Text-mining strategies to support computational research in chemical toxicity

Nancy Baker
Leidos, contractor to US EPA
ACS National Meeting, San Francisco
April 4, 2017

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

Acknowledgements

- Tom Knudsen
- Kevin Crofton
- Antony Williams
- EPA's National Center for Computational Toxicology (NCCT)

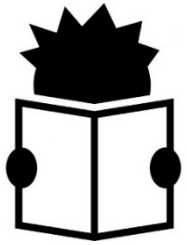
Goal today

- Literature informatics in a scientific organization
  - Five years of experience at NCCT
- Outline
  - Context, definitions, and motivation
  - Our work

Why literature informatics?

- Use the literature more effectively
- Find things you couldn’t find otherwise
- Fun

Approaches to Textual Information



Reading

Curation

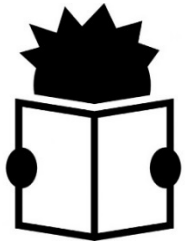
Computer-assisted curation

Indexing and article retrieval

Text-mining

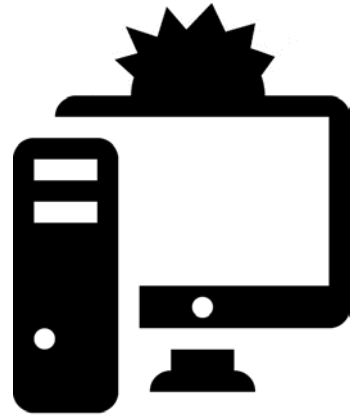
Literature Informatics

Approaches to Text



Reading

Text-mining



Literature Informatics

PubMed Abstract Sifter

Extraction of chemical properties from patents

High-throughput Text Mining (HTTM) : EPA LitDB

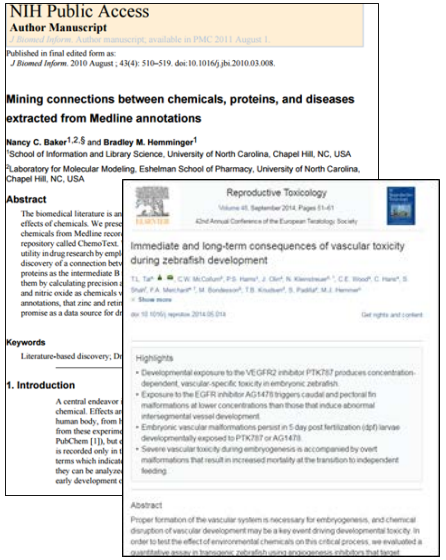
We're presenting more of this work in other sessions!

Text-mining

My definition: turning unstructured text into structured data

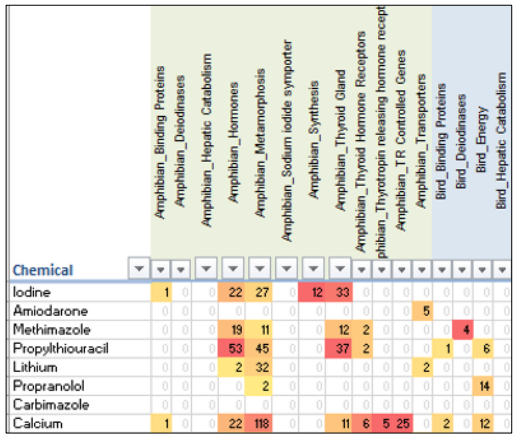
AND

Using that data to answer a question



Read it.

Why?



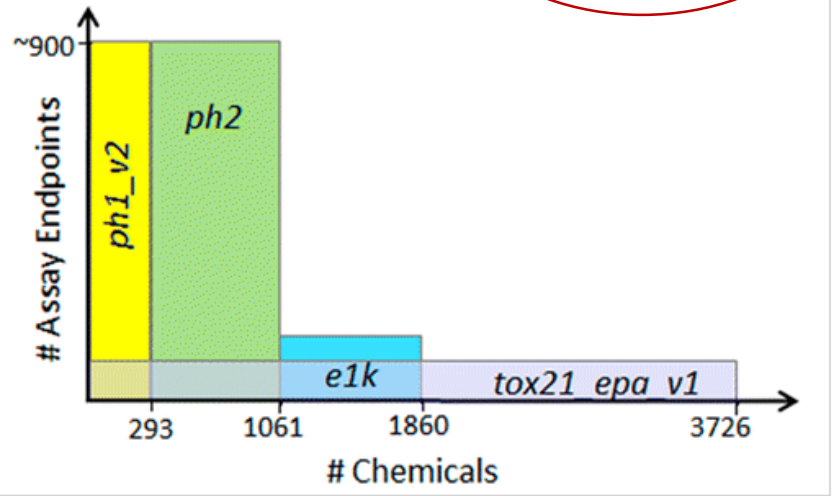
Chemical	Amphibian_Binding Proteins	Amphibian_Deiodinases	Amphibian_Hepatic Catabolism	Amphibian_Hormones	Amphibian_Metamorphosis	Amphibian_Sodium iodide symporter	Amphibian_Synthesis	Amphibian_Thyroid Gland	Amphibian_Thyroid Hormone Receptors	Amphibian_Thyrotropin releasing hormone recept	Amphibian_TR Controlled Genes	Amphibian_Transporters	Bird_Binding Proteins	Bird_Deiodinases	Bird_Energy	Bird_Hepatic Catabolism
Iodine	1	0	0	22	27	0	12	33	0	0	0	0	0	0	0	0
Amiodarone	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0
Methimazole	0	0	0	19	11	0	0	12	2	0	0	0	0	4	0	0
Propylthiouracil	0	0	0	53	45	0	0	37	2	0	0	0	1	0	6	0
Lithium	0	0	0	2	32	0	0	0	0	0	0	2	0	0	0	0
Propranolol	0	0	0	0	2	0	0	0	0	0	0	0	0	0	14	0
Carbimazole	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Calcium	1	0	0	22	118	0	0	11	6	5	25	0	2	0	12	0

Integrate it
Measure it
Formalize it
Analyze it
Compare it
Visualize it

First steps – analyze the needs

- Let's talk about our needs at the National Center for Computational Toxicology
- In response to NRC "Toxicity Testing in the 21st Century"
  - screen large sets of chemicals using in vitro assay with the goal of improving toxicity testing and prioritizing for testing the thousands of chemicals in commerce
- ToxCast and Tox21

Testing Phase	Chemical Set	Unique Chemicals	Assay Endpoints
ToxCast Phase I	*ph1_v1*	310	~700
ToxCast Phase II	*ph1_v2*	293	~200
	ph2	768	~900
	e1k	799	~50
Tox21	*tox21_epa_v1*	3726	~80



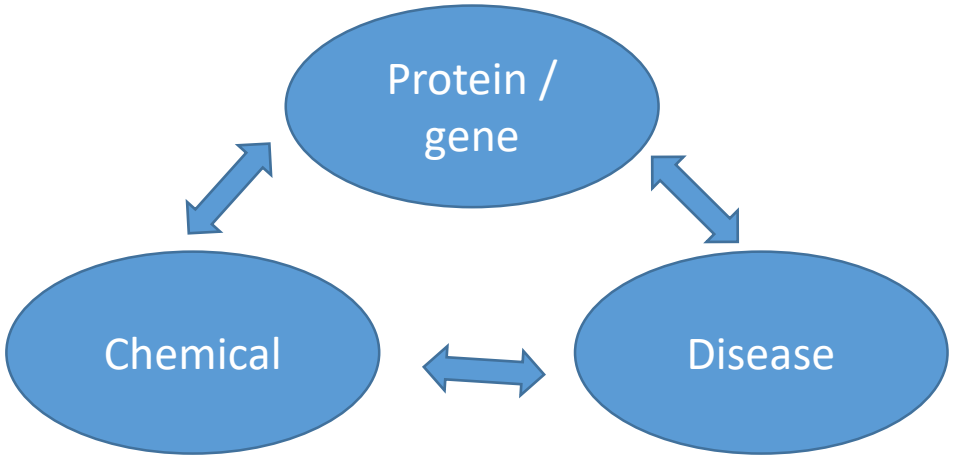
Richard AM, Judson RS, Houck KA, et al. ToxCast Chemical Landscape: Paving the Road to 21st Century Toxicology. *Chemical research in toxicology.* 2016;29(8):1225-1251.

Text-mining requirements – sample questions

- These 700 chemicals are all hits in this assay. What do these chemicals do?
- Generate a list of 30 chemicals that are kidney toxicants …
- What chemicals are described as 5-alpha reductase inhibitors in the literature?
- What genes are associated with this list of chemicals that cause liver cancer?
- What are the genes and proteins involved in the development of the embryonic heart?

Over 5 years … more than 150 such questions …

What we need – in a nutshell



Context

- Species
- Life stage
- Type of observation
- When

Methods

- Corpus – PubMed
- Strategy – take advantage of MeSH terms assigned to articles by NLM annotators
- Turn these annotations into data

N. C. Baker, B. M. Hemminger, Mining connections between chemicals, proteins, and diseases extracted from Medline annotations. *J Biomed Inform* **43**, 510 (Aug, 2010).

MeSH indexing terms become data

Format: Abstract

Send to

Biochem Pharmacol. 1993 Oct 19;46(8):1385-91.

Hexachlorobenzene-induced hypothyroidism. Involvement of different mechanisms by parent compound and metabolite.

van Raaij JA[1], Frijters CM, van den Berg KJ.

Author information

Abstract

Rats received repeated oral treatment with different doses of hexachlorobenzene (HCB) (0-3.5 mmol/kg) for 2 or 4 weeks. Measurements of thyroid hormone status after 2 weeks showed a dose-dependent decrease of total thyroxine (TT4) levels, decreased free thyroxine (FT4) levels and little change of total triiodothyronine (TT3) levels. The effects on thyroid hormone status were more pronounced after 4 weeks and also included increased thyroid stimulating hormone (TSH) levels. These conditions suggest that HCB had induced hypothyroidism in these animals. Indications for occupation of thyroid hormone binding proteins were found in serum of exposed animals. The major metabolite pentachlorophenol (PCP) also caused, by competitive interactions with thyroid hormone binding proteins in serum, a rapid and dose-dependent decrease of TT4 and FT4 levels, but not of TT3 levels in serum. The decrease of serum TT4 levels by repeated dosing with 3.5 mmol HCB/kg for 4 weeks could be attributed to competitive interactions of PCP with hormone serum binding proteins and to increased metabolism induced by HCB to an equal degree. At lower dose levels or with shorter dosing periods, increased metabolism of T4 is the main cause of decreased TT4 serum levels. This is the first indication that a similar effect is caused simultaneously by the parent compound and its metabolite through different and independent mechanisms.

PMID: 8240387

[PubMed - indexed for MEDLINE]

Publication Types, MeSH Terms, Substances

Publication Types, MeSH Terms, Substances

Publication Types

Research Support, Non-U.S. Gov't

MeSH Terms

- Animals
- Binding, Competitive
- Blood Proteins/metabolism
- Body Temperature/drug effects
- Dose-Response Relationship, Drug
- Hexachlorobenzene/blood
- Hexachlorobenzene/metabolism
- Hexachlorobenzene/toxicity*
- Hypothyroidism/chemically induced*
- Hypothyroidism/metabolism
- Male
- Pentachlorophenol/administration & dosage
- Pentachlorophenol/blood
- Rats
- Rats, Wistar
- Thyroid Hormones/metabolism
- Thyroxine/blood

National Library of Medicine Indexers

Indexing terms → data

PubMed ID	MeSH heading	Qualifier / subheading	Major topic?
8240387	Hexachlorobenzene	Toxicity	Y

PubMed ID	MeSH heading	Qualifier / subheading	Major topic?	Score
8240387	Hypothyroidism	Chemically induced	Y	2
8240387	Body Temperature	Drug effects	N	2
8240387	Thyroid Hormones	Metabolism	N	1
8240387	Thyroxine	Blood	N	1

Score reflects confidence.

We call this High-throughput text-mining (HTTM): a few readouts per article, but it adds up …

Hexachlorobenzene – 1485 articles

348 biological processes

Biological processes	Article Count
Organ Size	73
Body Weight	62
Enzyme Induction	36
Reproduction	17
Immunity	11
Birth Weight	6
Oxygen Consumption	5
Phagocytosis	5
Overweight	5
Motor Activity	4
Weight Gain	4
Cell Proliferation	4
Oxidative Stress	4
Oxidative Phosphorylation	4
Phosphorylation	4
Gluconeogenesis	4
Fertility	4
Apoptosis	4
Child Development	3
Obesity	3
Homeostasis	3
Lipid Peroxidation	3
Gene Expression	3

269 Proteins / genes

Protein / gene	Article Count
Cytochrome P-450 Enzyme System	81
Uroporphyrinogen Decarboxylase	54
Carboxy-Lyases	39
Cytochrome P-450 CYP1A1	24
5-Aminolevulinate Synthetase	21
porphyrinogen carboxy-lyase	18
Glutathione	17
Thyroxine	16
Mixed Function Oxygenases	15
Aryl Hydrocarbon Hydroxylases	15
Receptors, Aryl Hydrocarbon	15
Glutathione Transferase	12
Oxygenases	11
Aminolevulinic Acid	11
Aminopyrine N-Demethylase	11
Triiodothyronine	11
Immunoglobulin M	11
Ferrochelatase	9
Immunoglobulin G	9
Receptors, Estrogen	8
Aniline Hydroxylase	8
7-Alkoxycoumarin O-Dealkylase	8
gamma-Glutamyltransferase	8
Alanine Transaminase	6

180 Diseases / conditions

Diseases	Article Count
Porphyrias	184
Body Weight	87
Drug-Induced Liver Injury	36
Prenatal Exposure Delayed Effects	30
Disease Models, Animal	27
es	26
asms, Experimental	22
ses	21
Cutanea Tarda	16
asms	14
nt	12
plasms	11
, Experimental	10
enesis	8
us Conditions	7
Hepatocellular	6
	6
t	5
ning	5
	5
Hepatic	5
al Diseases	5
	5
eases	5
ies, Drug-Induced	5
n	5
pontaneous	5
Diseases	4
Neoplasms	4
	3
Tract Infections	3

185 Anatomical terms

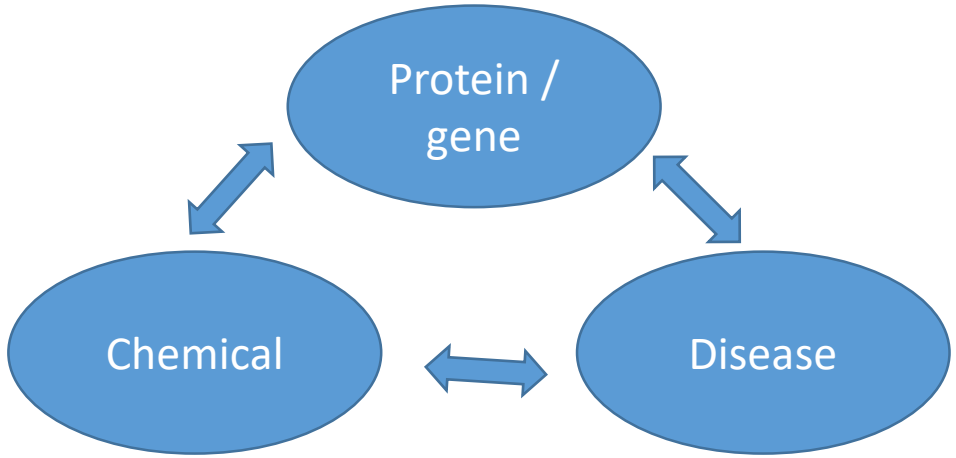
Anatomy Terms	Article Count
Liver	286
Adipose Tissue	124
Milk, Human	74
Microsomes, Liver	67
Feces	45
Kidney	39
Milk	27
Thyroid Gland	23
Skin	23
Brain	22
Lung	21
Fetal Blood	20
Muscles	19
Spleen	19
Mitochondria, Liver	17
Fetus	14
Bile	14
Ovary	12
Ovum	11
Chick Embryo	11
Placenta	11
T-Lymphocytes	11
Macrophages	10
Erythrocytes	10
Thymus Gland	9
Intestines	9
Lymph Nodes	8
Myocardium	8

How big is the data?

- 26 million articles in PubMed
- 12+ million articles have chemical annotations
- 200 million MeSH annotations
- Growth rate: 1 million / month
- ~238K chemicals
- ~141K small molecule chemicals

How we use the data

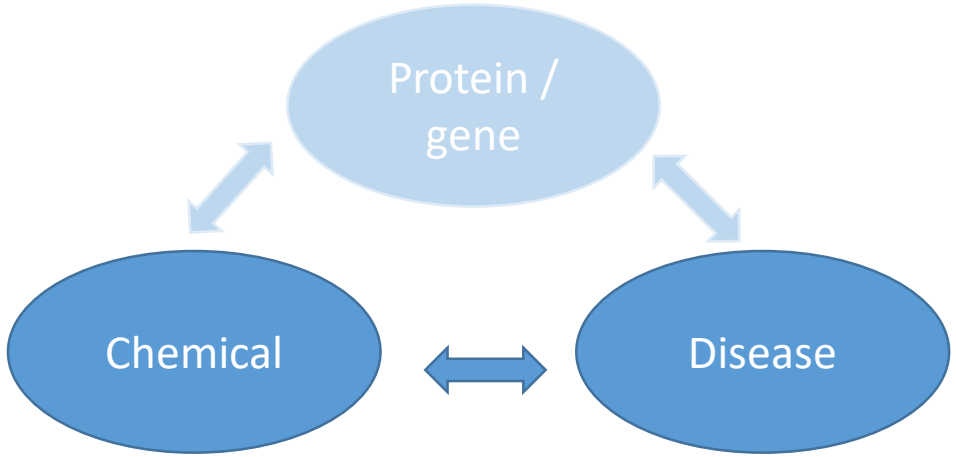
- Simple queries – simple lists – binary relationships



Context

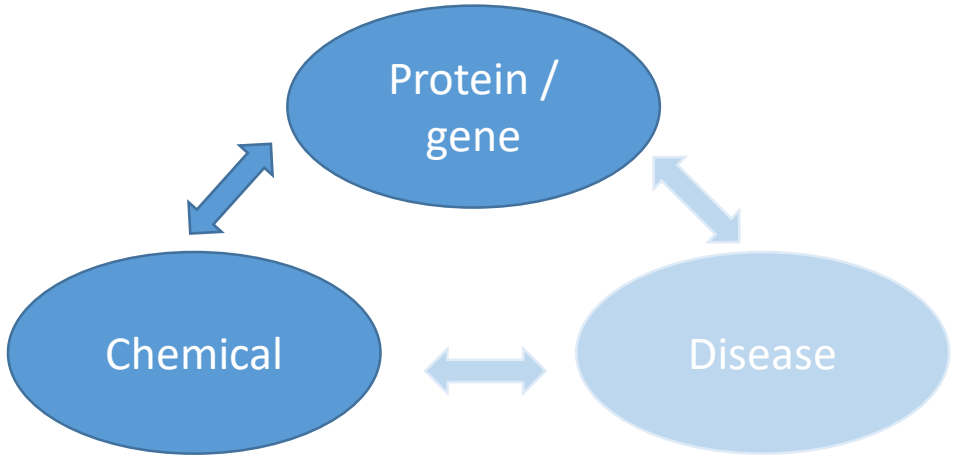
- Species
- Life stage
- Type of observation
- When

Example 1.



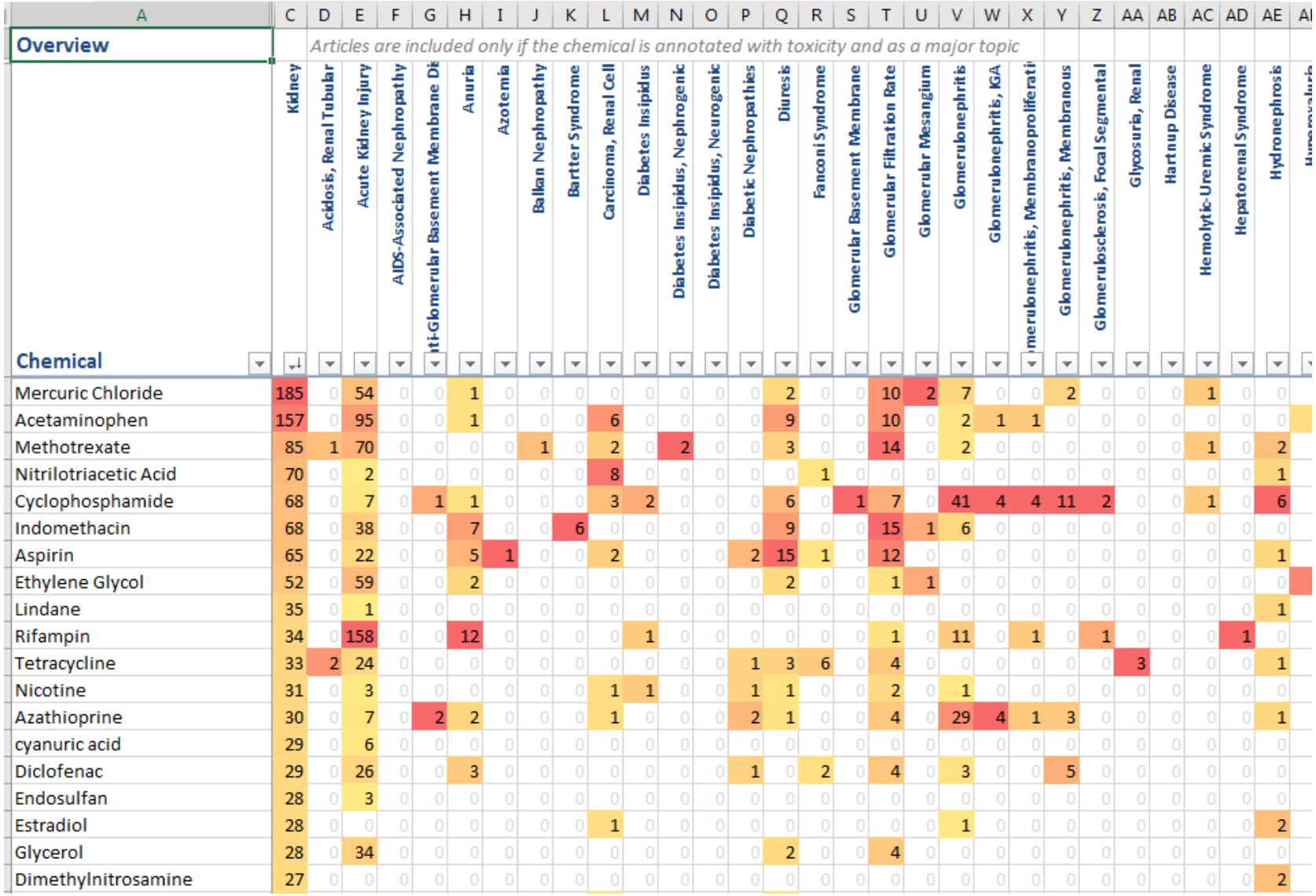
	A	T	AN
1	Chemicals associated with causing infertility		
5	Chemical	Article Ct	
6	Sulfasalazine	168	
7	Lead	152	
8	Cyclophosphamide	88	
9	Propane	88	
10	Cadmium	80	
11	Diethylstilbestrol	74	
12	1,2-dibromo-3-chloropropane	68	
13	Gossypol	48	
14	Testosterone	42	
15	Ethanol	40	
16	Cisplatin	38	
17	Finasteride	36	
18	bisphenol A	32	
19	Clomiphene	28	
20	Mercury	24	
21	Endosulfan	24	
22	Estradiol	24	
23	Ethylnitrosourea	22	
24	Colchicine	20	
25	Methotrexate	20	
26	Diethylhexyl Phthalate	20	
27	Doxorubicin	20	
28	Sirolimus	20	
29	Busulfan	19	
30	Valproic Acid	16	
31	Chlorambucil	16	
32	alpha-Chlorohydrin	16	
33	Cottonseed Oil	16	
34	DDT	16	
35	Methoxychlor	16	

Example 2.



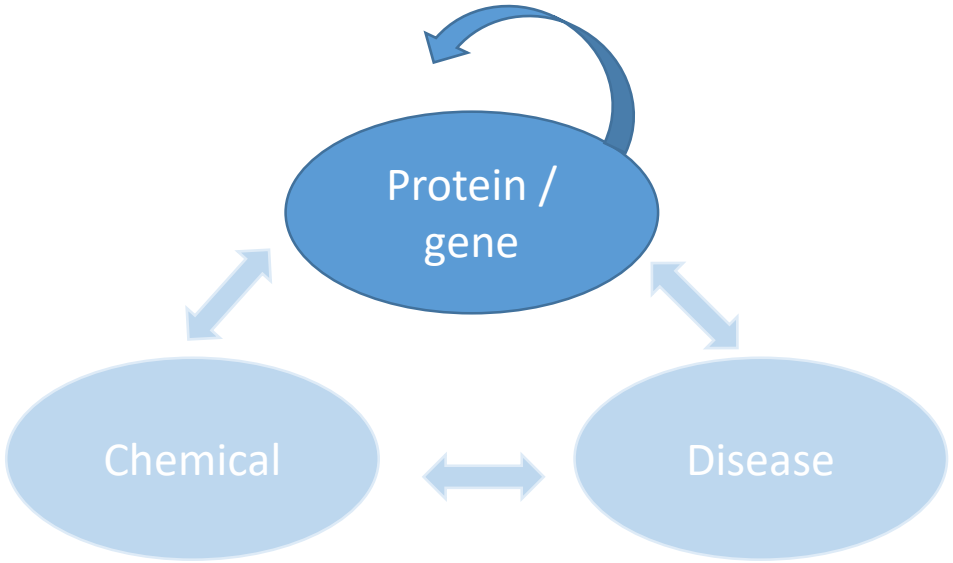
	A	AE
1	**Chemicals associated with sodium iodide symporter**	
5	**Chemical**	**Article Ct**
6	Iodine	81
7	perchlorate	21
8	Tretinoin	18
9	Sodium Iodide	14
10	Sodium	12
11	Astatine	10
12	Sodium Pertechnetate Tc 99m	9
13	Technetium	8
14	Rhenium	8
15	thiocyanate	5
16	resveratrol	4
17	Valproic Acid	3
18	Amiodarone	3
19	sunitinib	3
20	Potassium Iodide	3
21	Water	3
22	sodium perchlorate	3
23	tetrafluoroboric acid	3
24	Doxorubicin	2
25	Hydrocortisone	2
26	Metformin	2
27	2,3',4,4',5-pentachlorobiphenyl	2
28	Dexamethasone	2
29	Nevirapine	2
30	Hydrogen Peroxide	2
31	Oxygen	2
32	(6-(4-(2-piperidin-1-ylethoxy)phenyl))-3-p	2
33	3-amino-1-methyl-5H-pyrido(4,3-b)indole	2
34	4,4'-Diisothiocyanostilbene-2,2'-Disulfonic	2
35	AZD 6244	2
36	Harmaline	2
37	Measles Vaccine	2
38	N-(oxo-5,6-dihydrophenanthridin-2-yl)-N,	2

- What chemicals are associated with kidney toxicity?



Overview	Articles are included only if the chemical is annotated with toxicity and as a major topic																														
Chemical	Kidney	Acidosis, Renal Tubular	Acute Kidney Injury	AIDS-Associated Nephropathy	ti-Glomerular Basement Membrane Di	Anuria	Azotemia	Balkan Nephropathy	Bartter Syndrome	Carcinoma, Renal Cell	Diabetes Insipidus	Diabetes Insipidus, Nephrogenic	Diabetes Insipidus, Neurogenic	Diabetic Nephropathies	Diuresis	Fanconi Syndrome	Glomerular Basement Membrane	Glomerular Filtration Rate	Glomerular Mesangium	Glomerulonephritis	Glomerulonephritis, IGA	merulonephritis, Membranoproliferati	Glomerulonephritis, Membranous	Glomerulosclerosis, Focal Segmental	Glycosuria, Renal	Hartnup Disease	Hemolytic-Uremic Syndrome	Hepatorenal Syndrome	Hydronephrosis	
Mercuric Chloride	185	0	54	0	0	1	0	0	0	0	0	0	0	0	2	0	0	10	2	7	0	0	2	0	0	0	1	0	0		
Acetaminophen	157	0	95	0	0	1	0	0	0	6	0	0	0	0	9	0	0	10	0	2	1	1	0	0	0	0	0	0	0		
Methotrexate	85	1	70	0	0	0	0	1	0	2	0	2	0	0	3	0	0	14	0	2	0	0	0	0	0	0	1	0	2		
Nitrilotriacetic Acid	70	0	2	0	0	0	0	0	0	8	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1		
Cyclophosphamide	68	0	7	0	1	1	0	0	0	3	2	0	0	0	6	0	1	7	0	41	4	4	11	2	0	0	1	0	6		
Indomethacin	68	0	38	0	0	7	0	0	6	0	0	0	0	0	9	0	0	15	1	6	0	0	0	0	0	0	0	0	0		
Aspirin	65	0	22	0	0	5	1	0	0	2	0	0	0	2	15	1	0	12	0	0	0	0	0	0	0	0	0	0	1		
Ethylene Glycol	52	0	59	0	0	2	0	0	0	0	0	0	0	0	2	0	0	1	1	0	0	0	0	0	0	0	0	0	0		
Lindane	35	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1		
Rifampin	34	0	158	0	0	12	0	0	0	0	1	0	0	0	0	0	0	1	0	11	0	1	0	1	0	0	0	1	0		
Tetracycline	33	2	24	0	0	0	0	0	0	0	0	0	0	1	3	6	0	4	0	0	0	0	0	0	3	0	0	0	1		
Nicotine	31	0	3	0	0	0	0	0	0	1	1	0	0	1	1	0	0	2	0	1	0	0	0	0	0	0	0	0	0		
Azathioprine	30	0	7	0	2	2	0	0	0	1	0	0	0	2	1	0	0	4	0	29	4	1	3	0	0	0	0	0	1		
cyanuric acid	29	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Diclofenac	29	0	26	0	0	3	0	0	0	0	0	0	0	1	0	2	0	4	0	3	0	0	5	0	0	0	0	0	0		
Endosulfan	28	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Estradiol	28	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2		
Glycerol	28	0	34	0	0	0	0	0	0	0	0	0	0	0	2	0	0	4	0	0	0	0	0	0	0	0	0	0	0		
Dimethylnitrosamine	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2		

Relationships in context



Overview	Action Potentials	Adipogenesis	Anaphase	Apoptosis	Asymmetric Cell Division	Autocrine Communication	Autophagy	Calcium Signaling	Cell Adhesion	Cell Aggregation	Cell Aging	Cell Communication	Cell Compartmentation	Cell Cycle	Cell Cycle Checkpoints
Gene - gene															
MAPK3_EGR1	0	0	0	6	0	1	0	3	3	0	0	0	0	2	0
MAPK3_MAPK1	0	0	0	4	0	1	0	2	2	0	0	0	0	1	1
AKT2_AKT1	0	0	0	9	0	0	1	0	1	0	1	0	0	1	1
AKT3_AKT1	0	0	0	9	0	0	1	0	1	0	1	0	0	1	1
AKT3_AKT2	0	0	0	9	0	0	1	0	1	0	1	0	0	1	1
EGR1_CREB3	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0
PRKCZ_EGR1	0	0	0	3	0	1	0	1	0	0	0	0	0	1	0
EGR1_EGFR	0	0	0	3	0	1	0	0	1	0	0	0	0	0	0
NFKB1_EGR1	0	0	0	9	0	0	0	1	1	0	0	0	0	2	0
JUN_EGR1	1	0	0	7	0	0	0	1	2	0	1	0	0	3	0
MAPK3_EGFR	0	0	0	0	0	0	0	0	0	1	1	0	0	0	3
PPARG_EGR1	0	0	0	3	0	0	0	1	0	0	0	0	0	0	6

Cell Size	Cell Survival	Cell Transdifferentiation	Cellular Microenvironment	Cellular Reprogramming	Chemotaxis	Chemotaxis, Leukocyte
0	7	0	0	0	0	0
0	8	0	0	0	0	0
0	6	0	0	2	0	0
0	6	0	0	2	0	0
0	6	0	0	2	0	0
0	2	0	0	0	0	0
0	2	0	0	0	0	0
0	3	0	0	0	1	0
0	6	0	0	0	0	0
0	3	0	0	0	0	0
0	2	0	0	2	0	0
0	2	0	0	2	0	0
0	2	0	0	2	0	0
0	3	0	0	0	0	0
0	3	0	0	0	0	0
0	3	0	0	0	0	0
0	3	0	0	0	0	0
0	2	0	0	0	0	1
0	4	6	2	84	0	0
0	0	0	0	0	0	0
0	4	0	1	0	0	0
0	0	0	0	0	0	0
0	3	0	0	0	0	0
0	1	0	0	0	0	0
0	2	0	0	1	0	0
0	1	0	0	0	0	0
0	1	0	0	0	0	0
0	1	0	0	0	0	0
0	3	0	0	0	0	0

Text-mining for inference

- In earlier examples, somebody wrote it down. But what about when people haven’t written it down?
- Don Swanson – undiscovered public knowledge
- Inference for hypothesis generation

Swanson DR. Fish oil, Raynaud's syndrome, and undiscovered public knowledge. *Perspectives in biology and medicine.* 1986;30(1):7-18.

Thyroid disruptors – very complex pathway

- If we could pull together observations on different species, we may have insight into what chemicals are true thyroid disruptors.

- Evidence

    - Over many years
    - Over wide variety of disciplines
    - Collected for many different reasons

- Mining that undiscovered public knowledge

Thyroid disruption – the inference famework

Overview — *How to use: double-click on scores in cells below.*

Amphibian

Row	Chemical	Amphibian_Binding Proteins	Amphibian_Deiodinases	Amphibian_Hepatic Catabolism	Amphibian_Hormones	Amphibian_Metamorphosis	Amphibian_Sodium iodide symporter	Amphibian_Synthesis	Amphibian_Thyroid Gland	Amphibian_Thyroid Hormone Receptors	Amphibian_Thyrotropin releasing hormone receptor	Amphibian_TR Controlled Genes	Amphibian_Transporters
5937	methylparaben	0	0	0	0	0	0	0	0	0	0	0	0
5938	methylselenic acid	0	0	0	0	0	0	0	0	0	0	0	0
5939	methylxanthine	0	0	0	0	2	0	0	0	0	0	0	0
5940	Methysergide	0	0	0	0	2	0	0	0	0	0	0	0
5941	Metiamide	0	0	0	0	0	0	0	0	0	0	0	0
5942	metocurine	0	0	0	0	0	0	0	0	0	0	0	0
5943	metolachlor	0	0	0	0	2	0	0	0	0	0	0	0
5944	Metolazone	0	0	0	0	0	0	0	0	0	0	0	0
5945	metomidate	0	0	0	0	1	0	0	0	0	0	0	0
5946	metralindole	0	0	0	0	0	0	0	0	0	0	0	0
5947	metrenperone	0	0	0	0	0	0	0	0	0	0	0	0
5948	Metribolone	0	0	0	0	0	0	0	0	0	0	0	0
5949	metribuzin	0	0	0	0	0	0	0	0	0	0	0	0
5950	Metrizamide	0	0	0	0	1	0	0	0	0	0	0	0
5951	Metrizoate	0	0	0	0	0	0	0	0	0	0	0	0
5952	metsulfuron methyl	0	0	0	0	0	0	0	0	0	0	0	0
5953	metylperon	0	0	0	0	0	0	0	0	0	0	0	0
5954	Mevalonic Acid	0	0	0	0	0	0	0	0	0	0	0	0
5955	Mevinphos	0	0	0	0	0	0	0	0	0	0	0	0
5956	mezerein	0	0	0	0	0	0	0	0	0	0	0	0
5957	Mezlocillin	0	0	0	0	0	0	0	0	0	0	0	0

Bird

Row	Chemical	Bird_Binding Proteins	Bird_Deiodinases	Bird_Energy	Bird_Hepatic Catabolism	Bird_Hormones	Bird_Sodium iodide symporter	Bird_Synthesis	Bird_Thyroid Gland	Bird_Thyroid Hormone Receptors	Bird_Thyrotropin releasing hormone receptor	Bird_TR Controlled Genes	Bird_Transporters
5937	methylparaben	0	0	0	0	0	0	0	0	0	0	0	0
5938	methylselenic acid	0	0	0	0	0	0	0	0	0	0	0	0
5939	methylxanthine	0	0	0	0	0	0	0	0	0	0	0	0
5940	Methysergide	0	0	0	0	0	0	0	0	0	0	0	0
5941	Metiamide	0	0	0	0	0	0	0	0	0	0	0	0
5942	metocurine	0	0	0	0	0	0	0	0	0	0	0	0
5943	metolachlor	0	0	0	0	0	0	0	0	0	0	0	0
5944	Metolazone	0	0	0	0	0	0	0	0	0	0	0	0
5945	metomidate	0	0	0	0	0	0	0	0	0	0	0	0
5946	metralindole	0	0	0	0	0	0	0	0	0	0	0	0
5947	metrenperone	0	0	0	0	0	0	0	0	0	0	0	0
5948	Metribolone	0	0	0	0	0	0	0	0	0	0	0	0
5949	metribuzin	0	0	0	0	0	0	0	0	0	0	0	0
5950	Metrizamide	0	0	0	0	0	0	0	0	0	0	0	0
5951	Metrizoate	0	0	0	0	0	0	0	0	0	0	0	0
5952	metsulfuron methyl	0	0	0	0	0	0	0	0	0	0	0	0
5953	metylperon	0	0	0	0	0	0	0	0	0	0	0	0
5954	Mevalonic Acid	0	0	0	0	0	0	0	0	0	0	0	0
5955	Mevinphos	0	0	0	0	0	0	0	0	0	0	0	0
5956	mezerein	0	0	0	0	0	0	0	0	0	0	0	0
5957	Mezlocillin	0	0	0	0	0	0	0	0	0	0	0	0

Fish

Row	Chemical	Fish_Binding Proteins	Fish_Deiodinases	Fish_Hepatic Catabolism	Fish_Hormones	Fish_Sodium iodide symporter	Fish_Synthesis	Fish_Thyroid Gland	Fish_Thyroid Hormone Receptors	Fish_TR Controlled Genes	Fish_Transporters
5937	methylparaben	0	0	0	0	0	0	0	0	0	0
5938	methylselenic acid	0	0	0	0	0	0	0	0	0	0
5939	methylxanthine	0	0	0	0	0	0	0	0	0	0
5940	Methysergide	0	0	0	0	0	0	0	0	0	0
5941	Metiamide	0	0	0	0	0	0	0	0	0	0
5942	metocurine	0	0	0	0	0	0	0	0	0	0
5943	metolachlor	0	0	0	0	0	0	3	2	0	0
5944	Metolazone	0	0	0	0	0	0	0	0	0	0
5945	metomidate	0	0	0	0	0	0	0	0	0	0
5946	metralindole	0	0	0	0	0	0	0	0	0	0
5947	metrenperone	0	0	0	0	0	0	0	0	0	0
5948	Metribolone	0	0	0	0	0	0	0	0	0	0
5949	metribuzin	0	0	0	0	0	0	0	0	0	0
5950	Metrizamide	0	0	0	0	0	0	0	0	0	0
5951	Metrizoate	0	0	0	0	0	0	0	0	0	0
5952	metsulfuron methyl	0	0	0	0	0	0	0	0	0	0
5953	metylperon	0	0	0	0	0	0	0	0	0	0
5954	Mevalonic Acid	0	0	0	0	0	0	0	0	0	0
5955	Mevinphos	0	0	0	0	0	0	0	0	0	0
5956	mezerein	0	0	0	0	0	0	0	0	0	0
5957	Mezlocillin	0	0	0	0	0	0	0	0	0	0

Mammal

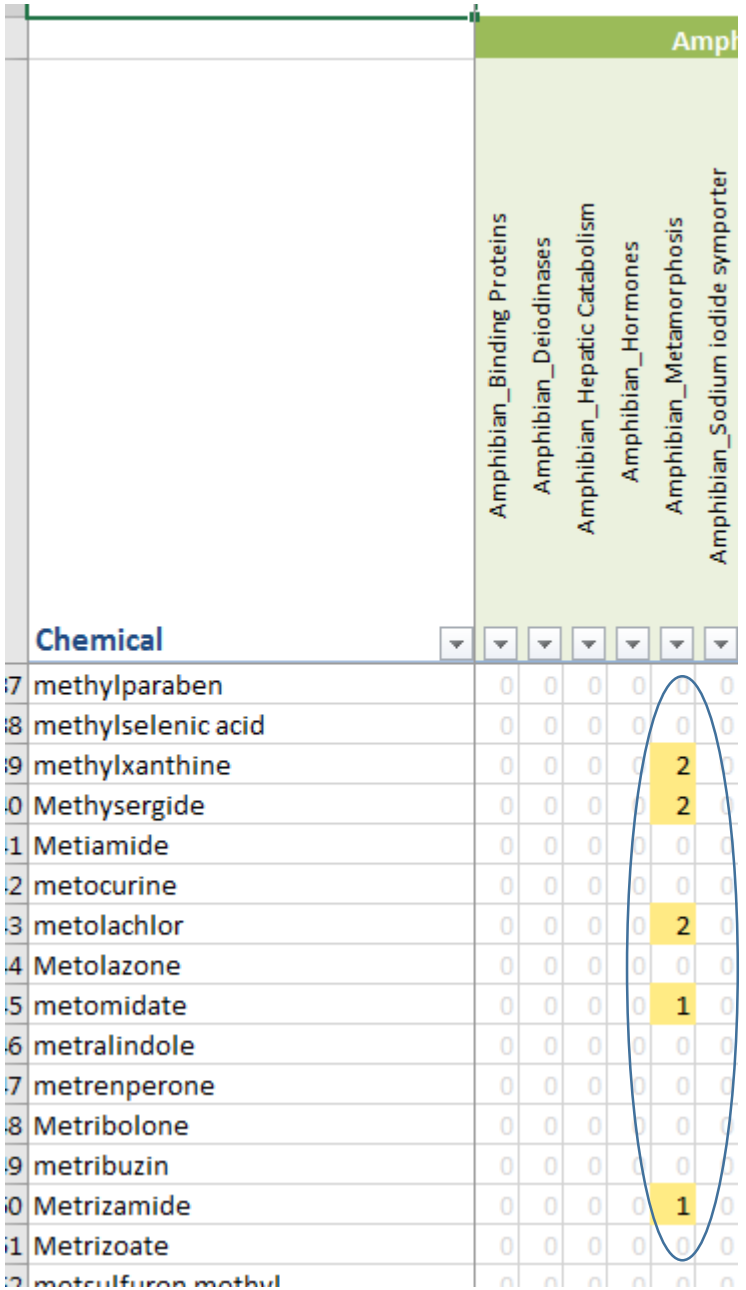
Row	Chemical	Mammal_Binding Proteins	Mammal_Deiodinases	Mammal_Energy	Mammal_Hepatic Catabolism	Mammal_Hormones	Mammal_Sodium iodide symporter	Mammal_Synthesis	Mammal_Thyroid Gland	Mammal_Thyroid Hormone Receptors	Mammal_Thyroid stimulating hormone receptor	Mammal_Thyrotropin releasing hormone receptor	Mammal_TR Controlled Genes	Mammal_Transporters
5937	methylparaben	0	0	0	2	5	0	0	3	0	0	0	0	0
5938	methylselenic acid	0	0	0	0	0	0	0	0	0	0	0	2	0
5939	methylxanthine	0	0	5	0	1	0	0	3	1	0	0	0	0
5940	Methysergide	1	0	11	0	25	0	2	7	0	0	0	0	0
5941	Metiamide	0	0	4	0	0	0	0	0	0	0	0	0	0
5942	metocurine	0	0	2	0	0	0	0	0	0	0	0	0	0
5943	metolachlor	0	0	0	1	1	0	0	3	0	0	0	0	0
5944	Metolazone	0	0	0	1	0	0	0	0	0	0	0	0	1
5945	metomidate	0	0	0	0	0	0	0	0	0	0	0	0	0
5946	metralindole	0	0	2	0	0	0	0	0	0	0	0	0	0
5947	metrenperone	0	0	3	0	0	0	0	0	0	0	0	0	0
5948	Metribolone	0	0	0	2	1	0	0	0	0	0	0	1	0
5949	metribuzin	0	0	0	1	0	0	0	0	0	0	0	0	0
5950	Metrizamide	1	0	8	0	2	0	0	4	0	0	0	0	0
5951	Metrizoate	0	0	3	0	0	0	0	0	0	0	0	0	0
5952	metsulfuron methyl	0	0	0	0	0	0	0	1	0	0	0	0	0
5953	metylperon	0	0	0	0	0	0	0	0	0	0	0	0	2
5954	Mevalonic Acid	0	0	0	1	4	0	1	0	0	0	0	0	0
5955	Mevinphos	0	0	0	0	0	0	1	0	0	0	0	0	0
5956	mezerein	0	0	0	0	4	0	1	6	0	0	0	2	2
5957	Mezlocillin	1	0	0	0	0	0	0	0	0	0	0	0	0

Human

Row	Chemical	Human_Binding Proteins	Human_Deiodinases	Human_Hepatic Catabolism	Human_Hormones	Human_Sodium iodide symporter	Human_Synthesis	Human_Thyroid Hormone Receptors	Human_Thyroid stimulating hormone receptor	Human_Thyrotropin releasing hormone receptor	Human_TR Controlled Genes	Human_Transporters	Human_Energy	Human_Thyroid Gland	Human_Cognition_IQ	Human_Clinical Conditions
5937	methylparaben	0	0	2	0	0	0	0	0	0	0	0	0	0	2	0
5938	methylselenic acid	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
5939	methylxanthine	0	0	0	1	0	0	1	0	0	0	0	5	3	3	0
5940	Methysergide	0	0	0	4	0	0	0	0	0	0	0	2	0	0	0
5941	Metiamide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5942	metocurine	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
5943	metolachlor	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5944	Metolazone	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
5945	metomidate	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5946	metralindole	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
5947	metrenperone	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5948	Metribolone	0	0	2	1	0	0	0	0	0	1	0	0	0	0	0
5949	metribuzin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5950	Metrizamide	0	0	0	0	0	0	0	0	0	0	0	8	0	21	0
5951	Metrizoate	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0
5952	metsulfuron methyl	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5953	metylperon	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
5954	Mevalonic Acid	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
5955	Mevinphos	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5956	mezerein	0	0	0	0	0	1	0	0	0	0	2	0	0	0	0
5957	Mezlocillin	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Inference process

If a chemical is associated with changes in amphibian metamorphoses …



	Chemical	Amphibian_Binding Proteins	Amphibian_Deiodinases	Amphibian_Hepatic Catabolism	Amphibian_Hormones	Amphibian_Metamorphosis	Amphibian_Sodium iodide symporter
7	methylparaben	0	0	0	0	0	0
8	methylselenic acid	0	0	0	0	0	0
9	methylxanthine	0	0	0	0	2	0
0	Methysergide	0	0	0	0	2	0
1	Metiamide	0	0	0	0	0	0
2	metocurine	0	0	0	0	0	0
3	metolachlor	0	0	0	0	2	0
4	Metolazone	0	0	0	0	0	0
5	metomidate	0	0	0	0	1	0
6	metralindole	0	0	0	0	0	0
7	metrenperone	0	0	0	0	0	0
8	Metribolone	0	0	0	0	0	0
9	metribuzin	0	0	0	0	0	0
0	Metrizamide	0	0	0	0	1	0
1	Metrizoate	0	0	0	0	0	0
2	metsulfuron methyl	0	0	0	0	0	0

If a chemical is associated with changes in amphibian metamorphoses

AND

If the same chemical is associated with thyroid activity in mammals …

	A	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV
1	Overview													
2		Mammal												
3	Chemical	Mammal_Binding Proteins	Mammal_Deiodinases	Mammal_Energy	Mammal_Hepatic Catabolism	Mammal_Hormones	Mammal_Sodium iodide symporter	Mammal_Synthesis	Mammal_Thyroid Gland	Mammal_Thyroid Hormone Receptors	Mammal_Thyroid stimulating hormone receptor	Mammal_Thyrotropin releasing hormone receptor	Mammal_TR Controlled Genes	Mammal_Transporters
5937	methylparaben	0	0	0	2	5	0	0	3	0	0	0	0	0
5938	methylselenic acid	0	0	0	0	0	0	0	0	0	0	0	2	0
5939	methylxanthine	0	0	5	0	1	0	0	3	1	0	0	0	0
5940	Methysergide	1	0	11	0	25	0	2	7	0	0	0	0	0
5941	Metiamide	0	0	4	0	0	0	0	0	0	0	0	0	0
5942	metocurine	0	0	2	0	0	0	0	0	0	0	0	0	0
5943	metolachlor	0	0	0	1	1	0	0	3	0	0	0	0	0
5944	Metolazone	0	0	0	1	0	0	0	0	0	0	0	0	1
5945	metomidate	0	0	0	0	0	0	0	0	0	0	0	0	0
5946	metralindole	0	0	2	0	0	0	0	0	0	0	0	0	0
5947	metrenperone	0	0	3	0	0	0	0	0	0	0	0	0	0
5948	Metribolone	0	0	0	2	1	0	0	0	0	0	0	1	0
5949	metribuzin	0	0	0	1	0	0	0	0	0	0	0	0	0
5950	Metrizamide	1	0	8	0	2	0	0	4	0	0	0	0	0
5951	Metrizoate	0	0	3	0	0	0	0	0	0	0	0	0	0

If a chemical is associated with changes in amphibian metamorphoses

AND

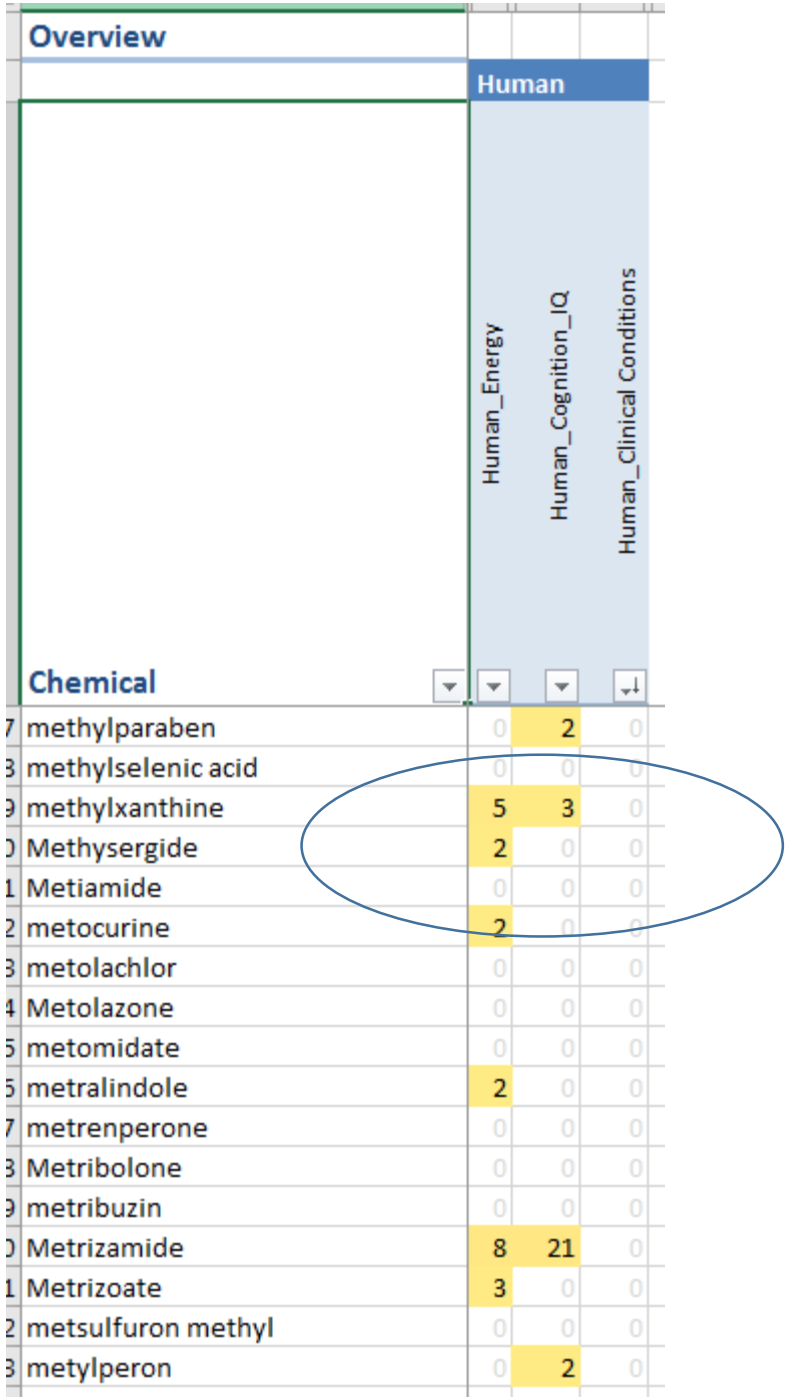
If the same chemical is associated with thyroid activity in mammals

AND

If the same chemical is associated with energy / cognition effects in humans …

MAYBE

It is a thyroid pathway disruptor.



Overview		Human		
	Chemical	Human_Energy	Human_Cognition_IQ	Human_Clinical Conditions
7	methylparaben	0	2	0
8	methylselenic acid	0	0	0
9	methylxanthine	5	3	0
0	Methysergide	2	0	0
1	Metiamide	0	0	0
2	metocurine	2	0	0
3	metolachlor	0	0	0
4	Metolazone	0	0	0
5	metomidate	0	0	0
6	metralindole	2	0	0
7	metrenperone	0	0	0
8	Metribolone	0	0	0
9	metribuzin	0	0	0
0	Metrizamide	8	21	0
1	Metrizoate	3	0	0
2	metsulfuron methyl	0	0	0
3	metylperon	0	2	0

Review the goals

- Use the literature more effectively
- Find things you couldn’t find otherwise
- Fun
- *People are asking questions they wouldn’t have asked before.*

Thank you!
… and if you want to hear more

- Tony Williams: EPA CompTox chemistry dashboard: An online resource for environmental chemists
  - Division of Chemical Health and Safety
  - Tuesday, April 4, 3:05-3:30 PM
- Drug repurposing: A bibliometric analysis by text-mining PubMed
  - Division of the History of Chemistry
  - Wednesday, April 5, 10:15, session from 8:30 – 11:45
- Supporting Read-across predictions of chemical toxicity using high-throughput text-mining
  - Division of Environmental Chemistry
  - Thursday, April 6, 10:50 (session from 8 – 12)