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

Integrating Mass Spectrometry Non-Targeted Analysis and Computational Toxicology to Characterize Chemicals

Antony Williams

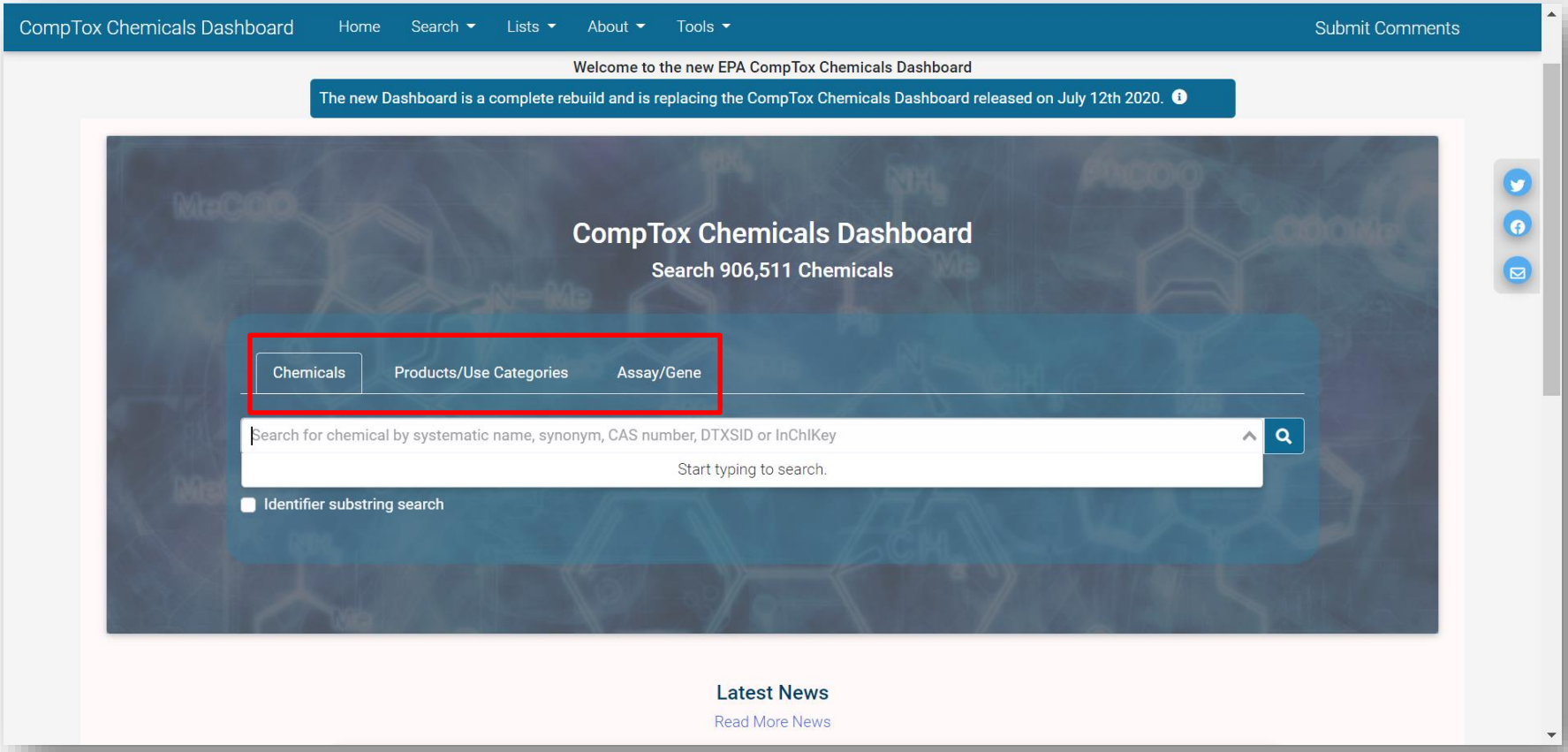
Center for Computational Toxicology and Exposure, US-EPA, RTP, NC

The views expressed in this presentation are those of the author and do not necessarily reflect the views or policies of the U.S. EPA

June 2022

Medical Device and Combination Product Specialty Section webinar

CompTox Chemicals Dashboard >906k chemicals

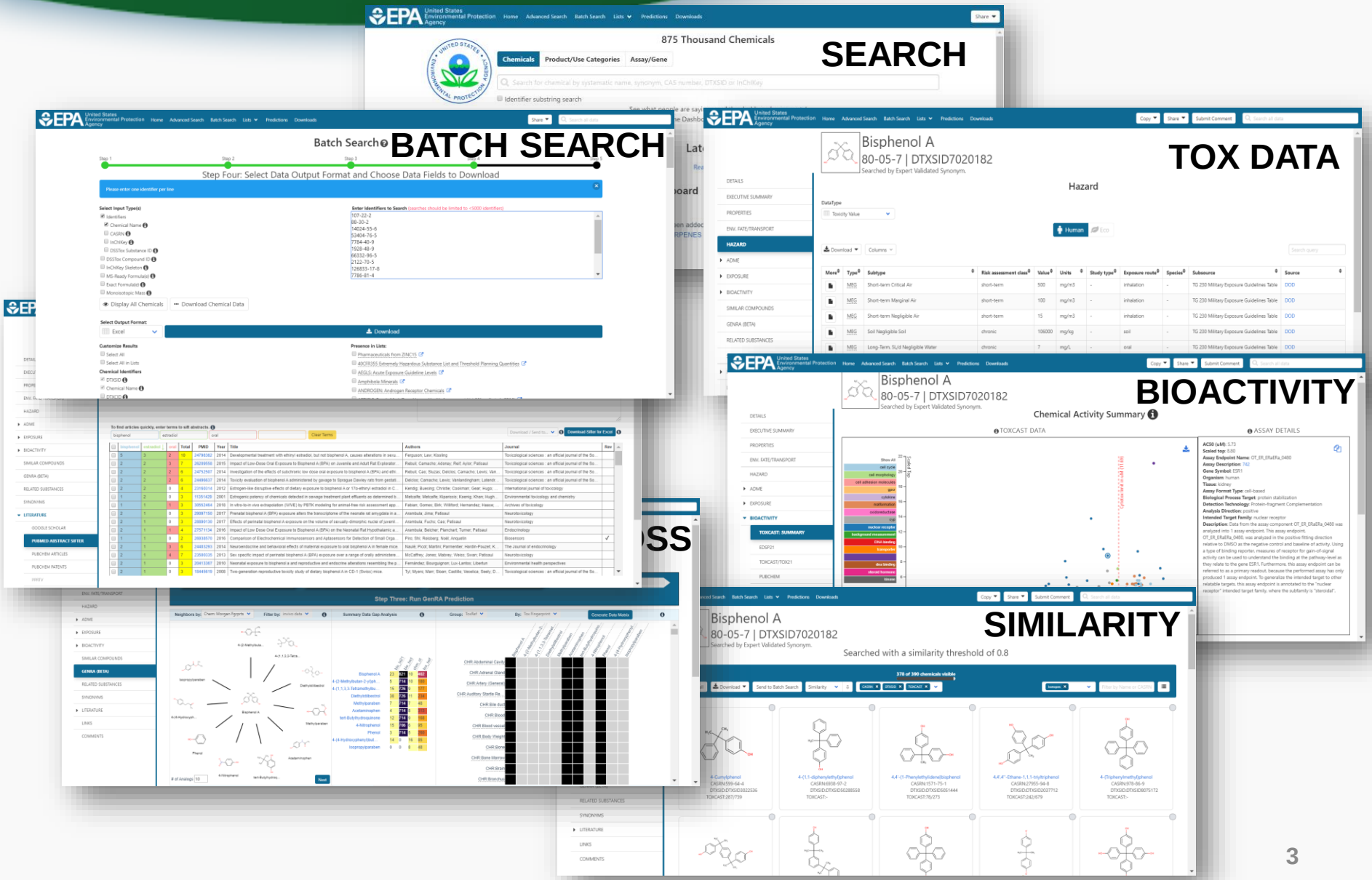


The Charge for the Dashboard

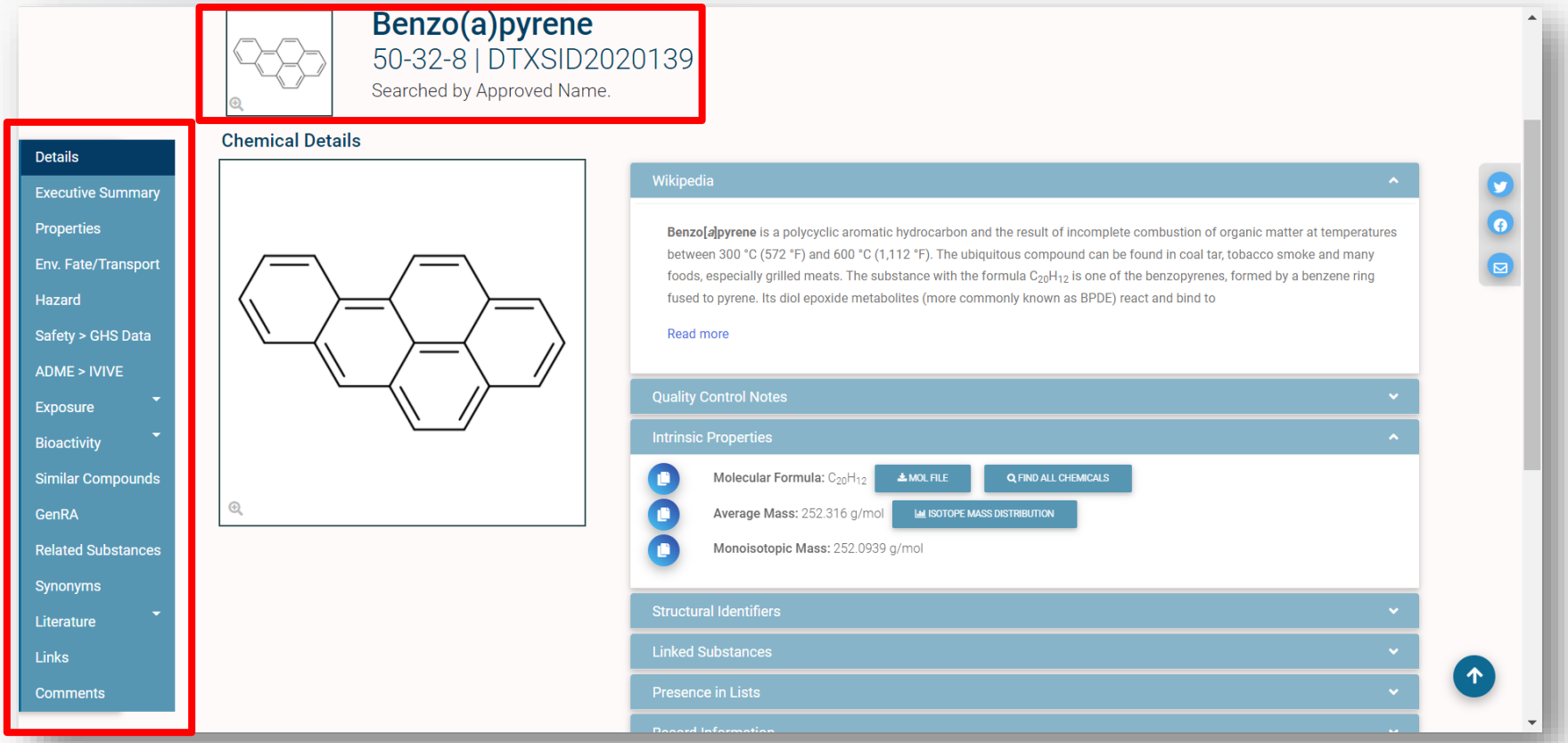
- Develop a “first-stop-shop” for environmental chemical data to support EPA and partner decision making:
  - **Centralized location** for relevant chemical data
  - Chemistry, exposure, hazard and dosimetry
  - Combination of existing data and predictive models
  - Publicly accessible, periodically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

CompTox Chemicals Dashboard

https://comptox.epa.gov/dashboard

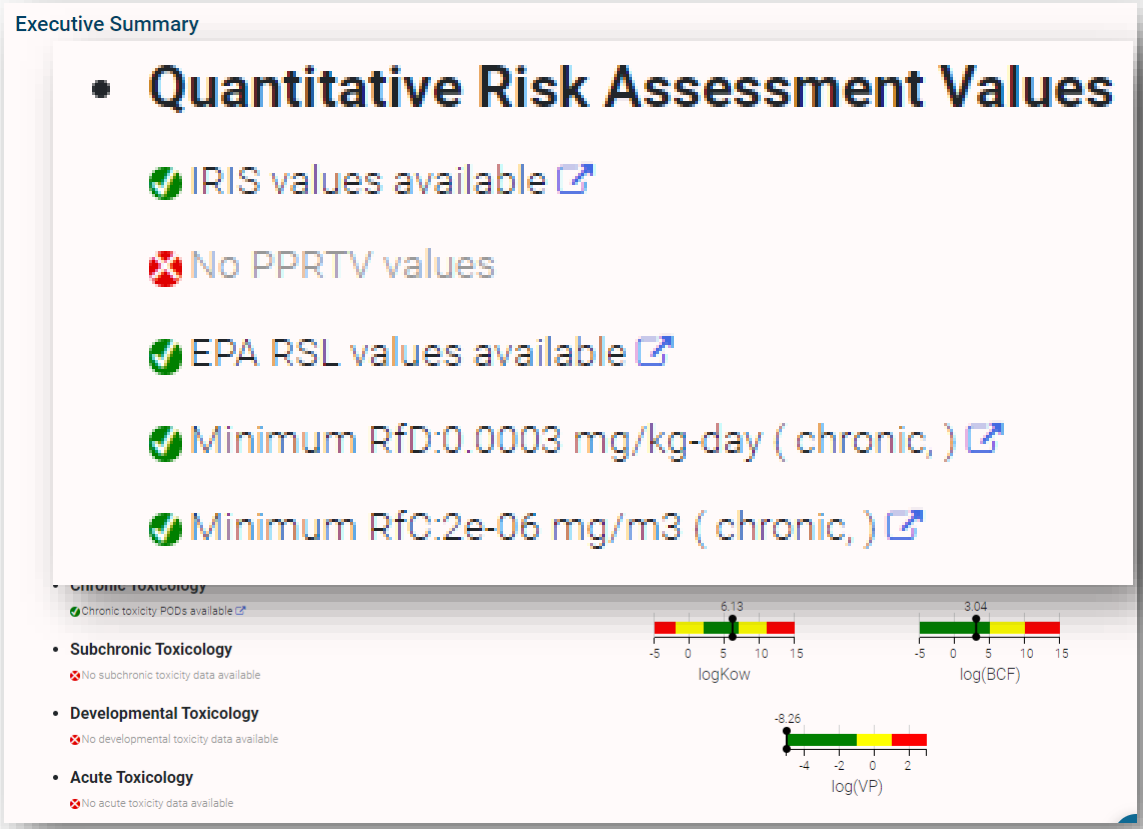


Detailed Chemical Pages



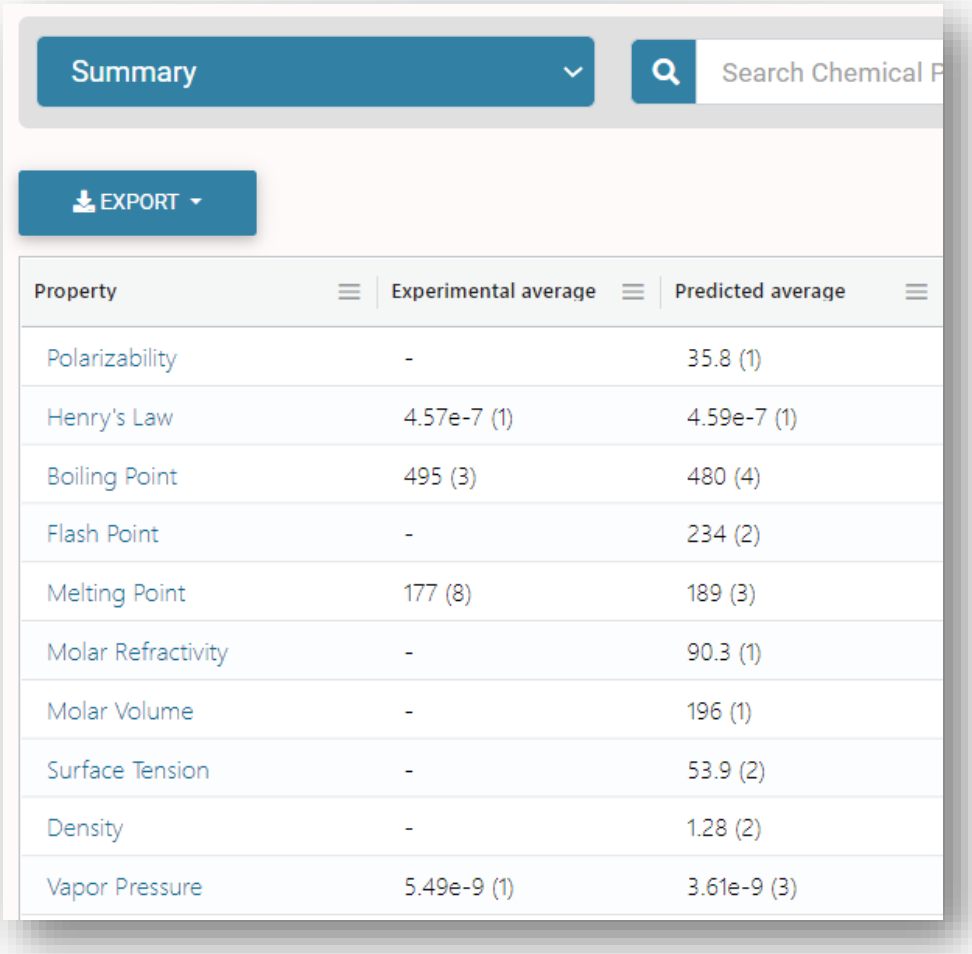
- Chemical page: Wikipedia snippet when available, intrinsic properties, structural identifiers, linked substances

“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



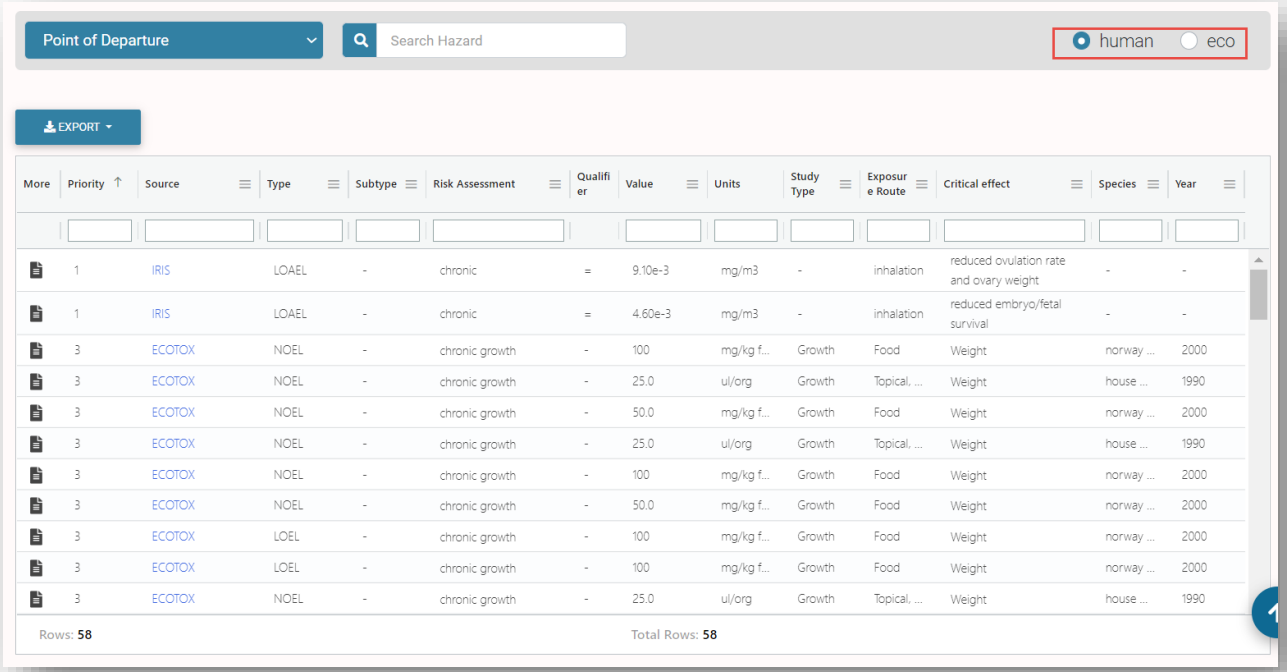
Property	Experimental average	Predicted average
Polarizability	-	35.8 (1)
Henry's Law	4.57e-7 (1)	4.59e-7 (1)
Boiling Point	495 (3)	480 (4)
Flash Point	-	234 (2)
Melting Point	177 (8)	189 (3)
Molar Refractivity	-	90.3 (1)
Molar Volume	-	196 (1)
Surface Tension	-	53.9 (2)
Density	-	1.28 (2)
Vapor Pressure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

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

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

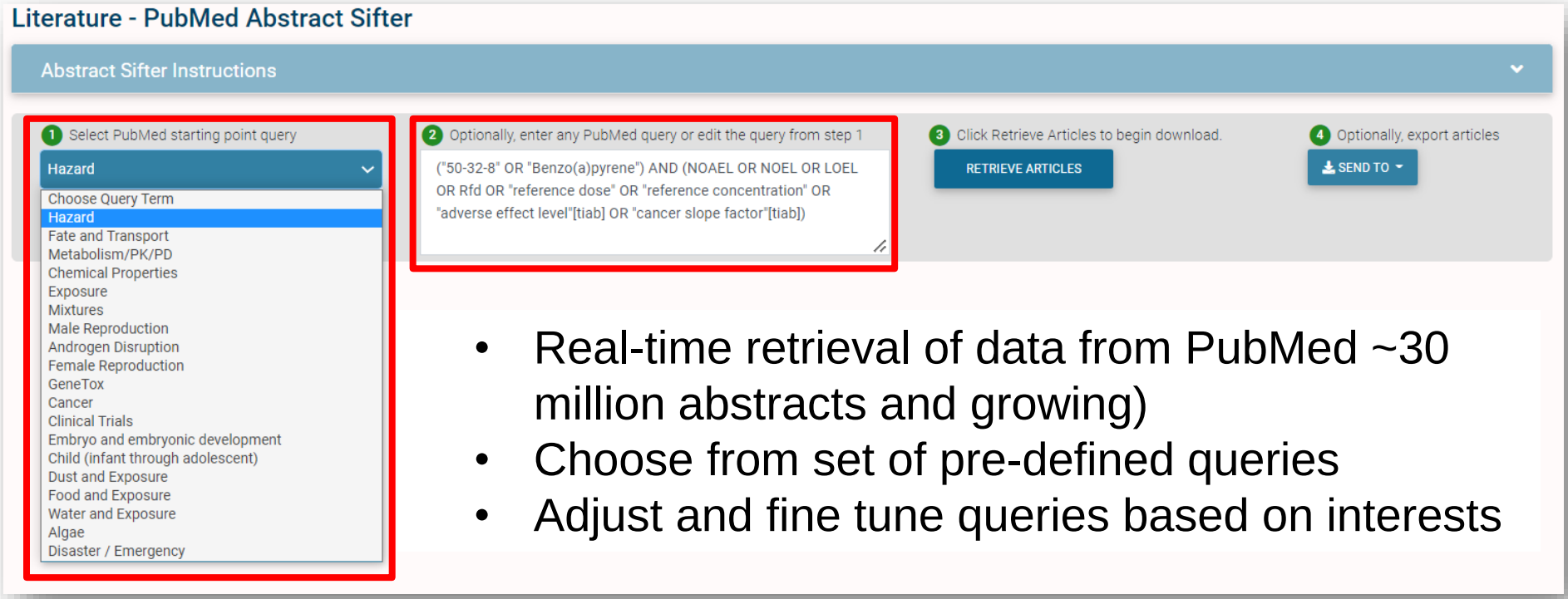
Product Name	Product Use Category	Categorization Subtype	Minimum Weight Fraction	Maximum Weight Fraction	Data Type	Source	Product Count
48743 pah mixture	Not yet Categorized		-	-	reported	SIRI	1
asphalt cement penetration 60-70	Not yet Categorized		-	-	reported	SIRI	1
base-neutral 4.1ml methylene chl...	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

EPA
United States
Environmental Protection
Agency

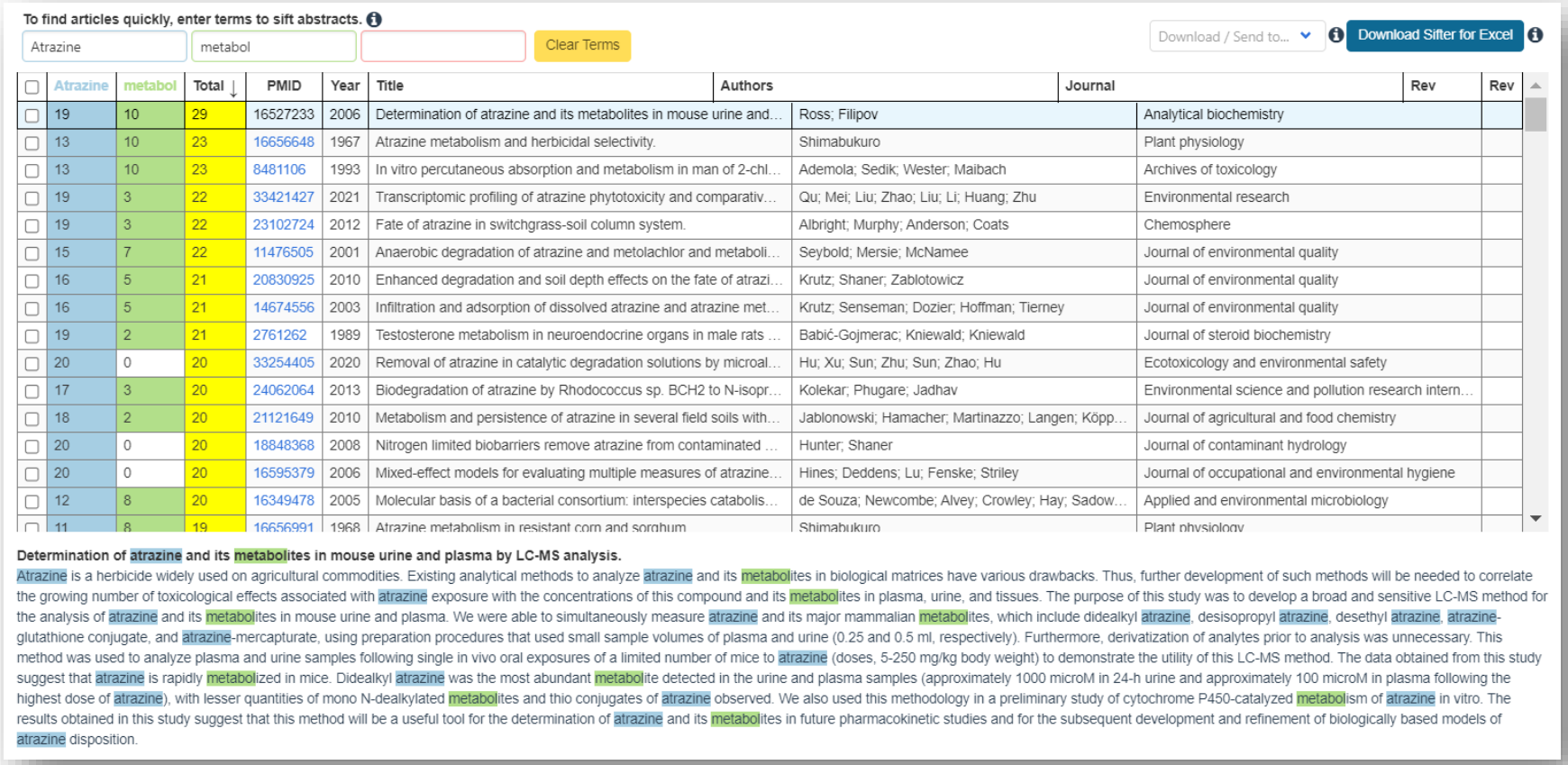
Searching Literature and the Internet

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

Markup of abstracts



To find articles quickly, enter terms to sift abstracts.

Atrazine | metabol | Clear Terms | Download / Send to... | Download Sifter for Excel

	Atrazine	metabol	Total	PMID	Year	Title	Authors	Journal	Rev	Rev
☐	19	10	29	16527233	2006	Determination of atrazine and its metabolites in mouse urine and…	Ross; Filipov	Analytical biochemistry		
☐	13	10	23	16656648	1967	Atrazine metabolism and herbicidal selectivity.	Shimabukuro	Plant physiology		
☐	13	10	23	8481106	1993	In vitro percutaneous absorption and metabolism in man of 2-chl…	Ademola; Sedik; Wester; Maibach	Archives of toxicology		
☐	19	3	22	33421427	2021	Transcriptomic profiling of atrazine phytotoxicity and comparativ…	Qu; Mei; Liu; Zhao; Liu; Li; Huang; Zhu	Environmental research		
☐	19	3	22	23102724	2012	Fate of atrazine in switchgrass-soil column system.	Albright; Murphy; Anderson; Coats	Chemosphere		
☐	15	7	22	11476505	2001	Anaerobic degradation of atrazine and metolachlor and metaboli…	Seybold; Mersie; McNamee	Journal of environmental quality		
☐	16	5	21	20830925	2010	Enhanced degradation and soil depth effects on the fate of atrazi…	Krutz; Shaner; Zablotowicz	Journal of environmental quality		
☐	16	5	21	14674556	2003	Infiltration and adsorption of dissolved atrazine and atrazine met…	Krutz; Senseman; Dozier; Hoffman; Tierney	Journal of environmental quality		
☐	19	2	21	2761262	1989	Testosterone metabolism in neuroendocrine organs in male rats …	Babić-Gojmerac; Kniewald; Kniewald	Journal of steroid biochemistry		
☐	20	0	20	33254405	2020	Removal of atrazine in catalytic degradation solutions by microal…	Hu; Xu; Sun; Zhu; Sun; Zhao; Hu	Ecotoxicology and environmental safety		
☐	17	3	20	24062064	2013	Biodegradation of atrazine by Rhodococcus sp. BCH2 to N-isopr…	Kolekar; Phugare; Jadhav	Environmental science and pollution research intern…		
☐	18	2	20	21121649	2010	Metabolism and persistence of atrazine in several field soils with…	Jablonowski; Hamacher; Martinazzo; Langen; Köpp…	Journal of agricultural and food chemistry		
☐	20	0	20	18848368	2008	Nitrogen limited biobarriers remove atrazine from contaminated …	Hunter; Shaner	Journal of contaminant hydrology		
☐	20	0	20	16595379	2006	Mixed-effect models for evaluating multiple measures of atrazine…	Hines; Deddens; Lu; Fenske; Striley	Journal of occupational and environmental hygiene		
☐	12	8	20	16349478	2005	Molecular basis of a bacterial consortium: interspecies catabolis…	de Souza; Newcombe; Alvey; Crowley; Hay; Sadow…	Applied and environmental microbiology		
☐	11	8	19	16656991	1968	Atrazine metabolism in resistant corn and sorghum	Shimabukuro	Plant physiology		

Determination of atrazine and its metabolites in mouse urine and plasma by LC-MS analysis.

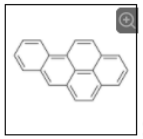
Atrazine is a herbicide widely used on agricultural commodities. Existing analytical methods to analyze atrazine and its metabolites in biological matrices have various drawbacks. Thus, further development of such methods will be needed to correlate the growing number of toxicological effects associated with atrazine exposure with the concentrations of this compound and its metabolites in plasma, urine, and tissues. The purpose of this study was to develop a broad and sensitive LC-MS method for the analysis of atrazine and its metabolites in mouse urine and plasma. We were able to simultaneously measure atrazine and its major mammalian metabolites, which include didealkyl atrazine, desisopropyl atrazine, desethyl atrazine, atrazine-glutathione conjugate, and atrazine-mercapturate, using preparation procedures that used small sample volumes of plasma and urine (0.25 and 0.5 ml, respectively). Furthermore, derivatization of analytes prior to analysis was unnecessary. This method was used to analyze plasma and urine samples following single in vivo oral exposures of a limited number of mice to atrazine (doses, 5-250 mg/kg body weight) to demonstrate the utility of this LC-MS method. The data obtained from this study suggest that atrazine is rapidly metabolized in mice. Didealkyl atrazine was the most abundant metabolite detected in the urine and plasma samples (approximately 1000 microM in 24-h urine and approximately 100 microM in plasma following the highest dose of atrazine), with lesser quantities of mono N-dealkylated metabolites and thio conjugates of atrazine observed. We also used this methodology in a preliminary study of cytochrome P450-catalyzed metabolism of atrazine in vitro. The results obtained in this study suggest that this method will be a useful tool for the determination of atrazine and its metabolites in future pharmacokinetic studies and for the subsequent development and refinement of biologically based models of atrazine disposition.

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

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

- What can WE do to help 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

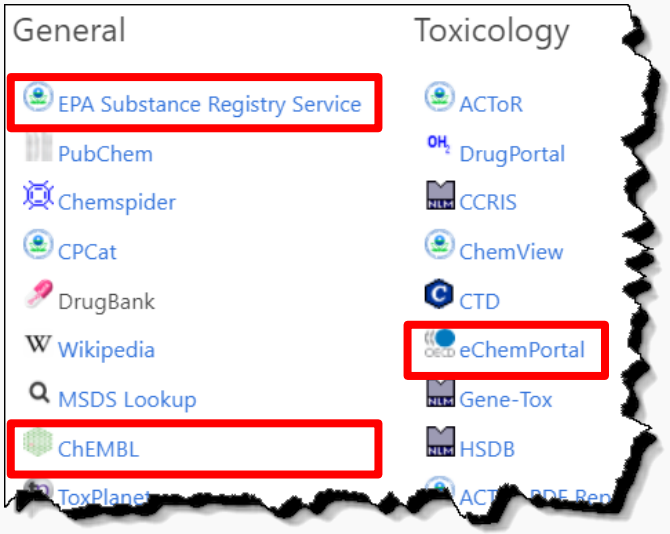
- RSC Analytical Abstracts
- Tox21 Analytical Data
- MONA: MassBank North America
- mzCloud
- NIST IR Spectrum
- NIST MS Spectrum
- MassBank
- NIST Antoine Constants
- IR Spectra on PubChem
- NIST Kovats Index values
- Protein DataBank
- National Environmental Methods Index

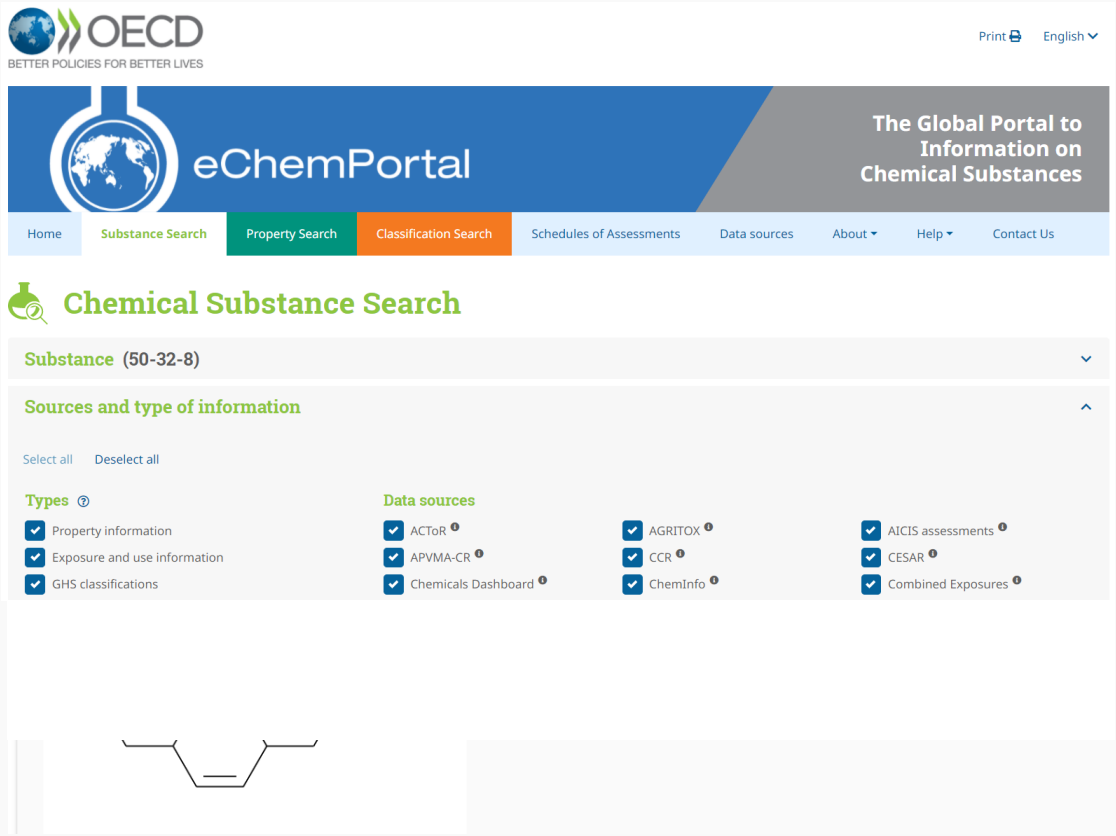
Prediction

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

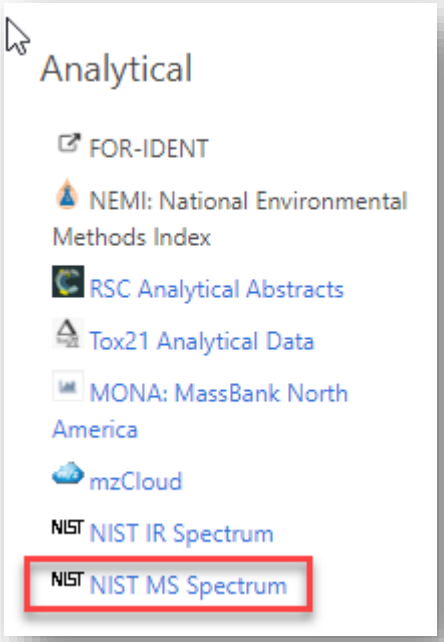
External Links

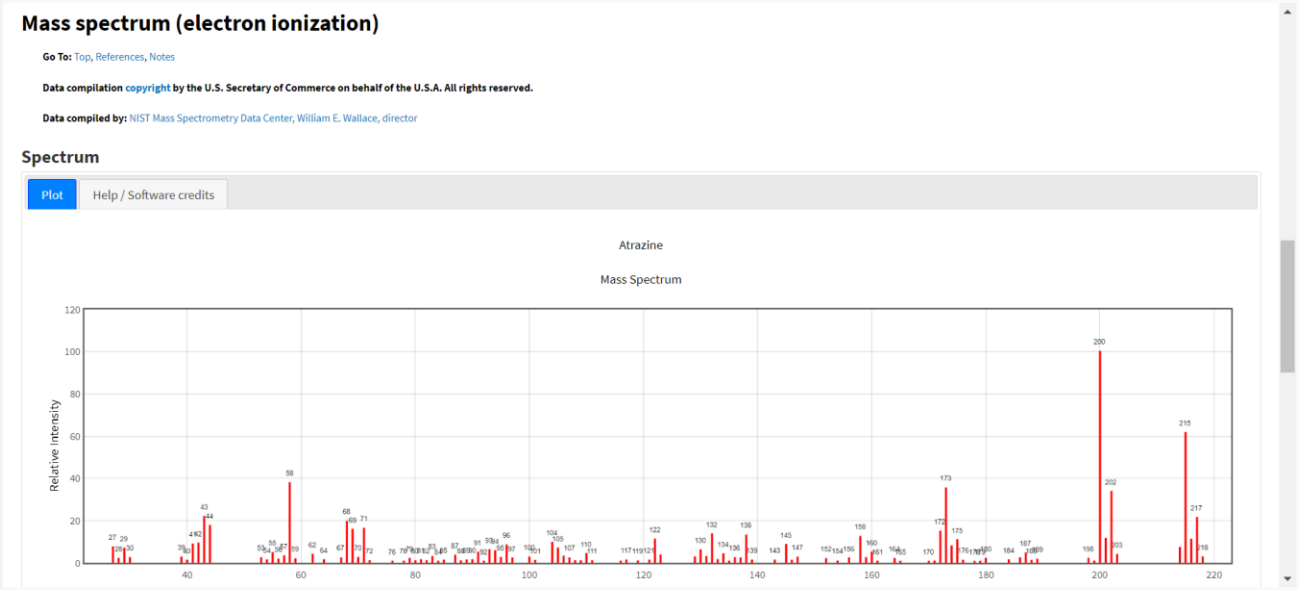
- Links to ~90 websites providing access to additional data on the chemical of interest



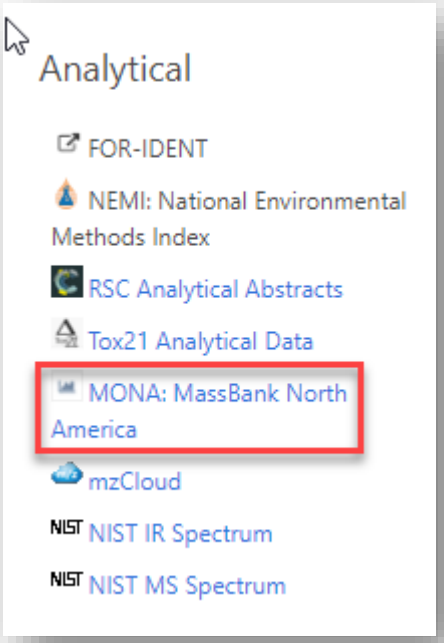


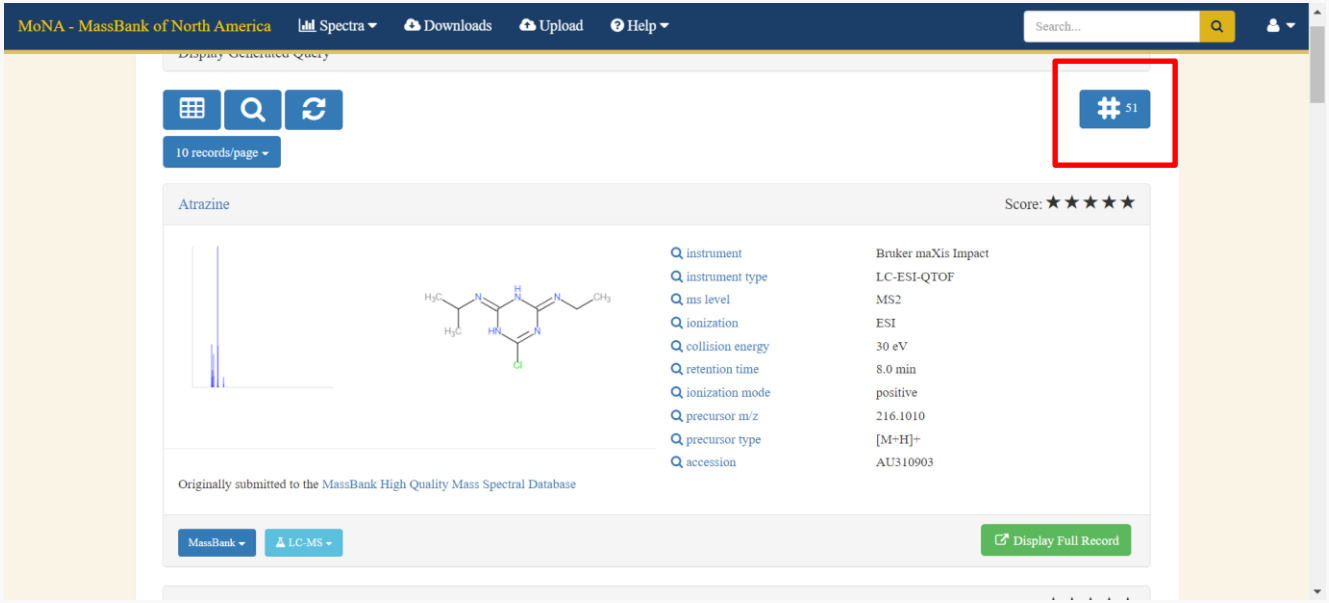
NIST WebBook
https://webbook.nist.gov/chemistry/





MassBank of North America
https://mona.fiehnlab.ucdavis.edu



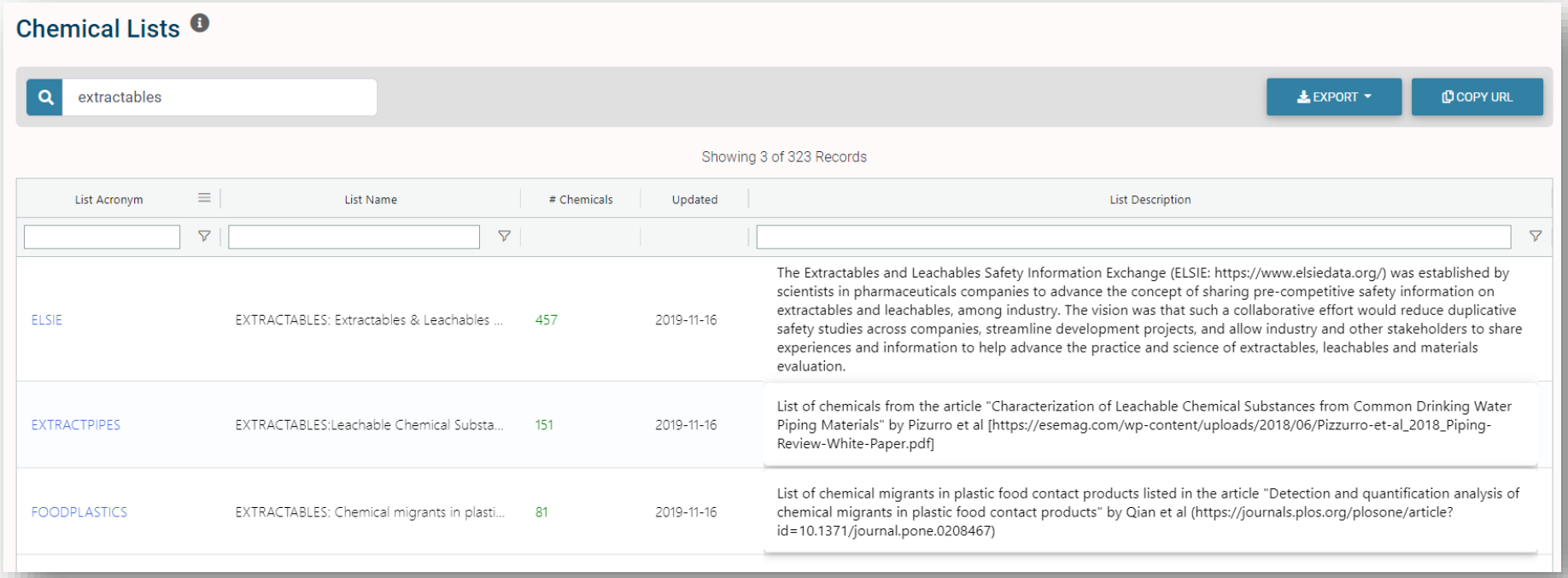


Chemical Lists of Interest…

Lists of Extractables and Leachables

https://comptox.epa.gov/dashboard/chemical-lists?filtered=&search=extractables

- Chemical lists with extractables & leachables
- Expands with literature extraction



The ELSIE Database

https://comptox.epa.gov/dashboard/chemical-lists/ELSIE

EXTRACTABLES: Extractables & Leachables Safety Information Exchange (ELSIE)

Search for chemical by systematic name, synonym, CAS number, DTXSID or InChIKey

Start typing to search.

Identifier substring search

List Details

Description: The Extractables and Leachables Safety Information Exchange (ELSIE: https://www.elsiedata.org/) was established by scientists in pharmaceuticals companies to advance the concept of sharing pre-competitive safety information on extractables and leachables, among industry. The vision was that such a collaborative effort would reduce duplicative safety studies across companies, streamline development projects, and allow industry and other stakeholders to share experiences and information to help advance the practice and science of extractables, leachables and materials evaluation.

Number of Chemicals: 457

Extractables

EXTRACTABLES:Leachable Chemical Substances from Common Drinking Water Piping Materials

Search for chemical by systematic name, synonym, CAS number, DTXSID or InChIKey

Start typing to search.

Identifier substring search

List Details

Description: List of chemicals from the article "Characterization of Leachable Chemical Substances from Common Drinking Water Piping Materials" by Pizurro et al [https://esemag.com/wp-content/uploads/2018/06/Pizzurro-et-al_2018_Piping-Review-White-Paper.pdf]

Number of Chemicals: 151

EXTRACTABLES: Chemical migrants in plastic food contact products

Search for chemical by systematic name, synonym, CAS number, DTXSID or InChIKey

Start typing to search.

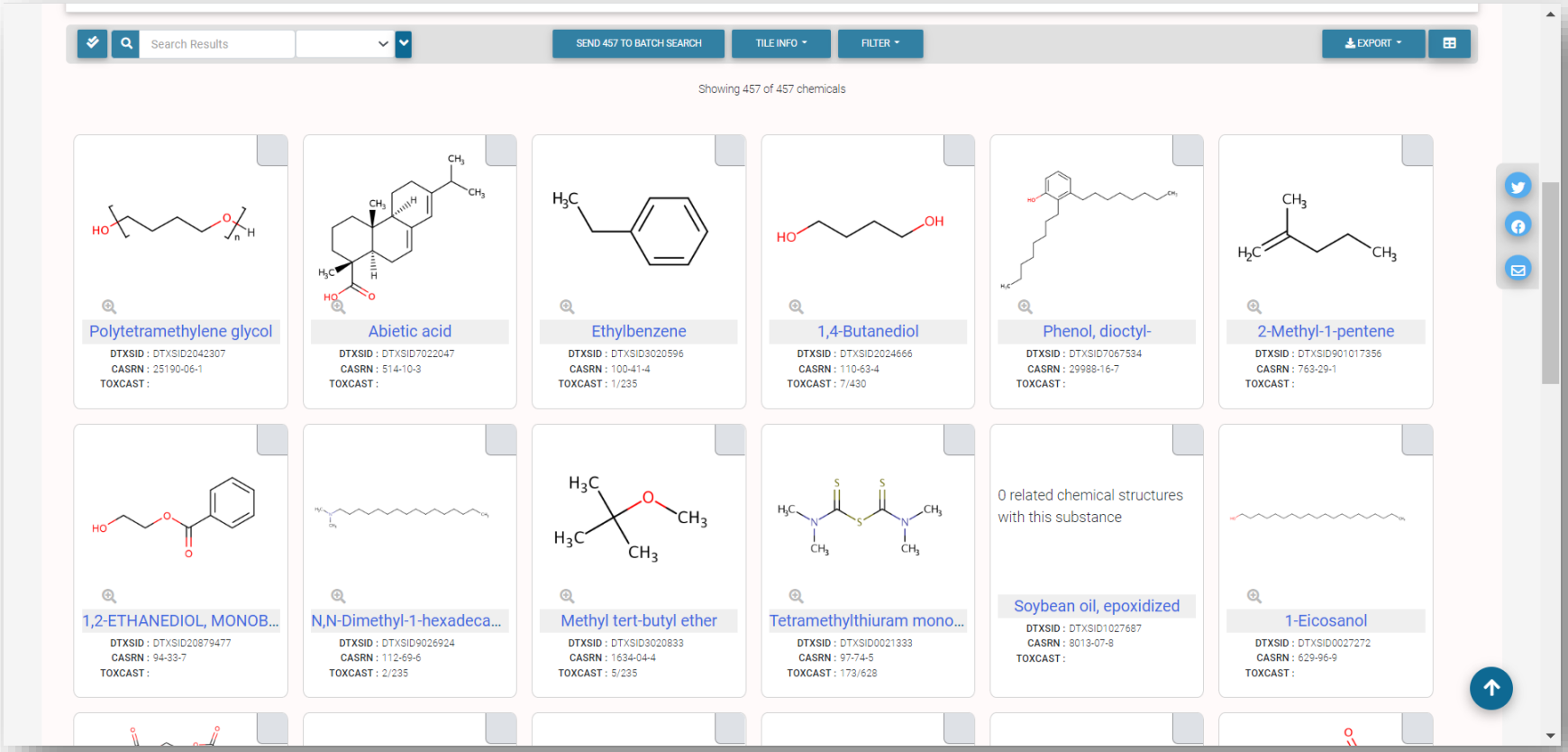
Identifier substring search

List Details

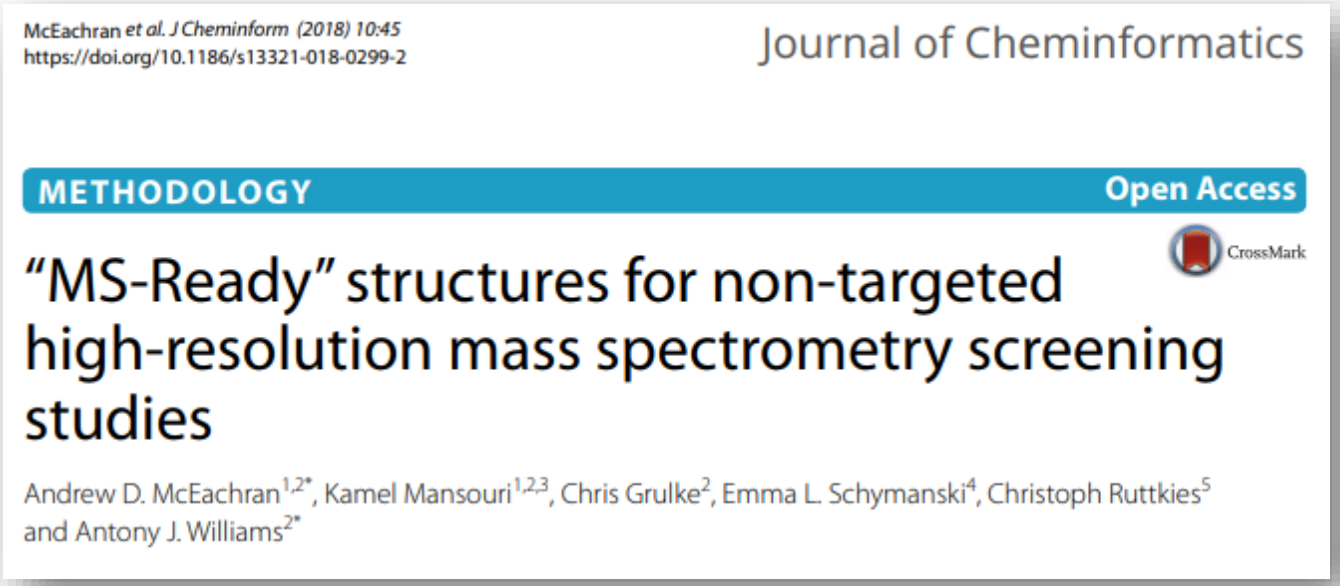
Description: List of chemical migrants in plastic food contact products listed in the article "Detection and quantification analysis of chemical migrants in plastic food contact products" by Qian et al (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0208467)

Number of Chemicals: 81

Download data, send to batch search for data harvesting



"MS-ready" structures



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

Journal of Cheminformatics

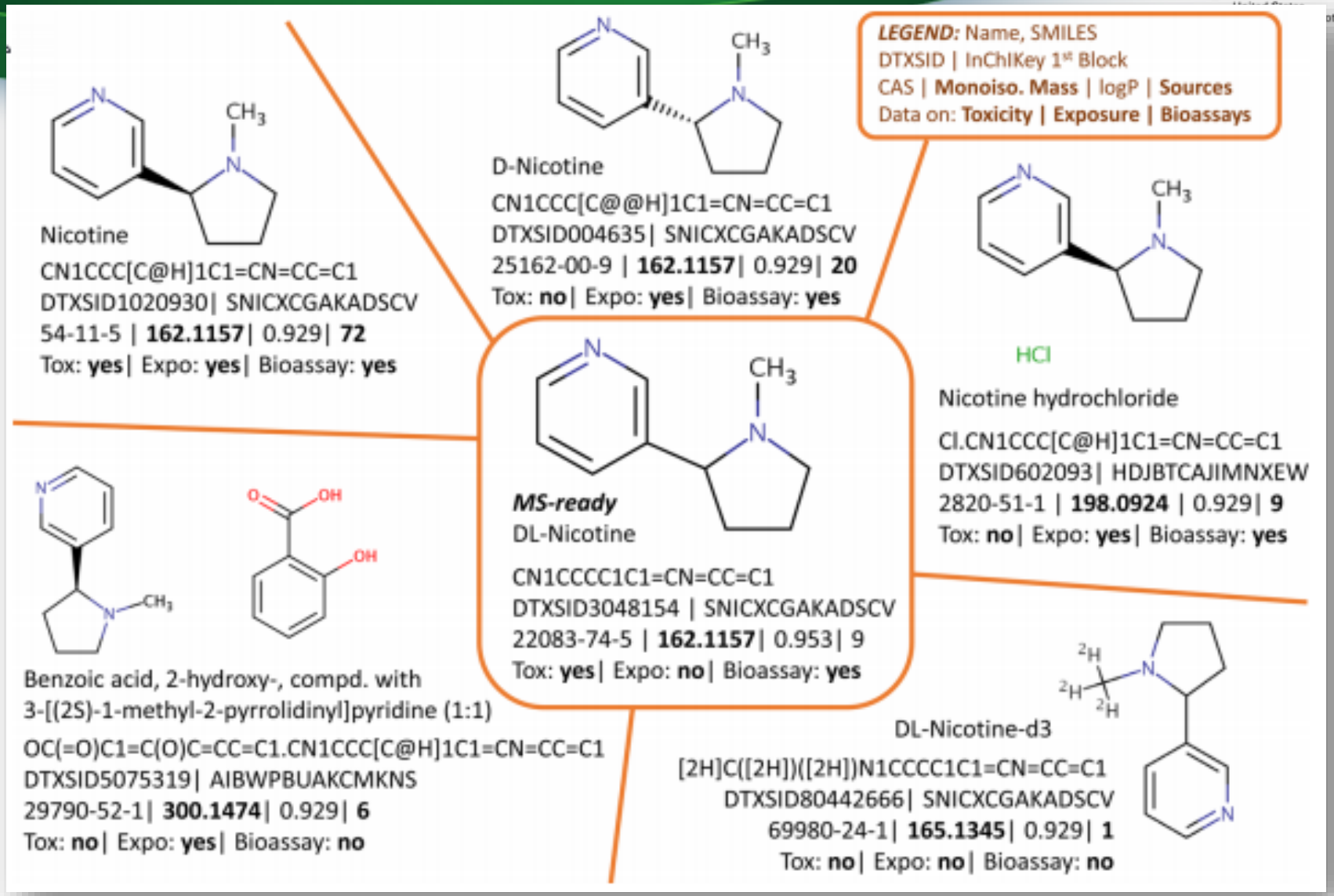
METHODOLOGY Open Access

"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*]

Overview of MS-Ready Structures

- All structure-based chemical substances are algorithmically processed to
  - Split multicomponent chemicals into individual structures
  - Desalt and neutralize individual structures
  - Remove stereochemical bonds from all chemicals
- MS-Ready structures are then mapped to original substances to provide a path between chemicals detected by mass spectrometry to original substances



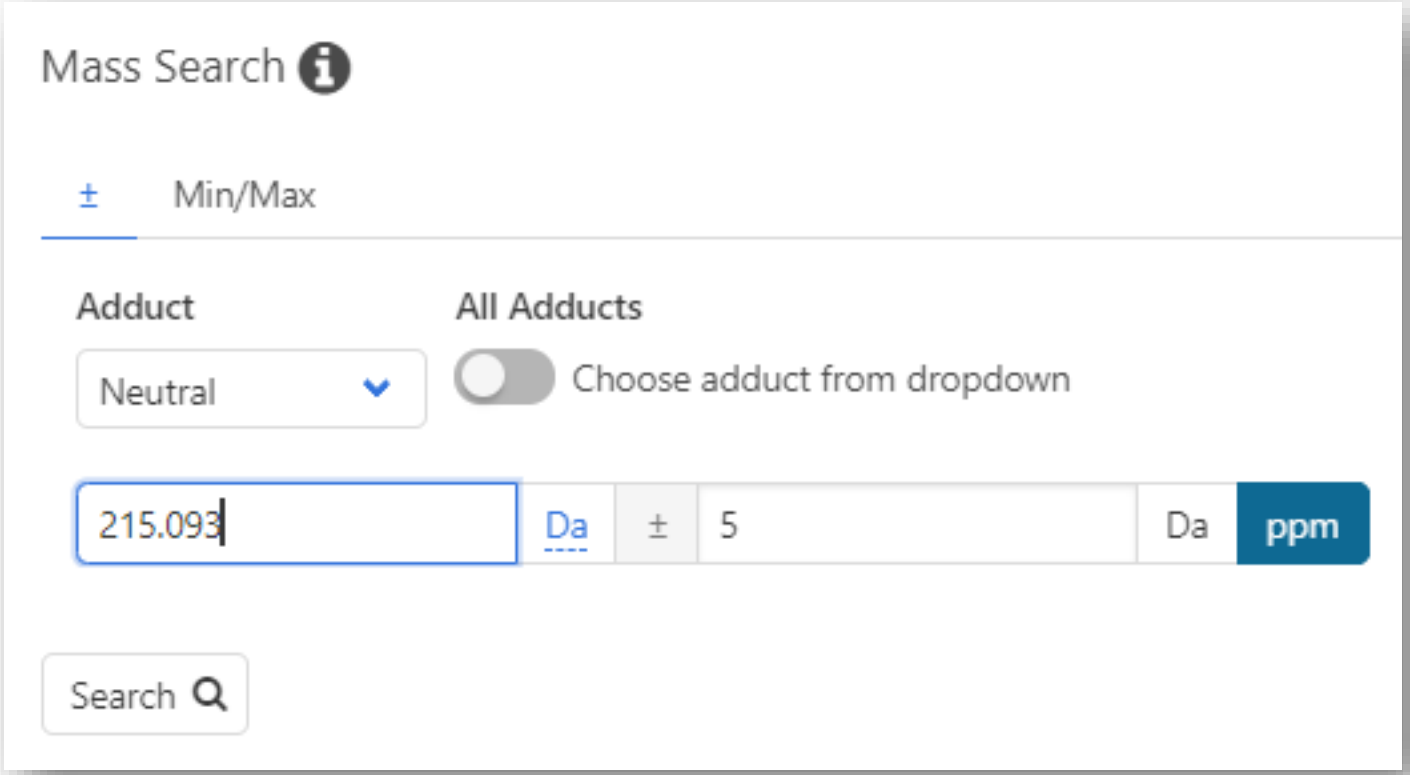
Environmental Science & Technology
Viewpoint
pubs.acs.org/est

Open Science for Identifying "Known Unknown" Chemicals

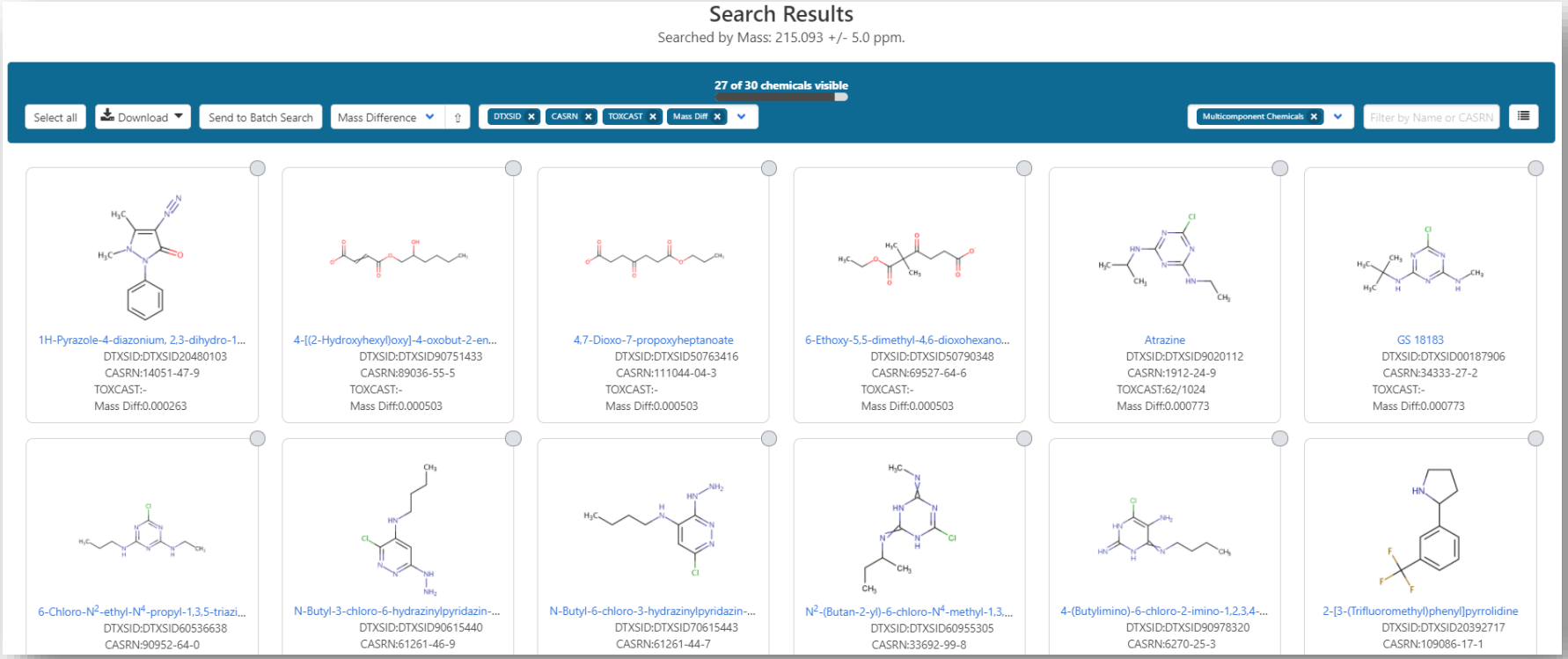
Emma L. Schymanski*,† and Antony J. Williams*,‡

Mass and Formula Searching

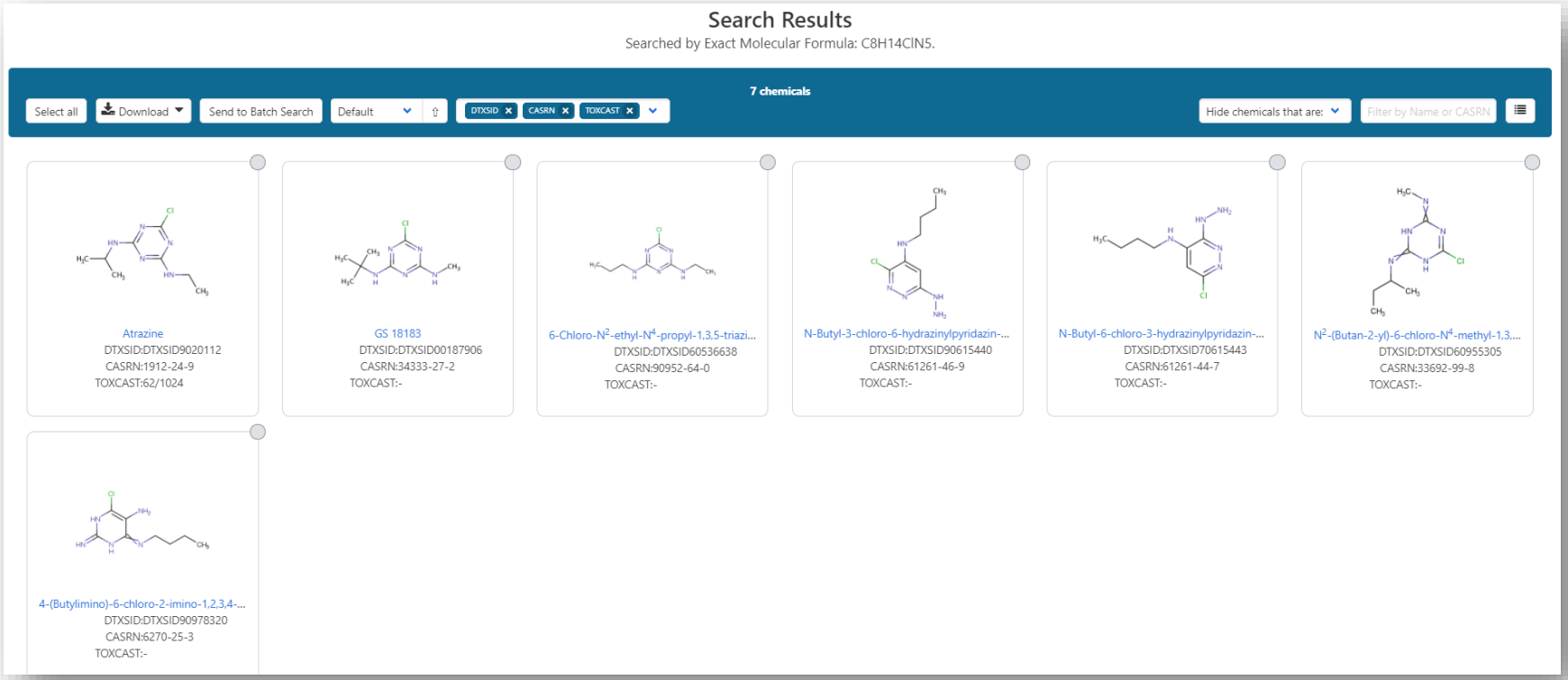
Advanced Searches: **Mass** Search



Advanced Searches: **Mass** Search



Advanced Searches: **Formula** Search



EPA
United States
Environmental Protection
Agency

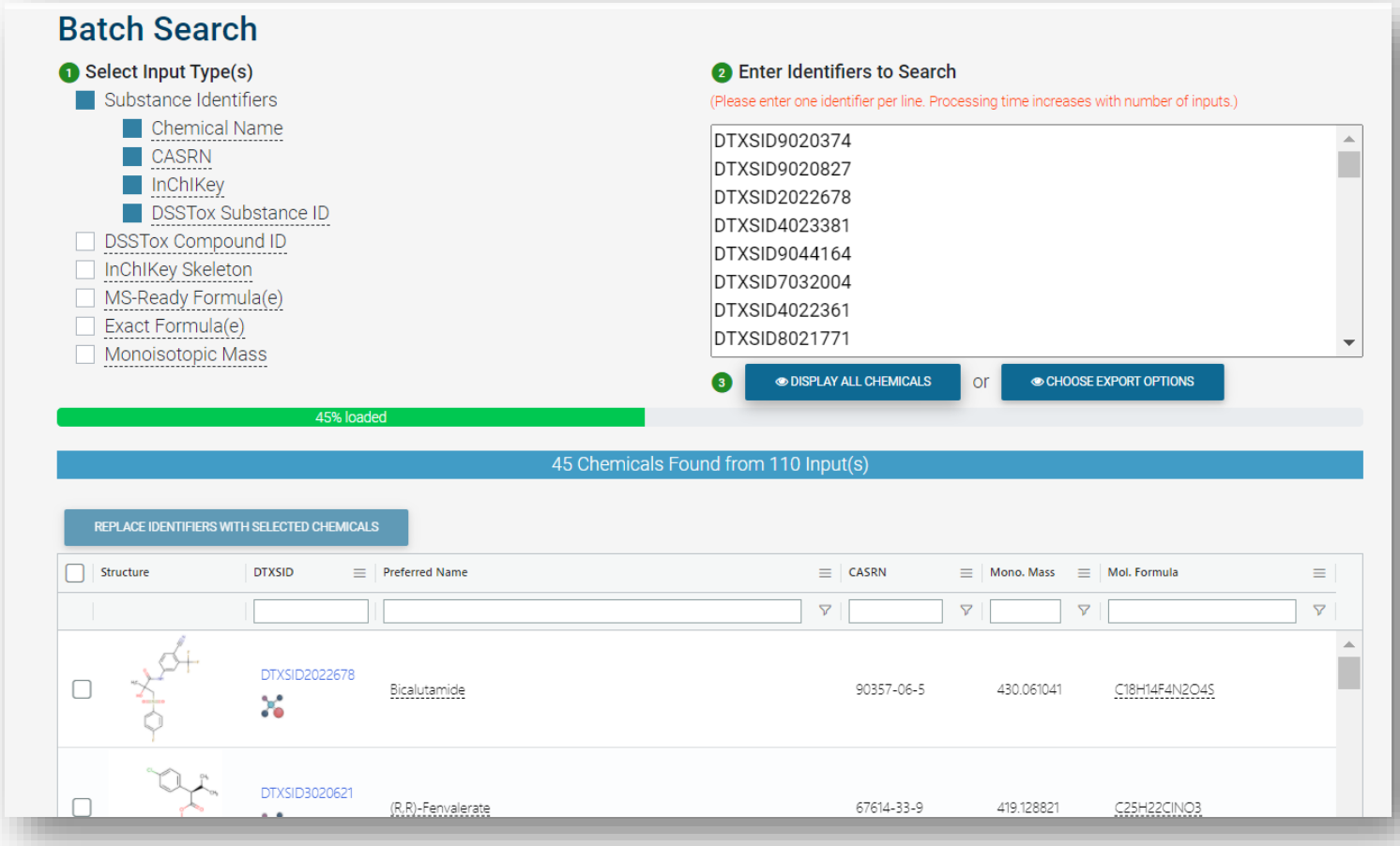
Batch Searching

Batch Searching

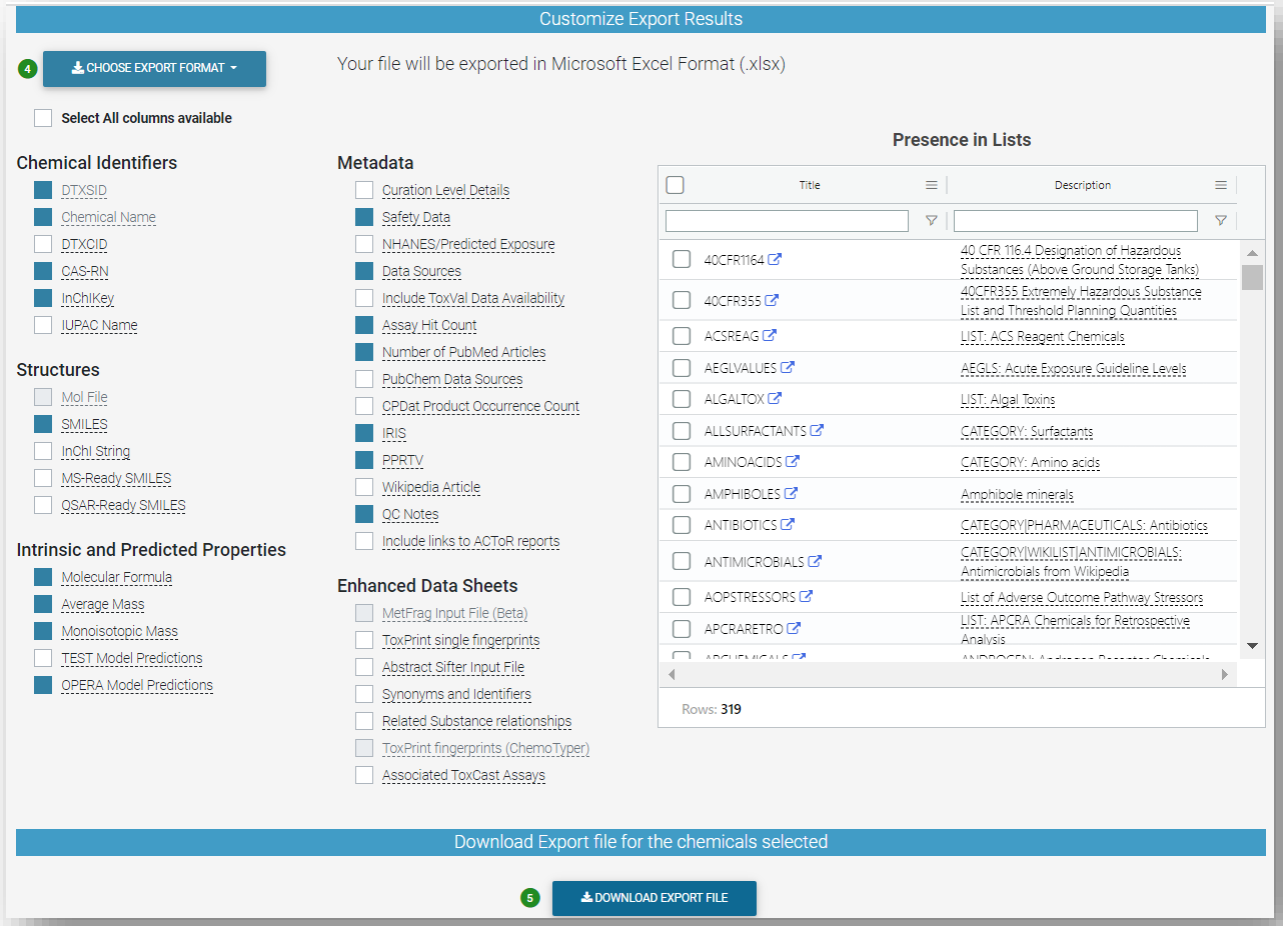
- Singleton searches are useful but we work with **thousands** of masses and formulae!

- Typical questions
  - What is the list of chemicals for the formula $C_xH_yO_z$
  - What is the list of chemicals for a mass +/- error
  - Can I get chemical lists in Excel files? In SDF files?
  - Can I include properties in the download file?

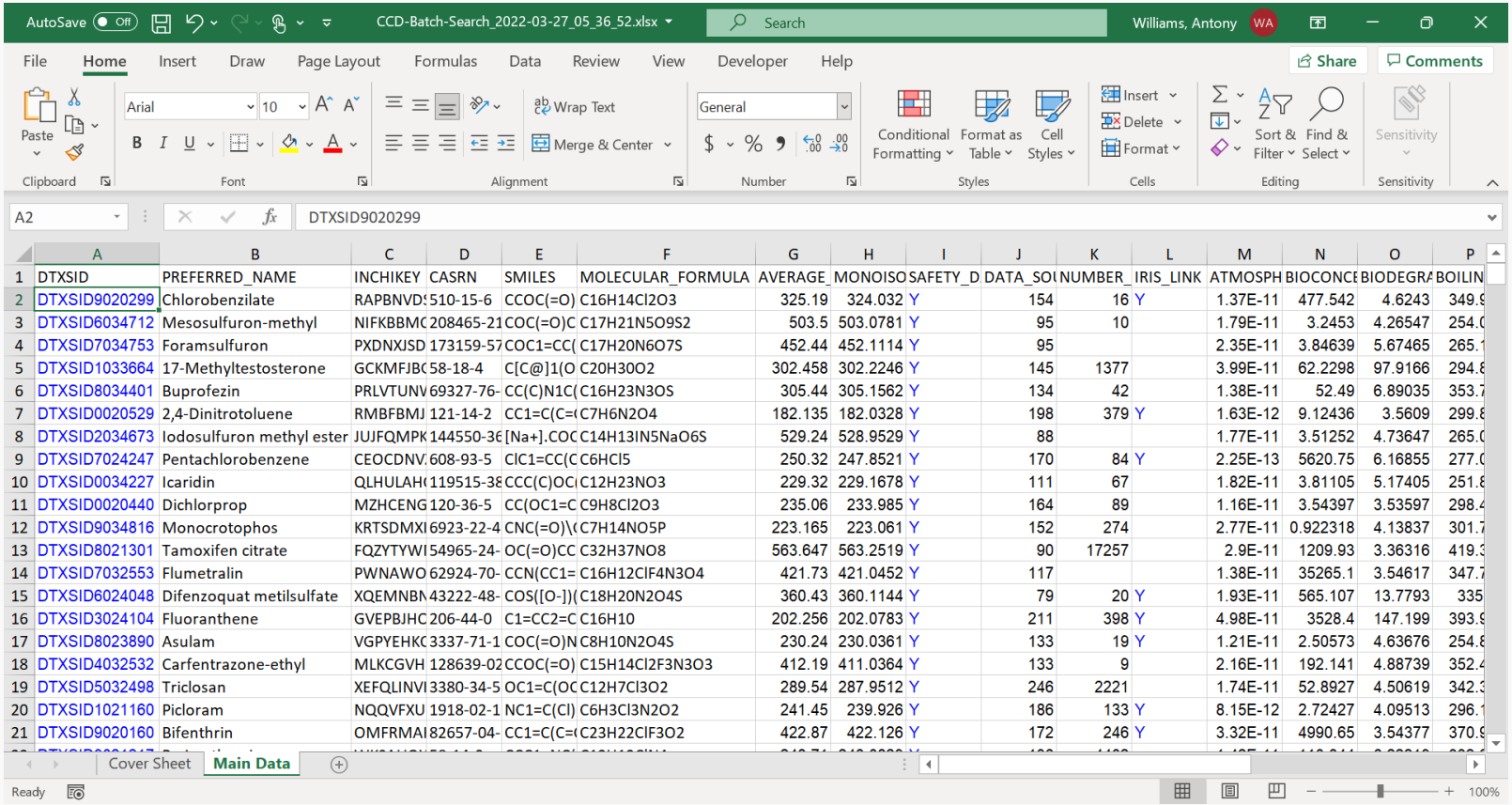
Batch Search



Batch Search – Excel, CSV, SDF file



Batch Search

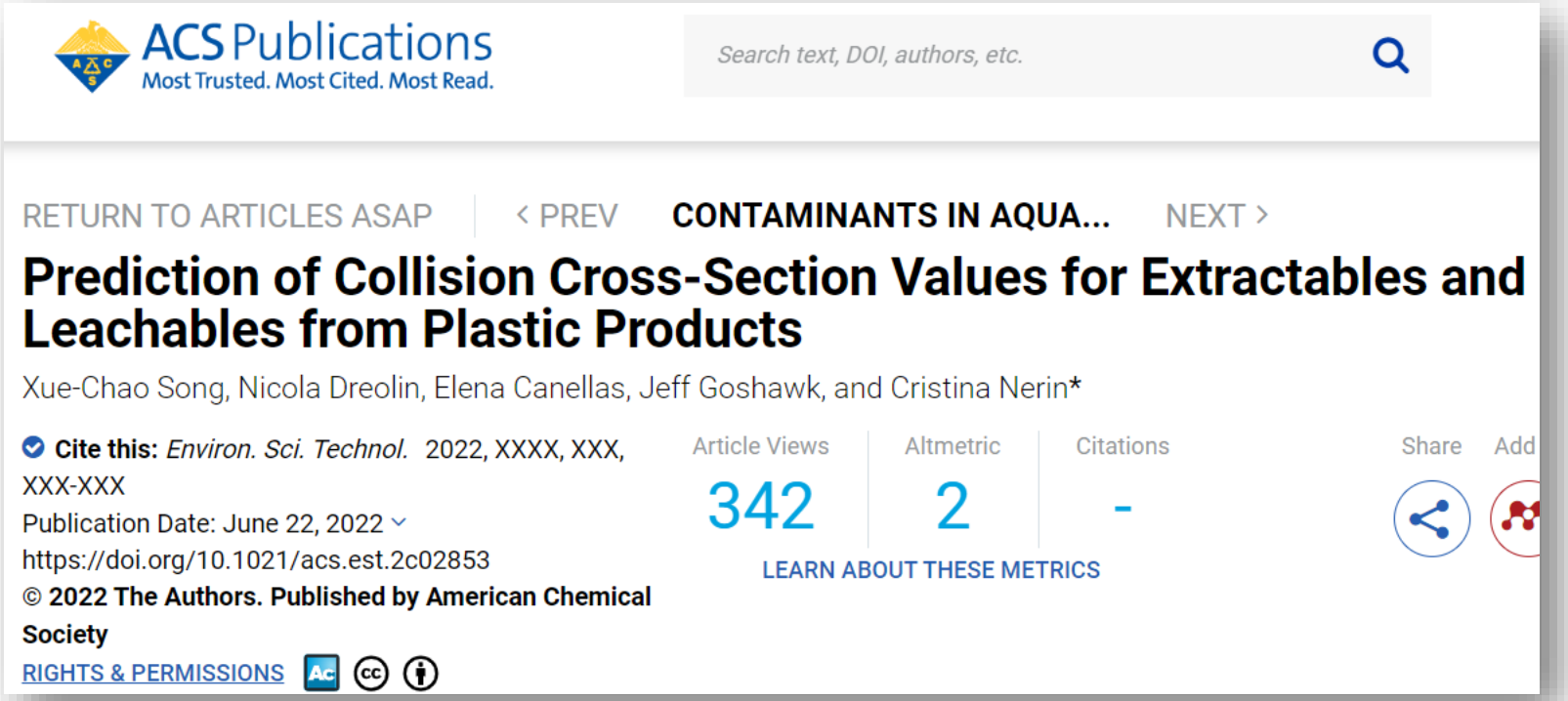


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	DTXSID	PREFERRED_NAME	INCHIKEY	CASRN	SMILES	MOLECULAR_FORMULA	AVERAGE_	MONOISO	SAFETY_D	DATA_SOI	NUMBER_	IRIS_LINK	ATMOSPH	BIOCONCE	BIODEGRA	BOILIN
2	DTXSID9020299	Chlorobenzilate	RAPBNVD	510-15-6	CCOC(=O)	C16H14Cl2O3	325.19	324.032	Y	154	16	Y	1.37E-11	477.542	4.6243	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	XEFQLINV	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

Work in Progress

New lists added regularly
In progress

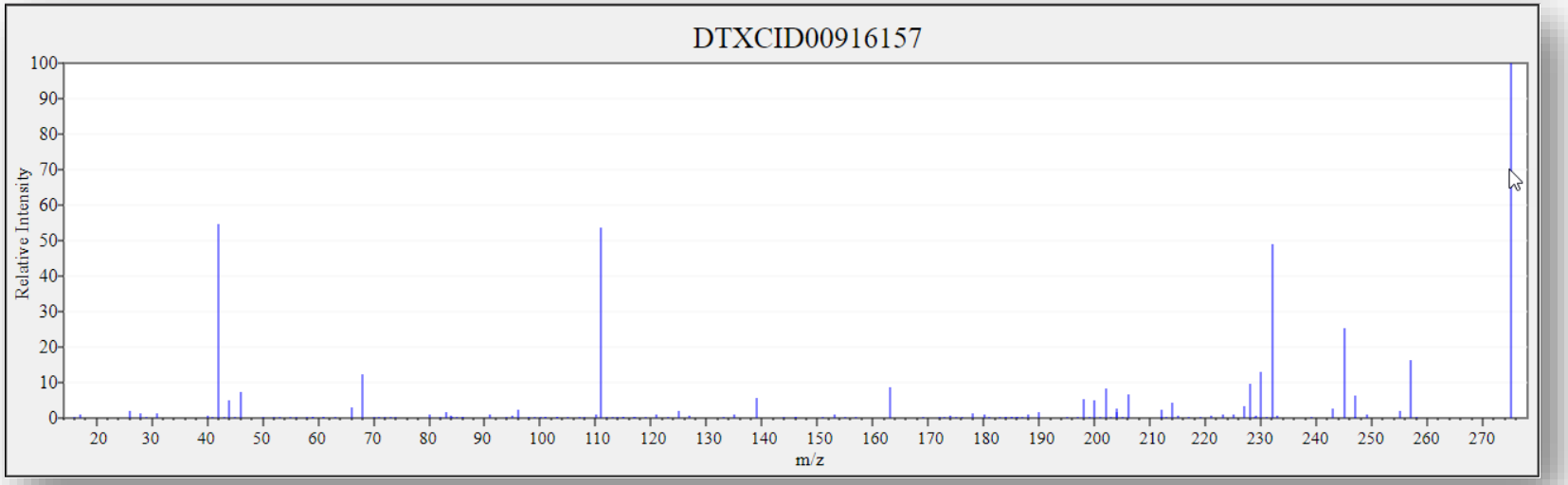
- New chemicals registered daily and released with each new version of the Dashboard



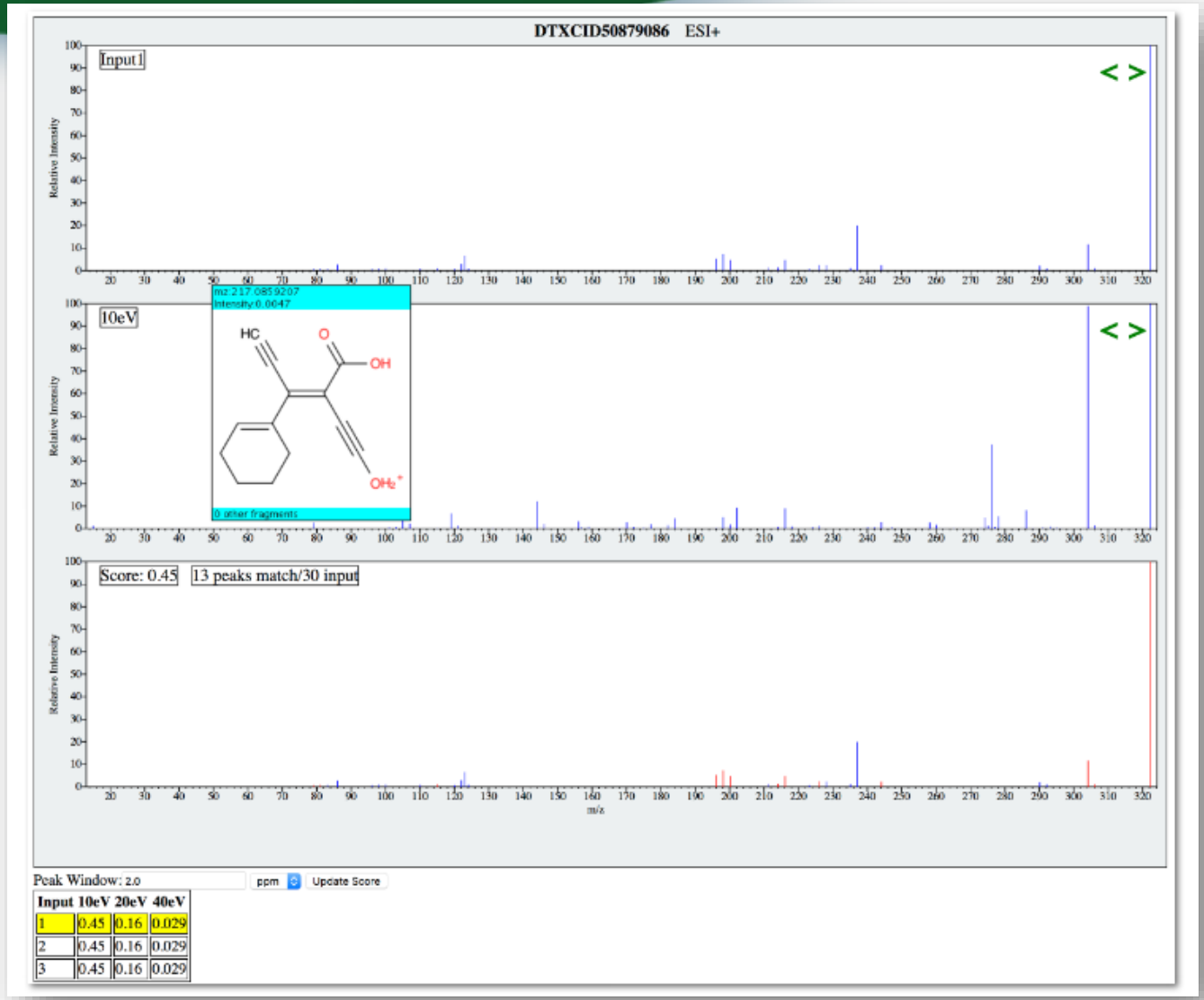
Predicted Mass Spectra

http://cfmid.wishartlab.com/

- MS/MS spectra prediction for ESI+, ESI-, and EI
- Predictions generated and stored for >800,000 structures, to be accessible via Dashboard



Spectral Viewer Comparison



Already published…

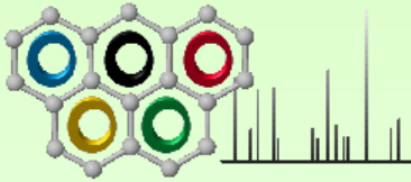
Analytical and Bioanalytical Chemistry
https://doi.org/10.1007/s00216-019-02351-7

In silico MS/MS spectra for identifying unknowns: a critical examination using CFM-ID algorithms and ENTACT mixture samples

Alex Chao [1,2] · Hussein Al-Ghoul [1,2] · Andrew D. McEachran [1,3] · Ilya Balabin [4] · Tom Transue [4] · Tommy Cathey [4] · Jarod N. Grossman [2,3] · Randolph Singh [1,5] · Elin M. Ulrich [2] · Antony J. Williams [6] · Jon R. Sobus [2]

Received: 4 October 2019 / Revised: 27 November 2019 / Accepted: 11 December 2019

CASMI



CASMI 2017

CASMI 2017
- Important Dates
- Contest Rules
- Challenge Data
- Solutions
- Preliminary results
- Results
- About the Team

CASMI 2016

CASMI 2014

CASMI 2013

CASMI 2012

Critical Assessment of Small Molecule Identification

The experimental and computational mass spectrometry communities are invited to participate in the fifth round of an open contest on the identification of small molecules from mass spectrometry data.

This year the contest will test the applicability of MS and MS/MS on natural products chemistry identifications. With 45 (Category 1) and up to 243 (Categories 2&3) natural products challenges - including a few tricky ones - there's something for everyone!

CASMI 2017 is organised by Dr. Dejan Nikolic (University of Illinois at Chicago, USA), Dr. Nir Shahaf (Weizmann Institute of Science, Israel), Dr. Emma Schymanski (Eawag, Switzerland) and Dr. Steffen Neumann (IPB Halle, Germany).

Mailing lists

CASMI

metabolites

MDPI

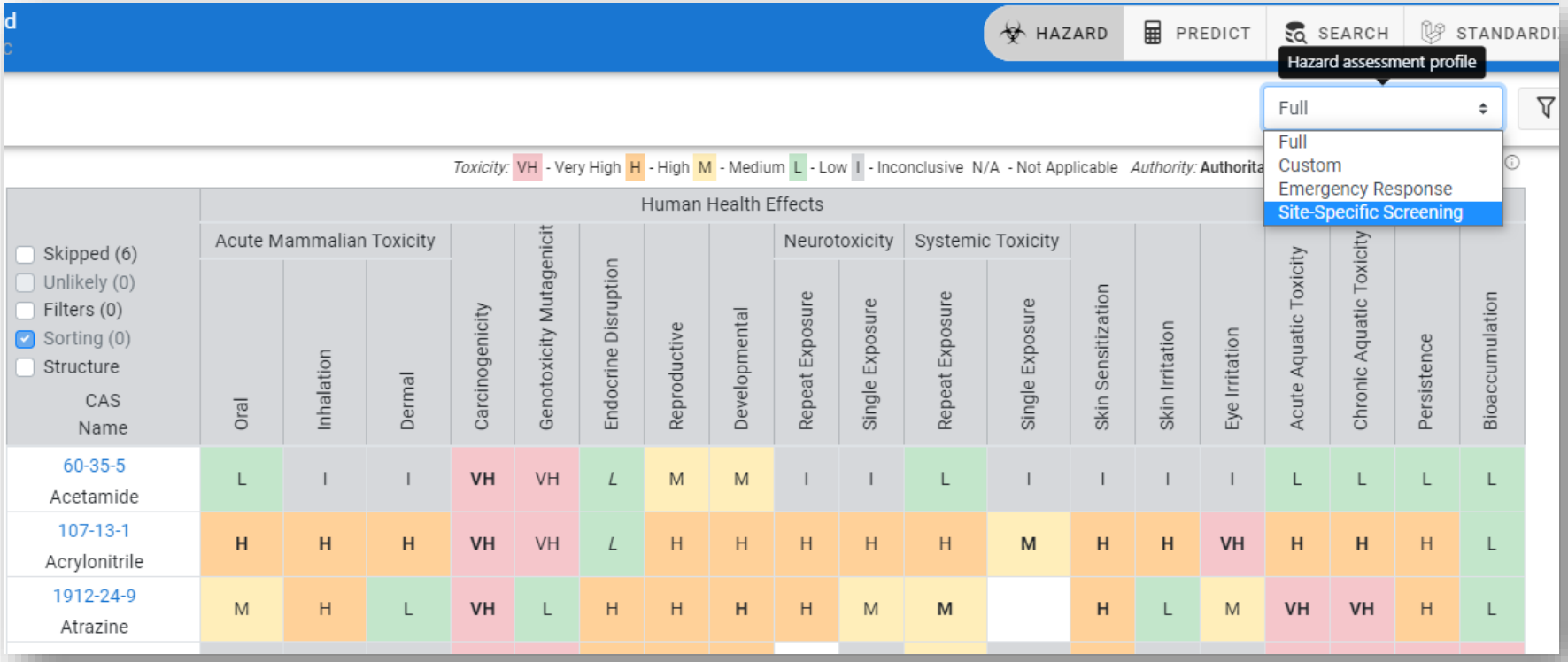
Article

Revisiting Five Years of CASMI Contests with EPA Identification Tools

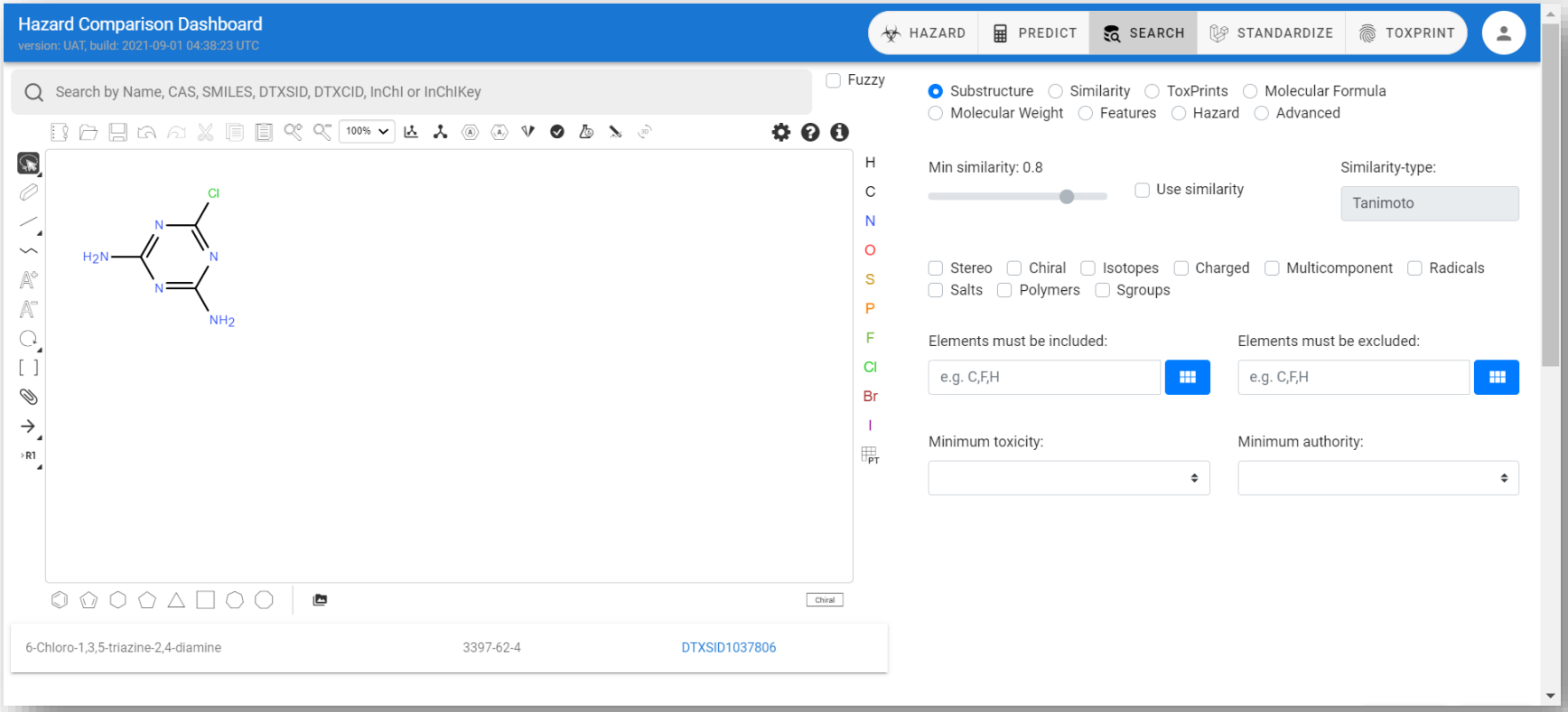
Table 2. Percentage of the total number of compounds from each CASMI contest year that were ranked in the top 5 by Competitive Fragmentation Modeling for Metabolite Identification (CFM-ID) only and by the summation of CFM-ID and DSSTox Data Source Counts (DS), alongside the percentage in the top 5 reported by the contest years' winning entry. Complete ranking results are provided in Supplemental File S1.

CASMI Year	CFM-ID Only	CFM-ID + DS	Winners' Results [1]	Total in DB/Total in Dataset [2]
2012	36%	64%	36%	14/14
2013	81%	88%	88%	16/16
2014	57%	76%	71%	42/42
2016-training	63%	96%		312/312
2016-challenge	66%	94%	81%	208/208
2017	59%	53%	74% [3]	227/243

Hazard Module



Structure/Substructure/Similarity



Database of Public Domain MS Data

- Metadata ranking of candidates based on mass/formula searching was a starting point
- Searching experimental spectra against *in silico* predictions was next.
- Searching experimental spectra against experimental spectra is the next phase.
- Assembling experimental mass spec data from the internet, homogenize formats and database – includes curation

Database of MS Methods

- There are hundreds of methods distributed across the EPA website – that can be enabled by cheminformatics
  - Focus: Aggregate MS method documents and extract chemicals to make methods structure searchable
  - Vision: Search by structure/substructure/similarity to find existing method(s) as a starting point

Work in Progress

https://acs.digitellinc.com/acs/sessions/445216/view

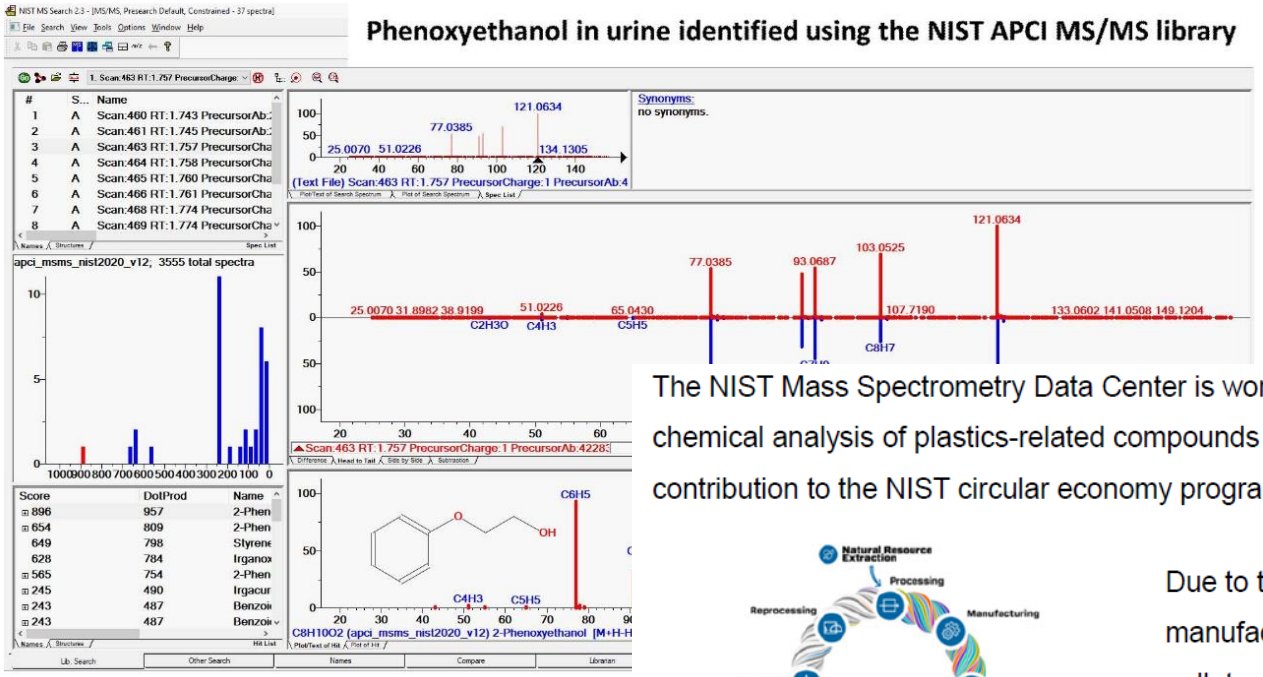
- Presently working with scientists from NIST

- Dataset to be released as a list
- Spectra will be added to our experimental spectral database for searching

Collaboration with NIST
>10,000 plastic related chemicals

We have identified phenoxyethanol, several phthalates and phthalate metabolites, and bisphenol A in pooled human urine samples (NIST/SRMs), together with other contaminants in need of further analysis. A similar study was conducted for extracts from orthopedic casts using 16 polymer extracts.

Phenoxyethanol in urine identified using the NIST APCI MS/MS library

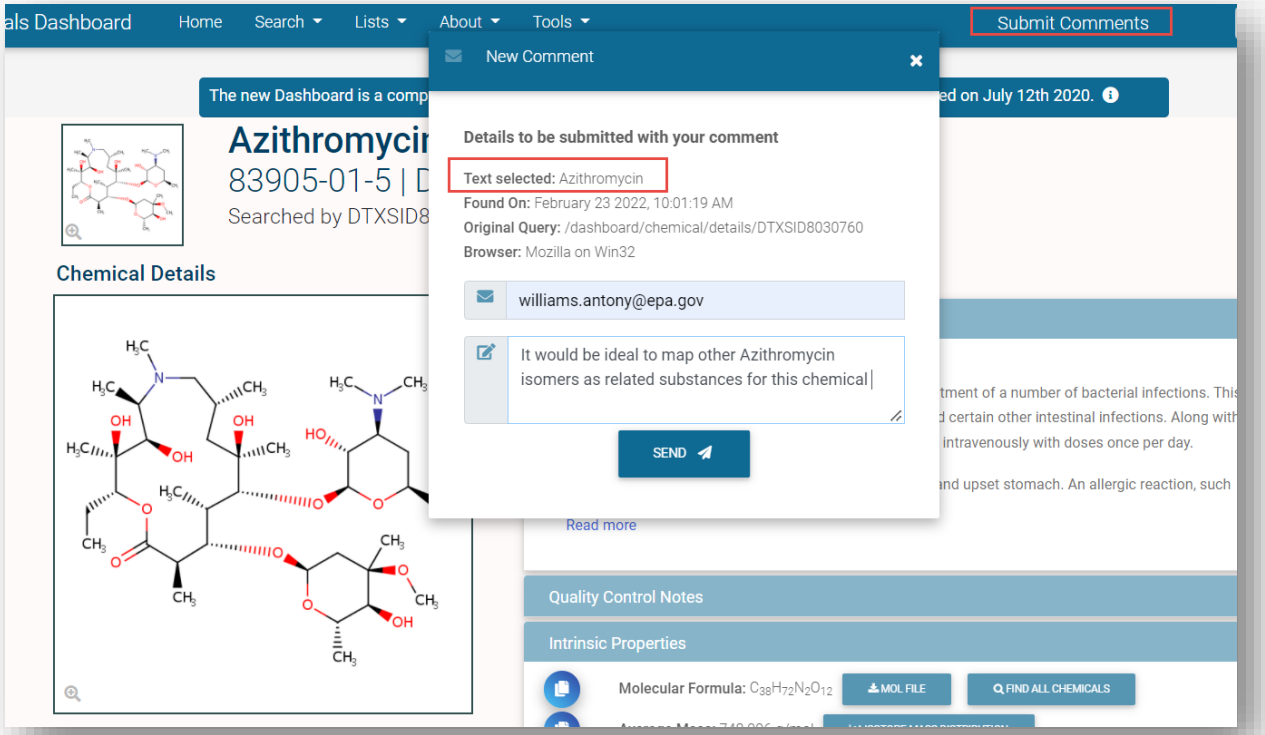


The NIST Mass Spectrometry Data Center is working on a comprehensive approach to the chemical analysis of plastics-related compounds (PRC) using mass spectrometry as a contribution to the NIST circular economy program (https://www.nist.gov/circular-economy).

Due to the versatility of plastics, many manufactured products and environmental pollutants are associated with PRC. 10,547 compounds including monomers, additives, and processing aids of the plastics industry (Environ. Sci. Technol. 2021, 55,9339-9351) and three types of ionizations has been used in this work to obtain mass spectra and build standard libraries of PRC.

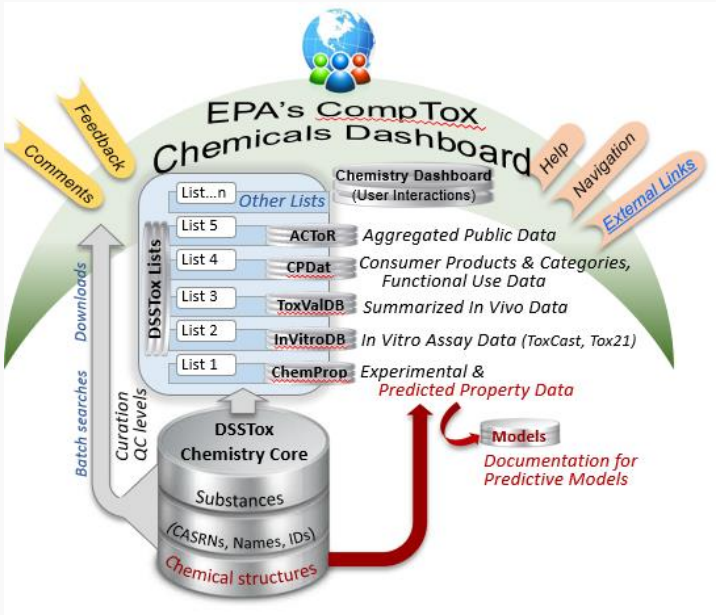
Request for participation

- Please point us to relevant datasets and articles
- Nothing is perfect – so please flag issues



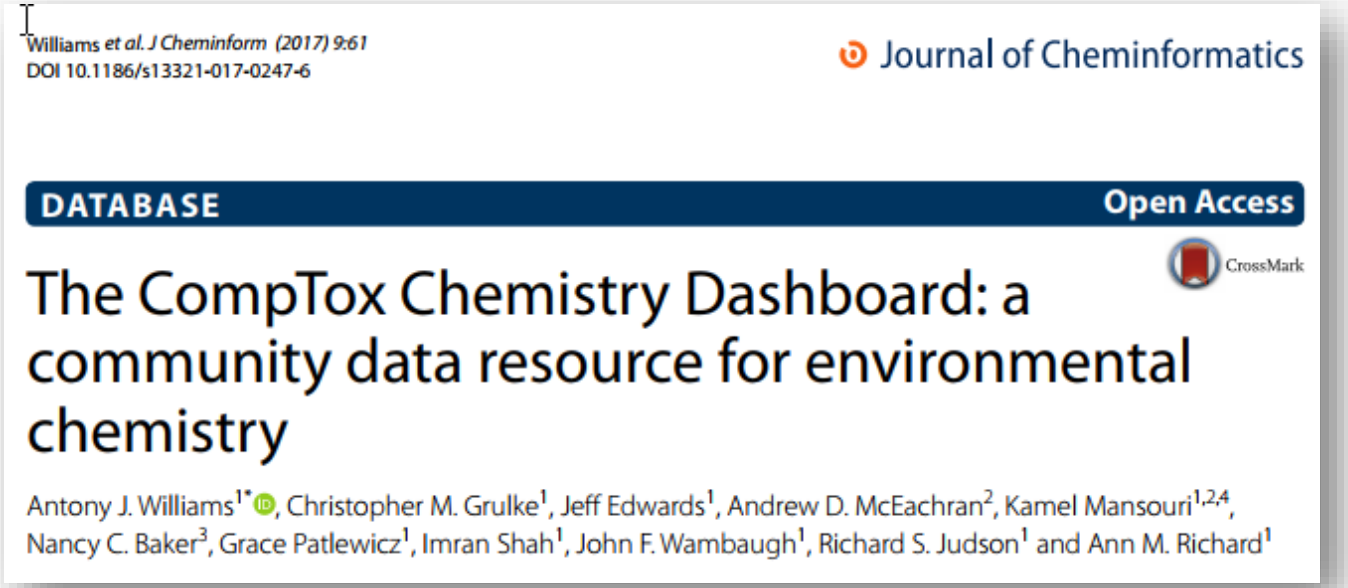
Conclusion

- Dashboard access to data for ~900,000 chemicals, next release will update to **1.2 million** substances
- Extractable and leachables lists continue to expand
- “MS-Ready data” facilitates structure identification
- Data continues to grow with ongoing curation activities
- Proof-of-concept developments will release in future versions
- ***Do you want to learn more?***



Contact

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



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