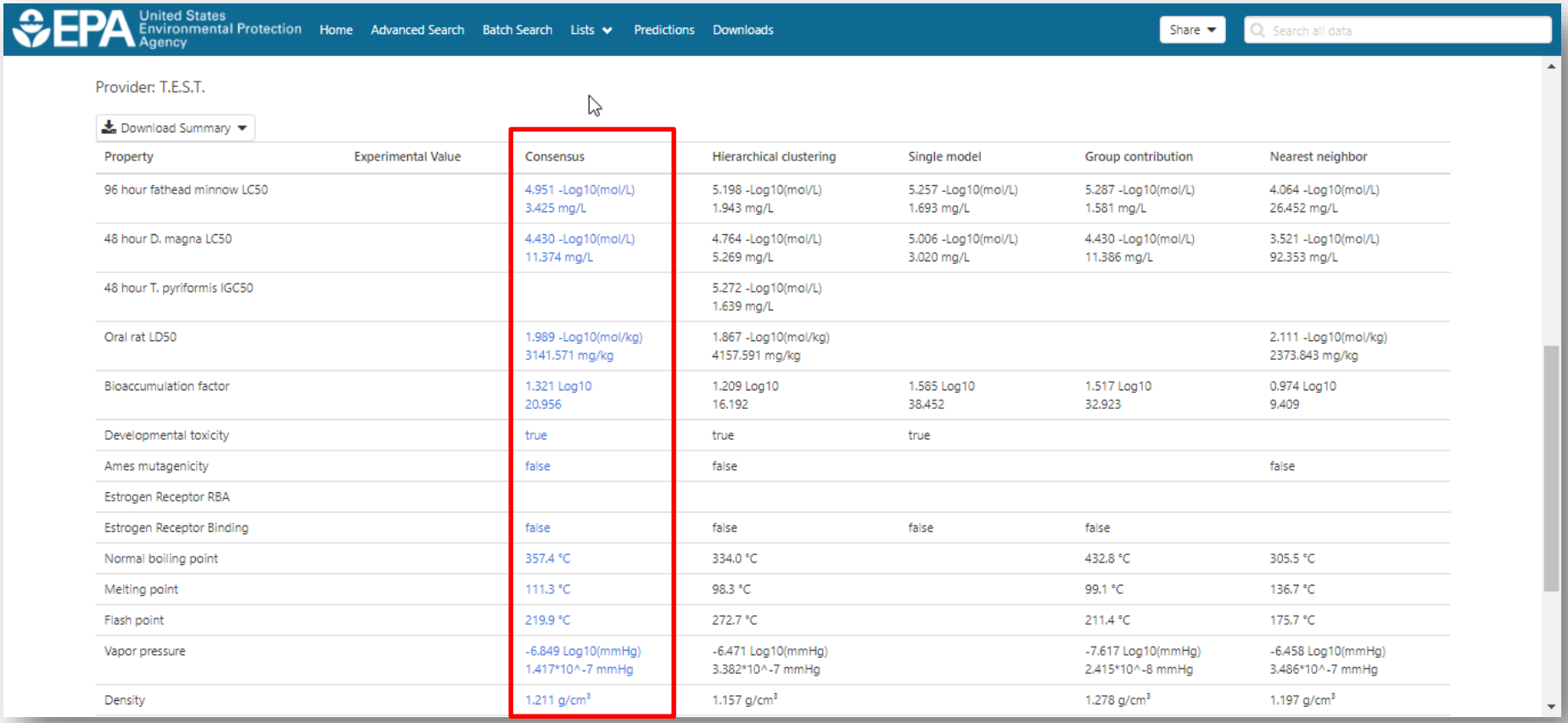
3-OH-2-(4-BIPHENYLYL)HEXANOIC ACID
Measured:3.75
Predicted: 3.70322

TEST Real-Time Predictions



TEST Real-Time Predictions with detailed calculation reports

EPA United States Environmental Protection Agency



Property	Experimental Value	Consensus	Hierarchical clustering	Single model	Group contribution	Nearest neighbor
96 hour fathead minnow LC50		4.951 -Log10(mol/L) 3.425 mg/L	5.198 -Log10(mol/L) 1.943 mg/L	5.257 -Log10(mol/L) 1.693 mg/L	5.287 -Log10(mol/L) 1.581 mg/L	4.064 -Log10(mol/L) 26.452 mg/L
48 hour D. magna LC50		4.430 -Log10(mol/L) 11.374 mg/L	4.764 -Log10(mol/L) 5.269 mg/L	5.006 -Log10(mol/L) 3.020 mg/L	4.430 -Log10(mol/L) 11.386 mg/L	3.521 -Log10(mol/L) 92.353 mg/L
48 hour T. pyriformis IGC50			5.272 -Log10(mol/L) 1.639 mg/L			
Oral rat LD50		1.989 -Log10(mol/kg) 3141.571 mg/kg	1.867 -Log10(mol/kg) 4157.591 mg/kg			2.111 -Log10(mol/kg) 2373.843 mg/kg
Bioaccumulation factor		1.321 Log10 20.956	1.209 Log10 16.192	1.585 Log10 38.452	1.517 Log10 32.923	0.974 Log10 9.409
Developmental toxicity		true	true	true		
Ames mutagenicity		false	false			false
Estrogen Receptor RBA						
Estrogen Receptor Binding		false	false	false	false	
Normal boiling point		357.4 °C	334.0 °C		432.8 °C	305.5 °C
Melting point		111.3 °C	98.3 °C		99.1 °C	136.7 °C
Flash point		219.9 °C	272.7 °C		211.4 °C	175.7 °C
Vapor pressure		-6.849 Log10(mmHg) 1.417*10^-7 mmHg	-6.471 Log10(mmHg) 3.382*10^-7 mmHg		-7.617 Log10(mmHg) 2.415*10^-8 mmHg	-6.458 Log10(mmHg) 3.486*10^-7 mmHg
Density		1.211 g/cm³	1.157 g/cm³		1.278 g/cm³	1.197 g/cm³

TEST Real-Time Predictions with detailed calculation reports

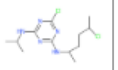
Predicted Vapor pressure at 25°C for ClC=1N=C(N=C(N1)NC(C)CCC(Cl)C)NC(C)C from Consensus method

Prediction results

Endpoint	Experimental value	Predicted value
Vapor pressure at 25°C Log10(mmHg)	N/A	-6.85
Vapor pressure at 25°C mmHg	N/A	1.42E-07

Individual Predictions

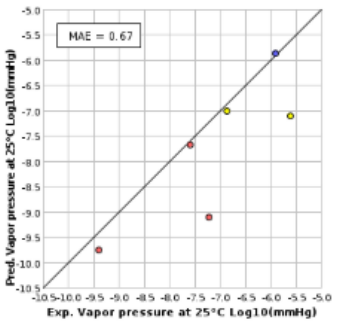
Method	Predicted value Log10(mmHg)
Hierarchical clustering	-6.47
Group contribution	-7.62
Nearest neighbor	-6.46



Predictions for the test chemical and for the most similar chemicals in the external tes

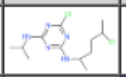
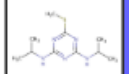
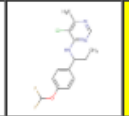
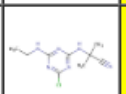
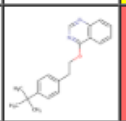
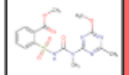
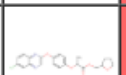
If the predicted value matches the experimental values for similar chemicals in the test set (and the similar che

Prediction results (colors defined in table below)



Chemicals	MAE*
Entire set	0.47
Similarity coefficient ≥ 0.5	0.67

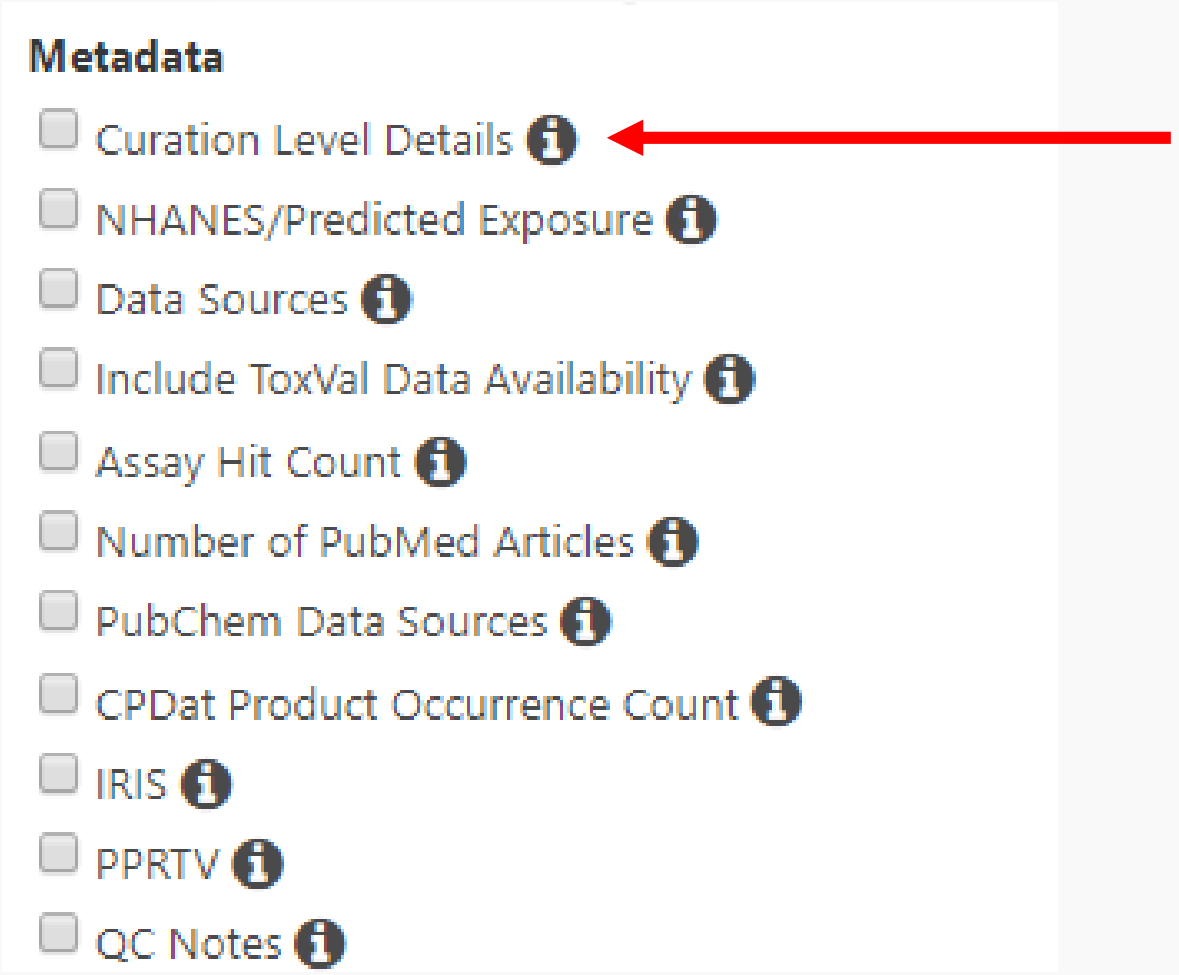
*Mean absolute error in Log10(mmHg)

CAS	Structure	Similarity Coefficient	Experimental value Log10(mmHg)	Predicted value Log10(mmHg)
ClC=1N=C(N=C(N1)NC(C)CCC(Cl)C)NC(C)C (test chemical)			N/A	-6.85
7287-19-6		0.83	-5.91	-5.86
130339-07-0		0.77	-5.62	-7.11
21725-46-2		0.76	-6.86	-7.01
120928-09-8		0.58	-7.59	-7.67
101200-48-0		0.56	-9.41	-9.76
119738-06-6		0.55	-7.23	-9.11

TEST/OPERA Batch Download

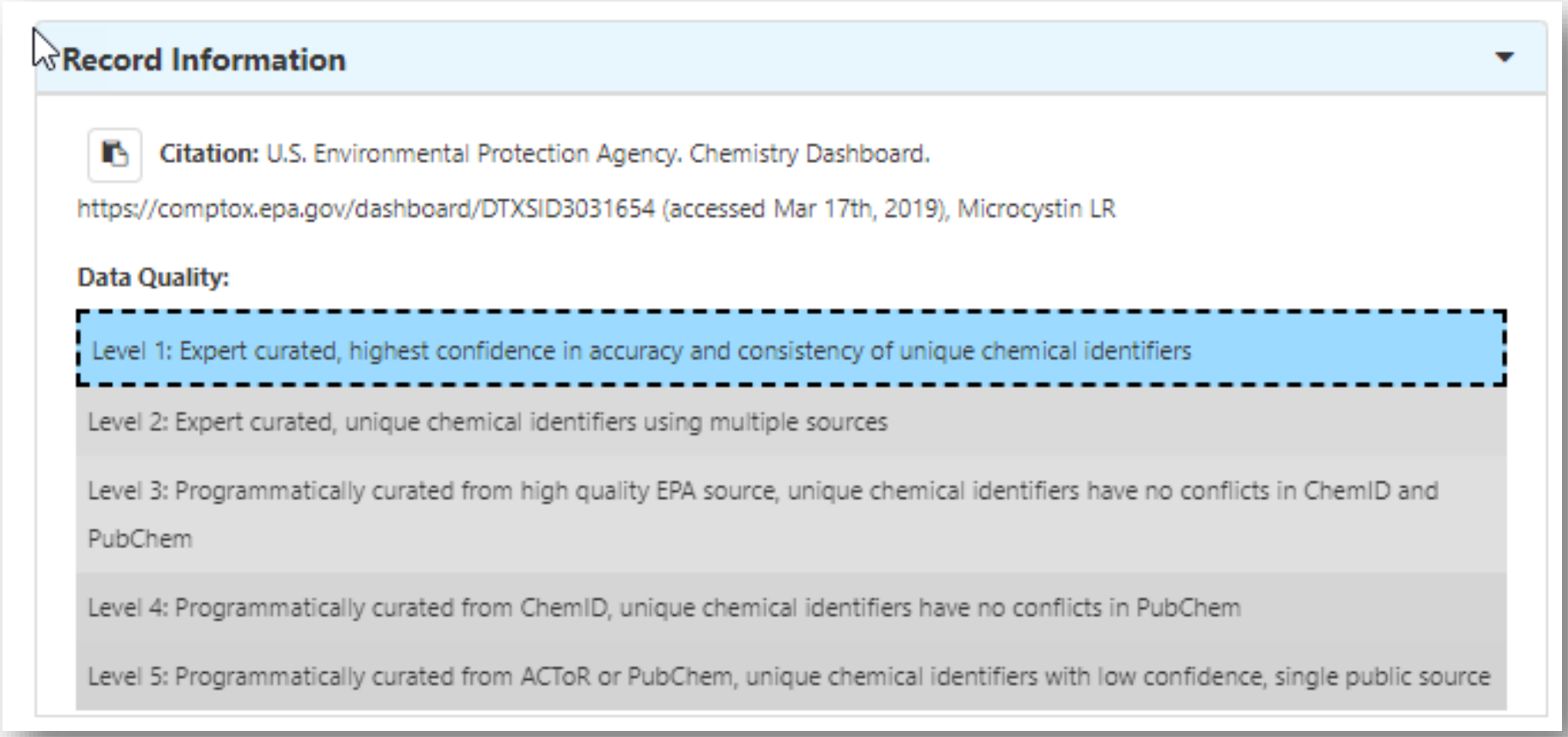
C	D	E	F	G	H	I	J	K	L	
DTXSID	PREFERRE	BIOCONCE	BOILING_F	48HR_DAP	DENSITY_(	DEVTOX_T	96HR_FATI	FLASH_PC	MELTING_I	A
DTXSID90	Methyl 2,4-(	4.30527	294.167	2.259E-05	1.253	0.673	9.29E-05	136.111	96.637	
DTXSID90	3-Phenyl-2-	7.2277	254.617	0.0003532	1.059	0.795	6.653E-05	124.886	57.333	
DTXSID90	alpha-Terpir	584.62	174.7	9.772E-06	0.86	0.78	1.462E-05	49.96	-39.21	
DTXSID90	3,3,5-Trimet	20.2302	197.088	0.0005224	0.885	0.974	0.0003819	76.508	21.347	
DTXSID90	Diethyl phos	0.796159	142.81	-	1.028	0.941	-	58.069	-42.817	
DTXSID90	4-Methylber	-	221.904	0.0001233	1.066	0.517	8.65E-05	87.033	23.053	
DTXSID90	Butylene ca	-	167.924	-	1.114	0.447	-	96.795	0.641	-
DTXSID90	Silane, trim	183.231	187.37	-	0.86	0.343	-	72.907	-60.671	
DTXSID90	Linoleic diet	5.87489	430.153	2.443E-05	0.96	0.454	4.875E-06	255.997	40.201	
DTXSID90	Cyclohexen	46.3447	197.405	4.613E-05	0.892	0.74	0.0001194	73.597	-10.519	
DTXSID90	C.I. Solvent	34.914	493.758	2.944E-06	1.36	0.87	1.104E-06	262.584	211.867	
DTXSID90	C.I. Pigmen	3.02691	-	-	1.43	1.074	-	423.613	249.017	
DTXSID90	C.I. Pigmen	3.2961	-	-	1.686	0.778	-	-	275.574	
DTXSID90	1-(Methylan	14.8936	434.621	5.129E-05	1.38	0.859	9.12E-06	292.795	212.665	
DTXSID90	Isoamyl sal	21.6272	285.877	3.681E-05	1.105	0.744	1.6E-05	117.032	44.513	
DTXSID90	3-Pyridinec	-	467.613	3.451E-08	1.267	0.983	-	317.894	182.082	

Possible Outputs…

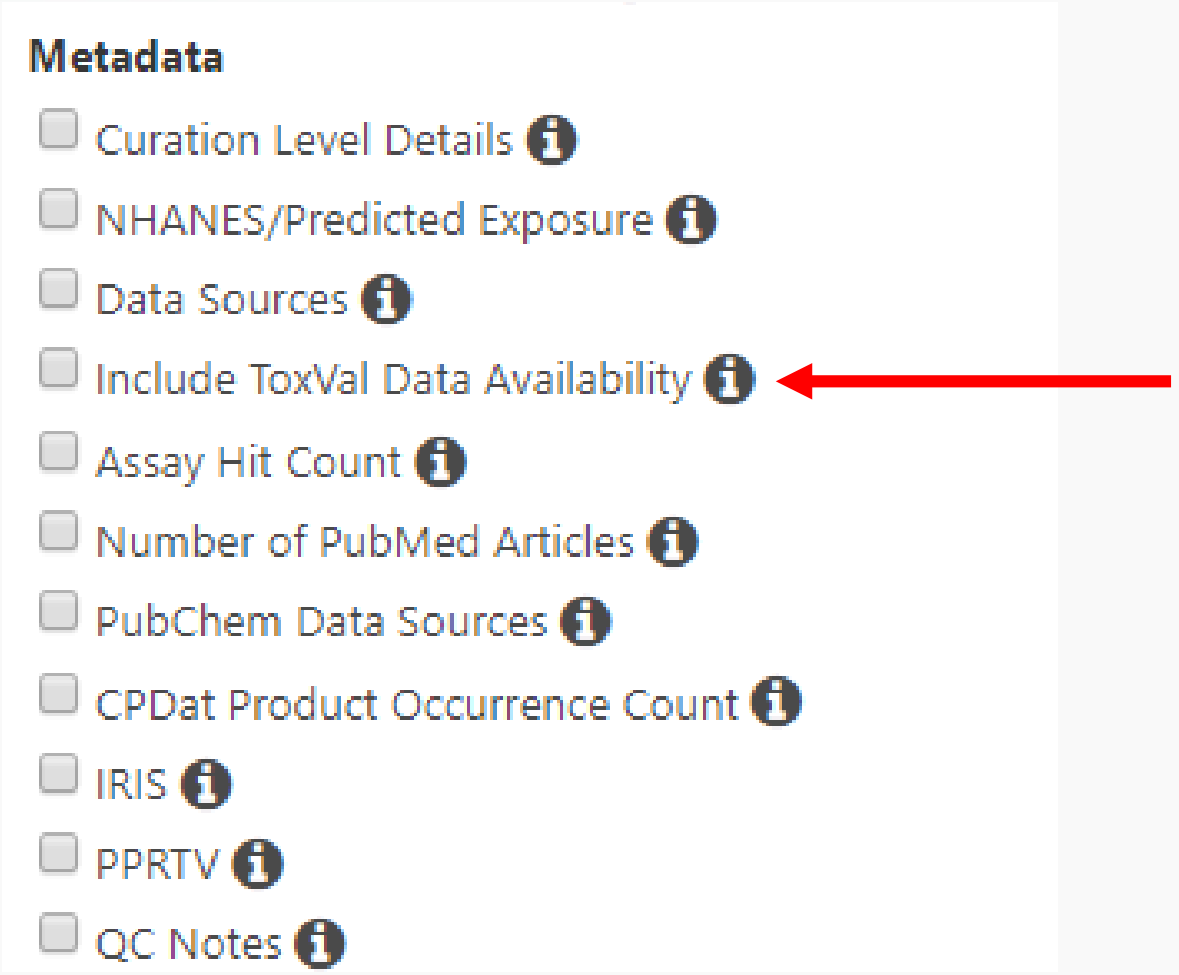


Curation Level Details

- We have full time curators checking data



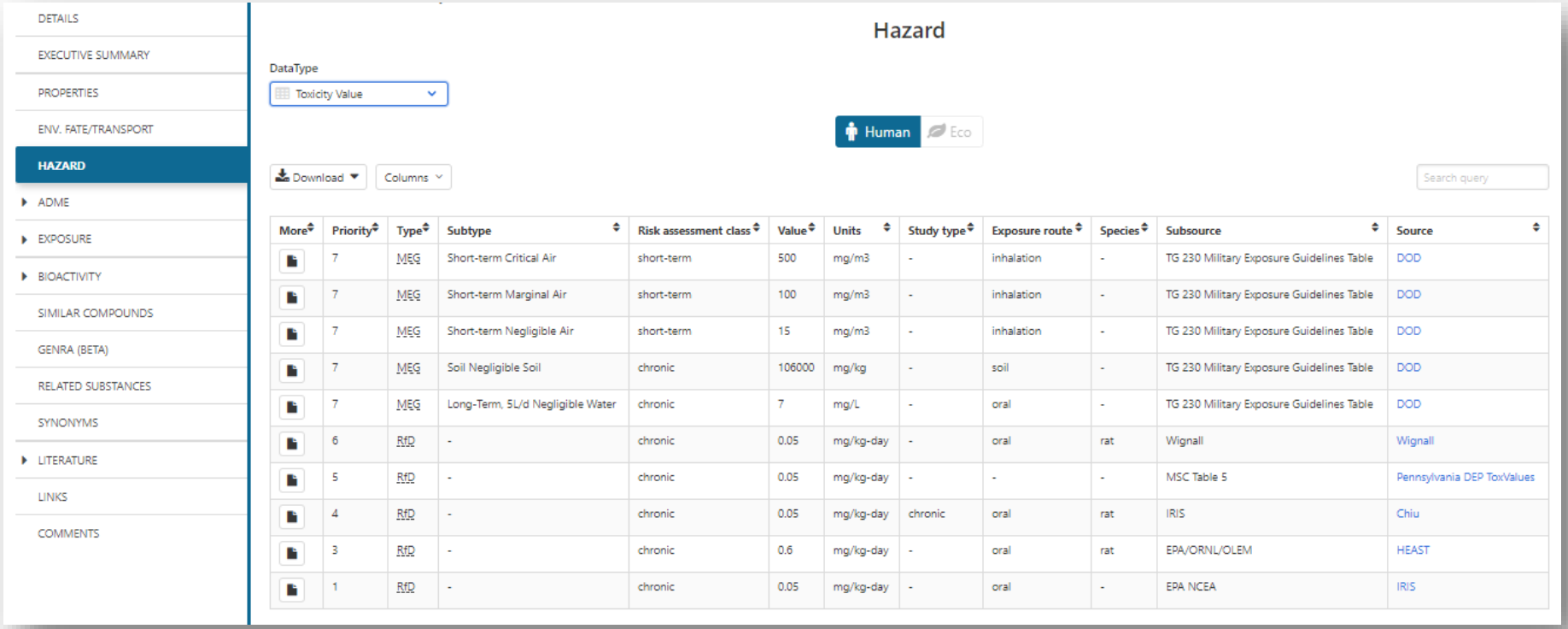
Possible Outputs…



Hazard Data from “ToxVal_DB”

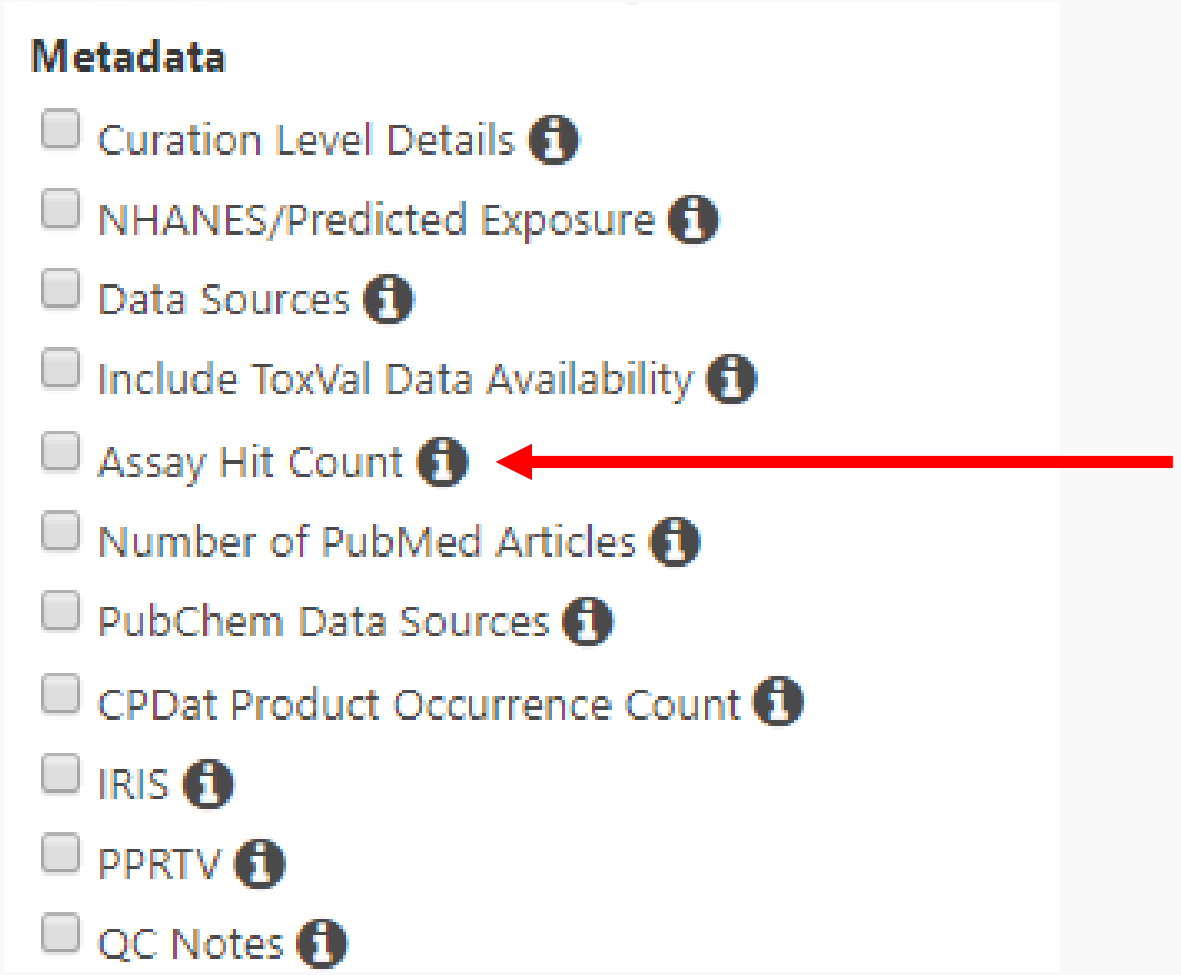
- ToxVal Database contains following data:
  - 30,050 chemicals
  - 772,721 toxicity values
  - 29 sources of data
  - 21,507 sub-sources
  - 4585 journals cited
  - 69,833 literature citations

Access to Chemical Hazard Data



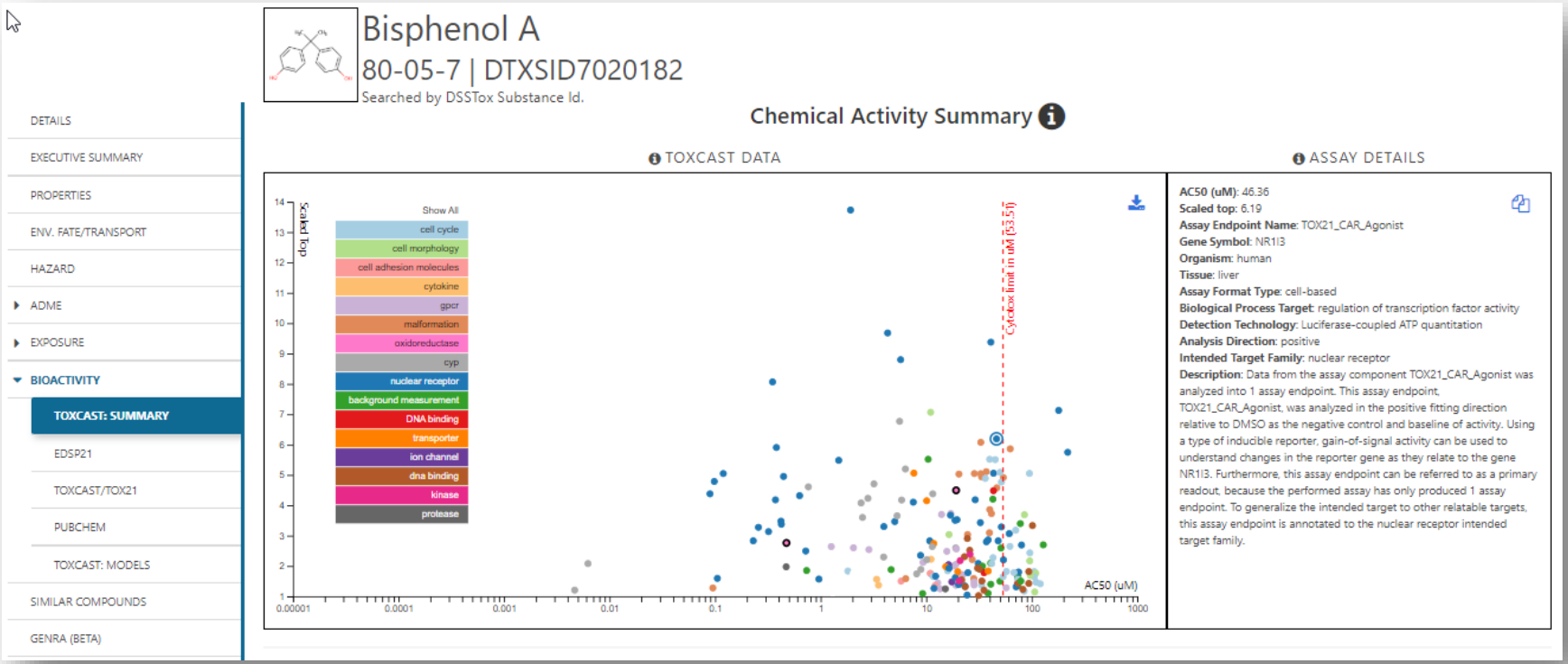
More	Priority	Type	Subtype	Risk assessment class	Value	Units	Study type	Exposure route	Species	Subsource	Source
	7	MEG	Short-term Critical Air	short-term	500	mg/m3	-	inhalation	-	TG 230 Military Exposure Guidelines Table	DOD
	7	MEG	Short-term Marginal Air	short-term	100	mg/m3	-	inhalation	-	TG 230 Military Exposure Guidelines Table	DOD
	7	MEG	Short-term Negligible Air	short-term	15	mg/m3	-	inhalation	-	TG 230 Military Exposure Guidelines Table	DOD
	7	MEG	Soil Negligible Soil	chronic	106000	mg/kg	-	soil	-	TG 230 Military Exposure Guidelines Table	DOD
	7	MEG	Long-Term, 5L/d Negligible Water	chronic	7	mg/L	-	oral	-	TG 230 Military Exposure Guidelines Table	DOD
	6	RfD	-	chronic	0.05	mg/kg-day	-	oral	rat	Wignall	Wignall
	5	RfD	-	chronic	0.05	mg/kg-day	-	-	-	MSC Table 5	Pennsylvania DEP ToxValues
	4	RfD	-	chronic	0.05	mg/kg-day	chronic	oral	rat	IRIS	Chiu
	3	RfD	-	chronic	0.6	mg/kg-day	-	oral	rat	EPA/ORNL/OLEM	HEAST
	1	RfD	-	chronic	0.05	mg/kg-day	-	oral	-	EPA NCEA	IRIS

Possible Outputs…

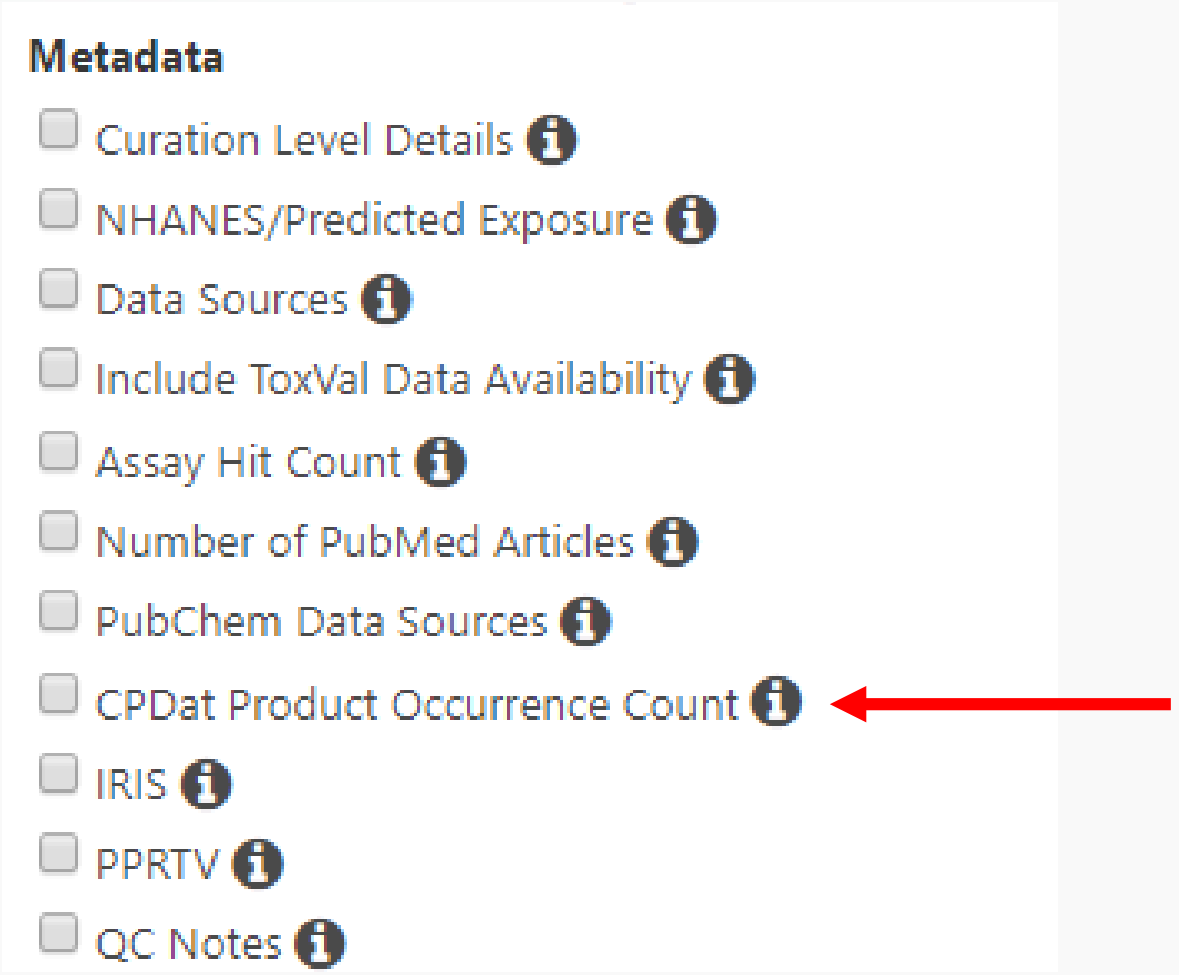


In Vitro Bioassay Screening

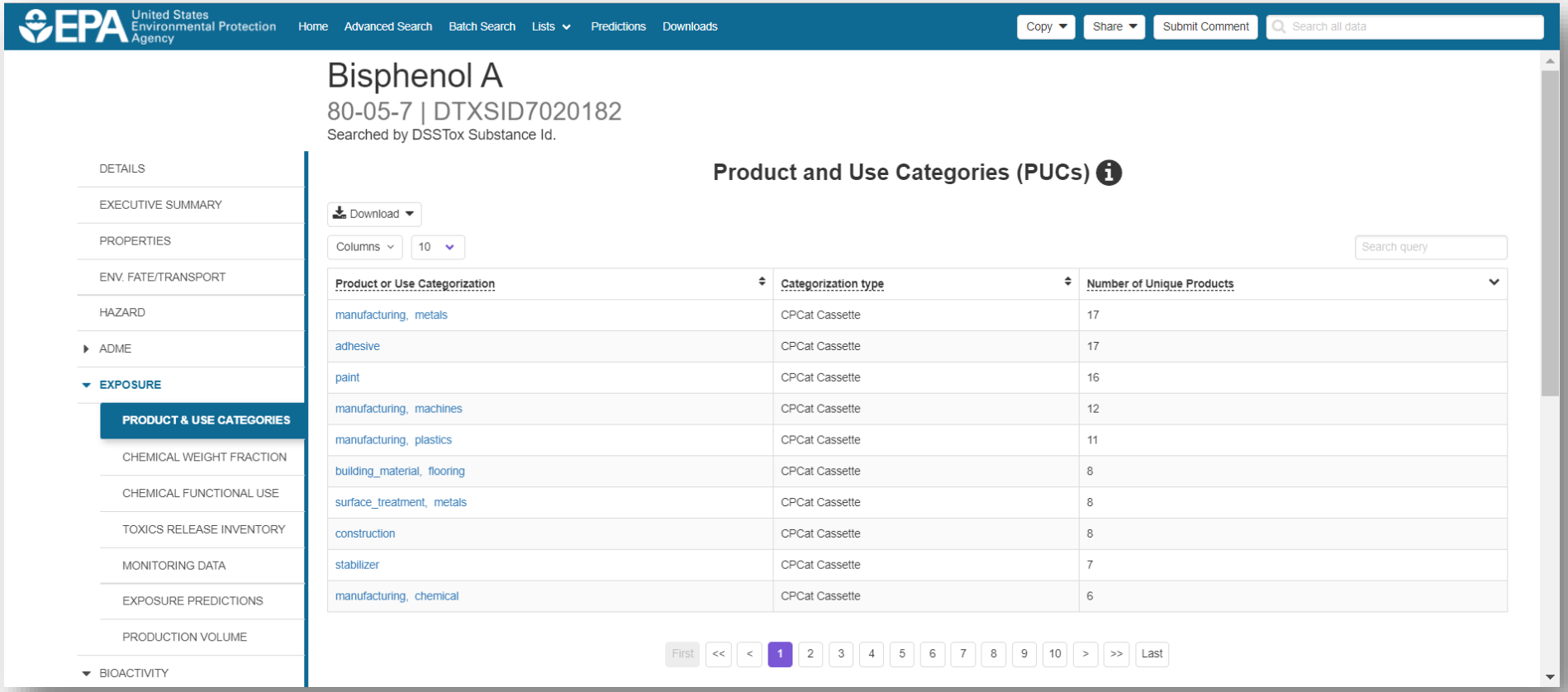
ToxCast and Tox21



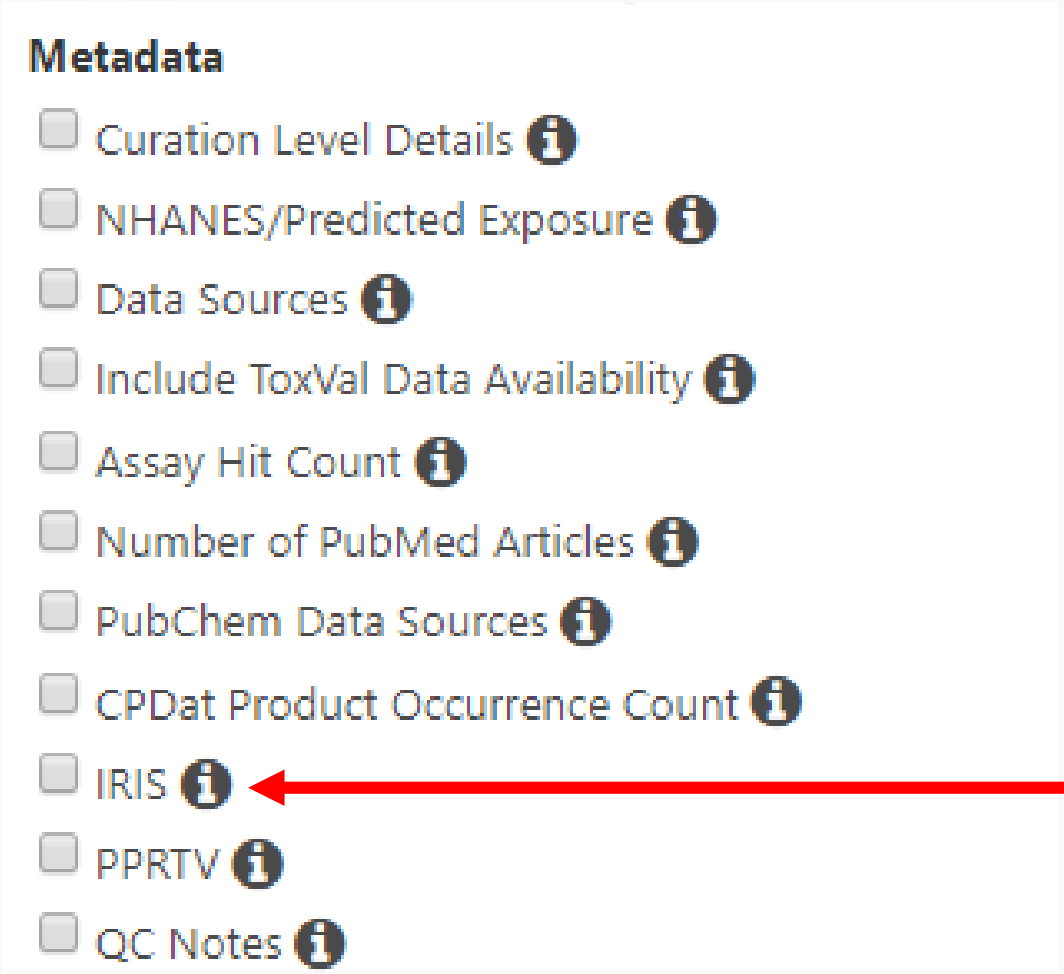
Possible Outputs…



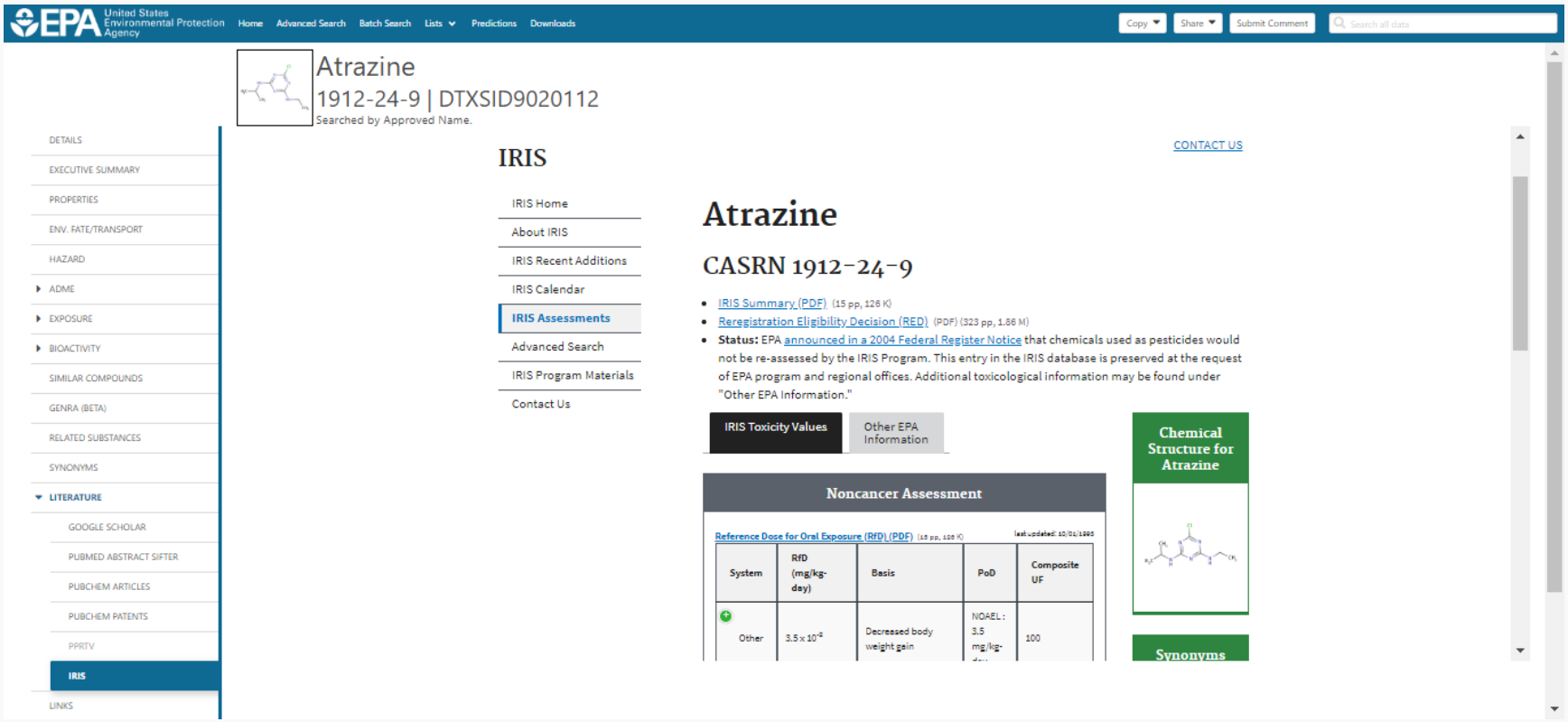
Sources of Exposure to Chemicals



Possible Outputs…



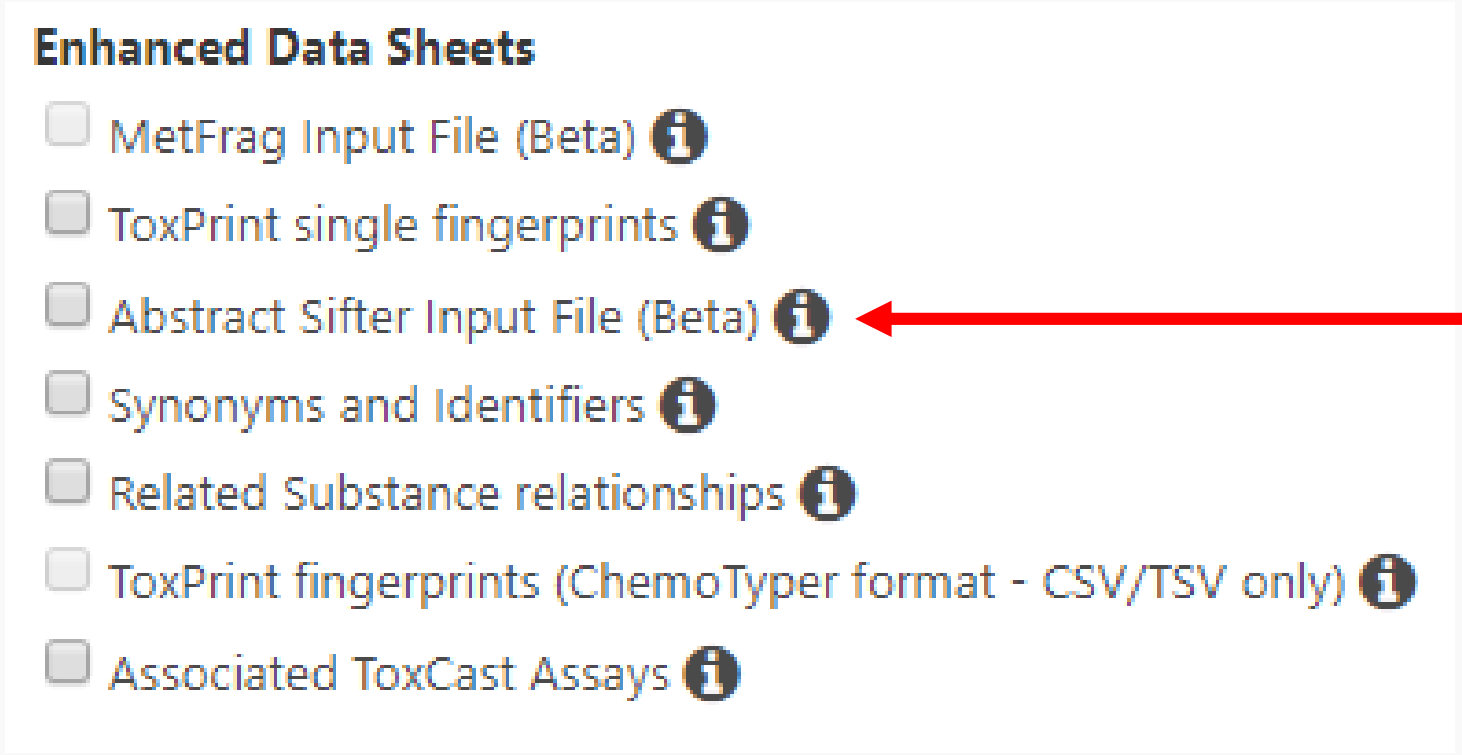
IRIS Report


EPA United States Environmental Protection Agency
Home
Advanced Search
Batch Search
Lists
Predictions
Downloads
Copy
Share
Submit Comment
Atrazine
1912-24-9 | DTXSID9020112
Searched by Approved Name.
DETAILS
EXECUTIVE SUMMARY
PROPERTIES
ENV. FATE/TRANSPORT
HAZARD
ADME
EXPOSURE
BIOACTIVITY
SIMILAR COMPOUNDS
GENRA (BETA)
RELATED SUBSTANCES
SYNONYMS
LITERATURE
GOOGLE SCHOLAR
PUBMED ABSTRACT SIFTER
PUBCHEM ARTICLES
PUBCHEM PATENTS
PPRTV
IRIS
LINKS
IRIS
IRIS Home
About IRIS
IRIS Recent Additions
IRIS Calendar
IRIS Assessments
Advanced Search
IRIS Program Materials
Contact Us
CONTACT US
Atrazine
CASRN 1912-24-9
IRIS Summary (PDF) (15 pp, 126 K)
Reregistration Eligibility Decision (RED) (PDF) (323 pp, 1.86 M)
Status: EPA announced in a 2004 Federal Register Notice that chemicals used as pesticides would not be re-assessed by the IRIS Program. This entry in the IRIS database is preserved at the request of EPA program and regional offices. Additional toxicological information may be found under "Other EPA Information."
IRIS Toxicity Values
Other EPA Information
Noncancer Assessment
Reference Dose for Oral Exposure (RfD) (PDF) (15 pp, 126 K)
last updated: 10/01/1995
System
RfD (mg/kg-day)
Basis
PoD
Composite UF
Other
3.5 x 10-2
Decreased body weight gain
NOAEL : 3.5 mg/kg-
100
Chemical Structure for Atrazine
Synonyms

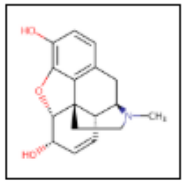
Integrated Data in Excel Sheet

DTXSID	PREFERRED_NAME	QC_LEVEL	TOXVAL_D	TOXCAST_	TOXCAST_	#PUBMED_ART	IRIS_LINK	QC_NOTES
DTXSID2026494	Dichlorodiphenylsilane	DSSTox_Low	Y	4.29	9/210	3	-	-
DTXSID5026621	Trichlorophenylsilane	DSSTox_Low	Y	-	-	-	-	-
DTXSID6026797	1,3,5-Trimethylbenzene	DSSTox_Low	Y	1.07	7/657	53	Y	-
DTXSID7026811	Benzenethiol	DSSTox_Low	Y	9.2	37/402	201	-	-
DTXSID6026872	Butyl isocyanate	DSSTox_Low	Y	0.95	2/211	21	-	-
DTXSID7026940	Ethyl silicon trichloride	DSSTox_Low	-	-	-	-	-	-
DTXSID3027033	Silane, dichloroethenylmethyl	DSSTox_Low	Y	-	-	1	-	-
DTXSID2027204	Methyl isothiocyanate	DSSTox_Low	Y	5.85	46/786	60	-	-
DTXSID5027273	Carbon monoxide	DSSTox_Low	Y	-	-	18454	-	-
DTXSID9027289	Diketene	DSSTox_Low	Y	-	-	22	-	-
DTXSID3027291	Tetramethyl orthosilicate	DSSTox_Low	Y	3.79	8/211	53	-	-
DTXSID0027355	Chlorodimethylsilane	DSSTox_Low	-	-	-	-	-	-
DTXSID3027407	Ethyl dichlorophosphate	DSSTox_Low	Y	2.93	11/376	-	-	-
DTXSID0027482	Trimethoxysilane	DSSTox_Low	Y	-	-	58	-	-
DTXSID4027529	2,2-Dimethylpropanoyl chloric	DSSTox_Low	Y	0.95	2/211	7	-	-
DTXSID7027627	Propylene dinitrate	DSSTox_Low	Y	-	-	14	-	-
DTXSID6027684	Kerosine	DSSTox_High	Y	-	-	712	-	A complex combi
DTXSID3027746	Pentacarbonyl iron	DSSTox_High	Y	-	-	155	-	-
DTXSID5029269	Trichloro(octyl)silane	DSSTox_Low	Y	-	-	7	-	-
DTXSID4029301	Trichlorosilane	DSSTox_Low	Y	-	-	46	-	-
DTXSID6029672	Sulfur dioxide	DSSTox_Low	Y	-	-	4646	-	-
DTXSID5029685	Nitric acid	DSSTox_Low	Y	-	-	1193	-	-

Possible Outputs…



Literature Searching



Morphine

57-27-2 | DTXSID9023336

Searched by Approved Name.

Abstract Sifter

1) Select PubMed starting point query then 2) click on Retrieve.

Select a Query Term

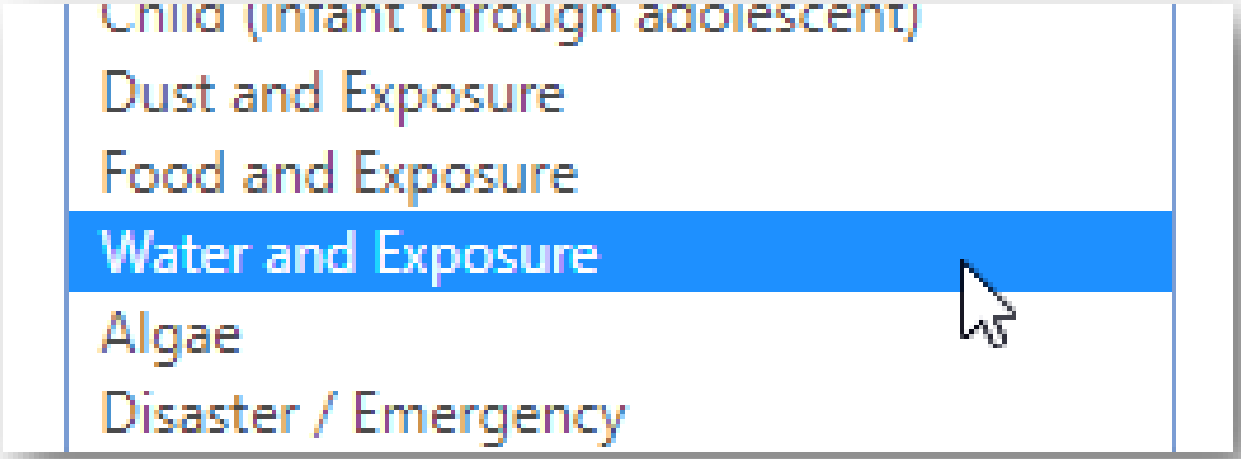
Retrieve Articles

- Select a Query Term
- Hazard
- Fate and Transport
- Metabolism/PK/PD
- Chemical Properties
- Exposure
- Mixtures
- Male Reproduction
- Androgen Disruption
- Female Reproduction
- GeneTox
- Cancer
- Clinical Trials
- Embryo and embryonic development
- Child (infant through adolescent)
- Dust and Exposure
- Food and Exposure
- Water and Exposure
- Algae
- Disaster / Emergency

Optionally, edit the query before retrieving.

"57-27-2" OR "Morphine"

Literature Searching



Optionally, edit the query before retrieving.

("57-27-2" OR "Morphine") AND ((water OR groundwater OR drinking water) AND Environmental Exposure)

Literature Searching

To find articles quickly, enter terms to sift abstracts.

wastewater | Spectrometry | EPA | Clear Terms | Download / Send to... | Download Sifter for Excel

	wastewater	Spectrometry	EPA	Total	PMID	Year	Title	Authors	Journal	Rev
	4	2	0	6	29274731	2017	Simultaneous analysis of opioid analgesics and thei...	Krizman-Matasic; Kostanjevecki; Ahel; Terzic	Journal of chromatography. A	
	0	1	0	1	25768972	2015	Evaluating external contamination of polybrominate...	Poon; Aleksa; Carnevale; Kapur; Goodyer; Koren	Therapeutic drug monitoring	
	0	1	0	1	22544551	2012	Spatial distribution of illicit drugs in surface waters o...	Vazquez-Roig; Andreu; Blasco; Morillas; Picó	Environmental science and pollution research inter...	
	1	1	0	2	20801487	2010	Analysis of llicit and illicit drugs in waste, surface an...	Berset; Brenneisen; Mathieu	Chemosphere	
	1	1	0	2	17935751	2007	Illicit drugs, a novel group of environmental contami...	Zuccato; Castiglioni; Bagnati; Chiabrando; Grassi; ...	Water research	
	2	1	1	4	17607391	2007	Using environmental analytical data to estimate lev...	Bones; Thomas; Paull	Journal of environmental monitoring : JEM	
	3	1	2	6	17180984	2006	Simultaneous determination of psychoactive drugs ...	Hummel; Löffler; Fink; Ternes	Environmental science & technology	
	6	0	0	6	30583189	2018	Assessment of drugs of abuse in a wastewater trea...	Kumar; Tscharke; O'Brien; Mueller; Wilkins; Padhye	The Science of the total environment	
	0	0	3	3	30488421	2018	Effect of enriched environment during adolescence ...	Mohammadian; Najafi; Miladi-Gorji	Developmental psychobiology	
	3	0	0	3	29574368	2018	Estimation of the consumption of illicit drugs during ...	Foppe; Hammond-Weinberger; Subedi	The Science of the total environment	
	1	0	0	1	28787791	2017	Evaluation of in-sewer transformation of selected illi...	Gao; Banks; Li; Jiang; Lai; Mueller; Thai	The Science of the total environment	
	9	0	0	9	28472697	2017	Occurrence and fate of illicit drugs and pharmaceuti...	Causanilles; Ruepert; Ibáñez; Emke; Hernández; d...	The Science of the total environment	
	0	0	0	0	28010888	2016	Dose-dependent effects of morphine on lipopolysac...	Mottaz; Schönenberger; Fischer; Eggen; Schirmer; ...	Environmental pollution (Barking, Essex : 1987)	
	0	0	0	0	27746311	2016	Effects of voluntary exercise on the viability, prolifer...	Haydari; Safari; Zarbakhsh; Bandegi; Miladi-Gorji	Neuroscience letters	
	0	0	0	0	27261879	2016	Genotoxic effects induced by the exposure to an en...	Parolini; Magni; Castiglioni; Binelli	Ecotoxicology and environmental safety	
	3	0	0	3	27179320	2016	Temporal trends in drug use in Adelaide, South Aus...	Tscharke; Chen; Gerber; White	The Science of the total environment	

Abstract Sifter for Excel

EPA United States Environmental Protection Agency

F1000Research Open for Science

Search

BROWSE GATEWAYS HOW TO PUBLISH ABOUT BLOG

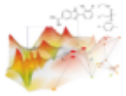
Check for updates

SOFTWARE TOOL ARTICLE

Abstract Sifter: a comprehensive front-end system to PubMed [version 1; referees: 2 approved]

Nancy Baker [1], Thomas Knudsen[2], Antony Williams [2]

Author details



This article is included in the Chemical Information Science gateway.

METRICS

629 VIEWS

118 DOWNLOADS

Get PDF

Get XML

Cite

Export

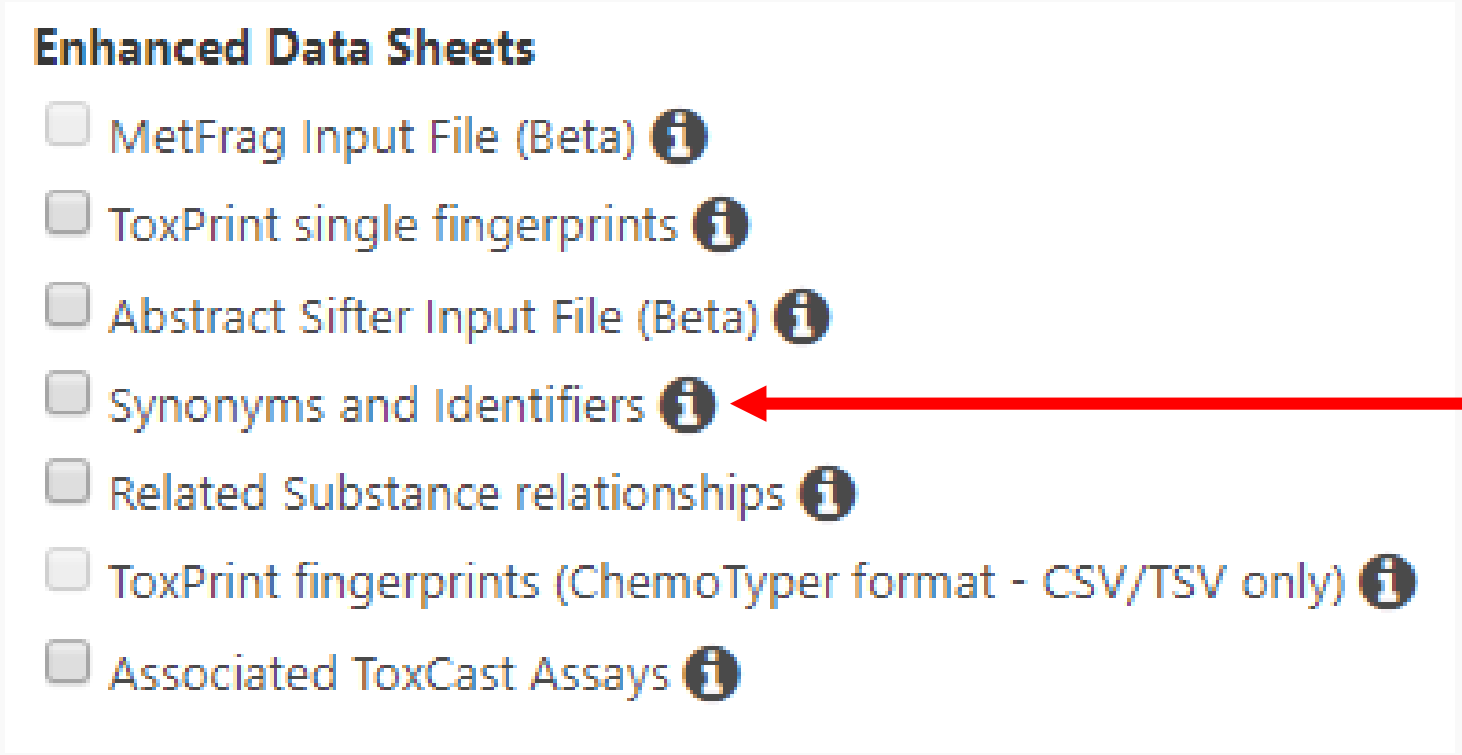
Track

Email

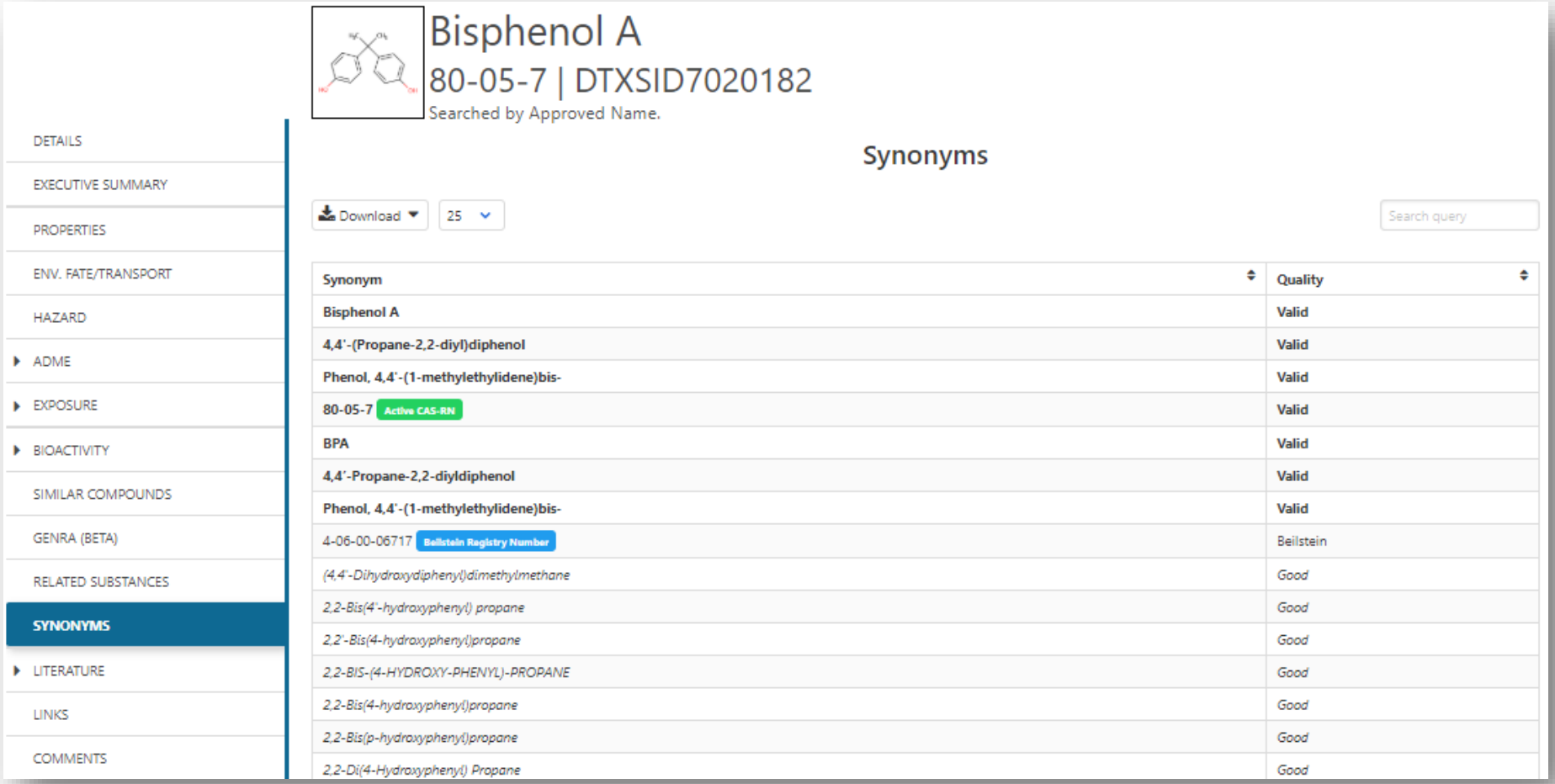
Abstract

The Abstract Sifter is a Microsoft Excel based application that enhances existing search capabilities of PubMed. The Abstract Sifter assists researchers to search effectively, triage results, and keep track of articles of interest. The tool implements an innovative “sifter” functionality for relevance ranking, giving the researcher a way to find articles of interest quickly. The tool also gives

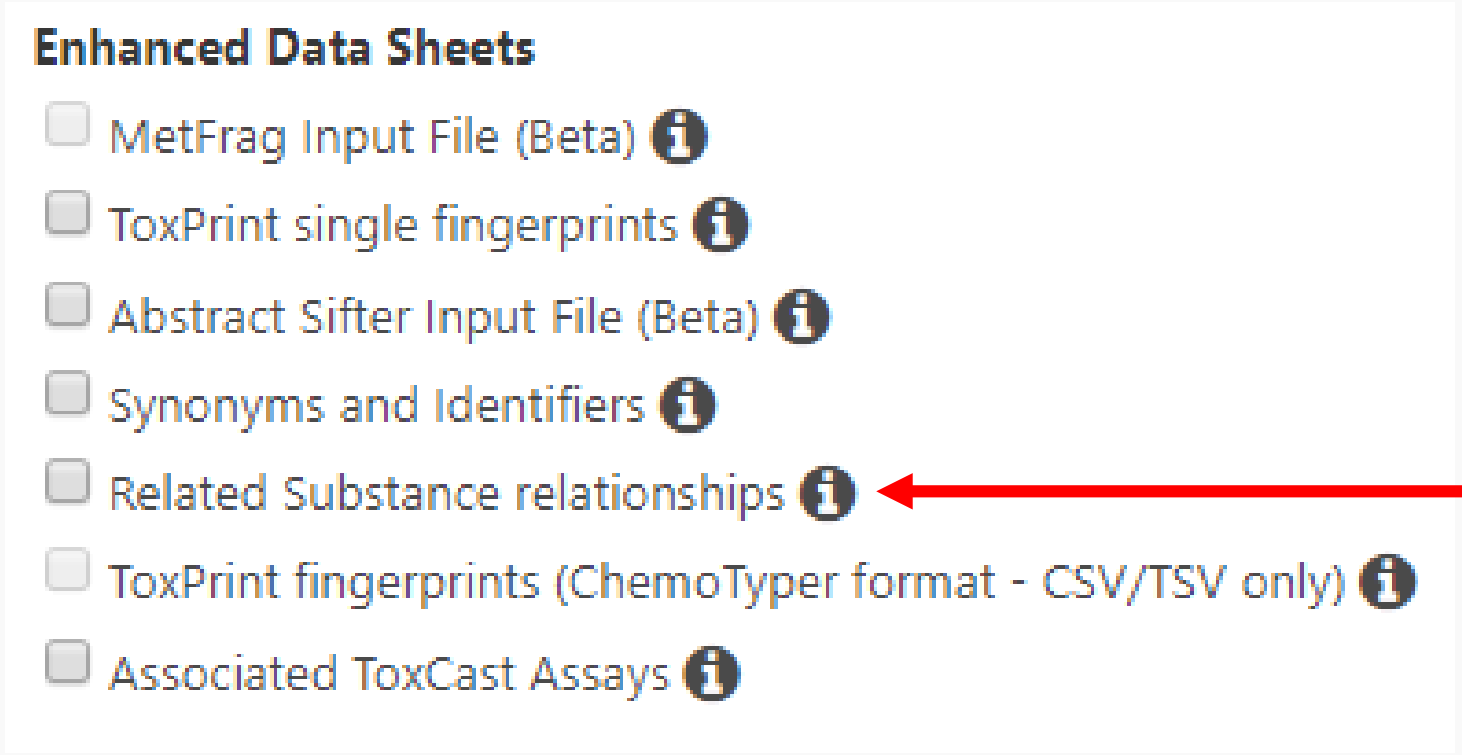
Possible Outputs…



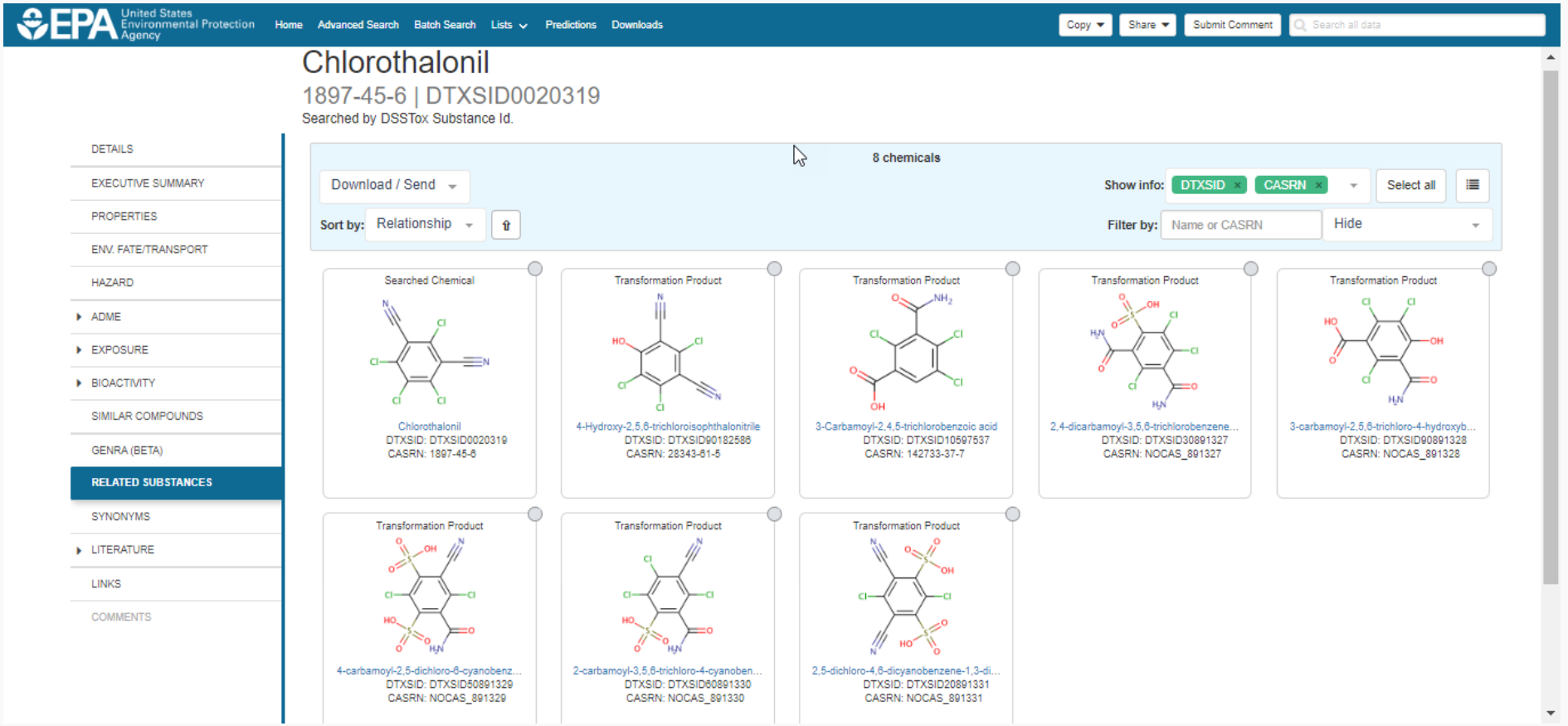
Identifiers to Support Searches



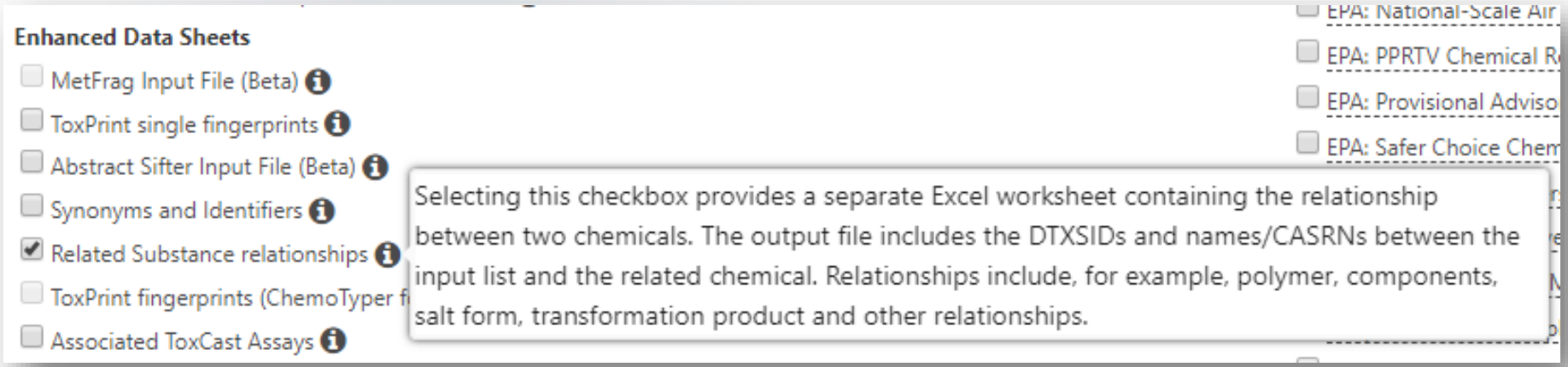
Possible Outputs…



Transformation Products



Related Substance Relationships

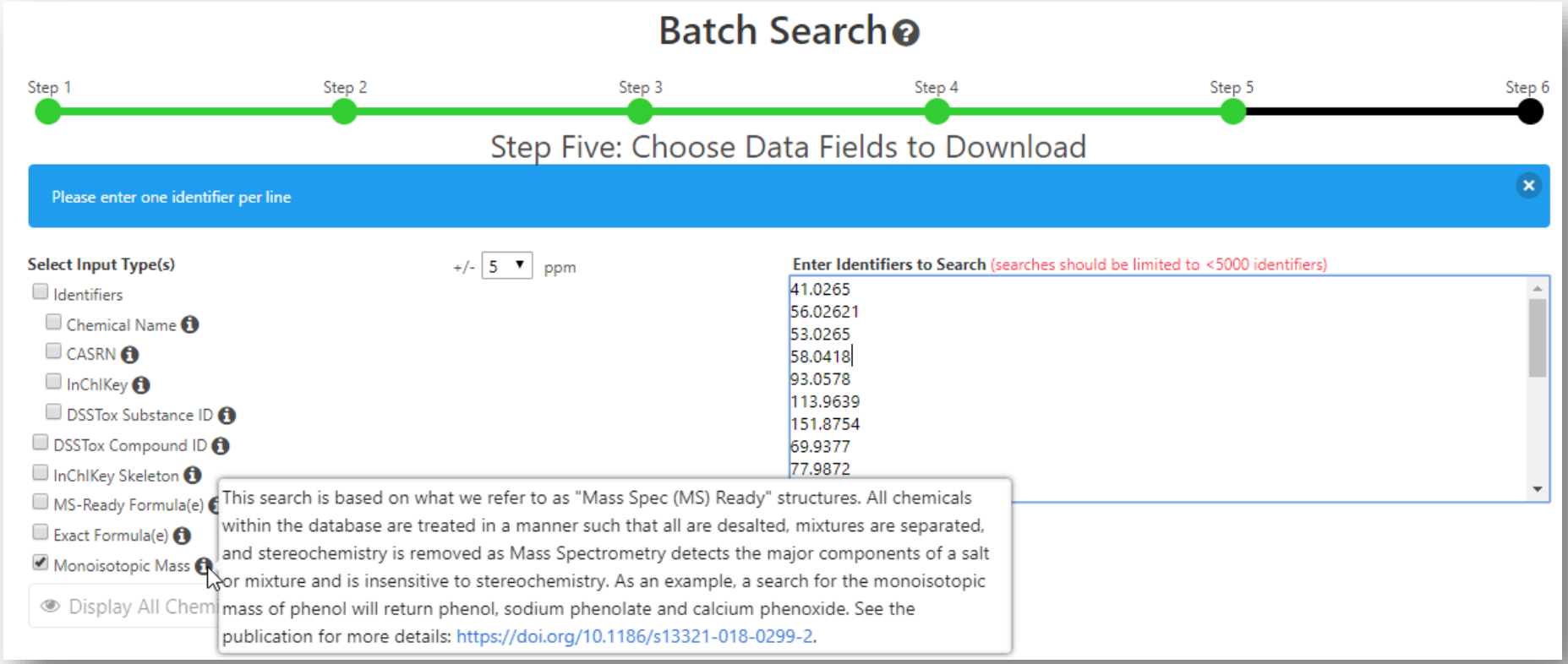


	A	B	C	D	E	
1	INPUT	DTXSID	PREFERRED_NAME	HAS_RELATIONSHIP_WITH	RELATED_DTXSID	RELATED_PREFERRED_NAM
2	xylenes	DTXSID2021446	Xylenes	Transformation Product	DTXSID40176394	N-Benzoylalanine
3	xylenes	DTXSID2021446	Xylenes	Component	DTXSID6026298	m-Xylene
4	xylenes	DTXSID2021446	Xylenes	Component	DTXSID3021807	o-Xylene
5	xylenes	DTXSID2021446	Xylenes	Component	DTXSID2021868	p-Xylene
6	xylenes	DTXSID2021446	Xylenes	Predecessor: Component	DTXSID9021421	Xylenes; defined mixture 1
7	xylenes	DTXSID2021446	Xylenes	Predecessor: Component	DTXSID7021447	Xylenes; defined mixture 2
8	xylenes	DTXSID2021446	Xylenes	Predecessor: Component	DTXSID30891529	Total Petroleum Hydrocarbons
9	xylenes	DTXSID2021446	Xylenes	Markush Child	DTXSID3021807	o-Xylene
10	xylenes	DTXSID2021446	Xylenes	Markush Child	DTXSID6026298	m-Xylene
11	xylenes	DTXSID2021446	Xylenes	Markush Child	DTXSID2021868	p-Xylene

CASRN Checking

	96949-21-2	NO_MATCH	-	-
	29066-34-0	Deleted CAS-RN	DTXSID9041817	dl-Men
	8039-[illegible]6	Deleted CAS-RN	DTXSID70897323	Ethoxyl
	76925-99-0	Deleted CAS-RN	DTXSID1044487	Methyl
	111-88-9	Checksum Failed	-	-
›is[[(1	187273-40-6	Checksum Failed	-	-
	24213-24-5	Checksum Failed	-	-
nd 1,2	25640-14-0	Checksum Failed	-	-
:hanec	68601-19-5	Checksum Failed	-	-
›pyl 2-	67990-40-3	CAS-RN	DTXSID90987412	N,N-Di
	68440-90-4	CAS-RN	DTXSID9098434	Siloxan
	68425-44-5	CAS-RN	DTXSID9098305	Amides
	68132-78-5	CAS-RN	DTXSID9096995	Amines
	68071-95-4	CAS-RN	DTXSID9096575	Quater

Batch Searching Formula/Mass



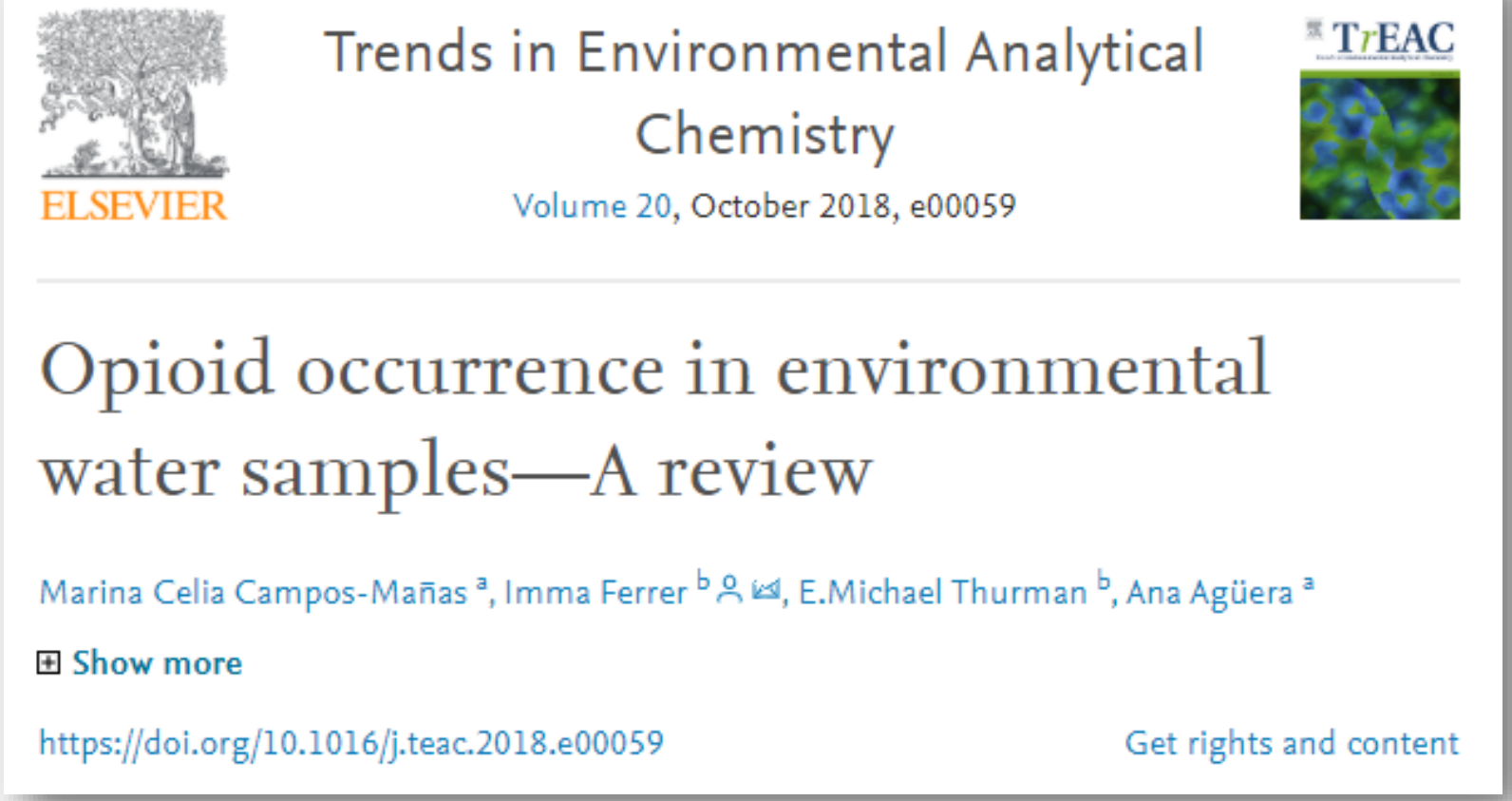
Searching batches using MS-Ready Formula (or mass) searching

EPA United States Environmental Protection Agency

	A	B	C	D	E	F	G
1	INPUT	DTXSID	CASRN	PREFERRED NAME	MOL FORMULA	MONOISOTOPIC MASS	DATA SOURCES
2	C14H22N2O3	DTXSID2022628	29122-68-7	Atenolol	C14H22N2O3	266.163042576	46
3	C14H22N2O3	DTXSID0021179	6673-35-4	Practolol	C14H22N2O3	266.163042576	32
4	C14H22N2O3	DTXSID4048854	841-73-6	Bucolome	C14H22N2O3	266.163042576	20
5	C14H22N2O3	DTXSID1045407	13171-25-0	Trimetazidine dihydrochloride	C14H24Cl2N2O3	338.116398	19
6	C14H22N2O3	DTXSID0045753	56715-13-0	R-(+)-Atenolol	C14H22N2O3	266.163042576	19
7	C14H22N2O3	DTXSID2048531	5011-34-7	Trimetazidine	C14H22N2O3	266.163042576	14
8	C14H22N2O3	DTXSID10239405	93379-54-5	Esatenolol	C14H22N2O3	266.163042576	12
9	C14H22N2O3	DTXSID50200634	52662-27-8	N-(2-Diethylaminoethyl)-2-(4-hydroxyphenoxy)acetamide	C14H22N2O3	266.163042576	7
10	C14H22N2O3	DTXSID4020111	51706-40-2	dl-Atenolol hydrochloride	C14H23ClN2O3	302.1397203	6
11	C14H22N2O3	DTXSID1068693	51963-82-7	Benzenamine, 2,5-diethoxy-4-(4-morpholinyl)-	C14H22N2O3	266.163042576	5
12	C18H34N2O6S	DTXSID3023215	154-21-2	Lincomycin	C18H34N2O6S	406.213757997	35
13	C18H34N2O6S	DTXSID7047803	859-18-7	Lincomycin hydrochloride	C18H35ClN2O6S	442.1904357	22
14	C18H34N2O6S	DTXSID20849438	1398534-62-7	PUBCHEM_71432748	C18H35ClN2O6S	442.1904357	1
15	C10H12N2O	DTXSID1047576	486-56-6	Cotinine	C10H12N2O	176.094963014	40
16	C10H12N2O	DTXSID8075330	50-67-9	Serotonin	C10H12N2O	176.094963014	22
17	C10H12N2O	DTXSID8044412	2654-57-1	4-Methyl-1-phenylpyrazolidin-3-one	C10H12N2O	176.094963014	18
18	C10H12N2O	DTXSID80165186	153-98-0	Serotonin hydrochloride	C10H13ClN2O	212.0716407	11
19	C10H12N2O	DTXSID2048870	29493-77-4	(4R,5S)-4-methyl-5-phenyl-4,5-dihydro-1,3-oxazol-2-amine	C10H12N2O	176.094963014	10
20	C10H12N2O	DTXSID10196105	443-31-2	6-Hydroxytryptamine	C10H12N2O	176.094963014	9
21	C10H12N2O	DTXSID90185693	31822-84-1	1,4,5,6-Tetrahydro-5-phenoxypyrimidine	C10H12N2O	176.094963014	7
22	C10H12N2O	DTXSID40178777	2403-66-9	2-Benzimidazolepropanol	C10H12N2O	176.094963014	7
23	C10H12N2O	DTXSID80157026	13140-86-8	N-Cyclopropyl-N'-phenylurea	C10H12N2O	176.094963014	6
24	C10H12N2O	DTXSID30205607	570-14-9	4-Hydroxytryptamine	C10H12N2O	176.094963014	6
25	C14H18N4O3	DTXSID5023900	17804-35-2	Benomyl	C14H18N4O3	290.137890456	68
26	C14H18N4O3	DTXSID3023712	738-70-5	Trimethoprim	C14H18N4O3	290.137890456	51
27	C14H18N4O3	DTXSID40209671	60834-30-2	Trimethoprim hydrochloride	C14H19ClN4O3	326.1145682	8
28	C14H18N4O3	DTXSID70204210	55687-49-5	Benzenemethanol, 4-((2,4-diamino-5-pyrimidinyl)methyl)-2,	C14H18N4O3	290.137890456	5
29	C14H18N4O3	DTXSID20152671	120075-57-2	6-Methoxy-4-(3-(N,N-dimethylamino)propylamino)-5,8-quin	C14H18N4O3	290.137890456	4
30	C14H18N4O3	DTXSID30213742	63931-79-3	1H-1,2,4-Benzotriazepine-3-carboxylic acid, 4,5-dihydro-4-	C14H18N4O3	290.137890456	3
31	C14H18N4O3	DTXSID30219608	69449-07-6	2,4-Pyrimidinediamine, 5-((3,4,5-trimethoxyphenyl)methyl)-	C14H20N4O4	308.14845514	3
32	C14H18N4O3	DTXSID20241155	94232-27-6	L-Aspartic acid, compound with 5-((3,4,5-trimethoxyphenyl	C18H25N5O7	423.175398165	3
33	C14H18N4O3	DTXSID80241156	94232-28-7	L-Glutamic acid, compound with 5-((3,4,5-trimethoxypheny	C19H27N5O7	437.191048229	3
34	C14H18N4O3	DTXSID20143781	101204-93-7	1H-Pyrido(2,3-e)-1,4-diazepine-2,3,5-trione, 4-(2-(diethylam	C14H18N4O3	290.137890456	3
35	C12H11N7	DTXSID6021373	396-01-0	Triamterene	C12H11N7	253.107593382	52
36	C12H11N7	DTXSID00204465	5587-93-9	Ampyrimine	C12H11N7	253.107593382	7
37	C12H11N7	DTXSID5064621	7300-26-7	Benzenamine, 4-azido-N-(4-azidophenyl)-	C12H9N7	251.091943318	4
38	C12H11N7	DTXSID00848025	90293-82-6	Sulfuric acid--6-phenylpteridine-2,4,7-triamine (1/1)	C12H13N7O4S	351.074973101	1
39	C12H11N7	DTXSID50575293	92310-83-3	(1E)-N-Phenyl-1,2-bis(1H-1,2,4-triazol-1-yl)ethan-1-imine	C12H11N7	253.107593382	1
40	C8H9NO2	DTXSID2020006	103-90-2	Acetaminophen	C8H9NO2	151.063328534	75
41	C8H9NO2	DTXSID6026567	134-20-3	Methyl 2-aminobenzoate	C8H9NO2	151.063328534	60

AN EXAMPLE APPLICATION

Aggregate data for a list of chemicals



ELSEVIER

Trends in Environmental Analytical Chemistry

Volume 20, October 2018, e00059

TrEAC

Opioid occurrence in environmental water samples—A review

Marina Celia Campos-Mañas [a], Imma Ferrer [b], E.Michael Thurman [b], Ana Agüera [a]

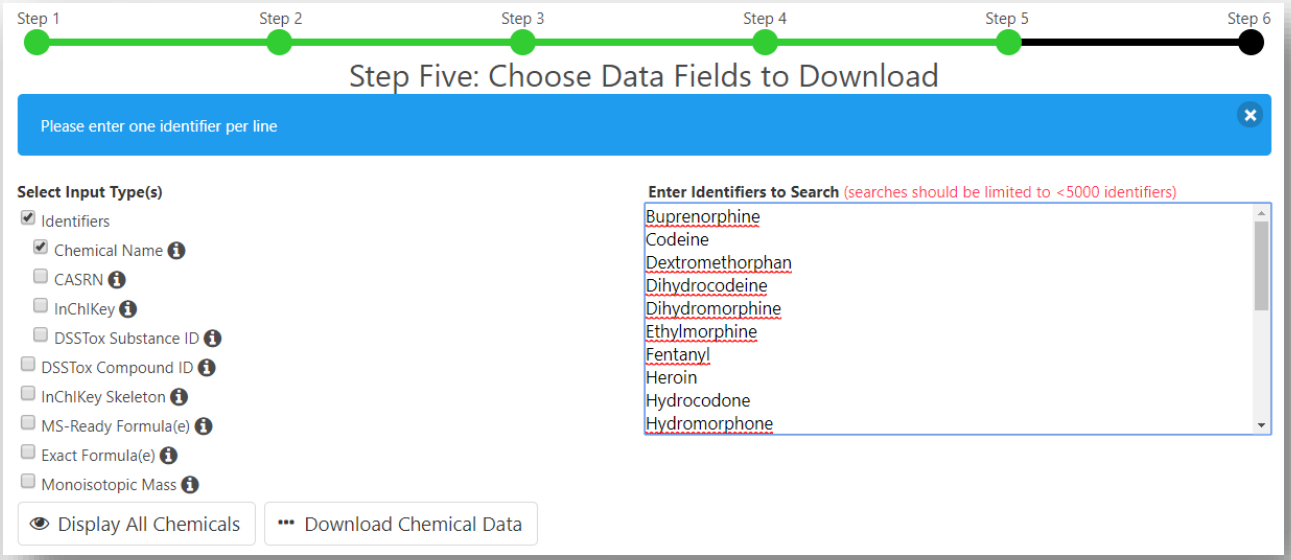
Show more

https://doi.org/10.1016/j.teac.2018.e00059

Get rights and content

Batch Search Names

Buprenorphine
Codeine
Dextromethorphan
Dihydrocodeine
Dihydromorphine
Ethylmorphine
Fentanyl
Heroin
Hydrocodone
Hydromorphone
Ketamine
Meperidine
Methadone
Morphine
Morphinone
Naloxone
Naltriben
Oxycodone
Oxymorphone
Propoxyphene
Sufentanil
Tramadol



Excel Download

INPUT	FOUND_BY	DTXSID
Buprenorphine	Approved Name	DTXSID2022705
Codeine	Approved Name	DTXSID2020341
Dextromethorphan	Approved Name	DTXSID3022908
Dihydrocodeine	Approved Name	DTXSID5022936
Dihydromorphine	Approved Name	DTXSID7048908
Ethylmorphine	Approved Name	DTXSID1046760
Fentanyl	Approved Name	DTXSID9023049
Heroin	Synonym	DTXSID6046761
Hydrocodone	Approved Name	DTXSID8023131
Hydromorphone	Approved Name	DTXSID8023133
Ketamine	Approved Name	DTXSID8023187
Meperidine	Approved Name	DTXSID9023253
Methadone	Approved Name	DTXSID7023273
Morphine	Approved Name	DTXSID9023336

Add Other Data of Interest

Chemical Identifiers

- [x] DTXSID
- [x] Chemical Name
- [ ] DTXCID
- [x] CAS-RN
- [x] InChIKey
- [ ] IUPAC Name

Structures

- [ ] Mol File
- [ ] SMILES
- [ ] InChI String
- [x] MS-Ready SMILES
- [ ] QSAR-Ready SMILES

Intrinsic And Predicted Properties

- [x] Molecular Formula
- [ ] Average Mass
- [x] Monoisotopic Mass
- [ ] TEST Model Predictions
- [ ] OPERA Model Predictions

INPUT	DTXSID	CASRN	MOLECULAR_F	MONOISOTOPIC	MS_READY_SMI
Buprenorph	DTXSID202	52485-79-7	C29H41NO4	467.3035588	[H]C12CC3=C4C(
Codeine	DTXSID202	76-57-3	C18H21NO3	299.1521435	[H]C12CC3=C4C(
Dextrometh	DTXSID302	125-71-3	C18H25NO	271.1936144	[H]C12CC3=C(C=
Dihydrocod	DTXSID502	125-28-0	C18H23NO3	301.1677936	[H]C12CC3=C4C(
Dihydromor	DTXSID704	509-60-4	C17H21NO3	287.1521435	[H]C12CC3=C4C(
Ethylmorph	DTXSID104	76-58-4	C19H23NO3	313.1677936	[H]C12CC3=C4C(
Fentanyl	DTXSID902	437-38-7	C22H28N2O	336.2201635	CCC(=O)N(C1CC
Heroin	DTXSID604	561-27-3	C21H23NO5	369.1576228	[H]C12CC3=C4C(
Hydrocodor	DTXSID802	125-29-1	C18H21NO3	299.1521435	[H]C12CC3=C4C(
Hydromorph	DTXSID802	466-99-9	C17H19NO3	285.1364935	[H]C12CC3=C4C(
Ketamine	DTXSID802	6740-88-1	C13H16ClNO	237.0920418	CNC1(CCCCC1=(
Meperidine	DTXSID902	57-42-1	C15H21NO2	247.1572289	CCOC(=O)C1(CC
Methadone	DTXSID702	76-99-3	C21H27NO	309.2092645	CCC(=O)C(CC(C)
Morphine	DTXSID902	57-27-2	C17H19NO3	285.1364935	[H]C12CC3=C4C(
Morphinone	DTXSID501	467-02-7	C17H17NO3	283.1208434	[H]C12CC3=C4C(
Naloxone	DTXSID802	465-65-6	C19H21NO4	327.1470582	[H]C12CC3=C4C(
Naltriben	-	-	-	-	-
Oxycodone	DTXSID502	76-42-6	C18H21NO4	315.1470582	[H]C12CC3=C4C(
Oxymorpho	DTXSID502	76-41-5	C17H19NO4	301.1314081	[H]C12CC3=C4C(
Propoxyphe	DTXSID102	469-62-5	C22H29NO2	339.2198292	CCC(=O)OC(CC1
Sufentanil	DTXSID602	56030-54-7	C22H30N2O2S	386.2027994	CCC(=O)N(C1=C
Tramadol	DTXSID908	27203-92-5	C16H25NO2	263.188529	COC1=CC=CC(=

Possible Outputs – Chemical Lists

Presence in Lists:

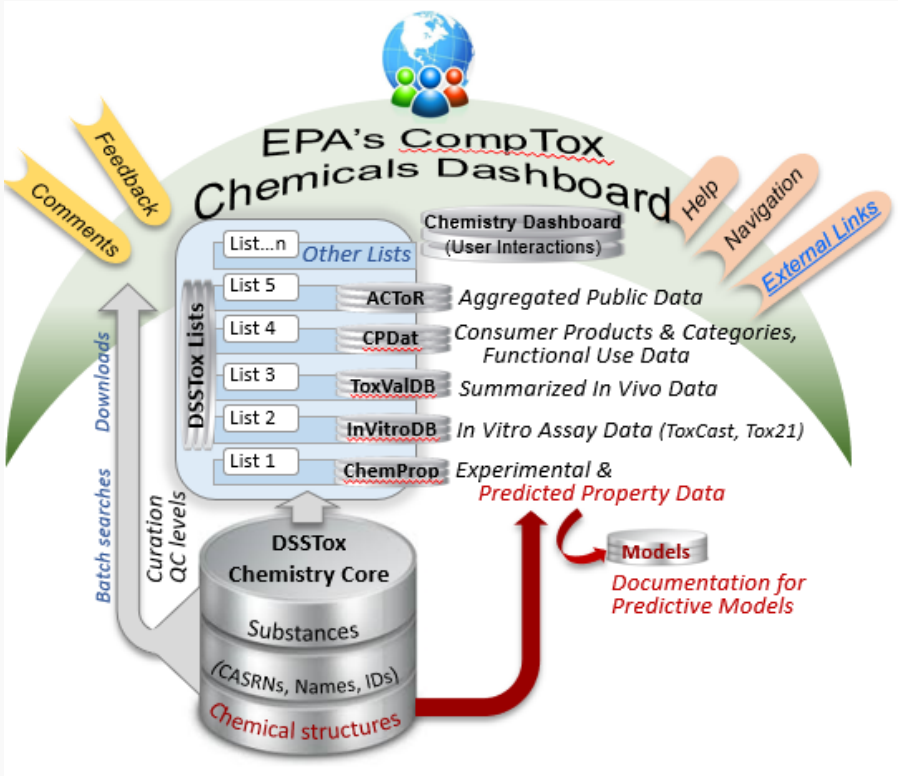
- [ ] 40CFR355 Extremely Hazardous Substance List and Threshold Planning Quantities
- [ ] AEGLS: Acute Exposure Guideline Levels
- [ ] ANDROGEN: Androgen Receptor Chemicals
- [ ] ARTICLE; Bench-Mark Dose Human Health Assessment List (Wignall et al., 2014)
- [ ] ATSDR: Minimal Risk Levels (MRLs) for Hazardous Substances
- [ ] ATSDR: Toxic Substances Portal Chemical List
- [ ] Barbara Wetmore's HTTK list
- [ ] California Office of Environmental Health Hazard Assessment
- [ ] CERAPP: Collaborative Estrogen Receptor Activity Prediction Project
- [ ] CHEMINV; ToxCast/Tox21 Chemical inventory available as DMSO solutions (20181123)
- [ ] CHEMINV: EPA Chemical Inventory for ToxCast
- [ ] CHEMINV: EPA ToxCast CHEMINV list of volatiles

Batch Search in specific lists

INPUT	DTXSID	MASSBANKREF	NEMILIST	WRTMSD	NORMANPRI	SUSDAT
Buprenorph	DTXSID202	-	-	Y	-	Y
Codeine	DTXSID202	Y	Y	Y	Y	Y
Dextrometh	DTXSID302	Y	Y	Y	-	Y
Dihydrocod	DTXSID502	Y	-	Y	Y	Y
Dihydromor	DTXSID704	-	-	-	-	Y
Ethylmorph	DTXSID104	-	-	Y	-	Y
Fentanyl	DTXSID902	Y	-	Y	-	Y
Heroin	DTXSID604	Y	-	Y	Y	Y
Hydrocodor	DTXSID802	Y	Y	Y	Y	Y
Hydromorpl	DTXSID802	-	-	Y	-	Y
Ketamine	DTXSID802	Y	-	Y	-	Y
Meperidine	DTXSID902	Y	-	Y	-	Y
Methadone	DTXSID702	Y	Y	Y	-	Y
Morphine	DTXSID902	Y	Y	Y	Y	Y
Morphinone	DTXSID501	-	-	-	-	Y
Naloxone	DTXSID802	-	-	Y	-	Y
Naltriben	-	-	-	-	-	-
Oxycodone	DTXSID502	Y	Y	Y	Y	Y
Oxymorpho	DTXSID502	-	-	Y	-	Y
Propoxyphe	DTXSID102	Y	Y	Y	-	Y
Sufentanil	DTXSID602	-	-	Y	-	Y
Tramadol	DTXSID908	Y	Y	Y	Y	Y

Conclusion

- Dashboard access to data for ~875,000 chemicals
- Batch search capability for integrated search for thousands of chemicals across diverse data
- Future capabilities will include batch prediction
- Ongoing data expansion continues day to day
- We are committed to open API development for data access and predictions



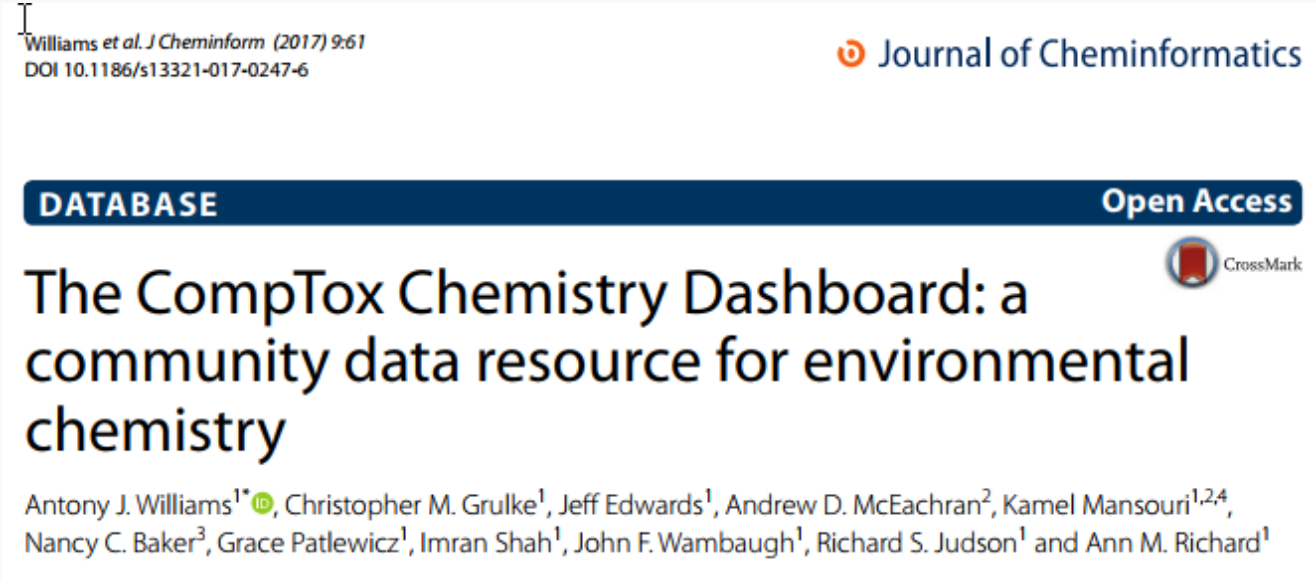
Acknowledgements

EPA-RTP

- *An enormous team of contributors from NCCT, especially the IT software development team*
- *Our curation team for their care and focus on data quality*
- *Multiple centers and laboratories across the EPA*
- *Many public domain databases and open data contributors*

Contact

Antony Williams
NCCT, US EPA Office of Research and Development,
Williams.Antony@epa.gov
ORCID: https://orcid.org/0000-0002-2668-4821



Williams et al. J Cheminform (2017) 9:61
DOI 10.1186/s13321-017-0247-6

Journal of Cheminformatics

DATABASE Open Access

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

Antony J. Williams[1*], Christopher M. Grulke[1], Jeff Edwards[1], Andrew D. McEachran[2], Kamel Mansouri[1,2,4], Nancy C. Baker[3], Grace Patlewicz[1], Imran Shah[1], John F. Wambaugh[1], Richard S. Judson[1] and Ann M. Richard[1]

https://doi.org/10.1186/s13321-017-0247-6