Abstract Sifter User Guide, Version 3.0

Nancy C. Baker[1], Thomas Knudsen[2], Antony J. Williams[2]

Availability: The Abstract Sifter and documentation is freely available for download at ftp://newftp.epa.gov/COMPTOX/Sustainable_Chemistry_Data/Chemistry_Dashboard/Abstract_Sifter/ .

[1]Leidos, Research Triangle Park, NC, 27711, USA and [2]National Center for Computational Toxicology, U.S. Environmental Protection Agency, Research Triangle Park, NC, 27711, USA.

Contact: baker.nancy@epa.gov

Disclaimer: The views expressed in this paper are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

Abstract Sifter User Guide – Version 3.0

This user guide describes the functionality of the PubMed Abstract Sifter. The reader is invited to download the tool from the freely accessible ftp site and follow along:

ftp://newftp.epa.gov/COMPTOX/Sustainable_Chemistry_Data/Chemistry_Dashboard/Abstract_Sifter/

This document provides guidance on the use of the Abstract Sifter through a series of screen shots showing the most common tasks in the tool followed by some helpful tips.

First open the Abstract Sifter file AbstractSifter_v3.xlsm. A security warning may appear. If so, be sure to enable content as shown in Figure 1.

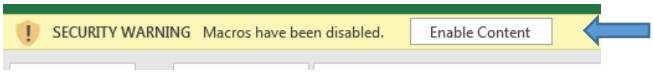


Figure 1. Enable macros upon opening

Once open you will see that the Abstract Sifter Excel file consists of seven sheets. Each sheet is described briefly in the table below.

Sheet name	**Sheet Function**
ReadMe	Basic information on the sifter with links to more documentation
Main	Starting point for PubMed queries and for sifting
Abstract	The sheet where the citation abstract is shown
Notes	Notes and tags are inserted here
Log	Log of every query run on Main sheet
PathwayQueries	Repository of sample queries to use in research disease or treatment pathways (e.g., AOPs)
SampleQueries	Repository of sample queries to use as starting points

Landscape	High level view of literature for entities

Main Sheet

The Main sheet is where the basic functionality of Abstract Shifter occurs, including functions we call "sifting". To begin using the Abstract Sifter, the end-user clicks on the *Query PubMed* button at the top of the screen in the Main sheet. A form is displayed in which the end-user types a PubMed query of interest (Figure 2). In the example, we are showing a very simple query: "chlorpyrifos", but these queries can be more complex. The text that the user enters into the box is sent to PubMed, so all PubMed syntax rules apply. (For a review of this syntax, visit https://www.ncbi.nlm.nih.gov/books/NBK3827/)

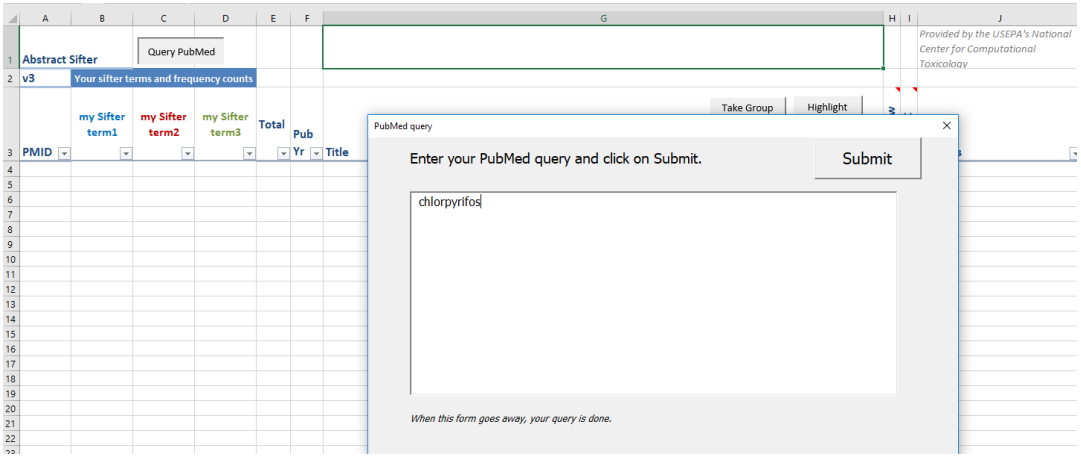


Figure 2. Running a PubMed query

When finished entering the query, the user clicks on *Submit* and the query is packaged by a Visual Basic Application (VBA) into an e-utility command that is passed to the NCBI (National Center for Biotechnology Information) web services. (Note that using Sifter Query PubMed capability requires internet access.) The first response returned by the utility is the number of articles found. (Figure 3) This

number is displayed, and the user is asked if he/she want to continue. If the number of articles is over 5,000, the query will not be run and the user is encouraged to refine the query to return fewer records.

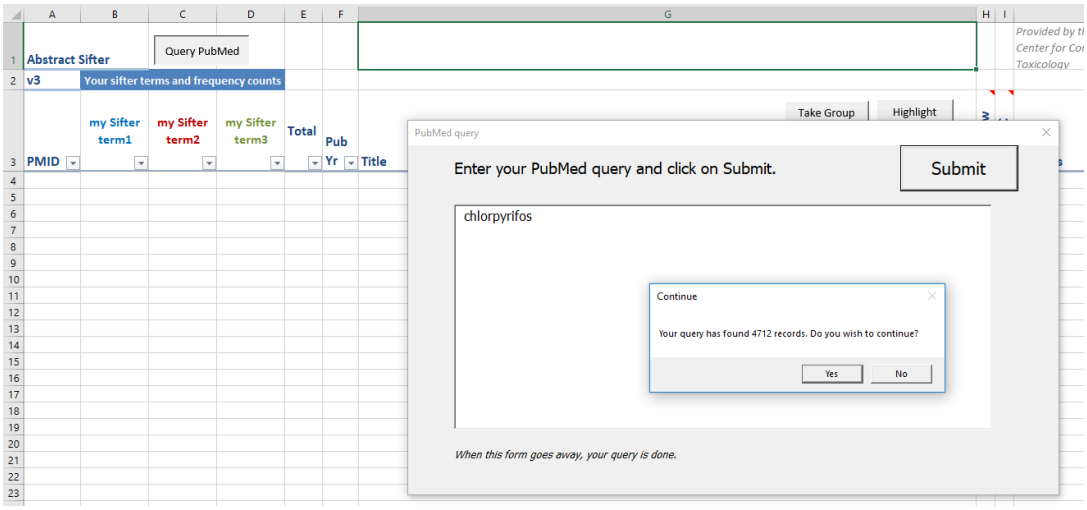


Figure 3. Responding to PubMed

If the returned results are fewer than 6,000 and the user indicates he/she wants to continue, the articles are downloaded from NCBI by Excel, and regular expressions are used to parse the citations for title, abstract, authors, publication year, journal, and PubMed identifier. Each record returned is inserted into a row in the Main sheet. Any rows in the Main sheet from a previous query are deleted.

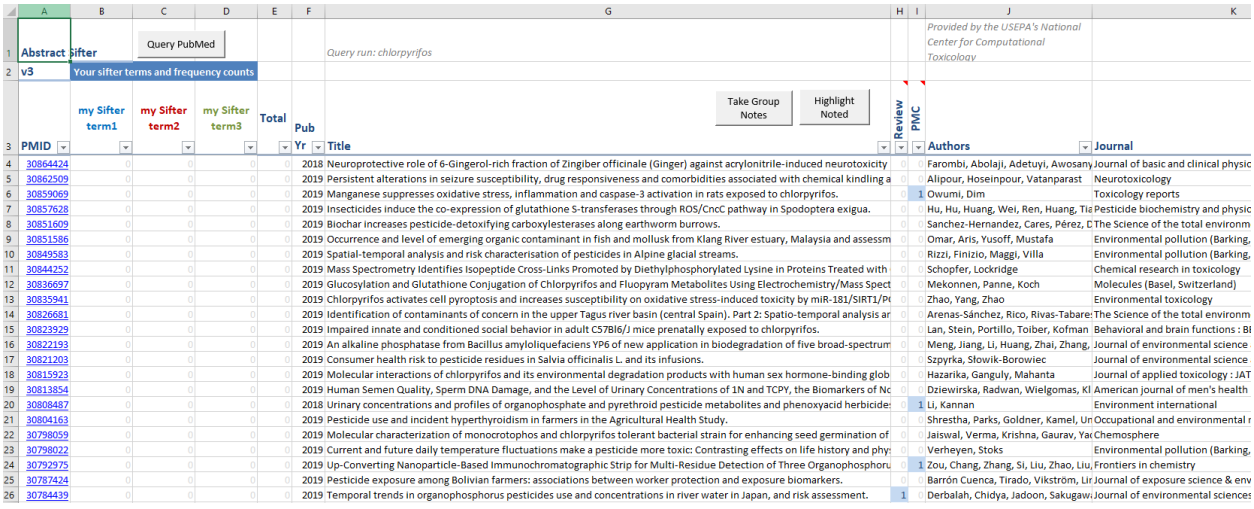


Figure 4.Results from PubMed query - before sifting

At this point the results of the query are stored in the Main sheet and can be browsed like any other data in a spreadsheet (Figure 4); however, the most effective way to find articles of interest is to use the innovative sifter functionality. To demonstrate this functionality, we will continue to use our example of chlorpyrifos.

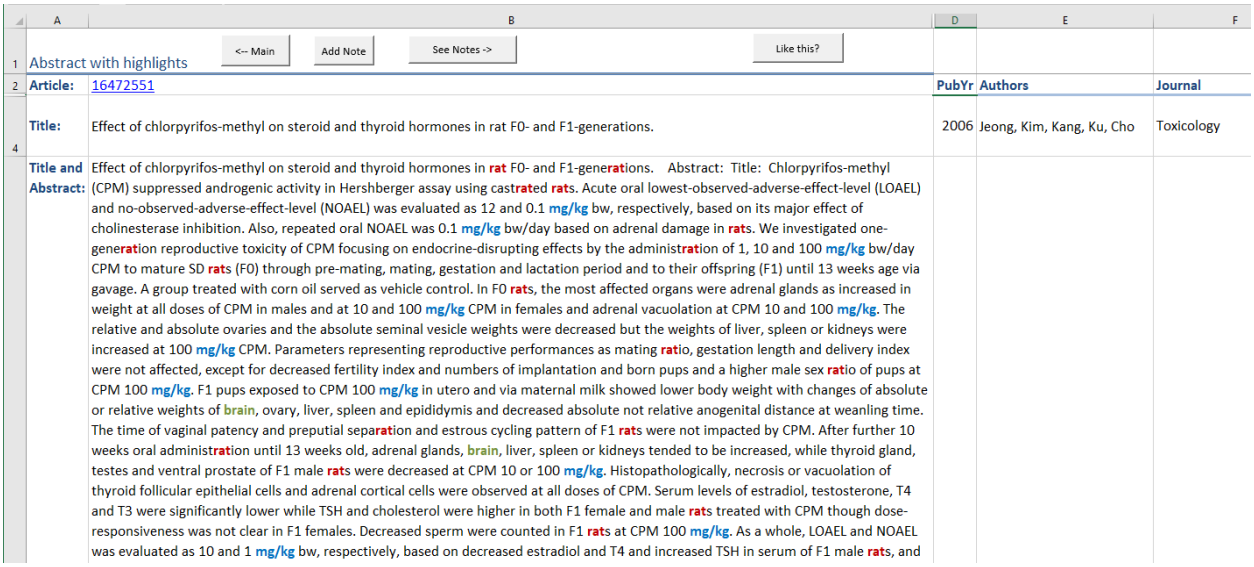
Let us suppose at this point that we are looking for dose-response data for chlorpyrifos in rat studies that focused on brain effects. We type the term “mg/kg” in cell B3, “rat” in C3, and “brain” in D3. As we finish typing and move to the next cell, the Abstract Sifter will count the occurrences of the terms in the title, abstract, and key words combined. The citations can then be sorted by these counts, either individually or by the total. Figure 5 shows what the Sifter looks like when these terms have been entered into cells B3, C3, and D3 and then the entries sorted by occurrence counts of “mg/kg” in descending order. PubMed 16472551 has 12 occurrences of “mg/kg”, 24 of “rat”, and two of “brain”. This article indeed describes a dose-response study in rats.

	A	B	C	D	E	F	G	H	I	J
1	Abstract Sifter	Query PubMed					*Query run: chlorpyrifos*			*Provided by the USEPA's National Center for Computational Toxicology*
2	v3	Your sifter terms and frequency counts								
3	PMID	mg/kg	rat	brain	Total	Pub Yr	Title (Take Group Notes / Highlight Noted)	Review	PMC	Authors
4	16472551	12	24	2	38	2006	Effect of chlorpyrifos-methyl on steroid and thyroid hormones in rat F0- and F1-generations.	0	0	Jeong, Kim, Kang, Ku, Cho
5	10653531	11	12	3	26	2000	Lack of carcinogenicity of chlorpyrifos insecticide in a high-dose, 2-year dietary toxicity study in Fischer 344 rats.	0	0	Yano, Young, Mattsson
6	94259	11	15	0	26	1979	Efficiency and safety of methidathion applied as a pour-on systemic insecticide for control of cattle lice.	0	0	Hart, Cavey, Moore, Strong
7	28260687	10	10	0	20	2017	Temperature influences the toxicity of deltamethrin, chlorpyrifos and dimethoate to the predatory mite Hypoaspis aculeif	0	0	Jegede, Owojori, Römbke
8	12521673	9	9	0	18	2003	Influence of gender on thermoregulation and cholinesterase inhibition in the long-evans rat exposed to diazinon.	0	0	Gordon, Mack
9	11829414	9	0	0	9	2001	Striatal dopaminergic pathways as a target for the insecticides permethrin and chlorpyrifos.	0	0	Karen, Li, Harp, Gillette, Bloomquist
10	3989223	9	4	0	13	1985	Toxicity of organophosphorus esters to laying hens after oral and dermal administration.	0	0	Francis, Metcalf, Hansen
11	30864424	8	7	6	21	2018	Neuroprotective role of 6-Gingerol-rich fraction of Zingiber officinale (Ginger) against acrylonitrile-induced neurotoxicity	0	0	Farombi, Abolaji, Adetuyi, Awosany
12	28962253	8	11	0	19	2014	Erythrocyte osmotic fragility and lipid peroxidation following chronic co-exposure of rats to chlorpyrifos and deltamethrin	1	1	Uchendu, Ambali, Ayo, Esievo, Umo
13	17454568	8	15	0	23	2007	Differential sensitivity to anticholinesterase insecticides in the juvenile rat: effects on thermoregulation.	0	0	Mack, Gordon
14	12830919	8	7	0	15	2003	Rapid multi-residue method for the determination of azinphos methyl, bromopropylate, chlorpyrifos, dimethoate, parath	0	0	Liapis, Aplada-Sarlis, Kyriakidis
15	9588346	8	9	0	17	1998	Effects of the pesticides carbofuran, chlorpyrifos, dimethoate, lindane, triallate, trifluralin, 2,4-D, and pentachlorophenol	0	0	Rawlings, Cook, Waldbillig
16	2017753	8	4	0	12	1991	Promotion of organophosphate-induced delayed polyneuropathy by phenylmethanesulfonyl fluoride.	0	0	Lotti, Caroldi, Capodicasa, Moretto
17	6201092	8	0	0	8	1984	Oral intubation of dogs with combinations of fertilizer, herbicide, and insecticide chemicals commonly used on lawns.	0	0	Yeary
18	30644245	7	3	0	10	2019	Protective Effect of N-Acetyl Cysteine on Chlorpyrifos-Induced Testicular Toxicity in Mice.	0	1	Kheradmandi, Jorsaraei, Feizi, Mogh
19	28280510	7	3	0	10	2017	Determination of Levels of Organochlorine, Organophosphorus, and Pyrethroid Pesticide Residues in Vegetables from Ma	0	1	Mahugija, Khamis, Lugwisha
20	26642910	7	11	1	19	2017	Decreased anxiety in juvenile rats following exposure to low levels of chlorpyrifos during development.	0	1	Carr, Armstrong, Buchanan, Eells, M
21	24609615	7	1	0	8	2014	Organophosphorus pesticide residues in vegetables from farms, markets, and a supermarket around Kwan Phayao Lake of	0	0	Sapbamrer, Hongsibsong
22	22504667	7	7	6	20	2012	Cholinesterase inhibition and toxicokinetics in immature and adult rats after acute or repeated exposures to chlorpyrifos	0	0	Marty, Andrus, Bell, Passage, Perala

Figure 5. After sifter terms were entered into cells B3, C3, D3 and sorting on B3

Abstract Sheet

To see the abstract for any of the retrieved articles, we can either click on the PubMed ID hyperlink to be taken to PubMed, or we can double-click on any other cell in the row for this article. This action brings us to the Abstract sheet where the abstract is displayed along with other article meta-data like title and authors (Figure 6).



	A	B	D	E	F
1	Abstract with highlights	<-- Main / Add Note / See Notes -> / Like this?			
2	**Article:**	16472551	**PubYr**	**Authors**	**Journal**
4	**Title:**	Effect of chlorpyrifos-methyl on steroid and thyroid hormones in rat F0- and F1-generations.	2006	Jeong, Kim, Kang, Ku, Cho	Toxicology
	Title and Abstract:	Effect of chlorpyrifos-methyl on steroid and thyroid hormones in rat F0- and F1-generations. Abstract: Title: Chlorpyrifos-methyl (CPM) suppressed androgenic activity in Hershberger assay using castrated rats. Acute oral lowest-observed-adverse-effect-level (LOAEL) and no-observed-adverse-effect-level (NOAEL) was evaluated as 12 and 0.1 mg/kg bw, respectively, based on its major effect of cholinesterase inhibition. Also, repeated oral NOAEL was 0.1 mg/kg bw/day based on adrenal damage in rats. We investigated one-generation reproductive toxicity of CPM focusing on endocrine-disrupting effects by the administration of 1, 10 and 100 mg/kg bw/day CPM to mature SD rats (F0) through pre-mating, mating, gestation and lactation period and to their offspring (F1) until 13 weeks age via gavage. A group treated with corn oil served as vehicle control. In F0 rats, the most affected organs were adrenal glands as increased in weight at all doses of CPM in males and at 10 and 100 mg/kg CPM in females and adrenal vacuolation at CPM 10 and 100 mg/kg. The relative and absolute ovaries and the absolute seminal vesicle weights were decreased but the weights of liver, spleen or kidneys were increased at 100 mg/kg CPM. Parameters representing reproductive performances as mating ratio, gestation length and delivery index were not affected, except for decreased fertility index and numbers of implantation and born pups and a higher male sex ratio of pups at CPM 100 mg/kg. F1 pups exposed to CPM 100 mg/kg in utero and via maternal milk showed lower body weight with changes of absolute or relative weights of brain, ovary, liver, spleen and epididymis and decreased absolute not relative anogenital distance at weanling time. The time of vaginal patency and preputial separation and estrous cycling pattern of F1 rats were not impacted by CPM. After further 10 weeks oral administration until 13 weeks old, adrenal glands, brain, liver, spleen or kidneys tended to be increased, while thyroid gland, testes and ventral prostate of F1 male rats were decreased at CPM 10 or 100 mg/kg. Histopathologically, necrosis or vacuolation of thyroid follicular epithelial cells and adrenal cortical cells were observed at all doses of CPM. Serum levels of estradiol, testosterone, T4 and T3 were significantly lower while TSH and cholesterol were higher in both F1 female and male rats treated with CPM though dose-responsiveness was not clear in F1 females. Decreased sperm were counted in F1 rats at CPM 100 mg/kg. As a whole, LOAEL and NOAEL was evaluated as 10 and 1 mg/kg bw, respectively, based on decreased estradiol and T4 and increased TSH in serum of F1 male rats, and			

Figure 6a. An abstract with highlighted sifter terms.

There are several aspects of the Abstract sheet that are important to note. First, the sifter terms in the abstract are highlighted. The font colors reflect the colors of the fonts in cells B3, C3, and D3 entered into the Main Sheet. This highlighting makes the reading the abstract easier by drawing attention to sentences that might be of more interest. It is also interesting to note that the counts and highlighting for "rat" also picked up "administration". Putting a spaces before or after search terms (e.g., " rat ") would eliminate some of these occurrences.

With version 3 of the Abstract Sifter, the *Like this?* Button was added to the top of the Abstract Sheet. Click on this button allows the user to find articles in PubMed that are similar to the article on the

Abstract sheet or find articles in PubMed Central that cite that article in question. (Figure 6b.) These functionalities are likely familiar as they are offered on the PubMed Entrez web site.

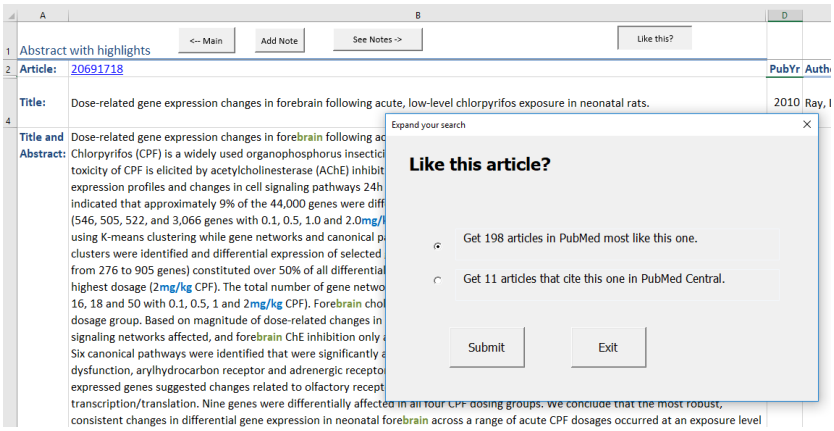


Figure 6b. Clicking on the Like this? Button on the Abstract Sheet.

Sifting the results through specifying sifter terms in B3, C3, and D3 can be repeated as many times as the user wishes. Similarly, new PubMed queries can be run, altered, rerun. There are no restrictions on either of these activities other than the 6000 record return limit.

Notes Sheet:

Given the dynamic nature of the sifter, many users find it helpful to be able to make notes on articles that they want to track. There are two ways using the Sifter to take notes: one way is through the Main page, and the other way starts with the Abstract page. To return to our case study, let us say that we have found a set of articles on the Main page that we know we need to read in depth. We can select these articles and then click on the *Take Group Notes* button. A form appears where we can enter information into fields called Tag and Notes. These elements are self-defined. We can also click on *yes*, *no*, or *maybe*. This set of variables is a quick way to associate articles with a note. Notice that these choices each come with a color (yes-green, no-red, and maybe-yellow). Entering any of these fields is optional. (Figures 7 and 8.) When we click on the *OK* button, each article selected will be inserted into the Notes page with the corresponding information (Figure 9) and the PubMed ID (PMID) on the Main

sheet will be colorized.

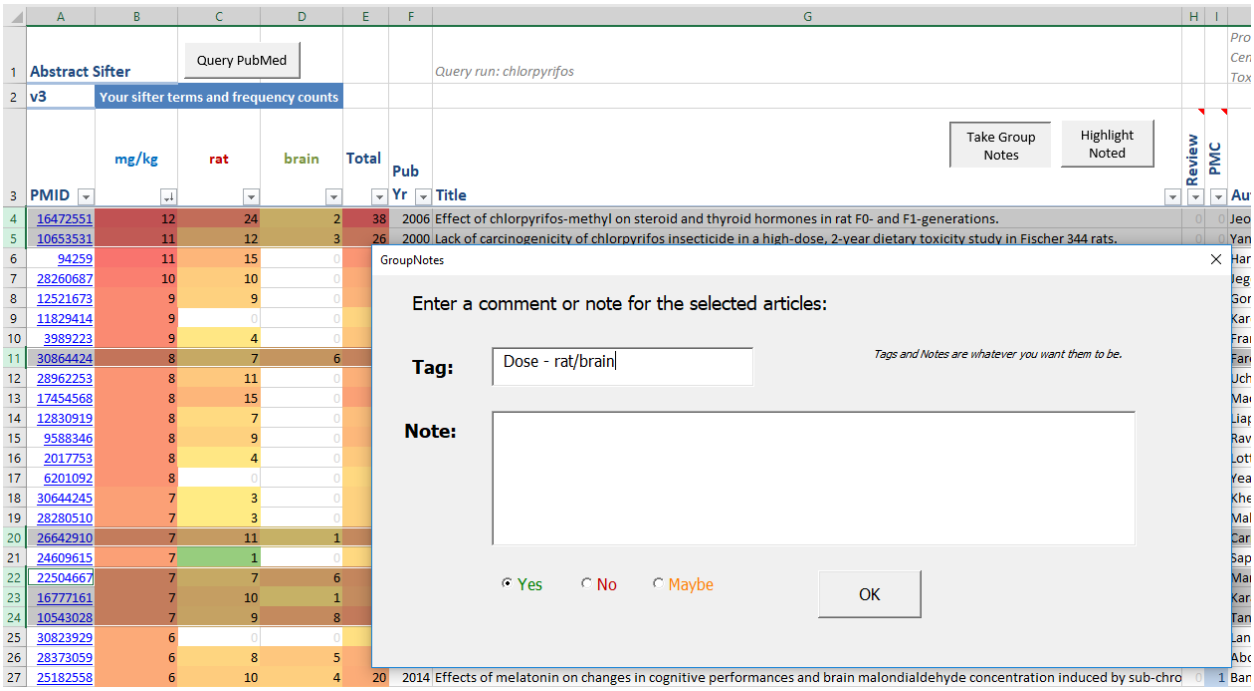


Figure 7. Taking group notes.

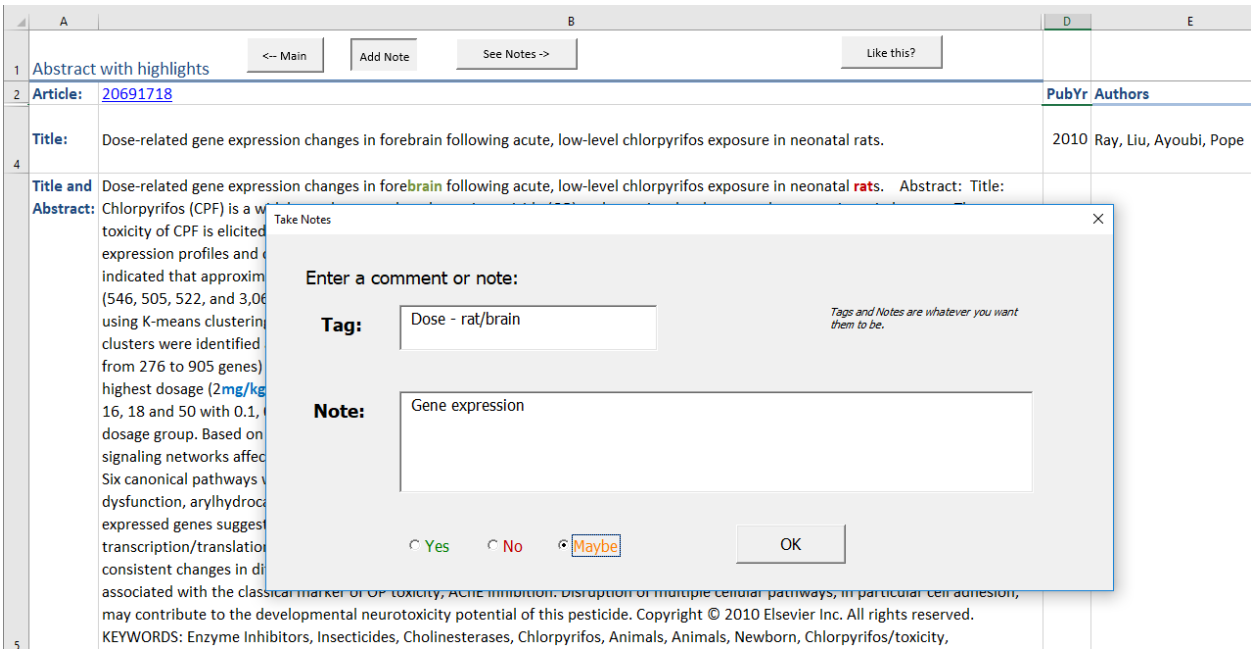
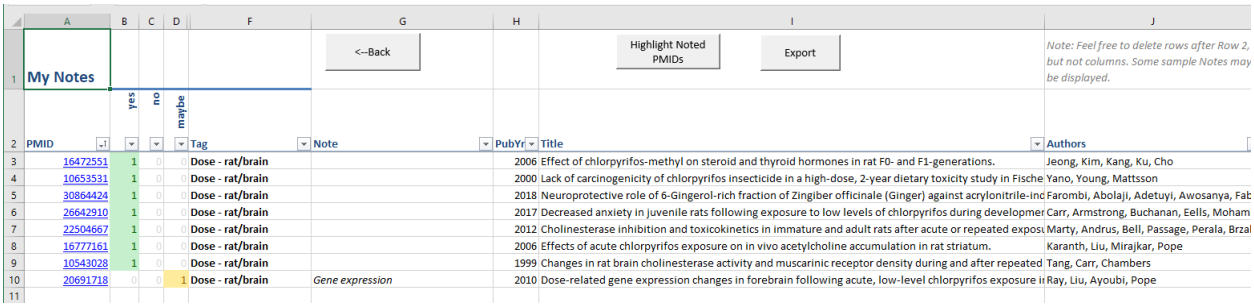


Figure 8. Taking single notes on the Abstract sheet.



	A	B	C	D	F	G	H	I	J
1	My Notes					<--Back		Highlight Noted PMIDs / Export	*Note: Feel free to delete rows after Row 2, but not columns. Some sample Notes may be displayed.*
2	PMID	yes	no	maybe	Tag	Note	PubYr	Title	Authors
3	16472551	1	0	0	**Dose - rat/brain**		2006	Effect of chlorpyrifos-methyl on steroid and thyroid hormones in rat F0- and F1-generations.	Jeong, Kim, Kang, Ku, Cho
4	10653531	1	0	0	**Dose - rat/brain**		2000	Lack of carcinogenicity of chlorpyrifos insecticide in a high-dose, 2-year dietary toxicity study in Fische	Yano, Young, Mattsson
5	30864424	1	0	0	**Dose - rat/brain**		2018	Neuroprotective role of 6-Gingerol-rich fraction of Zingiber officinale (Ginger) against acrylonitrile-inc	Farombi, Abolaji, Adetuyi, Awosanya, Fab
6	26642910	1	0	0	**Dose - rat/brain**		2017	Decreased anxiety in juvenile rats following exposure to low levels of chlorpyrifos during developmer	Carr, Armstrong, Buchanan, Eells, Moham
7	22504667	1	0	0	**Dose - rat/brain**		2012	Cholinesterase inhibition and toxicokinetics in immature and adult rats after acute or repeated exposu	Marty, Andrus, Bell, Passage, Perala, Brzal
8	16777161	1	0	0	**Dose - rat/brain**		2006	Effects of acute chlorpyrifos exposure on in vivo acetylcholine accumulation in rat striatum.	Karanth, Liu, Mirajkar, Pope
9	10543028	1	0	0	**Dose - rat/brain**		1999	Changes in rat brain cholinesterase activity and muscarinic receptor density during and after repeated	Tang, Carr, Chambers
10	20691718	0	0	1	**Dose - rat/brain**	*Gene expression*	2010	Dose-related gene expression changes in forebrain following acute, low-level chlorpyrifos exposure i	Ray, Liu, Ayoubi, Pope
11									

Figure 9. The Notes page. Remember to save your workbook.

The second option for note taking starts with the Abstract Sheet. (Figure 8) The “Add Note” button in the top row allows notes to be inserted into the Notes Sheet using the same form used to add Group Notes described above.

The note-taking can be used to help keep track of which citations have been read and evaluated and which have not. By clicking on the *Highlight Noted* button on the Main page or on the Notes page, the color of each noted PubMed ID will be modified as shown in Figure 10. On the Main sheet the PMIDs can be sorted by the noted color using the built-in Excel sorting functionality.

The user can make changes to the notes sheet by editing, adding or deleting rows below row 2.

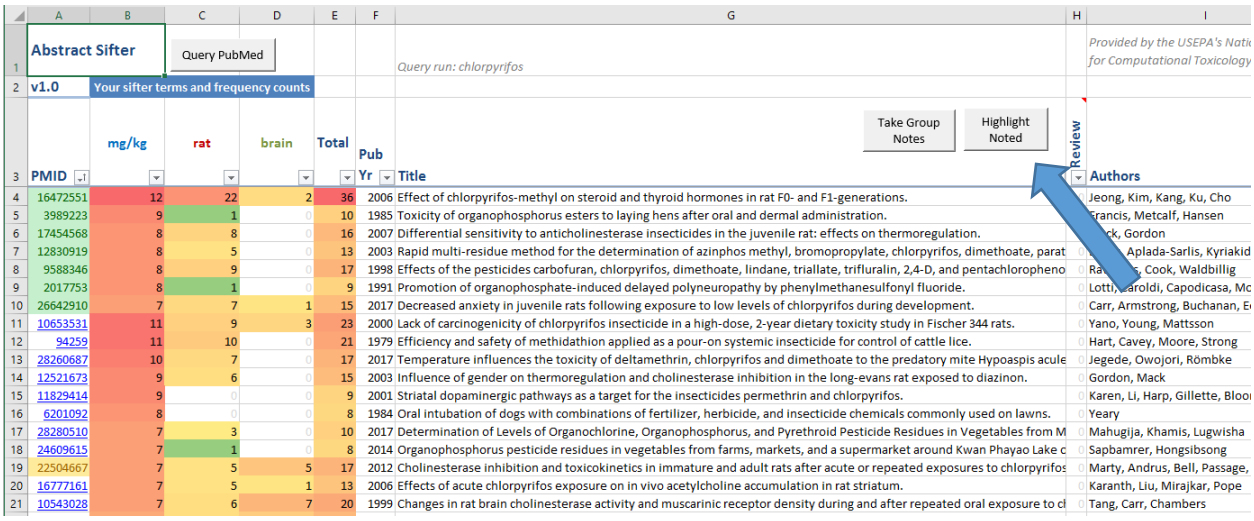


Figure 10. After clicking on Highlight Notes and then sorting by color

Log sheet

The Log sheet keeps track of the queries you have run. The Abstract Sifter routines insert a row into the sheet every time you complete a query. These queries can be viewed and rerun. To rerun a query, simply double-click on it. (Figure 11.) Delete any or all rows after Row 2 if you want to clear old entries.

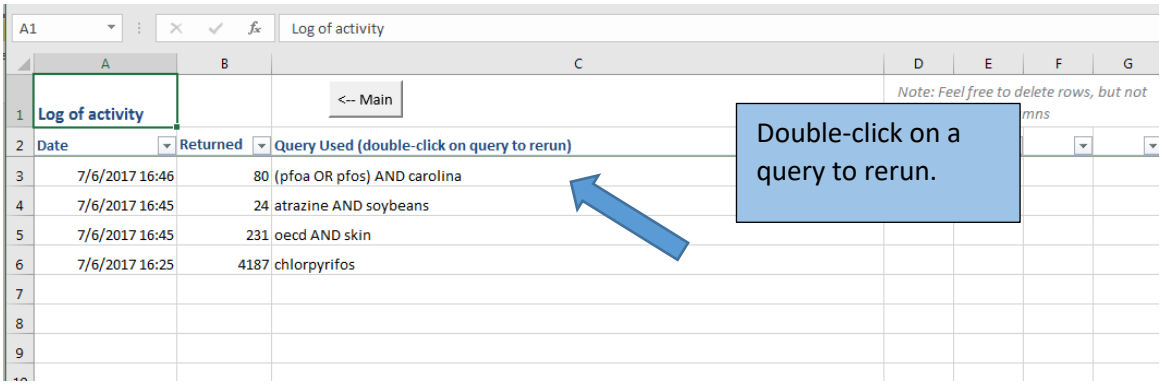


Figure 11. View of the Log sheet

Landscape sheet

The Landscape sheet provides an overview of the literature to the user for a set of entities, for example, a list of chemicals or genes. Figure 12 shows an example of a Landscape sheet built by a researcher interested in the toxicity of a particular set of chemicals. Let's take a look at that first. Queries designed to find the chemicals of interest are entered into Column C and in this case, a short version of the chemical name is in Column A. The queries in Row 3 are typical ones used in searching for articles about different kinds of chemical toxicity. We will refer to these queries as subject matter queries.

The premise behind the design of the Landscape sheet is very simple: PubMed queries will be built by taking the values in Column C (in this example chemical names and corresponding CAS numbers) and appending this query text to the subject matter query text in Row 3 with an " AND " in between the two query parts.

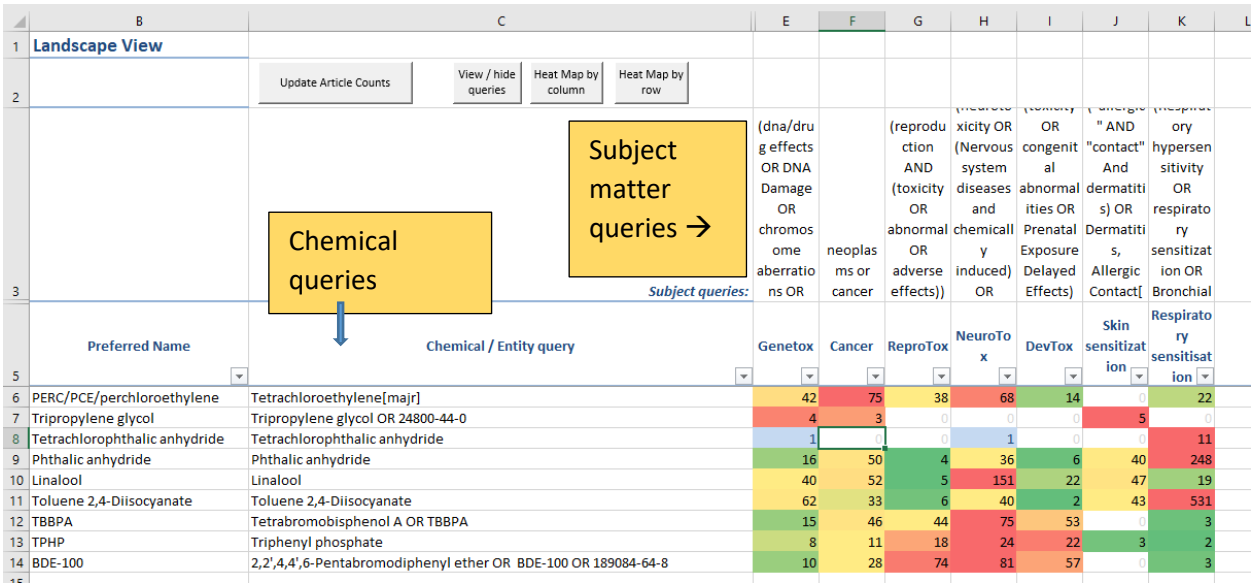
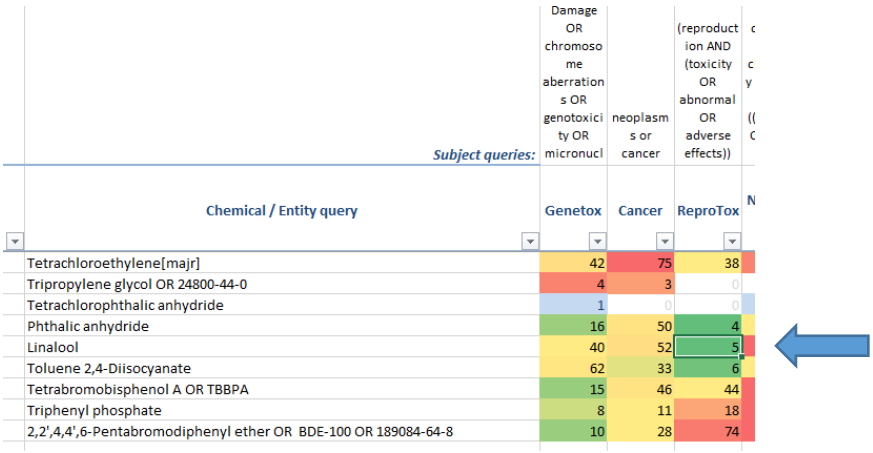


Figure 12. Example of Landscape sheet use

To illustrate, we will double-click on the cell with the arrow pointer in Figure 13. When we double-click on this cell this tells the Abstract Sifter to take the query text in Column C about Linalool and append it to query text designed to find citations about reproductive toxicity. Figure 14 shows the constructed query. We can then click on *Submit* and the query gets sent to PubMed and we can then see the results on the Main page. The number of articles retrieved from PubMed is 5. That count is placed in the corresponding Landscape cell that we just clicked on.



Subject queries:	Damage OR chromosome aberrations OR genotoxicity OR micronucl	neoplasms or cancer	(reproduction AND (toxicity OR abnormal OR adverse effects))	c c y ((C
Chemical / Entity query	Genetox	Cancer	ReproTox	N
Tetrachloroethylene[majr]	42	75	38	
Tripropylene glycol OR 24800-44-0	4	3	0	
Tetrachlorophthalic anhydride	1	0	0	
Phthalic anhydride	16	50	4	
Linalool	40	52	5	
Toluene 2,4-Diisocyanate	62	33	6	
Tetrabromobisphenol A OR TBBPA	15	46	44	
Triphenyl phosphate	8	11	18	
2,2',4,4',6-Pentabromodiphenyl ether OR BDE-100 OR 189084-64-8	10	28	74	

Figure 13. Double-click on article count cells

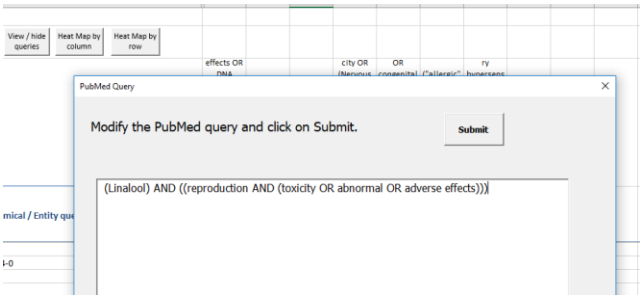


Figure 14. Constructed query

Now let's add to the Landscape sheet. Figure 15 shows how we added a new chemical to the list: aspirin. To find out the article counts for aspirin, select empty cells on the same row as aspirin, then click on *Update Article Counts* button. Excel will build each query from the aspirin part and the subject matter

part and send each query to PubMed to find out how many citations satisfy the query. The article counts are placed in the corresponding cells. To run the query and retrieve the results, just double-click on any of the article count cells.

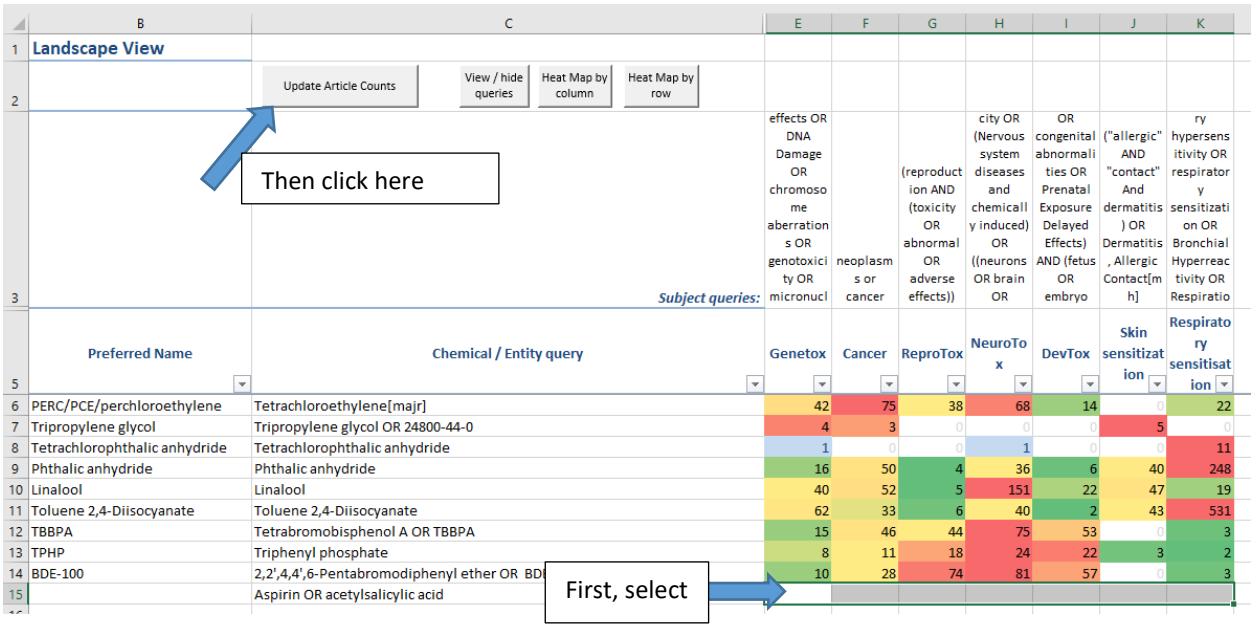


Figure 15. Adding rows to the Landscape sheet

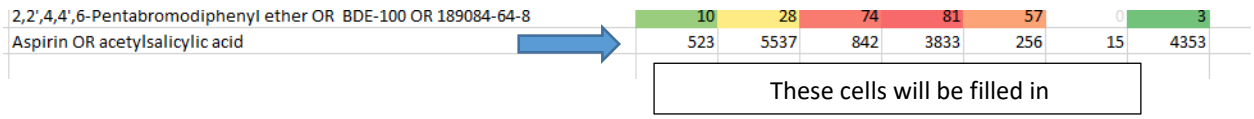


Figure 16. After clicking on Update Article Counts

New subject matter queries can be entered as well. The query part goes into Row 3 and a heading (of your choice) goes into Row 5. See the example below where the PubMed query part: skin OR dermatitis is entered with the heading skin. Next highlight the cells underneath and click on the *Update Article Counts* button. The counts of articles satisfying the queries are placed in the cells. What's happening behind the scenes? For each cell, a query is being built by the Abstract Sifter and sent to PubMed to retrieve a record count. That record count is then inserted into the corresponding cell. (Figure 17.)

Keep in mind that our examples revolve around chemicals, but that does not mean anyone is limited to chemicals. The entries in Column C and in Row 3 can be whatever you the end-user want them to be: genes, proteins, diseases, authors …

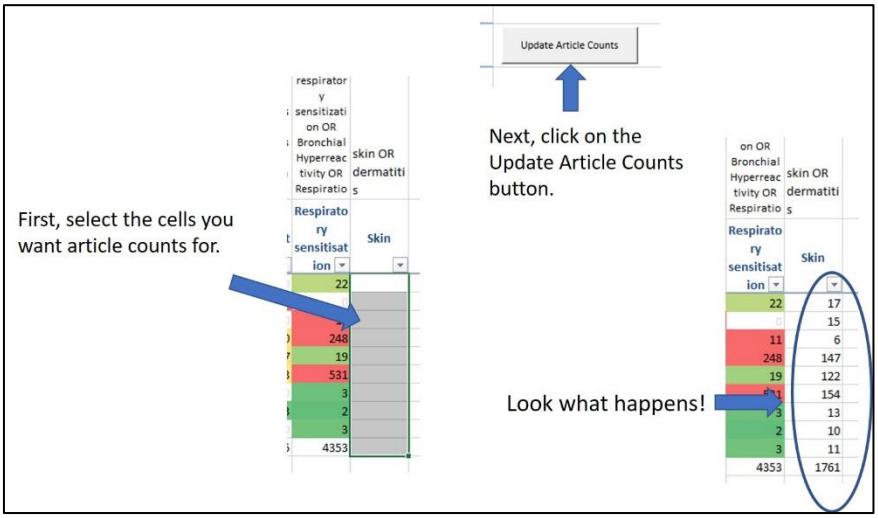


Figure 17. Steps for retrieving counts

Making things look good

The Landscape sheet has three buttons that make formatting easy (Figure 18). You can choose to hide the query row or show it. The heat map buttons will quickly apply heat map coloring to the cells with article counts either by column or by row. Try them out!

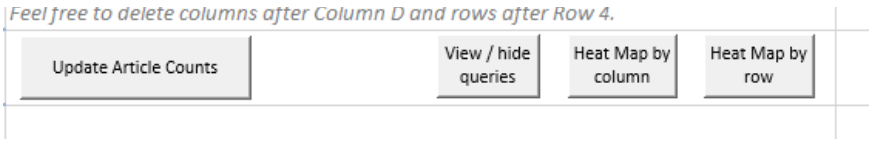
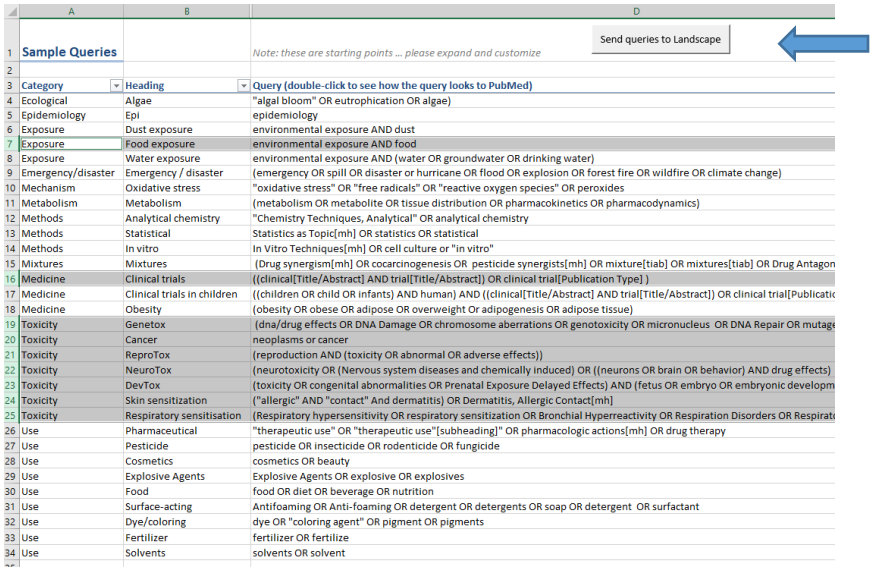


Figure 18. Buttons on the Landscape page include formatting actions

Sample_Queries and Pathway_Queries Sheets

These two sheets function in a very similar way. We will use the Sample_queries sheet as an example in this user guide. Both sheets contain a number of sample subject matter queries that the end user can use as a starting point for building a Landscape view of a set of entities. Let's see how. First, we will clean off the old subject matter queries by deleting columns E-L on the Landscape sheet. (You can let the previous work stay if you wish.) Next, on the Sample_Queries sheet we will select rows with queries of interest then we click on the button *Send Queries to Landscape* (Figure 19).



	A	B	D
1	**Sample Queries**		*Note: these are starting points ... please expand and customize* — Send queries to Landscape
2			
3	**Category**	**Heading**	**Query (double-click to see how the query looks to PubMed)**
4	Ecological	Algae	"algal bloom" OR eutrophication OR algae)
5	Epidemiology	Epi	epidemiology
6	Exposure	Dust exposure	environmental exposure AND dust
7	Exposure	Food exposure	environmental exposure AND food
8	Exposure	Water exposure	environmental exposure AND (water OR groundwater OR drinking water)
9	Emergency/disaster	Emergency / disaster	(emergency OR spill OR disaster or hurricane OR flood OR explosion OR forest fire OR wildfire OR climate change)
10	Mechanism	Oxidative stress	"oxidative stress" OR "free radicals" OR "reactive oxygen species" OR peroxides
11	Metabolism	Metabolism	(metabolism OR metabolite OR tissue distribution OR pharmacokinetics OR pharmacodynamics)
12	Methods	Analytical chemistry	"Chemistry Techniques, Analytical" OR analytical chemistry
13	Methods	Statistical	Statistics as Topic[mh] OR statistics OR statistical
14	Methods	In vitro	In Vitro Techniques[mh] OR cell culture or "in vitro"
15	Mixtures	Mixtures	(Drug synergism[mh] OR cocarcinogenesis OR pesticide synergists[mh] OR mixture[tiab] OR mixtures[tiab] OR Drug Antagon
16	Medicine	Clinical trials	((clinical[Title/Abstract] AND trial[Title/Abstract]) OR clinical trial[Publication Type])
17	Medicine	Clinical trials in children	((children OR child OR infants) AND human) AND ((clinical[Title/Abstract] AND trial[Title/Abstract]) OR clinical trial[Publicatic
18	Medicine	Obesity	(obesity OR obese OR adipose OR overweight Or adipogenesis OR adipose tissue)
19	Toxicity	Genetox	(dna/drug effects OR DNA Damage OR chromosome aberrations OR genotoxicity OR micronucleus OR DNA Repair OR mutage
20	Toxicity	Cancer	neoplasms or cancer
21	Toxicity	ReproTox	(reproduction AND (toxicity OR abnormal OR adverse effects))
22	Toxicity	NeuroTox	(neurotoxicity OR (Nervous system diseases and chemically induced) OR ((neurons OR brain OR behavior) AND drug effects)
23	Toxicity	DevTox	(toxicity OR congenital abnormalities OR Prenatal Exposure Delayed Effects) AND (fetus OR embryo OR embryonic developm
24	Toxicity	Skin sensitization	("allergic" AND "contact" And dermatitis) OR Dermatitis, Allergic Contact[mh]
25	Toxicity	Respiratory sensitisation	(Respiratory hypersensitivity OR respiratory sensitization OR Bronchial Hyperreactivity OR Respiration Disorders OR Respirat
26	Use	Pharmaceutical	"therapeutic use" OR "therapeutic use"[subheading]" OR pharmacologic actions[mh] OR drug therapy
27	Use	Pesticide	pesticide OR insecticide OR rodenticide OR fungicide
28	Use	Cosmetics	cosmetics OR beauty
29	Use	Explosive Agents	Explosive Agents OR explosive OR explosives
30	Use	Food	food OR diet OR beverage OR nutrition
31	Use	Surface-acting	Antifoaming OR Anti-foaming OR detergent OR detergents OR soap OR detergent OR surfactant
32	Use	Dye/coloring	dye OR "coloring agent" OR pigment OR pigments
33	Use	Fertilizer	fertilizer OR fertilize
34	Use	Solvents	solvents OR solvent

Figure 19. Selecting rows with queries of interest

Our Landscape sheet then looks like Figure 20.

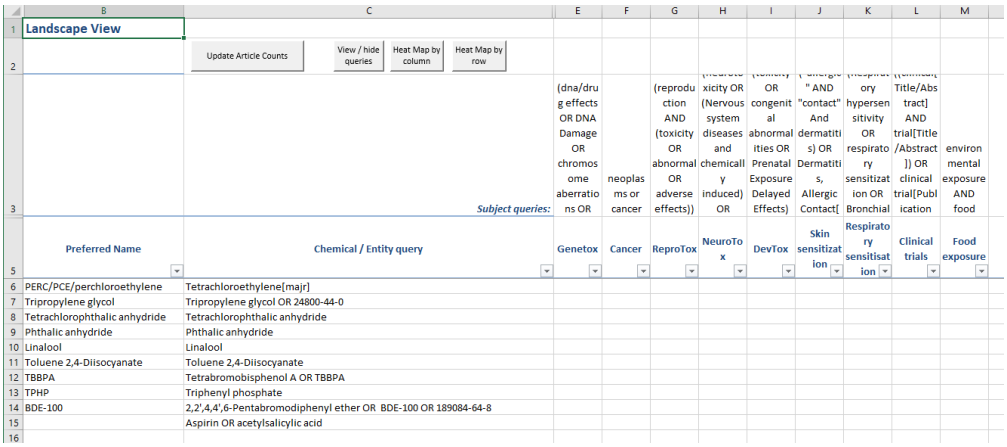


Figure 20. New queries on Landscape sheet

Next, we select the article count area and then click on *Update Article Counts*.

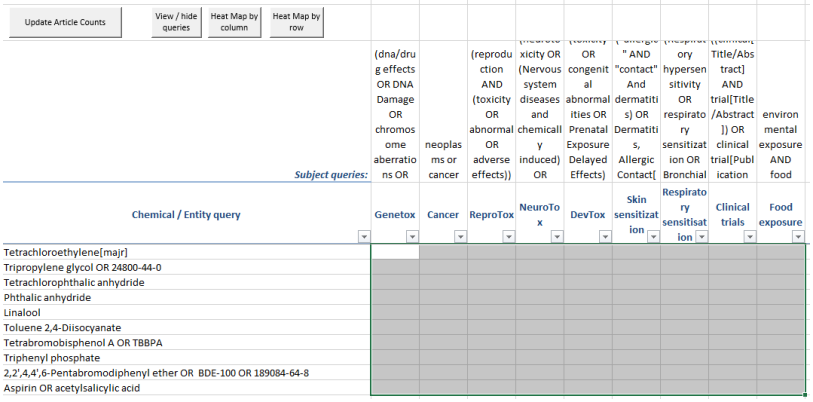
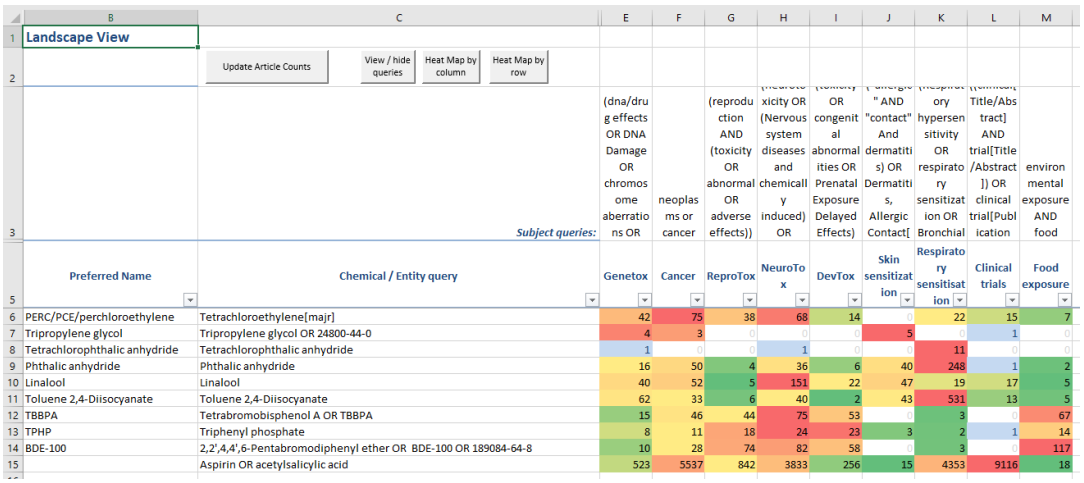


Figure 21. Selecting the cells for article counts

Once the article counts are populated, we click on Heat Map by Row and then on Hide queries. Our resulting Landscape view looks like Figure 22. To run the query and retrieve the results, just double-click on any of the article count cells.

Note that the Pathway_queries sheet works the same way: select the rows of interest and click on the button *Send queries to Landscape*. To add your own queries to either the pathway or sample queries sheets, just enter in new rows. Use the existing rows as a template to indicate where headers go and where the query text goes. Be mindful of parentheses and be sure to test the queries.



Landscape View

Update Article Counts | View / hide queries | Heat Map by column | Heat Map by row

	Preferred Name	Chemical / Entity query	Genetox	Cancer	ReproTox	NeuroTox	DevTox	Skin sensitization	Respiratory sensitisation	Clinical trials	Food exposure
Subject queries:			(dna/drug effects OR DNA Damage OR chromosome aberrations OR	neoplasms or cancer	(reproduction AND (toxicity OR abnormal OR adverse effects))	xicity OR (Nervous system diseases and chemicall y induced) OR	OR congenital abnormalities OR Prenatal Exposure Delayed Effects)	" AND "contact" And dermatitis) OR Dermatitis, Allergic Contact[	ory hypersensitivity OR respiratory sensitization OR Bronchial	Title/Abstract] AND trial[Title/Abstract]) OR clinical trial[Publication	environmental exposure AND food
6	PERC/PCE/perchloroethylene	Tetrachloroethylene[majr]	42	75	38	68	14	0	22	15	7
7	Tripropylene glycol	Tripropylene glycol OR 24800-44-0	4	3	0	0	0	5	0	1	0
8	Tetrachlorophthalic anhydride	Tetrachlorophthalic anhydride	1	0	0	1	0	0	11	0	0
9	Phthalic anhydride	Phthalic anhydride	16	50	4	36	6	40	248	1	2
10	Linalool	Linalool	40	52	5	151	22	47	19	17	5
11	Toluene 2,4-Diisocyanate	Toluene 2,4-Diisocyanate	62	33	6	40	2	43	531	13	5
12	TBBPA	Tetrabromobisphenol A OR TBBPA	15	46	44	75	53	0	3	0	67
13	TPHP	Triphenyl phosphate	8	11	18	24	23	3	2	1	14
14	BDE-100	2,2',4,4',6-Pentabromodiphenyl ether OR BDE-100 OR 189084-64-8	10	28	74	82	58	0	3	0	117
15		Aspirin OR acetylsalicylic acid	523	5537	842	3833	256	15	4353	9116	18

Figure 22. New Landscape view

Exporting to other applications from the Notes sheet

The Abstract Sifter allows the user to export articles from the Notes sheet to outside applications. On the Notes sheet there is a button labeled *Export*. By clicking on this button, the form in Figure 23 appears. The first set of radio buttons allows the user to select how the PubMed IDs are to be formatted. Next the user can choose to export all entries / rows on the Notes sheet or just selected rows. By clicking on Next Step, the PMIDs are copied to the clipboard and ready to be pasted to the next application. In the case of PubMed, they should be pasted into the query box in PubMed Entrez. (Figure 24) From PubMed, the citations can be downloaded in a variety of formats, including a format that can be imported into citation management software (Figures 25 and 26).

Figure 23. Form that appears after clicking on Export button

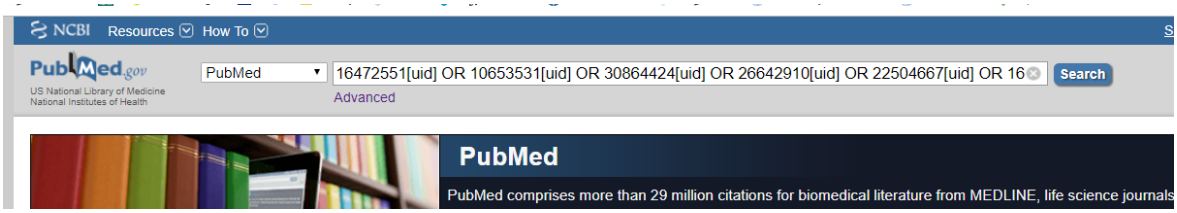


Figure 24. Paste PMIDs in the query box to retrieve the records.

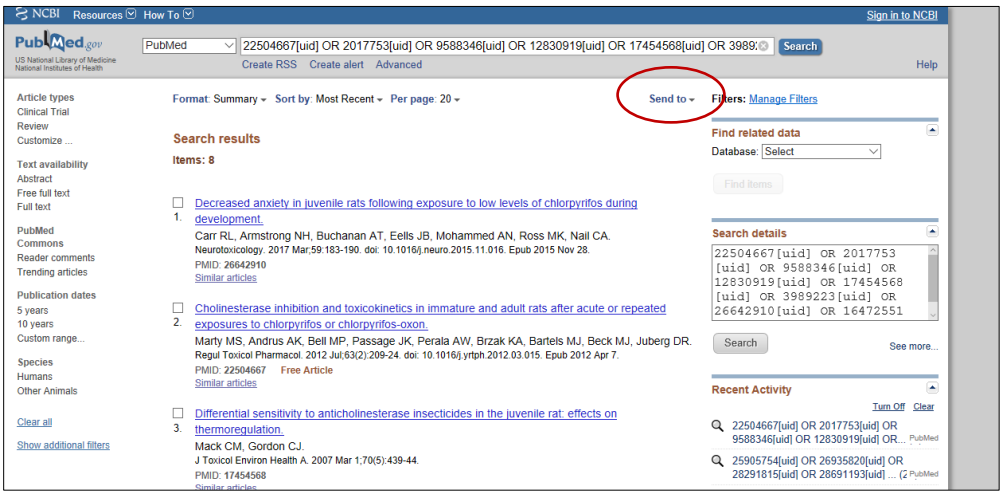


Figure 24. In PubMed, click on Send to

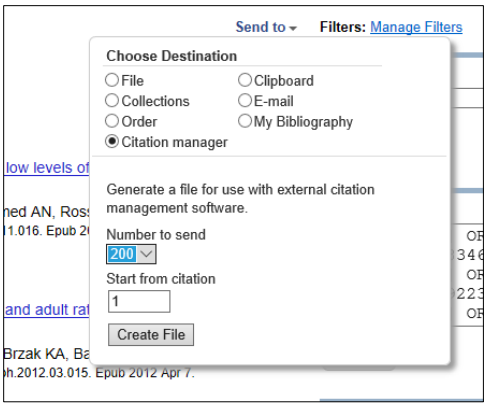


Figure 25. Dialog box for exporting to citation manager

When the user clicks on the Create File button, a file is created and downloaded in nbib format and can be imported into most common reference manager software.

Helpful Tips and Guidelines

Tip 1 – checking quality of results

The Landscape sheet is a great way to explore a set of chemicals, but some chemical names are long, complex, and a challenge to PubMed. If you get unexpected results from a chemical query, it's a good idea to check it in PubMed. You can take any query generated by the Abstract Sifter and copy and paste it into PubMed using Ctl-C to copy and Ctl-V to paste. For example, the query in the box shown in Figure 28 is selected and copied (with Ctl-C). Then in PubMed the query is pasted into the query line at the top as shown in Figure 29. After we click on search we see that PubMed brings back 51 records. On the right side of the page is a box entitled *Search Details*. Click on the *See More* … link to expand this box. (Figure 30) Figure 31 shows the information provided by PubMed about how it expands the query. If you need to learn more about PubMed queries, click on *Help* on the PubMed home page.

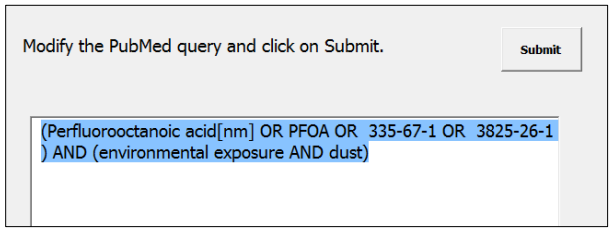


Figure 26. Select and Ctl-C to copy

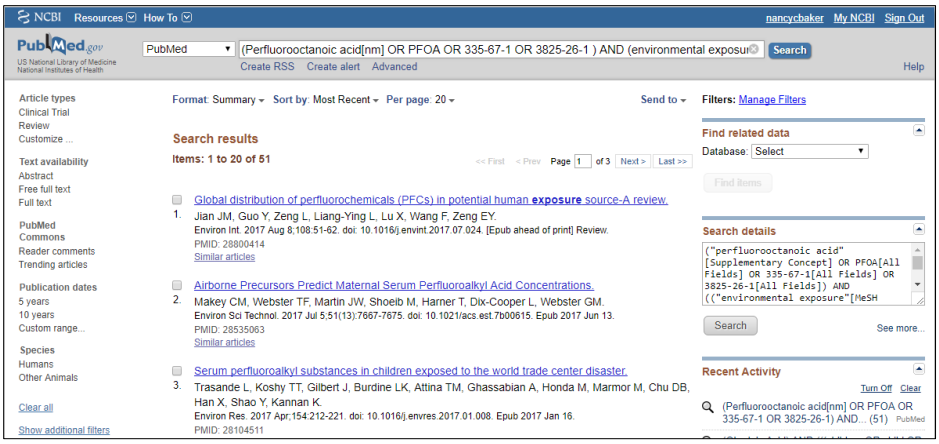


Figure 27. Ctl-V to paste in PubMed then search

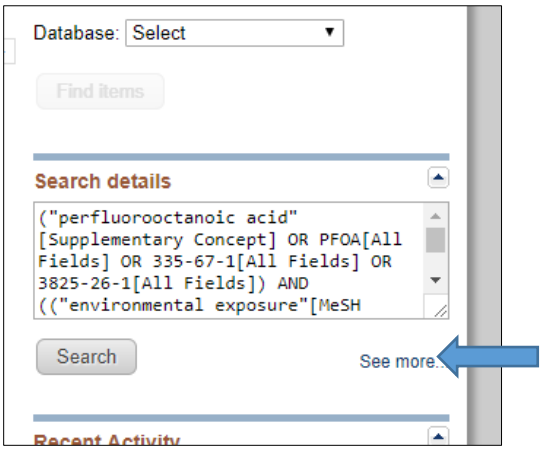


Figure 28. See what PubMed does to expand your search

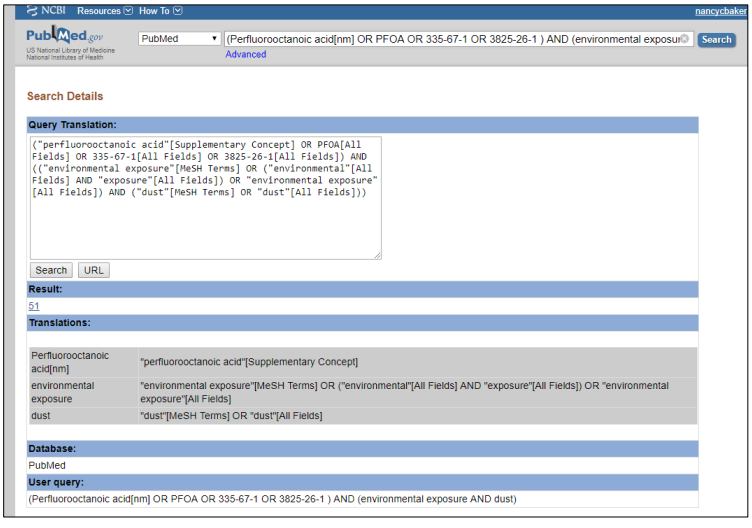


Figure 29. PubMed query breakdown and expansion

Tip 2 – Sifting the chemical literature

It can be very helpful in chemical research to include the chemical name in the sifting process. This is because a chemical can be mentioned in an abstract even in cases where the article is not really about the chemical and will be retrieved in the PubMed query (depending on how the query is worded). Counting the occurrences of the chemical name in the abstract through the sifting process can help the user discriminate between articles mentioning a chemical or those that are actually about the chemical.

Tip 3 – cleanup and customization

The Abstract Sifter can be cleaned up by deleting rows and columns from previous work, but the Abstract Sifter programming requires certain columns and rows to be in certain places. To learn how to clean up your sifter without disrupting the behind-the-scenes coding, consult the table below.

Sheet name	**Sheet Function**
Main	Do not add columns. Rows will be added and deleted by the Sifter.
Abstract	Do not add or delete rows or columns. The Sifter software updates this sheet.
Notes	Delete any unwanted rows after Row 2. Do not add or delete columns.
Log	Delete any unwanted rows after Row 2. Do not add or delete columns.
Sample_Queries	Delete unwanted rows after Row 3. Modify and add rows as desired, following the pattern of current rows. (That is, keep the heading in column B and the query text in column C.) *Hint: use this sheet to keep queries important to your organization.*
Pathway_queries	Delete unwanted rows after Row 3. Modify and add rows as desired, following the pattern of current rows.
Landscape	Delete or modify rows after Row 4 and columns after Column D.

Keep in mind that the Abstract Sifter is an Excel file. You can rename it, mail it, and of course, if you want to keep your Log, Notes, and Landscape entries, you should save it. The Sample_queries sheet provides an opportunity for you and your organization to start collecting and organizing queries that you have found useful.

Tip 4 – Connections to the EPA Chemicals Dashboard

The Environmental Protection Agency's Chemicals Dashboard is a great place to find chemical information to enhance your chemical search queries with synonyms and CAS numbers. Future releases of the Dashboard will offer opportunities to download a list of chemicals formatted for easy insertion into the Landscape sheet. You'll find the Chemistry Dashboard here: https://comptox.epa.gov/dashboard.

The EPA Chemical Dashboard also contains its own (slightly different) version of the Abstract Sifter. It works on the same basic premise as the Excel version, but has some interesting differences. To see it, start with a chemical search. Let's look at the chemical fipronil by entering the name in the search box and clicking on the search icon (magnifying glass) (Figure 32).

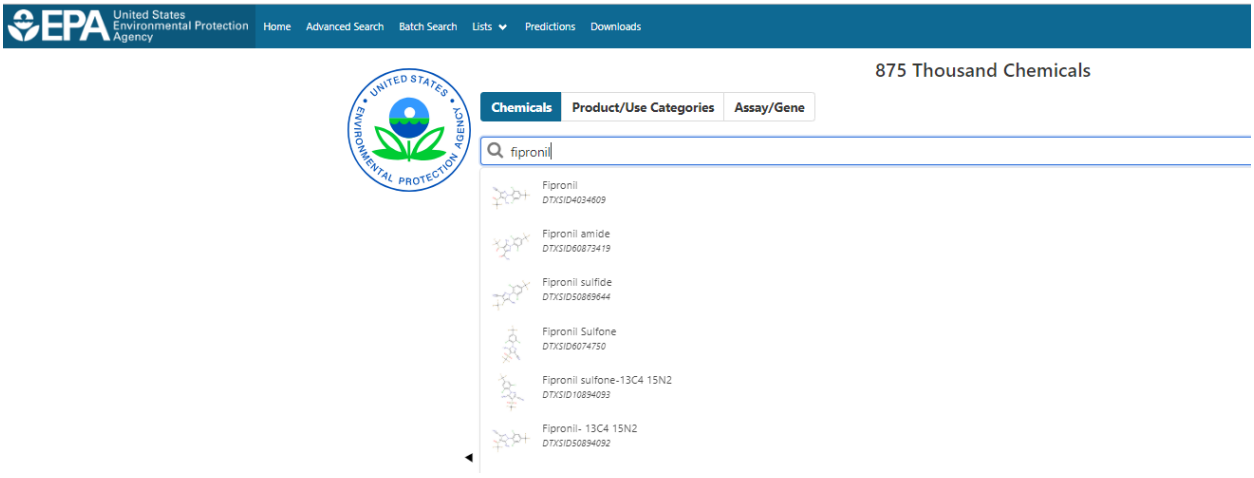


Figure 32. Searching for fipronil on the EPA's Chemistry Dashboard entry form

The main page for fipronil is displayed with the structure diagram and a selection of tabs below that lead to other information about the chemical. Click on the Literature tab as shown in Figure 33.

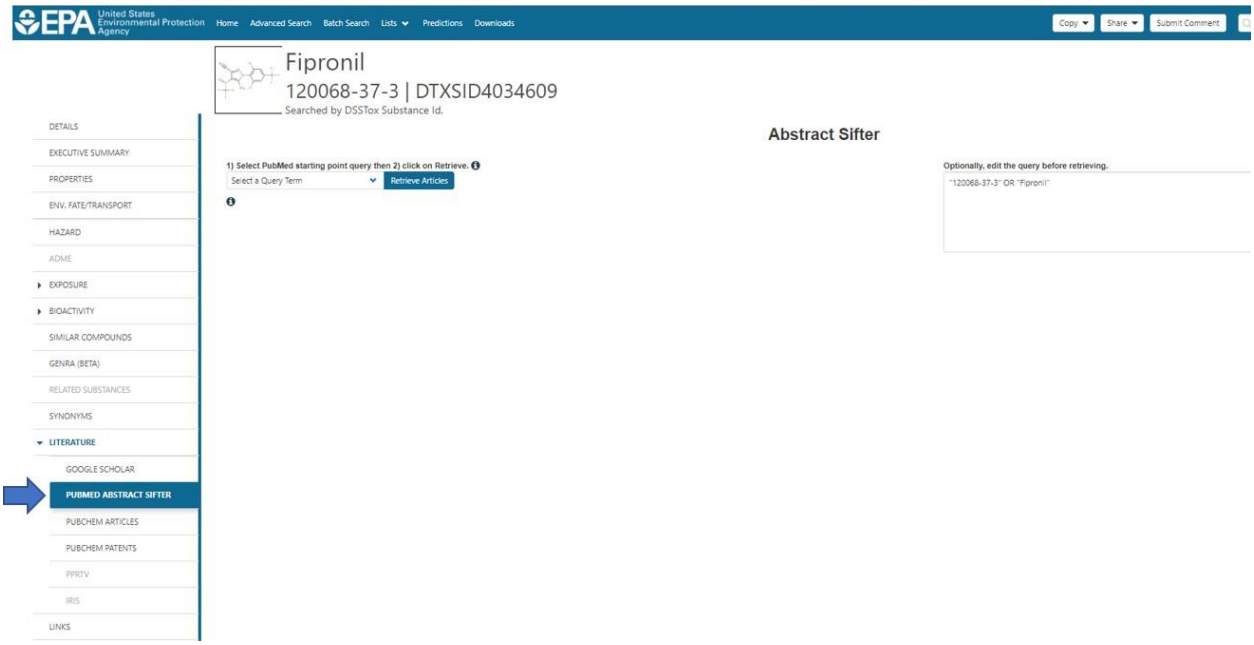


Figure 33. Select the Literature tab then on PubMed Abstract Sifter (see below).

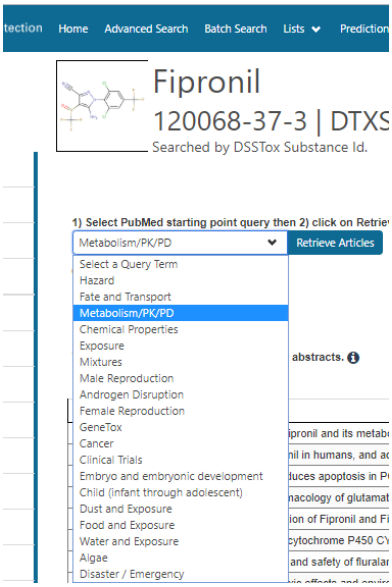


Figure 34. How to select prepared queries.

Select PubMed Abstract Sifter on the left set of buttons. The Dashboard helps you to build queries for this chemical. The chemical identifier part of the query is prepopulated on the right with name and CAS number. The subject matter part of the query is determined by selecting a topic area in the pull-down box in the center of the form. The user has several pre-composed queries to choose from. When one of them is chosen, the query is modified by appending the subject matter text. Figure 35 shows that when Metabolism/PK/PD is chosen, the text (metabolism OR metabolite OR tissue distribution OR pharmacokinetics OR pharmacodynamics) is appended to the chemical identifiers. The query can be modified manually as well. When ready, the user clicks on Retrieve Articles.

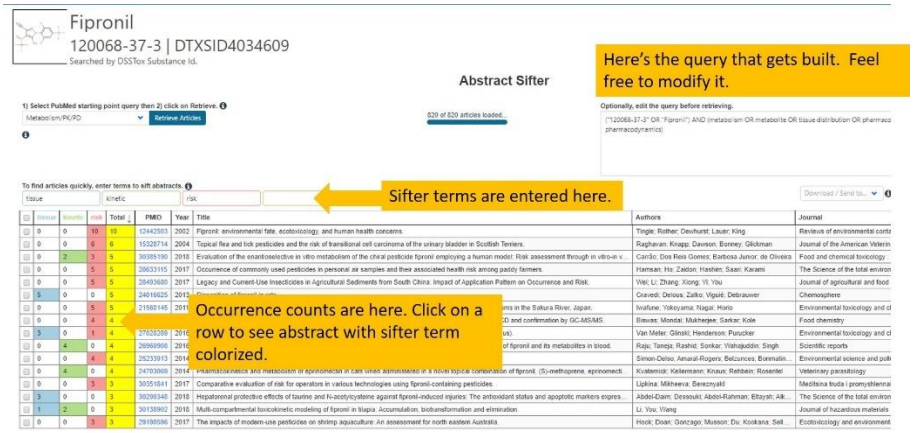


Figure 35. Sifting on the EPA Chemicals Dashboard's PubMed Abstract Sifter.

After the user clicks on Retrieve Articles, the article information is retrieved from PubMed and inserted into the results table. The articles can be sifted by entering terms into the boxes shown. In the example in Figure 35, the user has entered tissue into one box and kinetic into the other. The occurrences of these terms are counted for each PubMed citation and displayed. The table can be sorted on these values. Clicking on a row tells the Dashboard Sifter to display the title and abstract below the table with the sifter terms highlighted.

A check box on the left of the table provides a way to select citation rows. Selected rows can be downloaded or sent to PubMed by clicking on the pull-down box to the right.

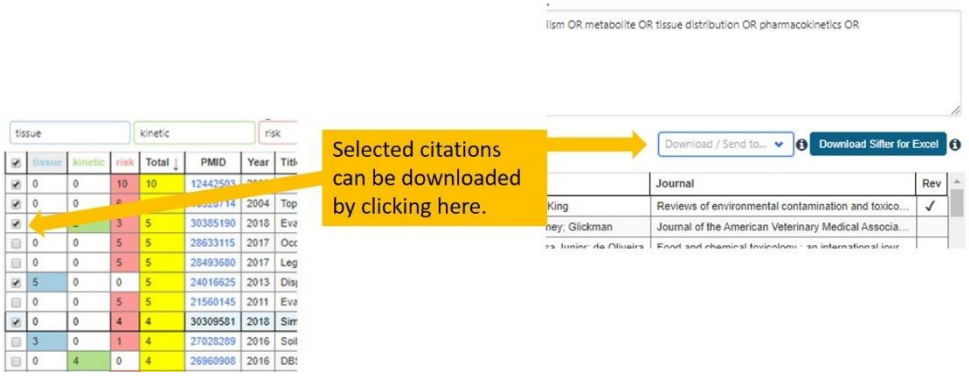


Figure 36. How to select and download citations.

Notice, too, that the blue button above lets the user download the Excel version of the Abstract Sifter. This button will always download the most up-to-date version. Check back periodically with the EPA's Chemicals Dashboard to learn about enhancements to PubMed Abstract Sifter.

Populating the Excel Abstract Sifter from the EPA CompTox Chemicals Dashboard

The Chemicals Dashboard can download chemicals in a variety of formats. One of those formats make it easy to use in the Excel Abstract Sifter. Here's an example to get you started. On the home page of the Dashboard, click on Lists, then chemical (Figure 37.)

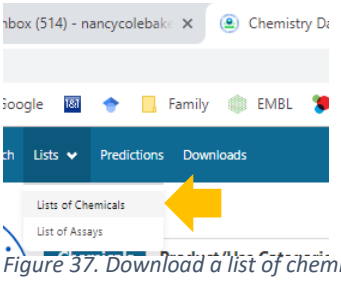


Figure 37. Download a list of chemicals.

Pick a list. We'll pick Algal Toxins as a sample. Click on the name, then, when the chemicals appear, click on Send to Batch Search. (Figure 38.)

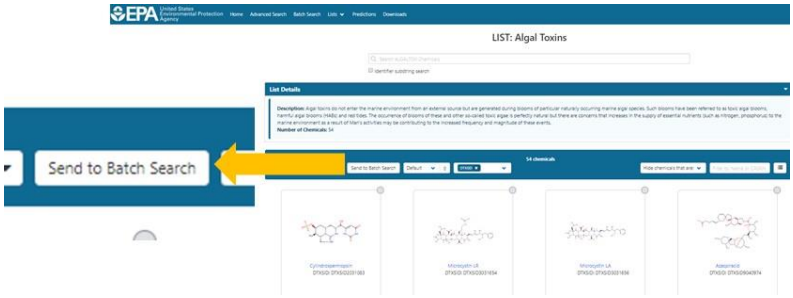


Figure 38. Send a list of chemicals to Batch Search.

The batch search page looks like Figure 39.

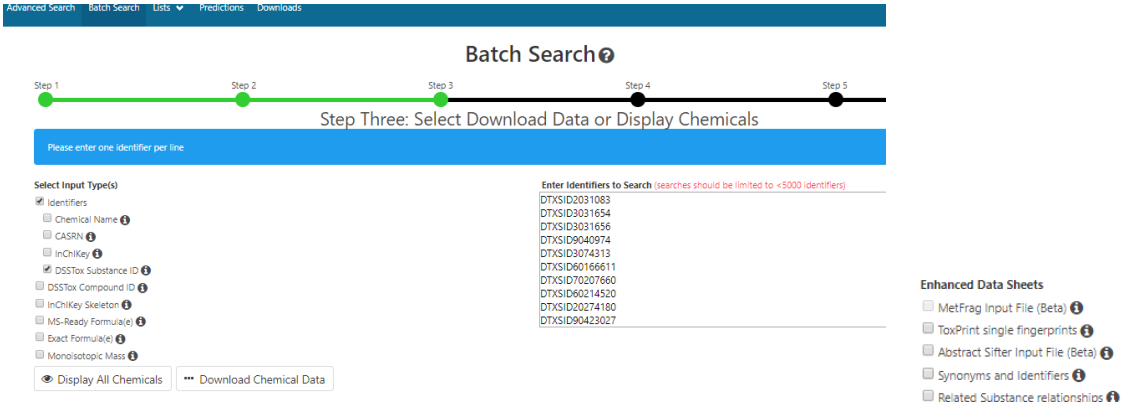


Figure 39. Send a list of chemicals to Batch Search.

Click on the following: Download Chemical Data, then Download as Excel, then Abstract Sifter Input File (Beta), then (finally) the Download bar. This action will download the chemicals to Excel. Open that file. It will have 2 sheets. Open the one that is called Abstract Sifter. It looks like Figure 40. On the Abstract Sifter Landscape sheet, unhide column A. This is done by clicking on the left border of Column B, then right-clicking to see the menu where you can click on *Unhide*. Paste rows from the downloaded spreadsheet onto the Landscape sheet as in Figure 42.

	A	B	C
1	DSSTOX_LINK	PREFERRED_NAME	CHEMICAL/ENTITY_QUERY
2	DTXSID2031083	Cylindrospermopsin	143545-90-8 OR Cylindrospermopsin
3	DTXSID3031654	Microcystin LR	101043-37-2 OR Microcystin LR OR cyanoginosin LR
4	DTXSID3031656	Microcystin LA	96180-79-9 OR Microcystin LA OR cyanoginosin-LA
5	DTXSID9040974	Azaspiracid	214899-21-5 OR Azaspiracid
6	DTXSID3074313	Saxitoxin	35523-89-8 OR Saxitoxin
7	DTXSID60166611	beta-N-Methylamino-L-alanine	15920-93-1 OR beta-N-Methylamino-L-alanine
8	DTXSID70207660	Decarbamylsaxitoxin	58911-04-9 OR Decarbamylsaxitoxin
9	DTXSID60214520	Gonyautoxin V	64296-25-9 OR Gonyautoxin V
10	DTXSID20274180	L-Domoic acid	14277-97-5 OR L-Domoic acid OR domoic acid
11	DTXSID90423027	palytoxin	77734-91-9 OR palytoxin
12	DTXSID50867064	Anatoxin a	64285-06-9 OR Anatoxin a OR anatoxin l
13	DTXSID60879996	Brevetoxin A	98112-41-5 OR Brevetoxin A
14	DTXSID20879997	Brevetoxin 2	79580-28-2 OR Brevetoxin 2
15	DTXSID40879999	Brevetoxin C	98225-48-0 OR Brevetoxin C
16	DTXSID40880000	Ciguatoxin 1	11050-21-8 OR Ciguatoxin 1
17	DTXSID00880001	Dinophysistoxin 1	81720-10-7 OR Dinophysistoxin 1
18	DTXSID60880002	Okadaic acid	78111-17-8 OR Okadaic acid
19	DTXSID10880012	Maitotoxin	59392-53-9 OR Maitotoxin
20	DTXSID90880015	Lyngbyatoxin-a	70497-14-2 OR Lyngbyatoxin-a
21	DTXSID10880017	Euglenophycin	1219817-69-2 OR Euglenophycin
22	DTXSID60880022	Nodularin	118399-22-7 OR Nodularin
23	DTXSID20880023	Yessotoxin	112514-54-2 OR Yessotoxin
24	DTXSID80880024	Azaspiracid 4	344422-49-7 OR Azaspiracid 4
25	DTXSID60880082	Aplysiatoxin	52659-57-1 OR Aplysiatoxin

Worksheet1 | Abstract Sifter

Figure 40. Excel view of downloaded chemicals on Abstract Sifter sheet.

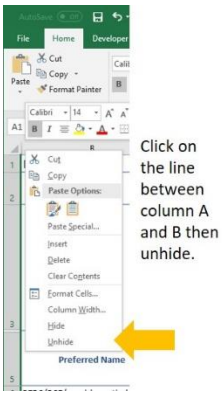
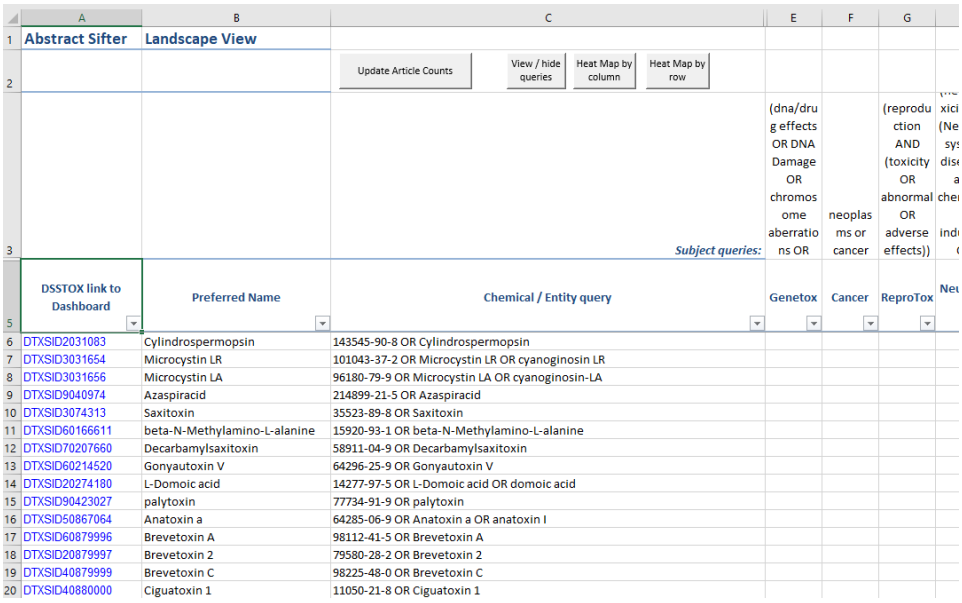


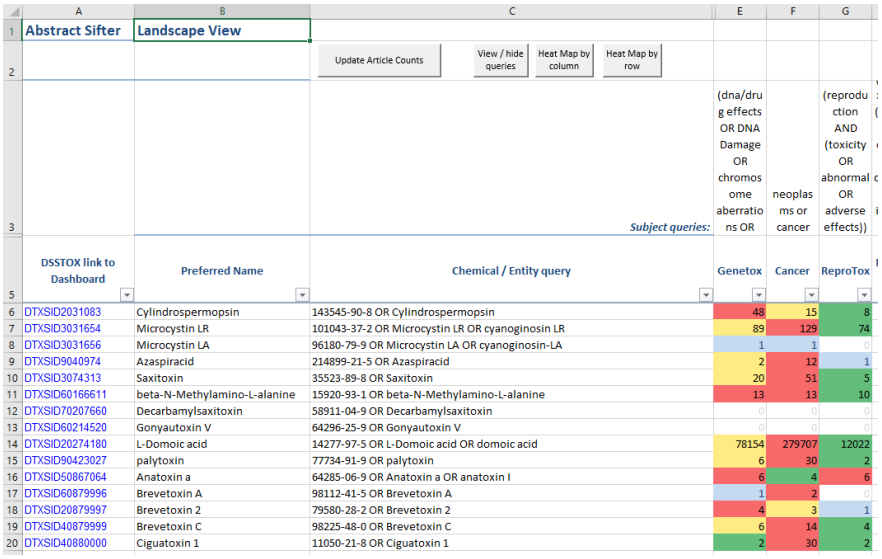
Figure 41. Unhide column A on the Landscape sheet in order to paste the DSSTox number there.



	A	B	C	E	F	G
1	Abstract Sifter	Landscape View				
2			Update Article Counts / View / hide queries / Heat Map by column / Heat Map by row			
3			*Subject queries:*	(dna/drug effects OR DNA Damage OR chromosome aberrations OR	neoplasms or cancer	(reproduction AND (toxicity OR abnormal OR adverse effects))
5	DSSTOX link to Dashboard	Preferred Name	Chemical / Entity query	Genetox	Cancer	ReproTox
6	DTXSID2031083	Cylindrospermopsin	143545-90-8 OR Cylindrospermopsin			
7	DTXSID3031654	Microcystin LR	101043-37-2 OR Microcystin LR OR cyanoginosin LR			
8	DTXSID3031656	Microcystin LA	96180-79-9 OR Microcystin LA OR cyanoginosin-LA			
9	DTXSID9040974	Azaspiracid	214899-21-5 OR Azaspiracid			
10	DTXSID3074313	Saxitoxin	35523-89-8 OR Saxitoxin			
11	DTXSID60166611	beta-N-Methylamino-L-alanine	15920-93-1 OR beta-N-Methylamino-L-alanine			
12	DTXSID70207660	Decarbamylsaxitoxin	58911-04-9 OR Decarbamylsaxitoxin			
13	DTXSID60214520	Gonyautoxin V	64296-25-9 OR Gonyautoxin V			
14	DTXSID20274180	L-Domoic acid	14277-97-5 OR L-Domoic acid OR domoic acid			
15	DTXSID90423027	palytoxin	77734-91-9 OR palytoxin			
16	DTXSID50867064	Anatoxin a	64285-06-9 OR Anatoxin a OR anatoxin I			
17	DTXSID60879996	Brevetoxin A	98112-41-5 OR Brevetoxin A			
18	DTXSID20879997	Brevetoxin 2	79580-28-2 OR Brevetoxin 2			
19	DTXSID40879999	Brevetoxin C	98225-48-0 OR Brevetoxin C			
20	DTXSID40880000	Ciguatoxin 1	11050-21-8 OR Ciguatoxin 1			

Figure 42. This is what the sheet will look like after unhiding Column A and pasting the chemicals downloaded from the Dashboard.

Now, enter subject matter queries, or, if you already have queries in place, select the intersecting cells and click on *Update Article Counts*. Click on one of the Heat Map buttons to make it pretty.



	A	B	C	E	F	G
1	Abstract Sifter	Landscape View				
2			Update Article Counts / View / hide queries / Heat Map by column / Heat Map by row			
3			*Subject queries:*	(dna/drug effects OR DNA Damage OR chromosome aberrations OR	neoplasms or cancer	(reproduction AND (toxicity OR abnormal OR adverse effects))
5	DSSTOX link to Dashboard	Preferred Name	Chemical / Entity query	Genetox	Cancer	ReproTox
6	DTXSID2031083	Cylindrospermopsin	143545-90-8 OR Cylindrospermopsin	48	15	8
7	DTXSID3031654	Microcystin LR	101043-37-2 OR Microcystin LR OR cyanoginosin LR	89	129	74
8	DTXSID3031656	Microcystin LA	96180-79-9 OR Microcystin LA OR cyanoginosin-LA	1	1	0
9	DTXSID9040974	Azaspiracid	214899-21-5 OR Azaspiracid	2	12	1
10	DTXSID3074313	Saxitoxin	35523-89-8 OR Saxitoxin	20	51	5
11	DTXSID60166611	beta-N-Methylamino-L-alanine	15920-93-1 OR beta-N-Methylamino-L-alanine	13	13	10
12	DTXSID70207660	Decarbamylsaxitoxin	58911-04-9 OR Decarbamylsaxitoxin	0	0	0
13	DTXSID60214520	Gonyautoxin V	64296-25-9 OR Gonyautoxin V	0	0	0
14	DTXSID20274180	L-Domoic acid	14277-97-5 OR L-Domoic acid OR domoic acid	78154	279707	12022
15	DTXSID90423027	palytoxin	77734-91-9 OR palytoxin	6	30	2
16	DTXSID50867064	Anatoxin a	64285-06-9 OR Anatoxin a OR anatoxin I	6	4	6
17	DTXSID60879996	Brevetoxin A	98112-41-5 OR Brevetoxin A	1	2	0
18	DTXSID20879997	Brevetoxin 2	79580-28-2 OR Brevetoxin 2	4	3	1
19	DTXSID40879999	Brevetoxin C	98225-48-0 OR Brevetoxin C	6	14	4
20	DTXSID40880000	Ciguatoxin 1	11050-21-8 OR Ciguatoxin 1	2	30	2

Figure 42. Downloaded chemicals and queries with subject matter queries.

Now you have an overview of your chemicals and what literature is out in PubMed for them. Take advantage of the iterative nature of the Abstract Sifter to query, sift, read, note as much as you need.

Contact:

Contact Nancy Baker at baker.nancy@epa.gov with issues, ideas, and feedback.