UNC ESHELMAN SCHOOL OF PHARMACY

MML UNC.EDU

In Silico Tools for Prediction and Mechanistic Interpretation of Systemic and Topical Toxicity

Alexander Tropsha

Laboratory for Molecular Modeling

UNC Chapel Hill, USA

Co-founder, Predictive, LLC

PREDICTIVE IN SILICO SOLUTIONS

OUTLINE

- Current NAMs (IMHO: New Alternative Methods) landscape
- Brief summary of current in silico approaches to chemical toxicity prediction
- Development of specific *in silico* tools
  - SToPTox for Systemic and Topical Toxicity Prediction
  - PreS/MD for Predicting Sensitization caused by Medical Devices
- Mechanistic interpretation of models
  - Chemical
  - Biological
- Summary and Outlook

NAMs Landscape: Key Directives

MML UNC.EDU

17TH CONGRESS
2D SESSION

S. 5002

AN ACT

To allow for alternatives to animal testing for purposes of drug and biological product applications.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

European Commission

English EN

Home > Press corner > Commission acts to accelerate phasing out of animal testing

Available languages: English

Press release | 25 July 2023 | Brussels

Commission acts to accelerate phasing out of animal testing in response to a European Citizens' Initiative

Page contents

Top

Quote(s)

Today, the Commission is responding to the European Citizens' Initiative (ECI) '*Save Cruelty-free Cosmetics - Commit to a Europe without Animal Testing*'. The response provides a comprehensive overview of

Focus

OECD
BETTER POLICIES FOR BETTER LIVES

» The OECD calls for urgent mobilisation of national and regional resources to support the validation of new methods for the safety testing of chemicals

» Omics technologies in chemical testing

» NANOMET: Towards tailored safety testing methods for nanomaterials

Guideline No. 497: Defined Approaches on Skin Sensitisation

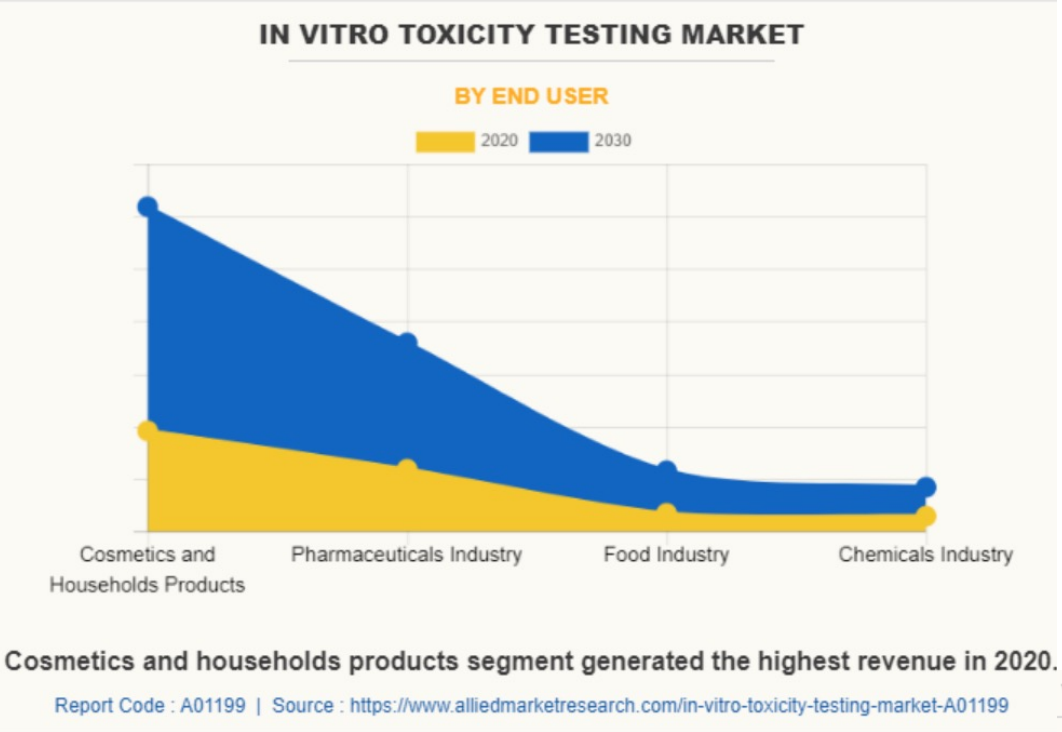
A Defined Approach (DA) consists of a selection of information sources (e.g in silico pre chemico, in vitro data) used in a specific combination, and resulting data are interpreted data interpretation procedure (DIP) (e.g. a mathematical, rule-based model). DAs use m

04 Jul 2023 | 53 pages | English | Also available in: French

https://doi.org/10.1787/b92879a4-en | 9789264903005 (PDF)

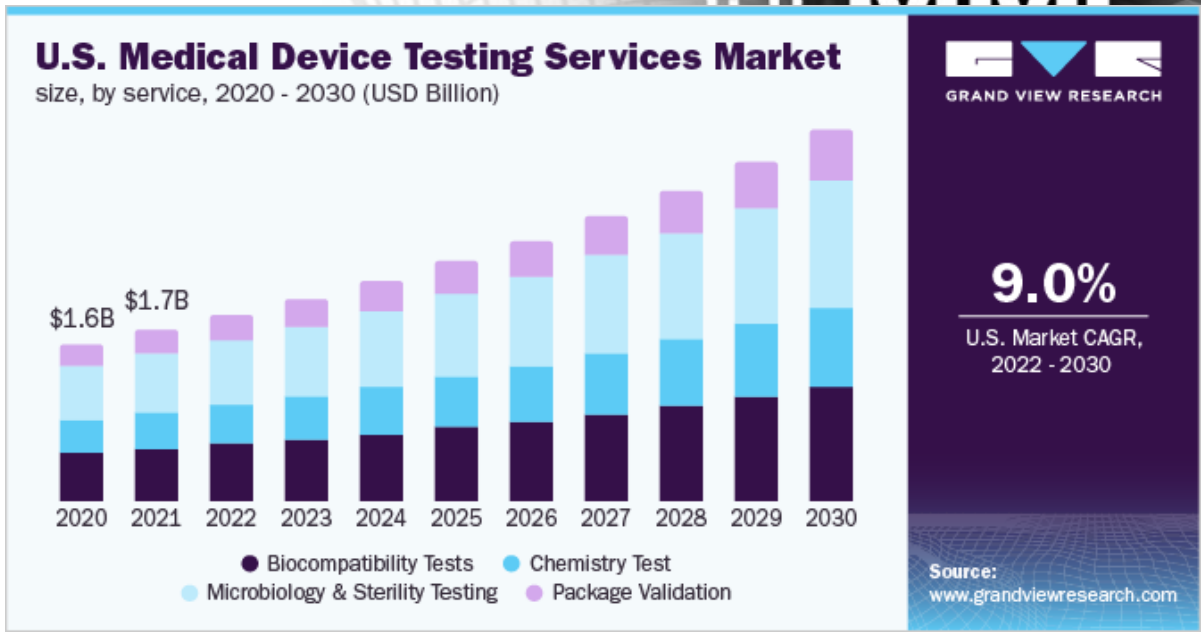
Author(s): OECD

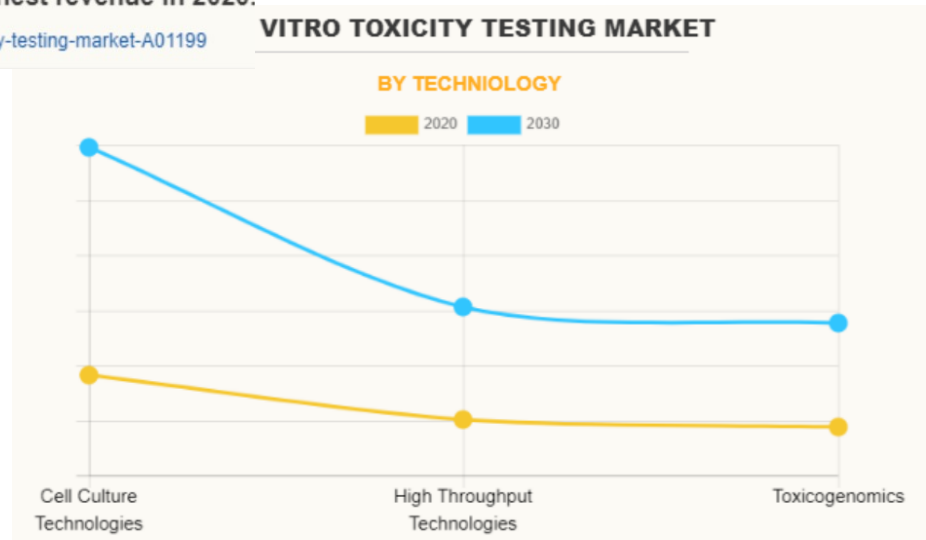
In Vitro Toxicity Testing Services: Market Trends



Cosmetics and households products segment generated the highest revenue in 2020.

Report Code : A01199 | Source : https://www.alliedmarketresearch.com/in-vitro-toxicity-testing-market-A01199





Cell culture technologies segment generated the highest revenue in 2020.

Report Code : A01199 | Source : https://www.alliedmarketresearch.com/in-vitro-toxicity-testing-market-A01199

NAMs Landscape: Tool Development

Web App for Using Defined Approaches to Predict Skin Sensitization Hazard and Potency

MML UNC.EDU

Welcome to the DASS App!

The DASS App applies defined approaches on skin sensitization (DASS) that are described in OECD Guideline No. 497 and the U.S. EPA's Interim Science Policy: Use of Alternative Approaches for Skin Sensitization as a Replacement for Laboratory Animal Testing . The defined approaches (DAs) predict skin sensitization hazard (either a sensitizer or non-sensitizer) and potency by integrating data from in vitro assays that represent key events in the Adverse Outcome Pathway (AOP) for Skin Sensitisation Initiated by Covalent Binding to Proteins and in silico hazard predictions.

More details are available in the User Guide .

For more information or to report a problem with the app, please contact NICEATM at ICE-support@niehs.nih.gov.

Step 1: Select the Defined Approaches to Apply

To begin, select the DAs to be implemented. Click on the green information buttons to view a description of the DA and the test mehods required to implement the DA.

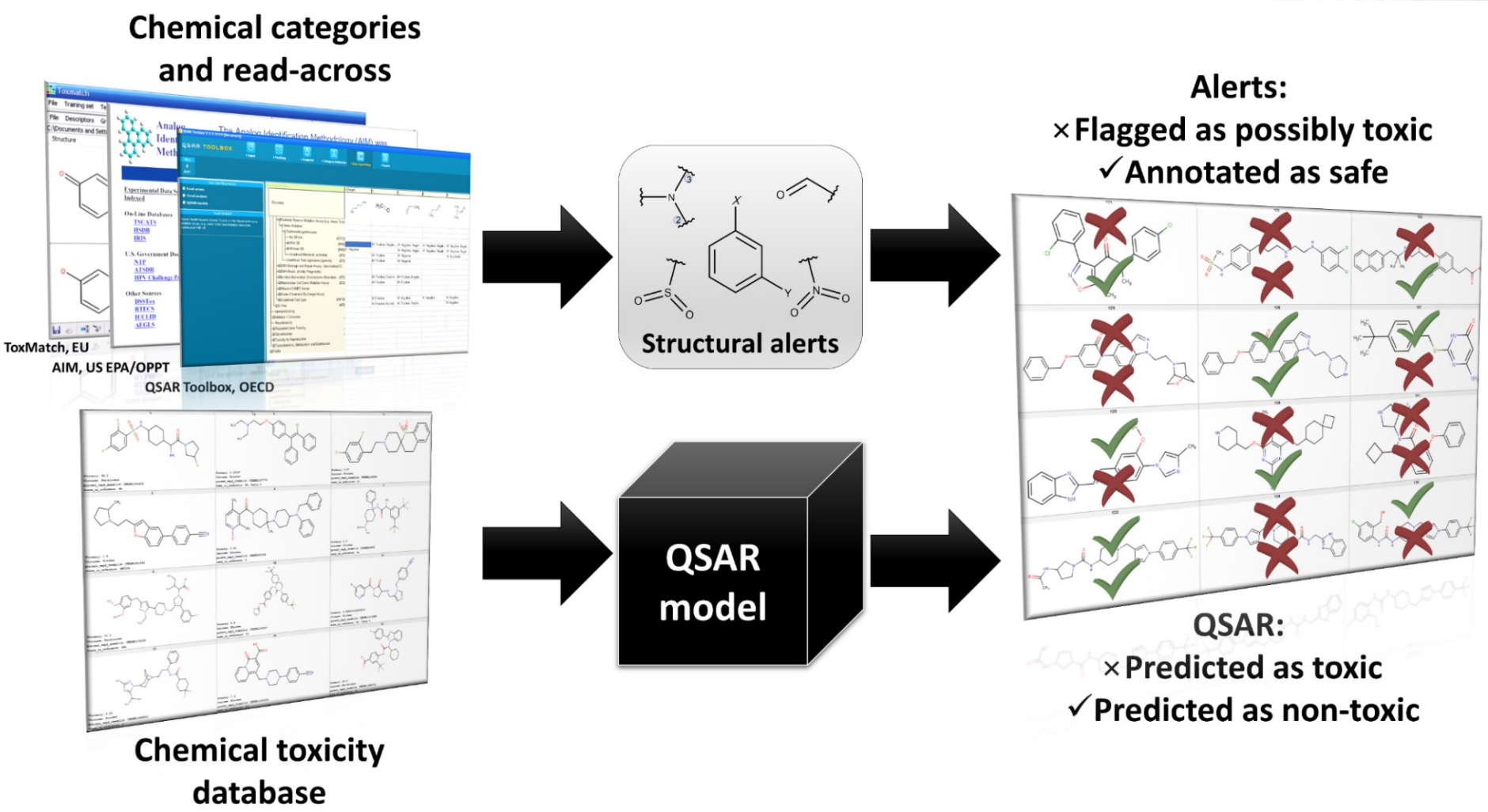
Select All | Deselect All

- [x] 2 out of 3 (2o3)
- [x] Integrated Testing Strategy (ITS)
- [x] Key Event 3/1 (KE 3/1) Sequential Testing Strategy (STS)

Step 2: Upload Data

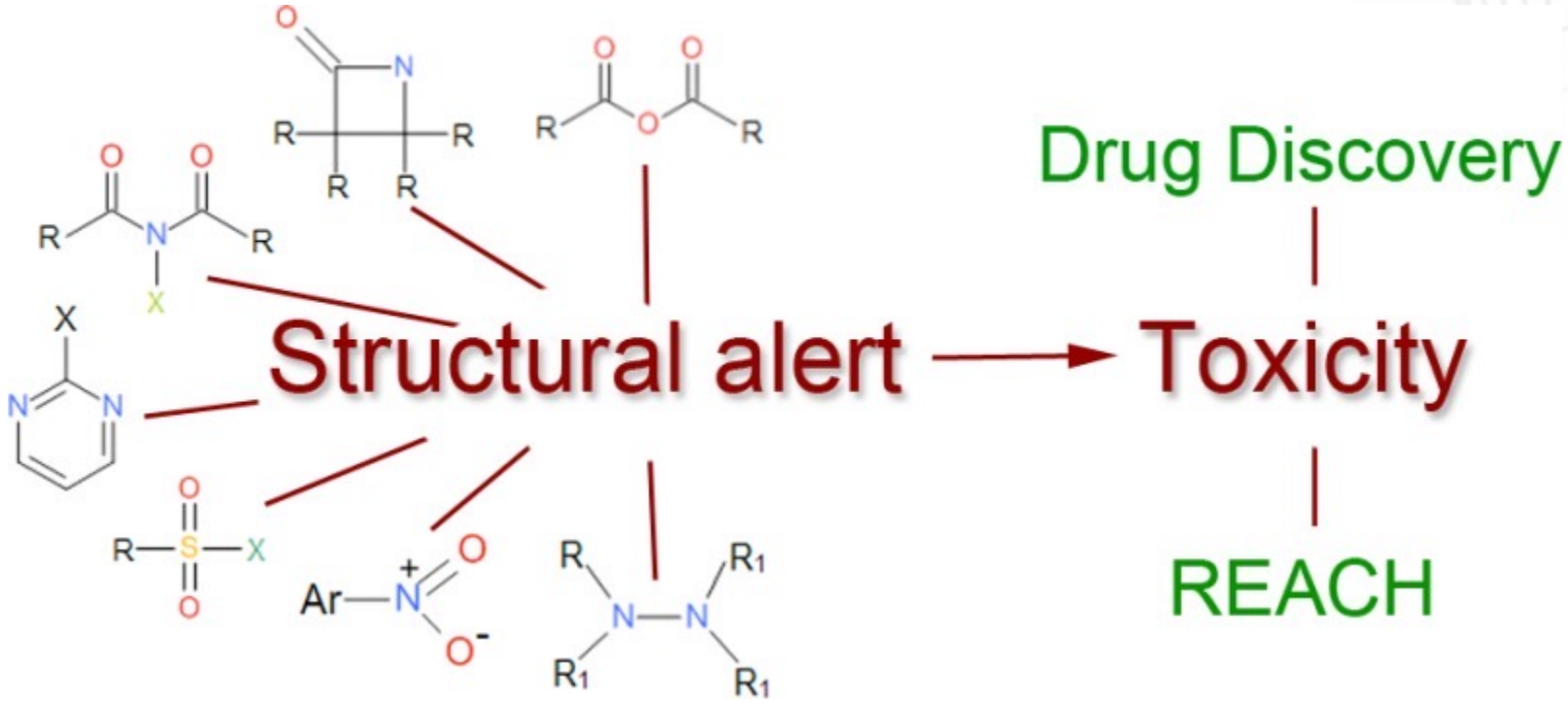
https://ntp.niehs.nih.gov/whatwestudy/niceatm/test-method-evaluations/skin-sens/da/dass-app

Major cheminformatics approaches to toxicity prediction: Structural alerts *vs*. Machine Learning (QSAR) methods



V. M. Alves et al., *Green Chem.* **2016**, *18* (16), 4348–4360.

Read-across: Structural Alerts



*Structural alerts are "molecular patterns that are associated with particular types of toxicity or ADRs either directly or after undergoing of a metabolic activation in vivo"**

*Image and definition from Sushko et al, J Chem Inf Model. 2012 Aug 27; 52(8): 2310–2316.

Alarms about Alerts: Many alerts cannot distinguish withdrawn vs. marketed drugs

Drug Name	State	QSAR prediction[2]*	Toxic hazard classification by Cramer (extension)	Toxic hazard classification by Cramer (original)	Carcinogenicity (genotox and nongenotox) alerts by ISS	DNA alerts for AMES, MN and CA by OASIS v.1.3	In vitro mutagenicity (Ames test) alerts by ISS	In vivo mutagenicity (Micronucleus) alerts by ISS
Amineptine	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	No alert found	No alert found	No alert	**Alerts**
Duract	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	No alert found	No alert found	No alert	**Alerts**
Vioxx	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	No alert found	No alert found	No alert	**Alerts**
Astemizole	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	**Alerts**	No alert found	No alert	**Alerts**
Cerivastatin	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	No alert found	No alert found	No alert	**Alerts**
Chlormezanone	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	**Alerts**	No alert found	No alert	**Alerts**
Fenfluramine	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	No alert found	No alert found	No alert	No alert
Flosequinan	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	**Alerts**	No alert found	**Alerts**	**Alerts**
Glafenine	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	**Alerts**	No alert found	No alert	**Alerts**
Grepafloxacin	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	**Alerts**	No alert found	No alert	**Alerts**
Mibefradil	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	**Alerts**	No alert found	No alert	**Alerts**
Troglitazone	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	No alert found	No alert found	No alert	**Alerts**
Ximelagatran	withdrawn	unsafe	**High (Class III)**	**High (Class III)**	No alert found	**Alerts**	No alert	**Alerts**
Aspirin	marketed	safe	Low (Class I)	Low (Class I)	No alert found	No alert found	No alert	**Alerts**
Ibuprofen	marketed	safe	Low (Class I)	Low (Class I)	No alert found	No alert found	No alert	**Alerts**
Valtrex	marketed	safe	**High (Class III)**	**High (Class III)**	No alert found	No alert found	No alert	**Alerts**
Microzide	marketed	safe	**High (Class III)**	**High (Class III)**	**Alerts**	No alert found	No alert	**Alerts**
Neurontin	marketed	safe	**High (Class III)**	**High (Class III)**	No alert found	No alert found	No alert	**Alerts**
Enoxaparin	marketed	safe	**High (Class III)**	**High (Class III)**	No alert found	No alert found	No alert	**Alerts**
Lyrica	marketed	safe	Low (Class I)	Low (Class I)	No alert found	No alert found	No alert	**Alerts**

*Zakharov, Lagunin, Poroikov. Chem. Res. Toxicol., 2012, 25, 2378–2385.

Criticism of chemical toxicity alerts

From the journal:
Green Chemistry

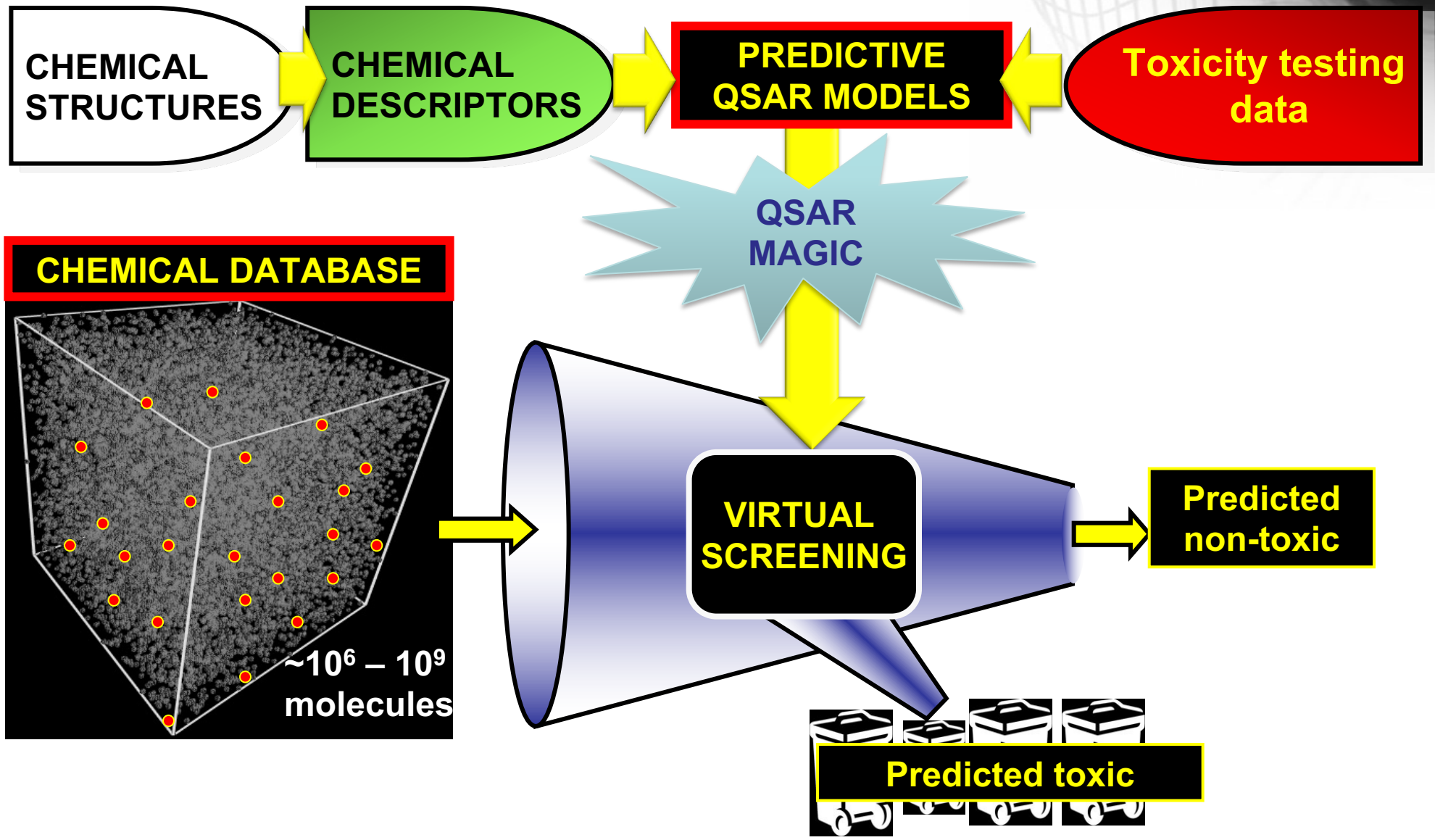
Alarms about structural alerts†

Vinicius M. Alves,[ab] Eugene N. Muratov,[ac] Stephen J. Capuzzi,[a] Regina Politi,[a] Yen Low,[d] Rodolpho C. Braga,[b] Alexey V. Zakharov,[e] Alexander Sedykh,[f] Elena Mokshyna,[g] Sherif Farag,[a] Carolina H. Andrade,[b] Victor E. Kuz'min,[g] Denis Fourches[h] and Alexander Tropsha*[a]

ACS Sustainable Chem. Eng. 2018, 6, 3, 2845–2859

Chemical Alerts of Toxicity: what are they for, really?

QSAR-based toxicity prediction



OECD principles of model validation

To facilitate the consideration of a QSAR model for regulatory purposes, it should be associated with the following information:

1) a defined endpoint
2) an unambiguous algorithm
3) a defined domain of applicability
4) appropriate measures of goodness-of-fit, robustness and predictivity *(R^2>0.8; Q^2>0.6; Q^2_{ext}>0.6)**
5) a mechanistic interpretation, if possible

** Model acceptance criteria used in our QSAR studies*

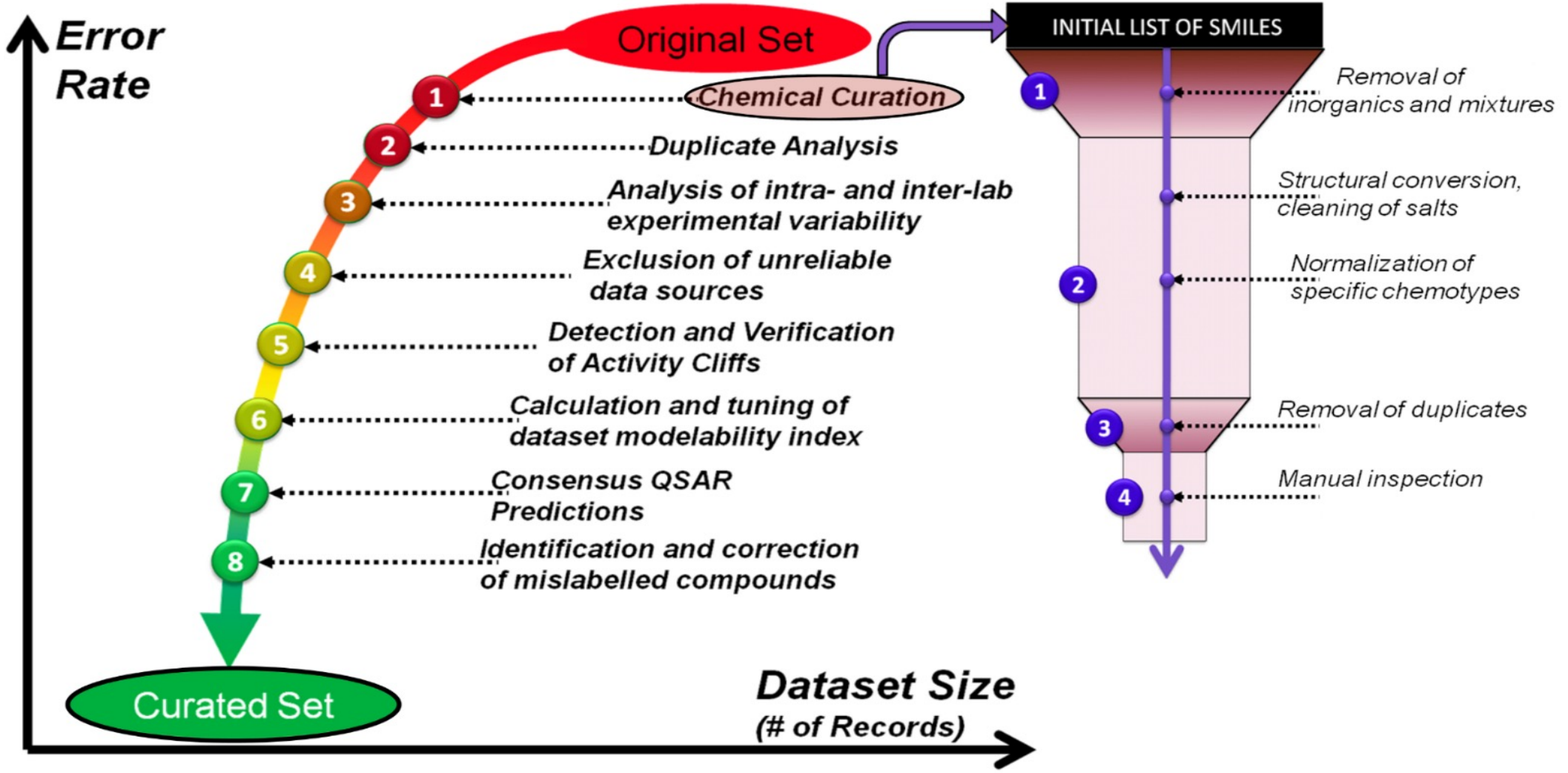
OECD = Organization of Economic Co-operation and Development
http://www.oecd.org/dataoecd/33/37/37849783.pdf

Our key principles of developing and using QSAR models as NAMs

- Data transparency and thorough curation
- Comprehensive model development and internal validation workflow
- Rigorous external validation
- Applicability domain and prediction confidence
- Mechanistic interpretation
  - Chemical functional group
  - Associated toxicity pathways

Comprehensive data curation pipeline

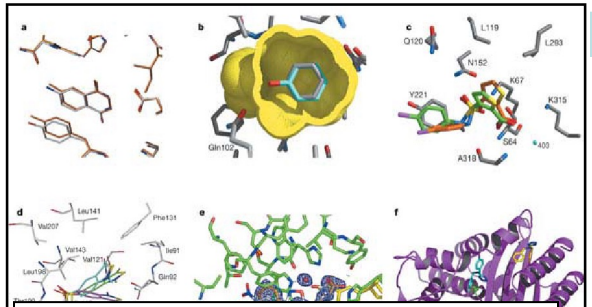
ECHA (OECD Test Guideline No. 406) + Literature; 1639 records



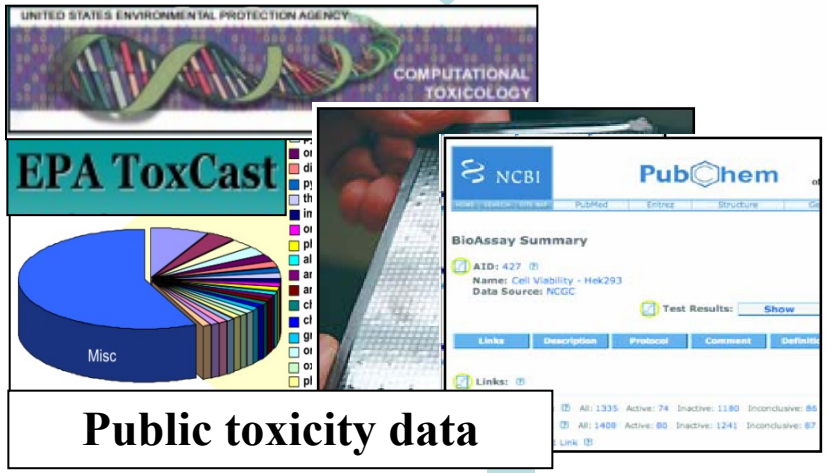
1238 compounds

Fourches, D., Muratov, E. & Tropsha, A. Curation of chemogenomics data. *Nat Chem Biol* **11**, 535 (2015). https://doi.org/10.1038/nchembio.1881

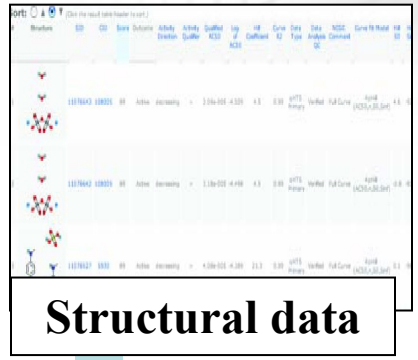
Data Continuum/Modeling workflow



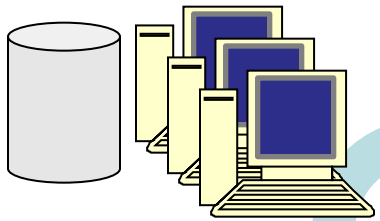
Chemical substances



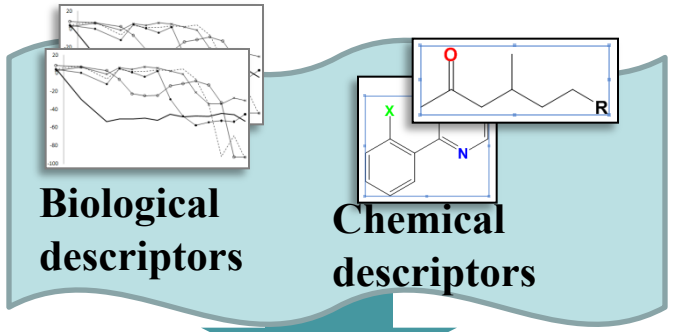
Public toxicity data



Structural data

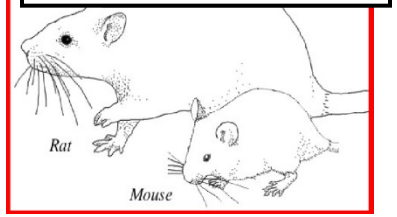


Data collection and curation



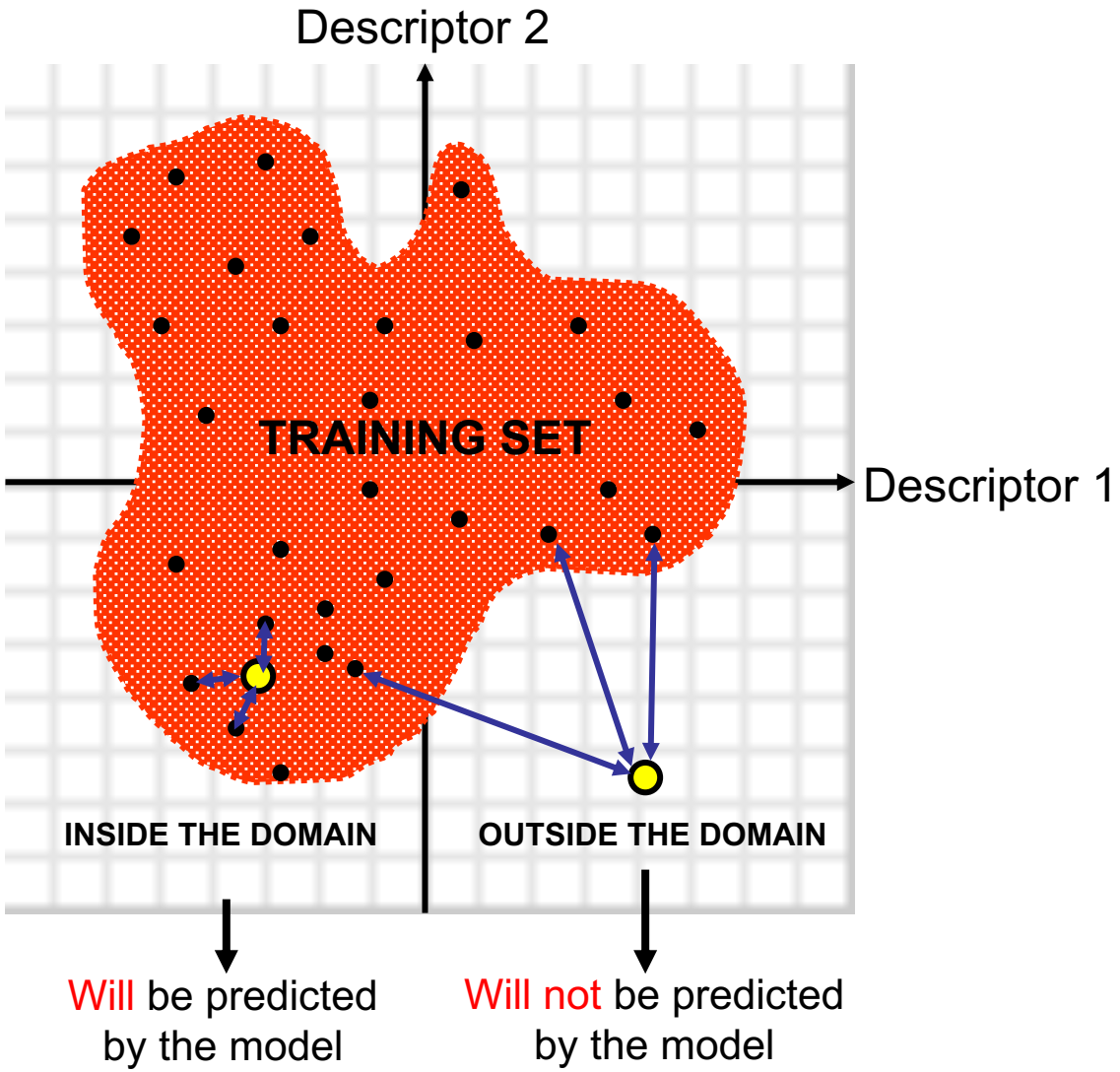
Combi-QSAR modeling

In vivo toxicity prediction



Accessible models

Applicability domain of QSAR models



For a given model, two parameters are calculated:

- $\langle D_k \rangle$: average euclidian distance between each compound of the training set and its *k* nearest neighbors in the descriptors space;
- s_k : standard deviation of the distances between each compound of the training set and its k nearest neighbors in the descriptors space.

= NEW COMPOUND

For each test compound *i*, the distance D_i is calculated as the average of the distances between i and its *k* nearest neighbors in the training set.

The new compound will be predicted by the model, only if :

$$D_i \leq \langle D_k \rangle + Z \times s_k$$

with Z, an empirical parameter (0.5 by default)

The “six-pack” battery of acute toxicity tests

Acute oral
(rat)
OECD TG 401, 420, 423, and 425

Acute dermal
(rat)
OECD TG 402

Ocular irritation
(rabbit)
OECD TG 405

Skin Sensitization
(mouse, guinea pig)
OECD TG 406, 429 and 442D

Skin irritation
(rabbit)
OECD TG 404

Acute inhalation
(rat)
OECD TG 403, and 436

PREDICTIVE
IN SILICO SOLUTIONS

UNC ESHELMAN SCHOOL OF PHARMACY

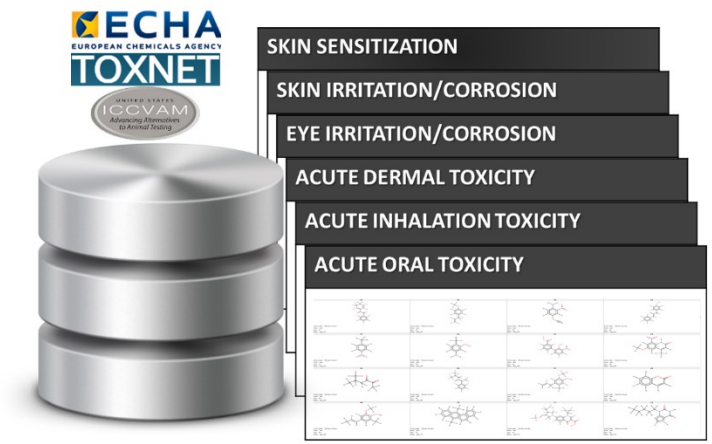
MML UNC.EDU

STopTox - web-based predictor of Systemic and Topical Toxicity

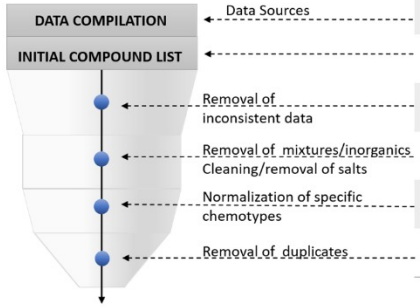
- The largest curated “6-pack” dataset in public domain
- Models developed and validated in compliance with OECD guidelines
- Comprehensive computational platform that can be used as an alternative to 6-pack regulatory animal assays (current cost ca. 10K/compound/assay).
- Implemented as Web Portal, http://stoptox.mml.unc.edu
- Commercial version available from Predictive, LLC

General workflow for STopTox

Data Retrieval

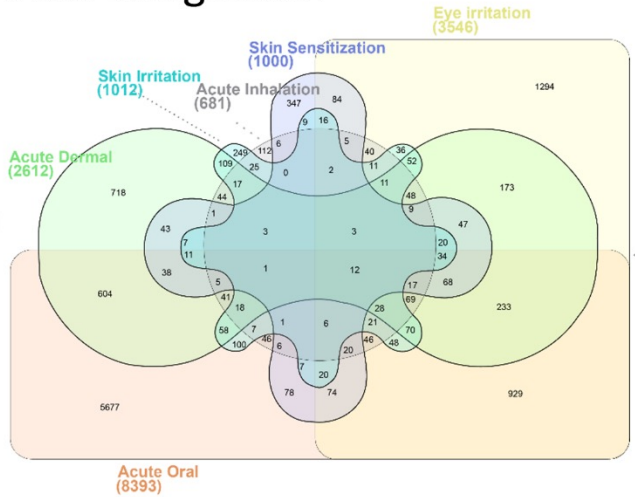


Data Curation

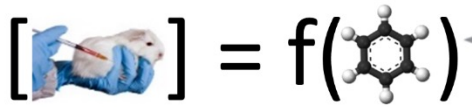


Skin Sensitization	Skin Irritation/ Corrosion	Eye Irritation/ Corrosion	Acute Dermal Toxicity	Acute Inhalation Toxicity	Acute Oral Toxicity
ECHA/ ICCVAM	ECHA	ECHA/ Literature	ECHA/ ToxValDB/ Creton et al.	ECHA/ ToxValDB	NICETATM
10,861	5,274	7,332	29,824	8,176	8,994
2,347	1,631	7,196	5,259	2,061	8,987
1,110	1,326	6,006	4,601	1,644	8,952
1,110	1,326	5,985	4,601	1,644	8,952
1,000	1,012	3,547	2,622	691	8,465

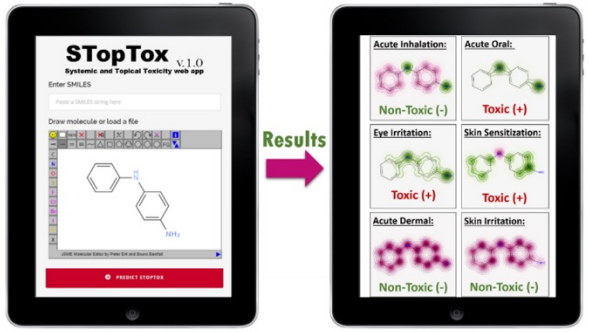
Data Integration



QSAR modeling



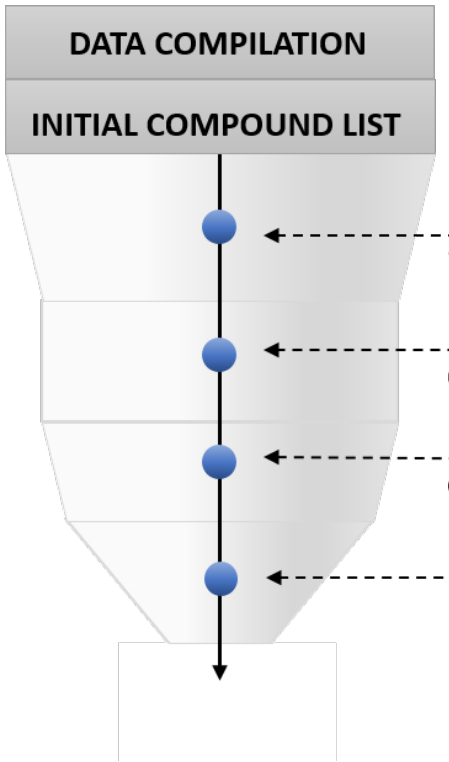
STopTox Server



Datasets for endpoints to develop STopTox

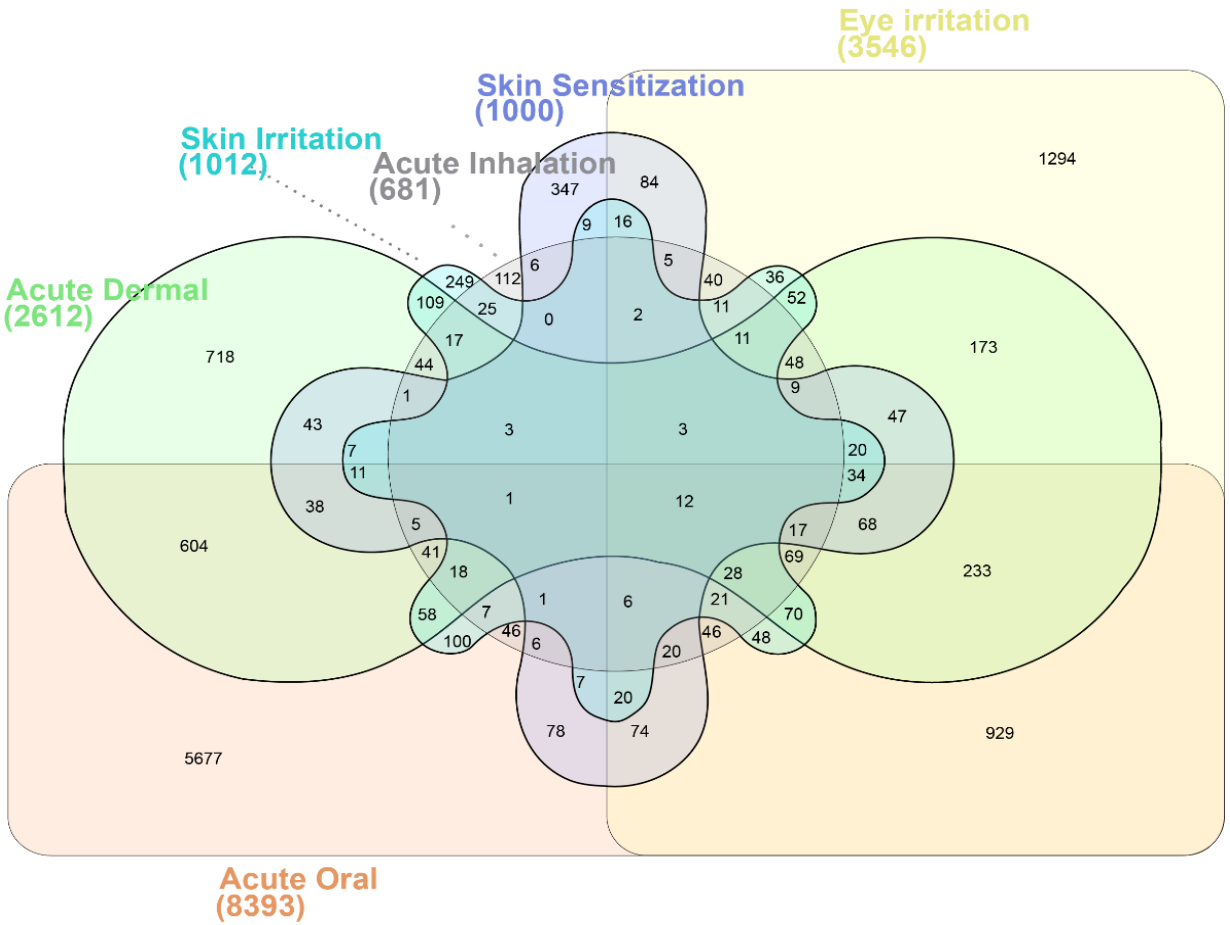
Endpoint	Assay type	# Chemicals before/after curation
Skin Sensitization	LLNA, DPRA, KeratinoSens, h-CLAT, human data	**11,648 / 1,000** **194 / 194** **190 / 190** **160 / 160** **302 / 138**
Skin irritation/corrosion	Draize test	**5,274 / 1012**
Eye irritation/corrosion	Draize test	**6,387 / 3,547**
Acute dermal	Acute dermal toxicity test	**29,824 / 2,622**
Acute inhalation	Acute inhalation toxicity test	**8,176 / 681**
Acute oral	Acute oral toxicity test	**8,994 / 8,795**

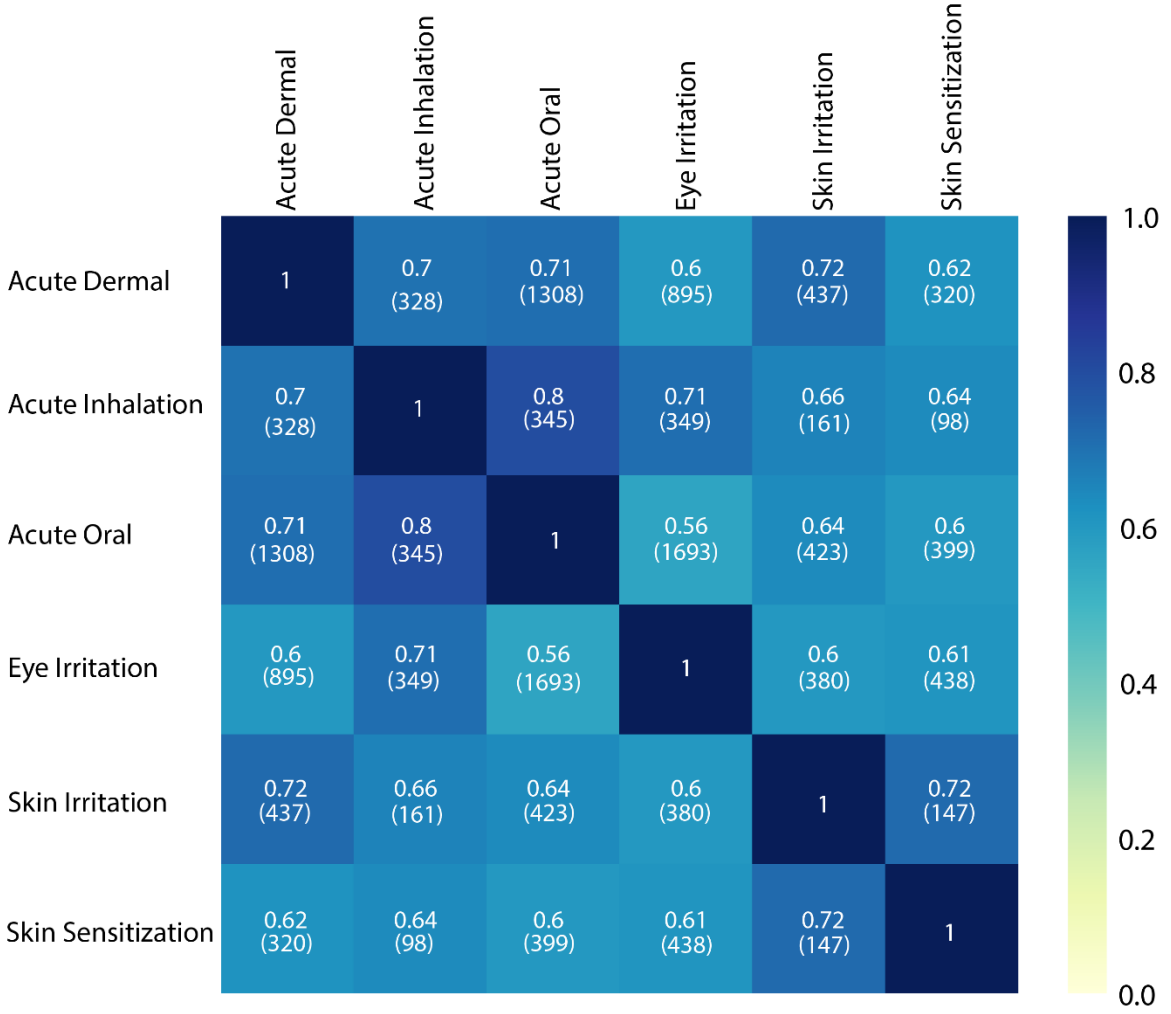
Data Curation



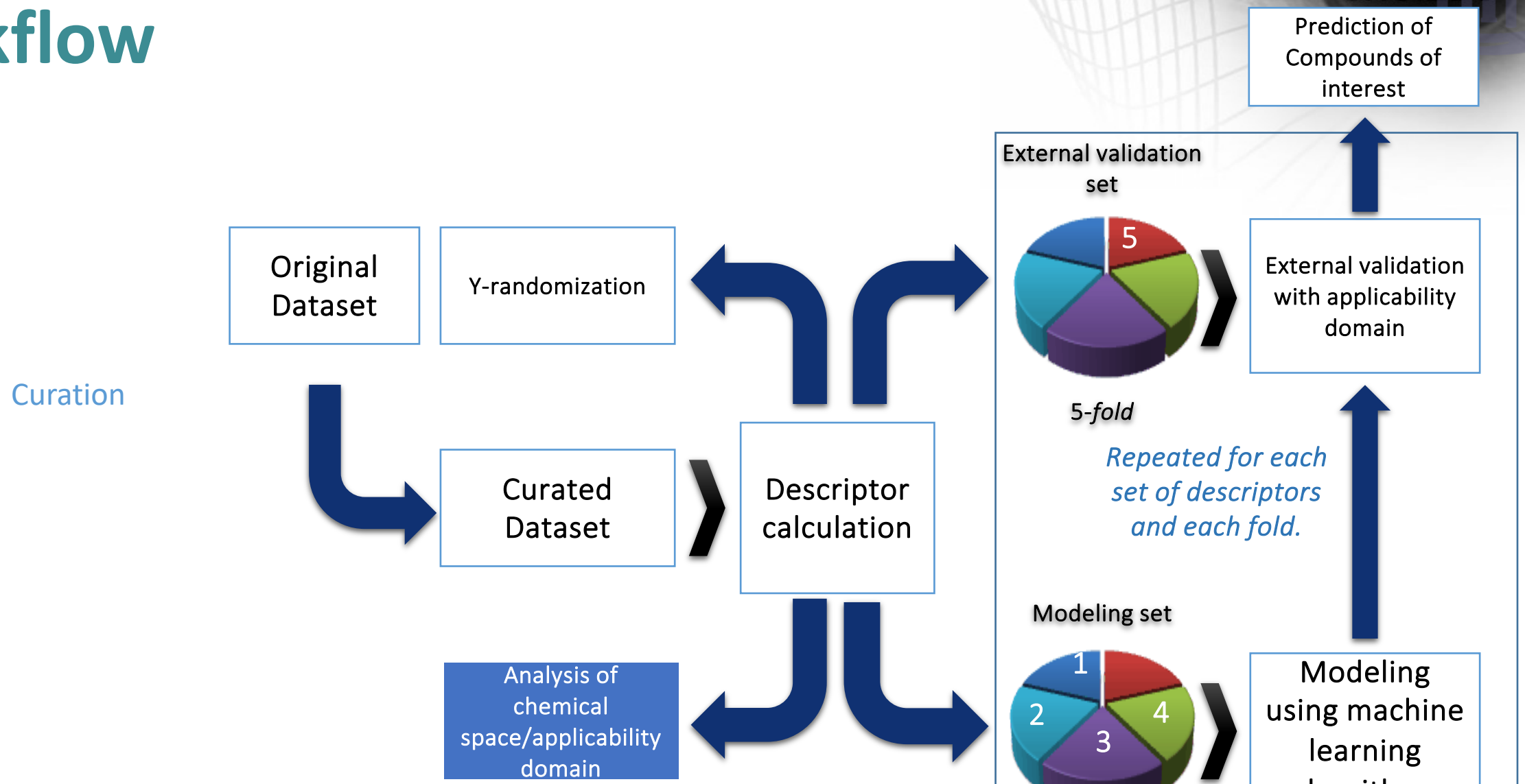
Data Sources	Skin Sensitization	Skin Irritation/ Corrosion	Eye Irritation/ Corrosion	Acute Dermal Toxicity	Acute Inhalation Toxicity	Acute Oral Toxicity
	ECHA/ ICCVAM	ECHA	ECHA/ Literature	ECHA/ ToxValDB/ Creton et al.	ECHA/ ToxValDB	NICETATM
INITIAL COMPOUND LIST	10,861	5,274	7,332	29,824	8,176	8,994
Removal of inconsistent data	2,347	1,631	7,196	5,259	2,061	8,987
Removal of mixtures/inorganics Cleaning/removal of salts	1,110	1,326	6,006	4,601	1,644	8,952
Normalization of specific chemotypes	1,110	1,326	5,985	4,601	1,644	8,952
Removal of duplicates	1,000	1,012	3,547	2,622	691	8,465

Cross-endpoint analysis of accessible data

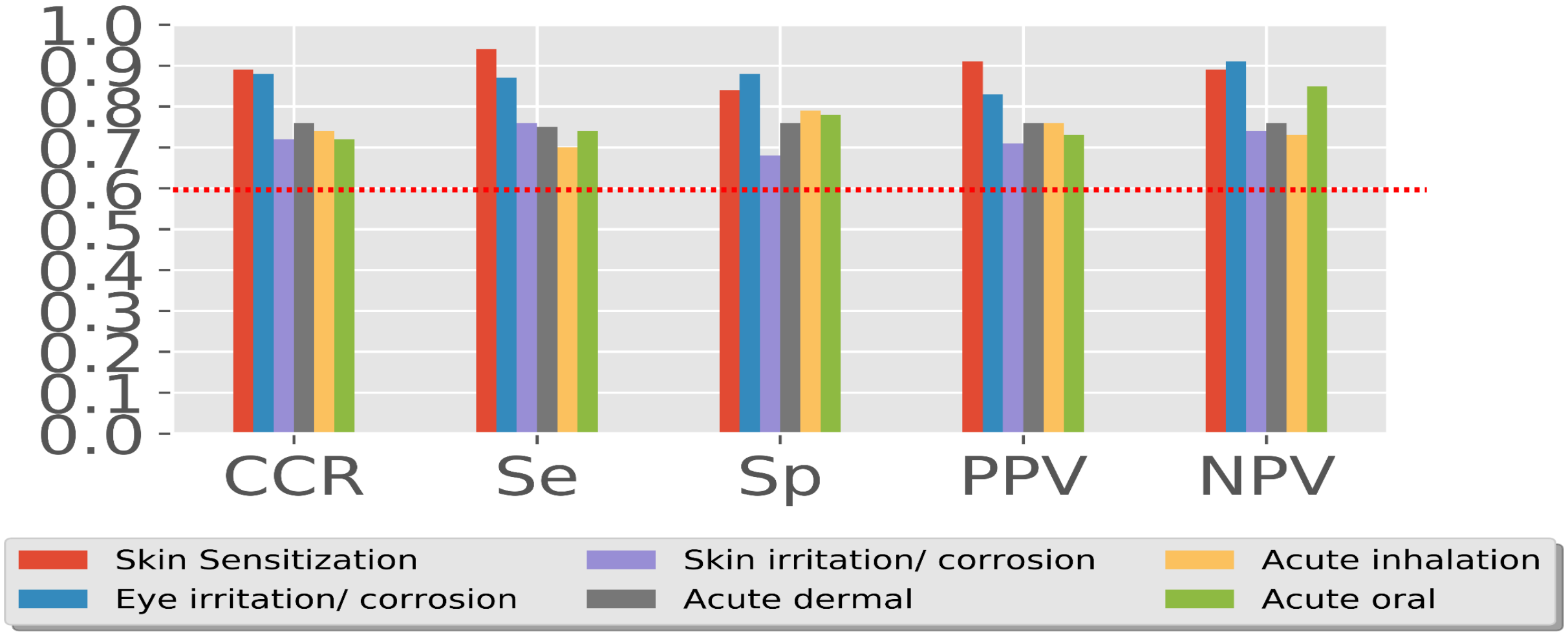




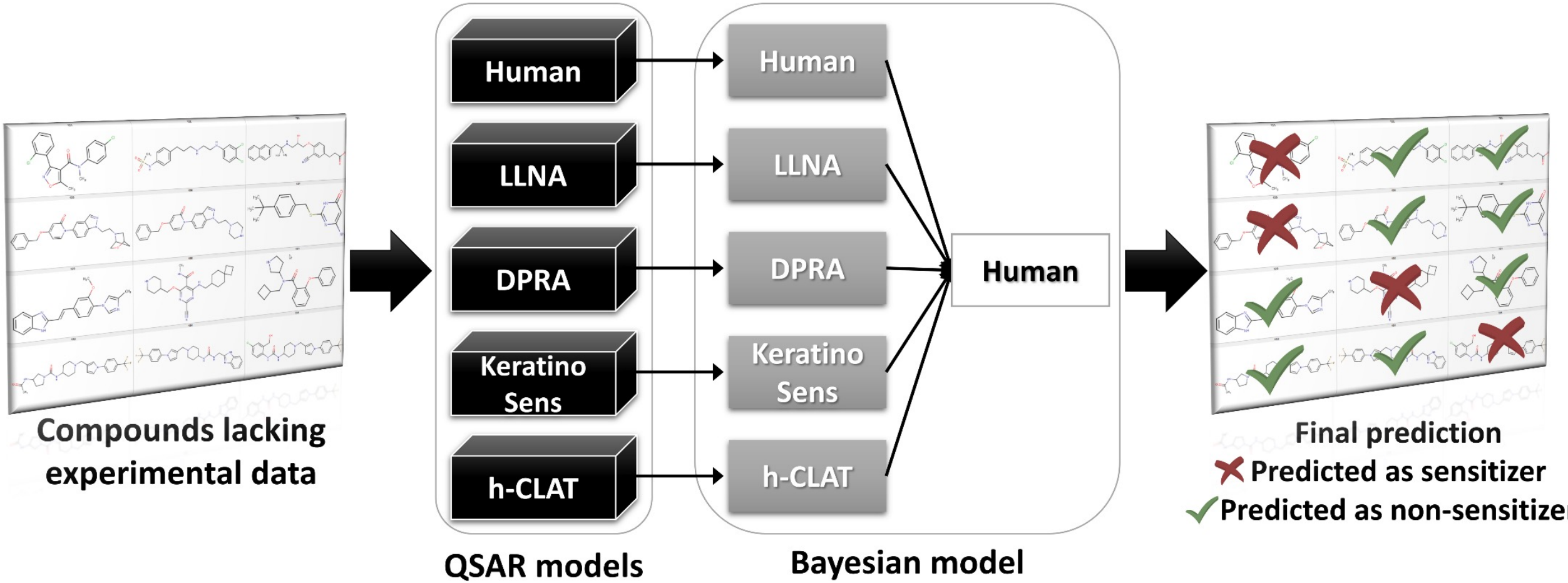
QSAR model development and validation workflow



Statistical characteristics of QSAR models

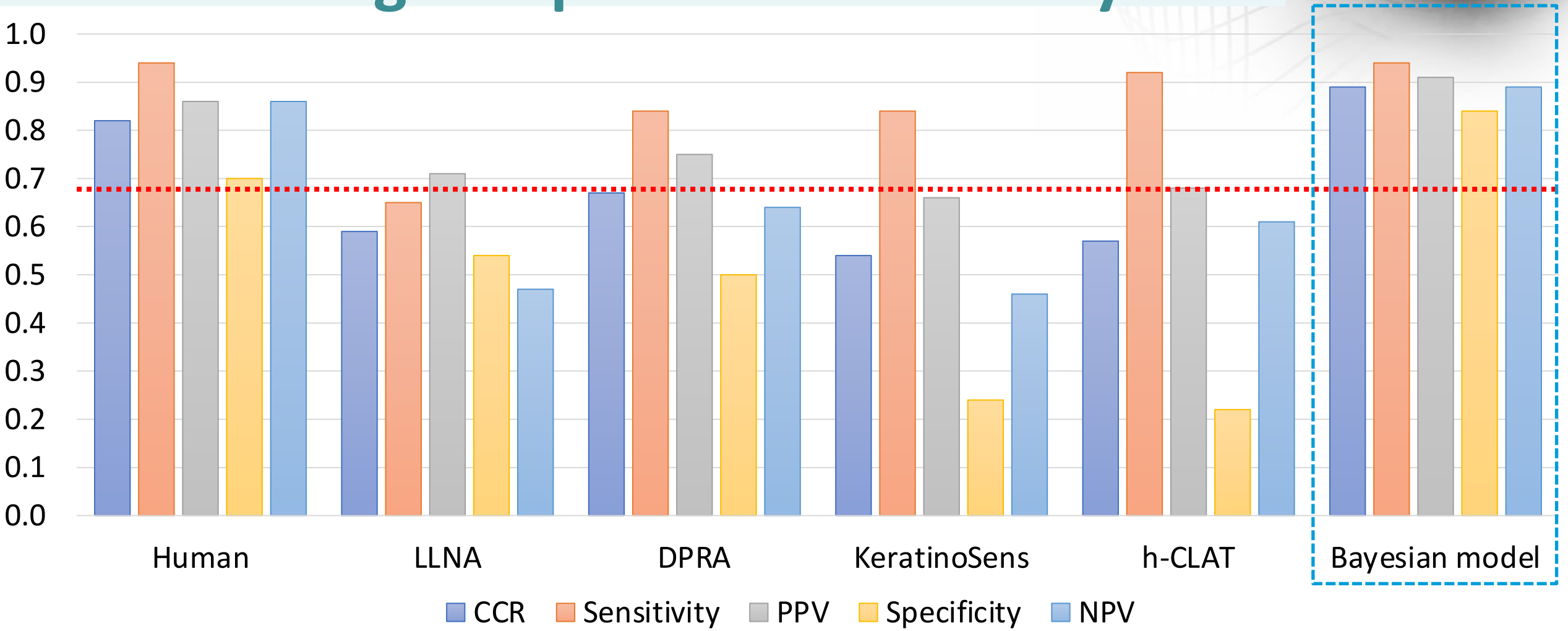


Consensus modeling with Bayesian approach



Alves, V. M. et al. *ACS Sustain. Chem. Eng.* **2018**, *6* (3), 2845–2859.

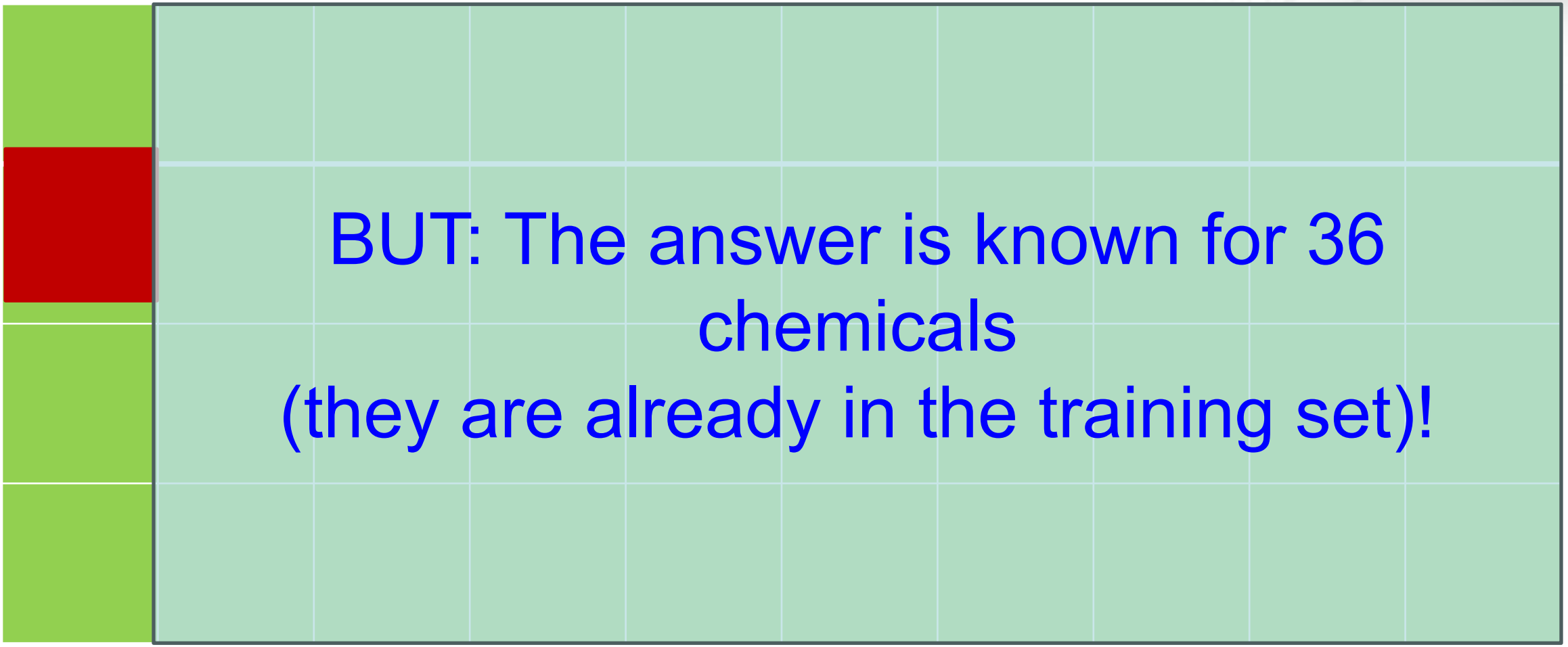
Bayesian model for skin sensitization shows the highest prediction accuracy



Alves, V. M. et al. *ACS Sustain. Chem. Eng.* **2018**, *6* (3), 2845–2859.

External validation accuracy of STopTox prediction (OECD DASS-me-too program)

External set of 40 chemicals; apparent accuracy of 97.5% (39/40)



Real External set of 4 chemicals; accuracy of 75% (3/4)

STopTox paper

ehp Environmental Health Perspectives

HOME ISSUE IN PROGRESS ARCHIVES COLLECTIONS AUTHORS REVIEWERS ABOUT INTRODUCING *JHP*

Open Access

Vol. 130, No. 2 | Research

STopTox: An *in Silico* Alternative to Animal Testing for Acute Systemic and Topical Toxicity

Joyce V.B. Borba, Vinicius M. Alves, Rodolpho C. Braga, Daniel R. Korn, Kirsten Overdahl, Arthur C. Silva, Steven U.S. Hall, Erik Overdahl, Nicole Kleinstreuer, Judy Strickland, David Allen, Carolina Horta Andrade, Eugene N. Muratov, and Alexander Tropsha

Published: 22 February 2022 | CID: 027012 | https://doi.org/10.1289/EHP9341 | **Cited by:** 18

Pilot STopTox Web-App: https://stoptox.mml.unc.edu/

STopTox

a machine learning app to predict the battery of acute toxicity test classically known as "6-pack".

LARGE DATA COLLECTION

STopTox was developed based on the largest publicly available data for the "6-pack" animal assays.

CURATED DATA

The collected data was carefully analyzed and curated to guarantee high quality models.

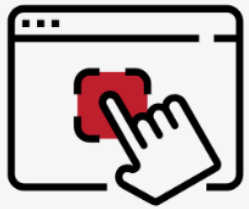
MACHINE LEARNING

QSAR models were developed using powerful machine learning algorithms.

RELIABLE PREDICTIONS

We follow the best practices for model development and validation for regulatory purposes recommended by OECD.

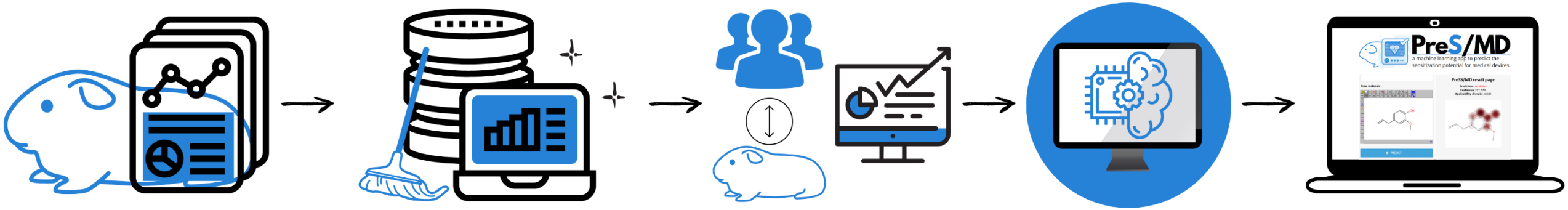
USER FRIENDLY



Models predictivity are assured by high quality curated data and proper external validation.

Commercial version: please visit https://www.predictive-llc.com/

Predictor of Sensitization for Medical Devices (PreS/MD): Study design



Collection of GPMT data from ECHA and literature

Data curation and integration

Cross-species analysis with human data

Development of QSAR models

Development of the PreS/MD web app

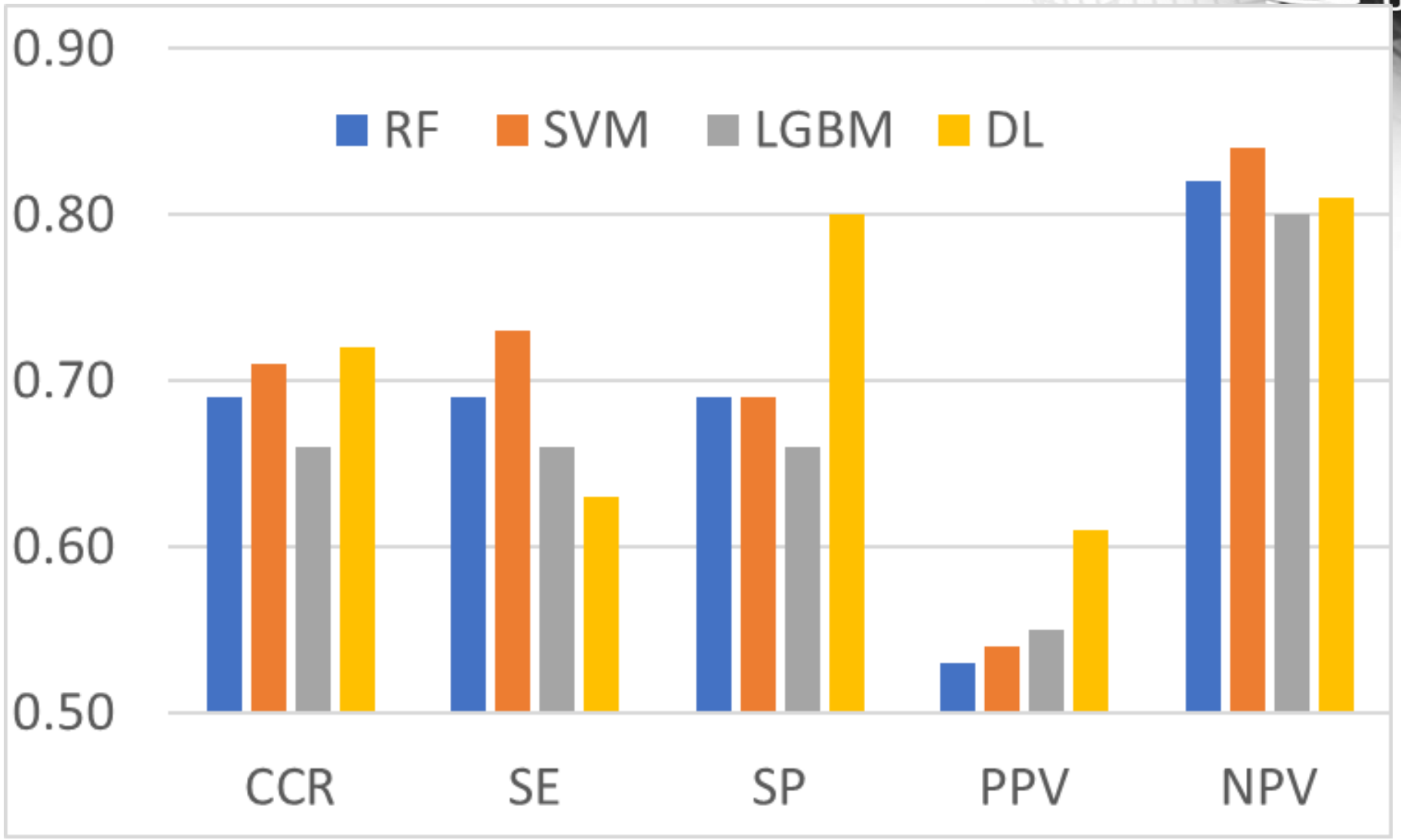
Borba, J. et al. PreS/MD: Predictor of Sensitization Hazard for Chemical Substances Released From Medical Devices, *Toxicological Sciences*, Volume 189, Issue 2, October 2022, Pages 250–259, https://doi.org/10.1093/toxsci/kfac078

Relevance of GPMT assay: comparison with human patch assay data for skin sensitization

	Human		
GPMT	**Sensitizer**	**Non-Sensitizer**	**Total**
Sensitizer	46	14	60
Non-sensitizer	8	41	49
Total	54	55	**109**

Alves et al., ACS Sustainable Chem. Eng. 2018, 6, 3, 2845–2859

Model accuracy



Borba, J. et al. PreS/MD: Predictor of Sensitization Hazard for Chemical Substances Released From Medical Devices, *Toxicological Sciences*, Volume 189, Issue 2, October 2022, Pages 250–259, https://doi.org/10.1093/toxsci/kfac078

Experimental results and predictions for external set compounds

Ingredient	GPMT	PreS/MD
Abietic Acid	Sensitizer	Sensitizer
Ethanol	Sensitizer	Non-sensitizer
Eugenol	Sensitizer	Sensitizer
Geraniol	Non-sensitizer	Non-sensitizer
Methylparaben	Non-sensitizer	Non-sensitizer
Sulfanilic Acid	Sensitizer	Sensitizer
1,2-Dibromo-2,4-dicyanobutane	Non-sensitizer	Sensitizer
2-Methyl-3(2H)-isothiazolone	Sensitizer	Sensitizer
4,5-Dichloro-2-methyl-4-isothiazolin-3-one	Non-sensitizer	Sensitizer

Borba, J. et al. PreS/MD: Predictor of Sensitization Hazard for Chemical Substances Released From Medical Devices, *Toxicological Sciences*, Volume 189, Issue 2, October 2022, Pages 250–259, https://doi.org/10.1093/toxsci/kfac078

Planned enhancements of PreS/md → PredTox/MD

Computational Software covering skin sensitization, irritation, and cytotoxicity endpoints.

Software	Endpoint			Computational Approach				Version	
	Skin Sensitization	Irritation	Cytotoxicity	QSAR	Structural Alerts	Read-Across	Bayesian	Free (Web)	Proprietary
PredTox/MD	✓▯	✓▯	✓▯	✓	X	✓▯	✓	✓	✓
TIMES-SS	✓▯	✓	X	✓▯	✓▯	X▯	X▯	X▯	✓▯
TOPKAT	✓▯	✓▯	X	✓▯	X▯	X▯	X▯	✓▯	✓▯
ACD/ Percepta	✓	✓▯	X	✓▯	X▯	X▯	X▯	X▯	✓▯
CASE Ultra	✓▯	✓▯	X	✓▯	✓▯	X▯	X▯	X▯	✓▯
DEREK NEXUS	✓▯	✓▯	X	X▯	✓▯	X▯	X▯	X▯	✓▯
REACHAcross (RASAR)	✓▯	✓▯	X	✓▯	X▯	✓▯	X▯	X▯	✓▯
Danish QSAR database	✓▯	✓▯	X	✓	✓	X	X	✓	X
ToxTree	✓▯	✓▯	X	X	✓	X	X	✓	X
OECD QSAR Toolbox	✓▯	✓▯	✓▯	✓	X	✓	X	✓	X
VEGA	✓▯	X	X	X	X	✓	X	✓	X
Pred-Skin	✓▯	X	X	✓	X	X	✓	✓	X

“Best Published Paper Advancing the Science of Risk Assessment Award” at the 2023 Society of Toxicology (SOT) annual meeting.

MML UNC.EDU

OXFORD

SOT | Society Toxicolo

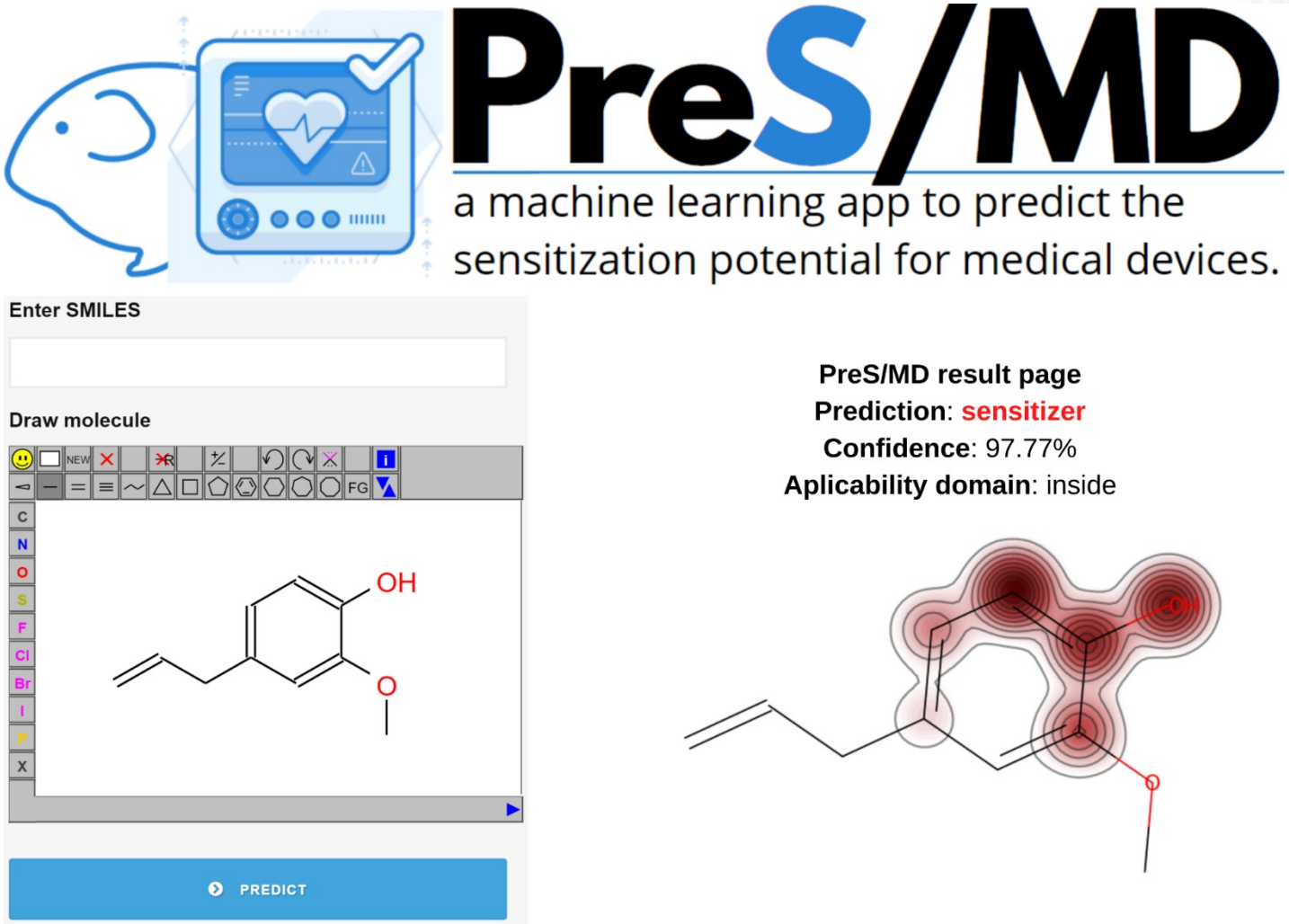
academic.oup.com/tox

PreS/MD: Predictor of

Chemical Substances

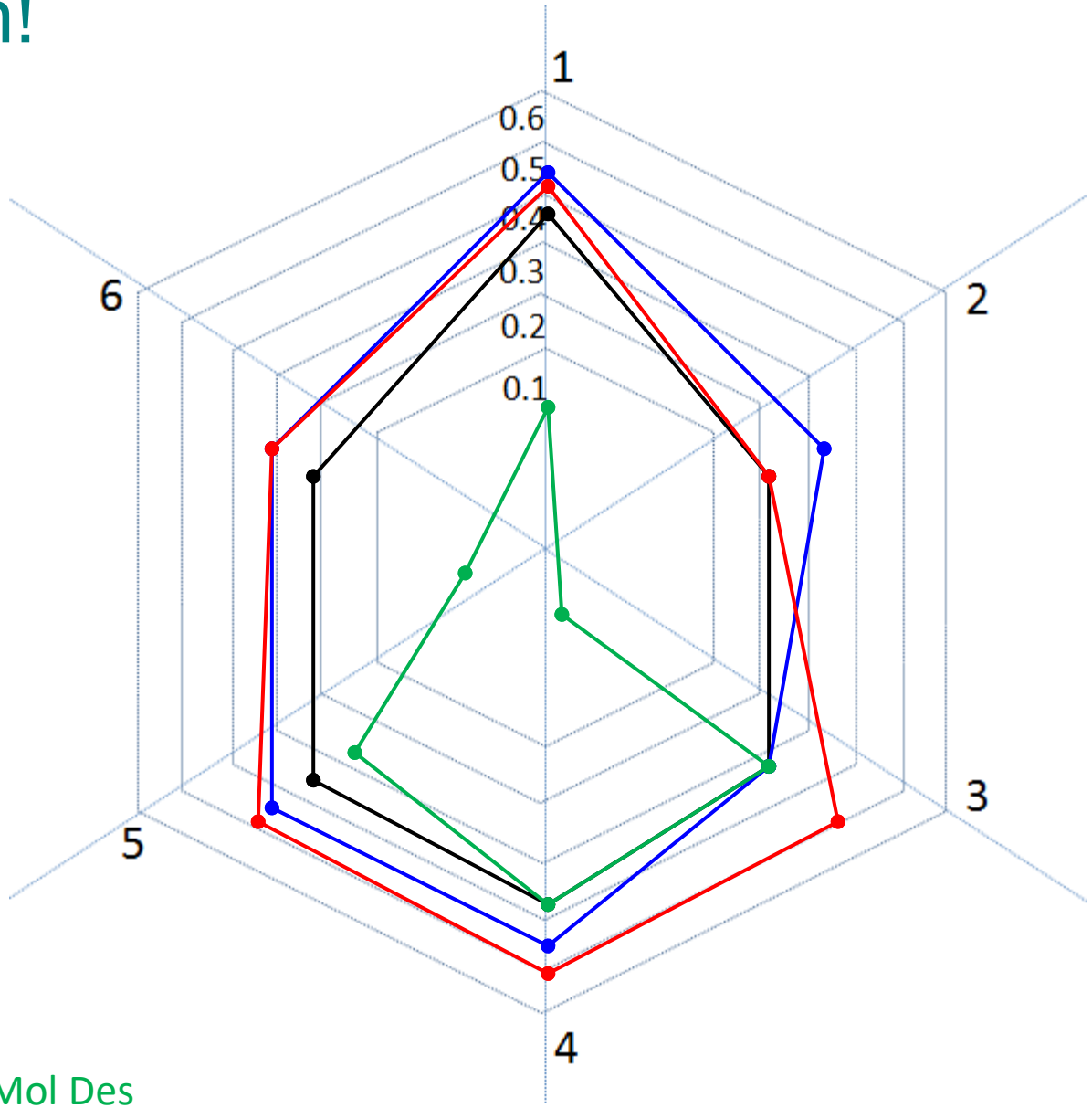
Vinicius M. Alves,* Joyce V. B.
Nicole Kleinstreuer ,‡ Kevin
and Eugene N. Muratov *,1

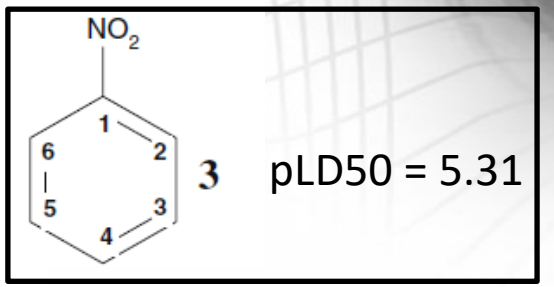
Pilot PRES/MD WEB-app: https://pressmd.mml.unc.edu/



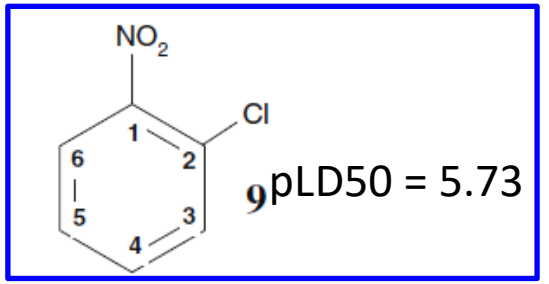
Commercial version: please visit https://www.predictive-llc.com/

Model interpretation: Chemical fragments don't act in isolation!

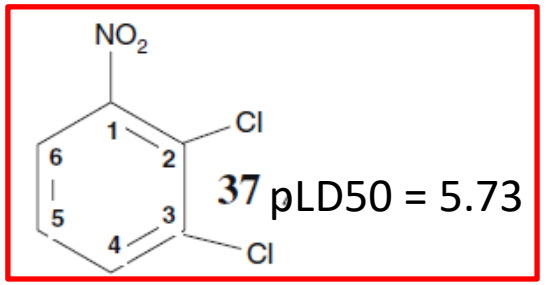




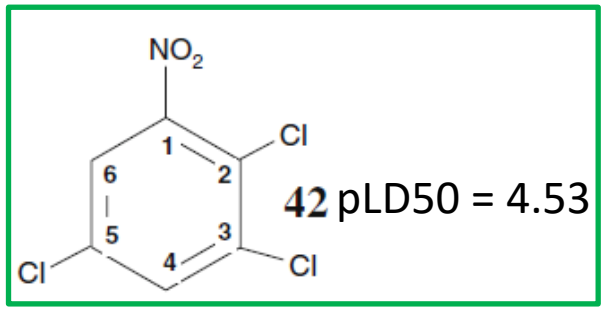
3 pLD50 = 5.31



9 pLD50 = 5.73



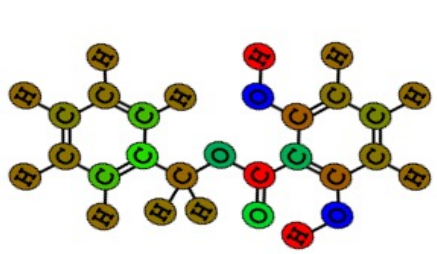
37 pLD50 = 5.73



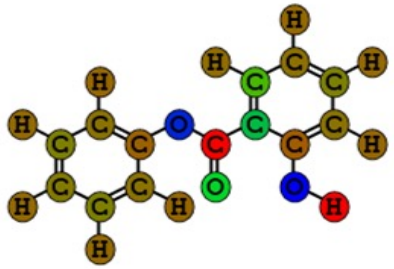
42 pLD50 = 4.53

Modified from
Kuz'min et al.
J Comp Aided Mol Des
2008, 22, 747–759

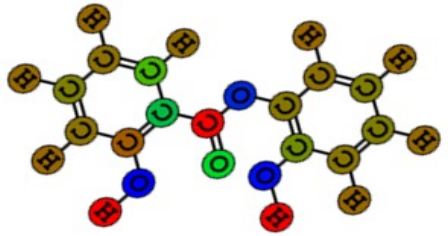
Model-driven Structure optimization



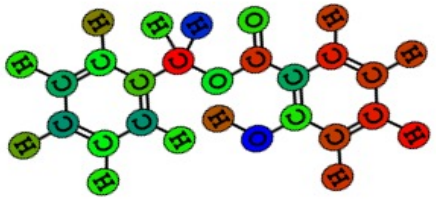
Benzyl 2,6-dihydroxybenzoate
QSAR: Non-sensitizer



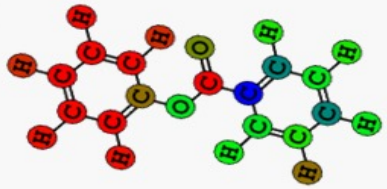
Phenyl salicylate
QSAR: Non-sensitizer



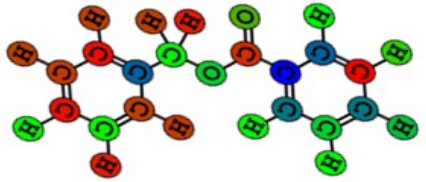
2-hydroxyphenyl 2-hydroxybenzoate
QSAR: Non-sensitizer



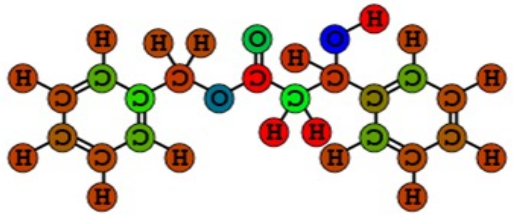
Benzyl salicylate
Human: Non-sensitizer
LLNA: Sensitizer
QSAR: Non-sensitizer



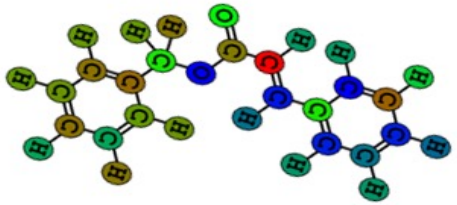
Phenyl benzoate
Human: Sensitizer
LLNA: Sensitizer
QSAR: Sensitizer



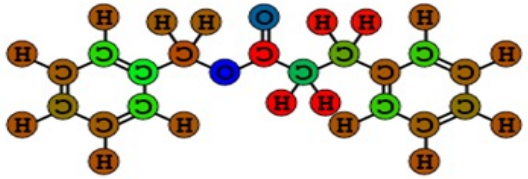
Benzyl benzoate
Human: Non-sensitizer
LLNA: Sensitizer
QSAR: Non-sensitizer



Benzyl 3-hydroxy-3-phenylpropanoate
QSAR: Non-sensitizer

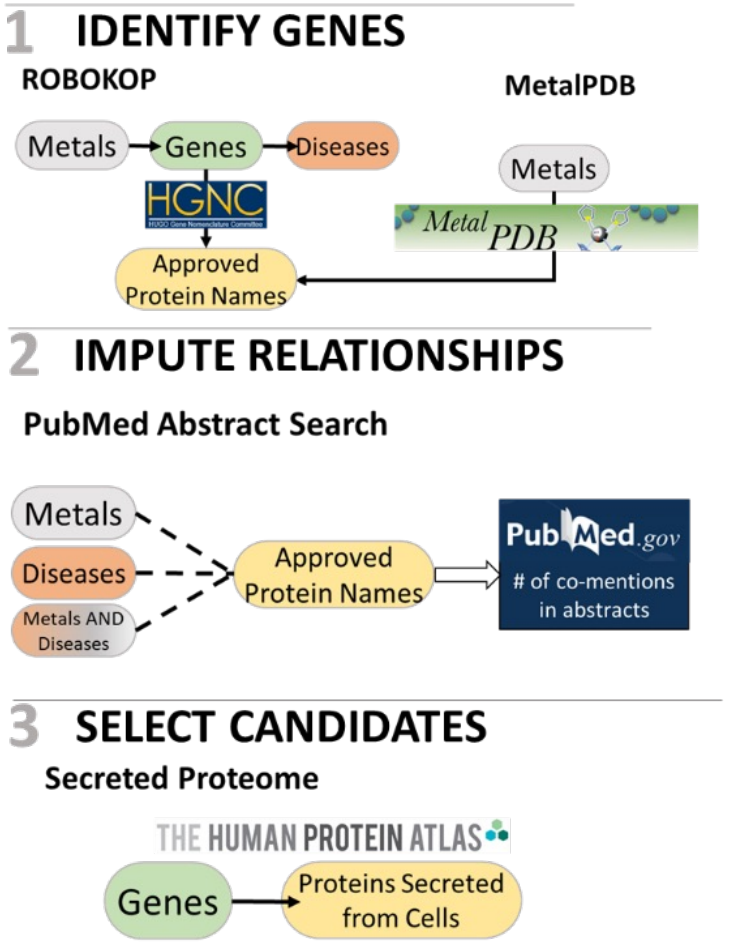


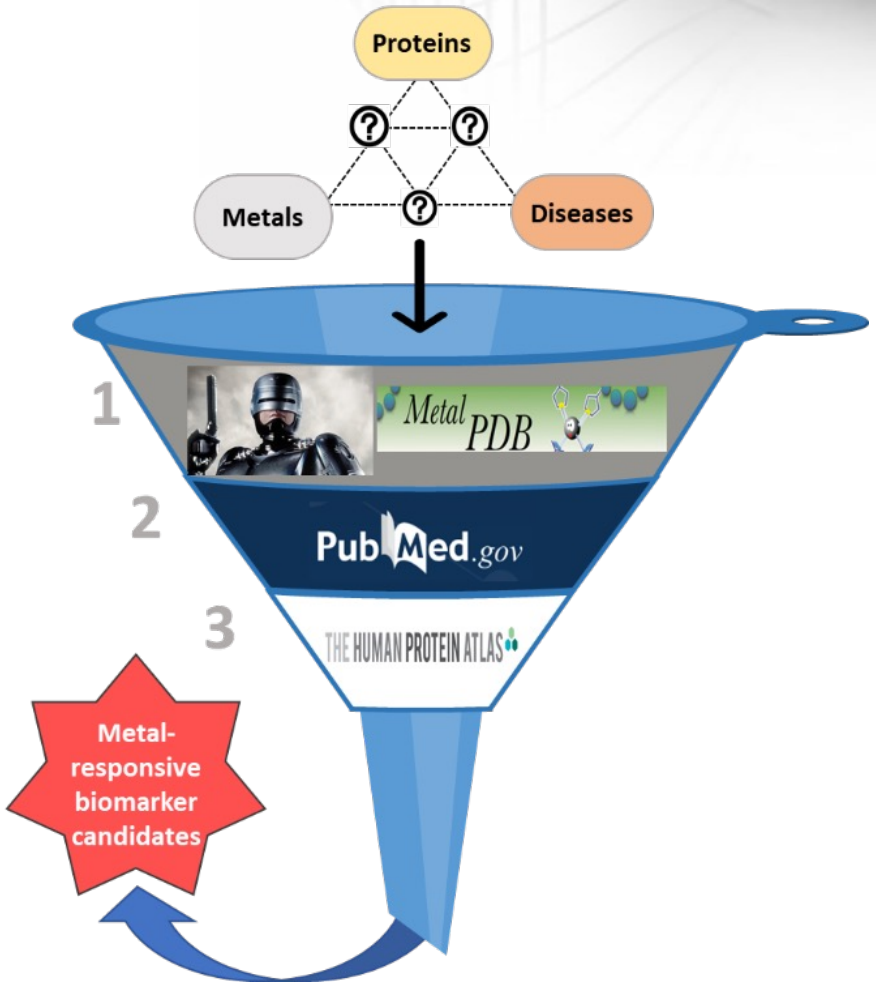
Benzyl cinnamate
Human: Non-sensitizer
LLNA: Sensitizer
QSAR: Non-sensitizer



Benzyl 3-phenylpropanoate
QSAR: Non-sensitizer

Computational metallomics: Elucidation of adverse outcome pathways of metal toxicity using knowledge graph mining





Beasley, JMB, et al. Integrated approach to elucidate metal-implant related adverse outcome pathways. Regul Toxicol Pharmacol . **2022** Dec;136:105277. doi: 10.1016/j.yrtph.2022.105277.

Metal Leaching from Implants Causes Toxicity

- Mechanical forces and corrosion release metal particles
- Adverse local tissue reaction (ALTR) -Severe inflammation, tissue necrosis, and pseudotumor formation at site of implant [1]
- Neuropathic pain disorders, hypersensitization/allergic responses, and cardiomyopathy among other reported adverse events [1,2]

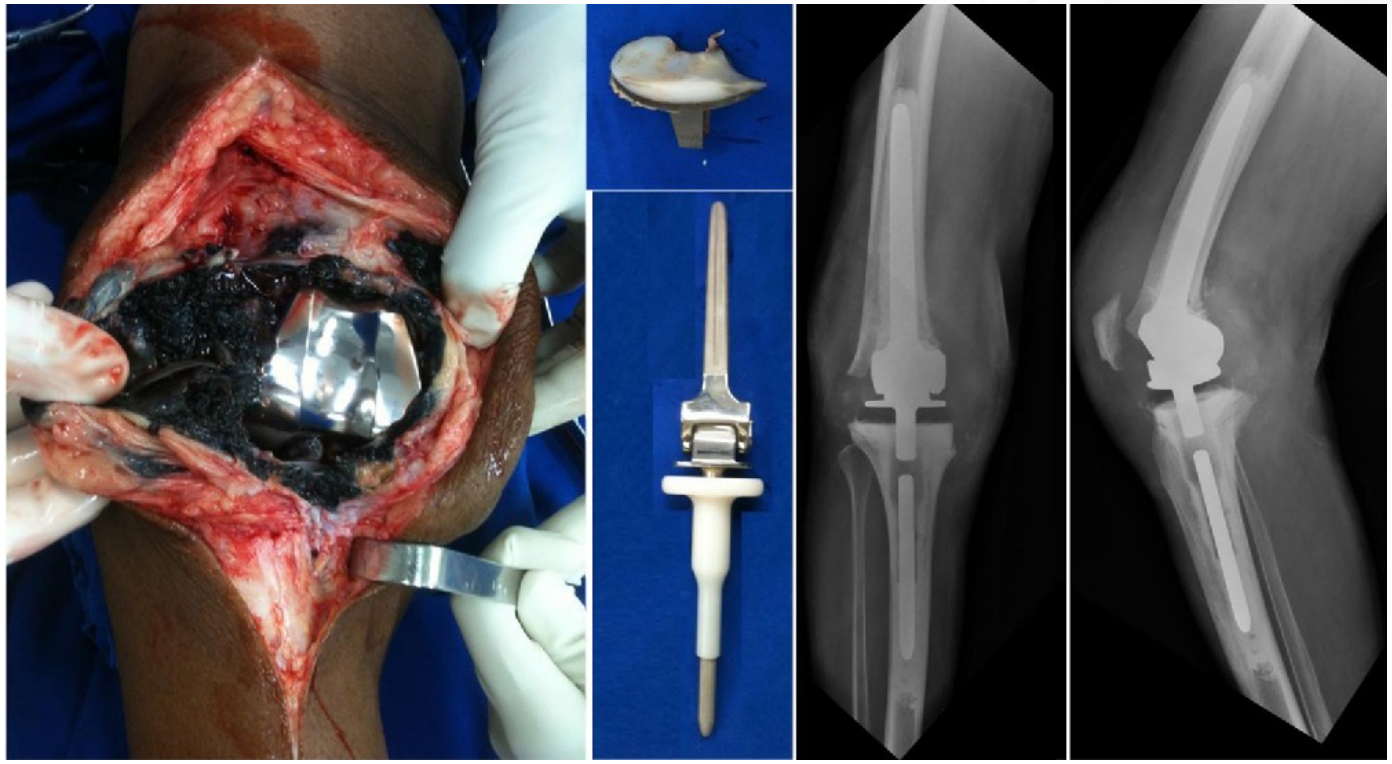
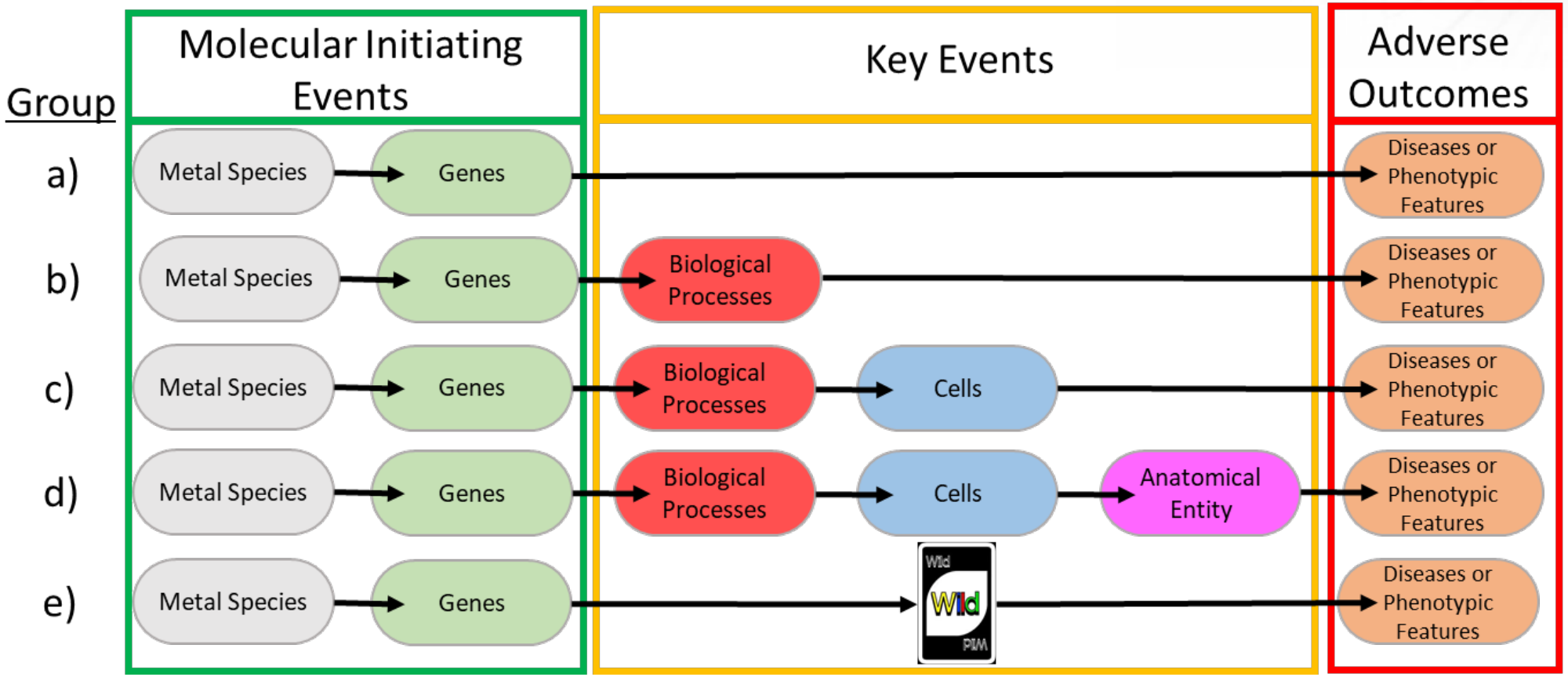


Photo: Anastasia Floyd. Case Study Early Failure of a Modular Hip Implant. *Univ. of Calif. Berkeley*. 2015.

[1] U.S. Food and Drug Administration Center for Devices and Radiological Health. *Biological Responses to Metal Implants*. www.fda.gov/media/131150/download (2019).

[2] Eliaz, N. Corrosion of Metallic Biomaterials: A Review. Materials (Basel). 12, (2019).

Adverse Outcome Pathway (AOP) Models can help explain Toxicity and identify disease biomarkers



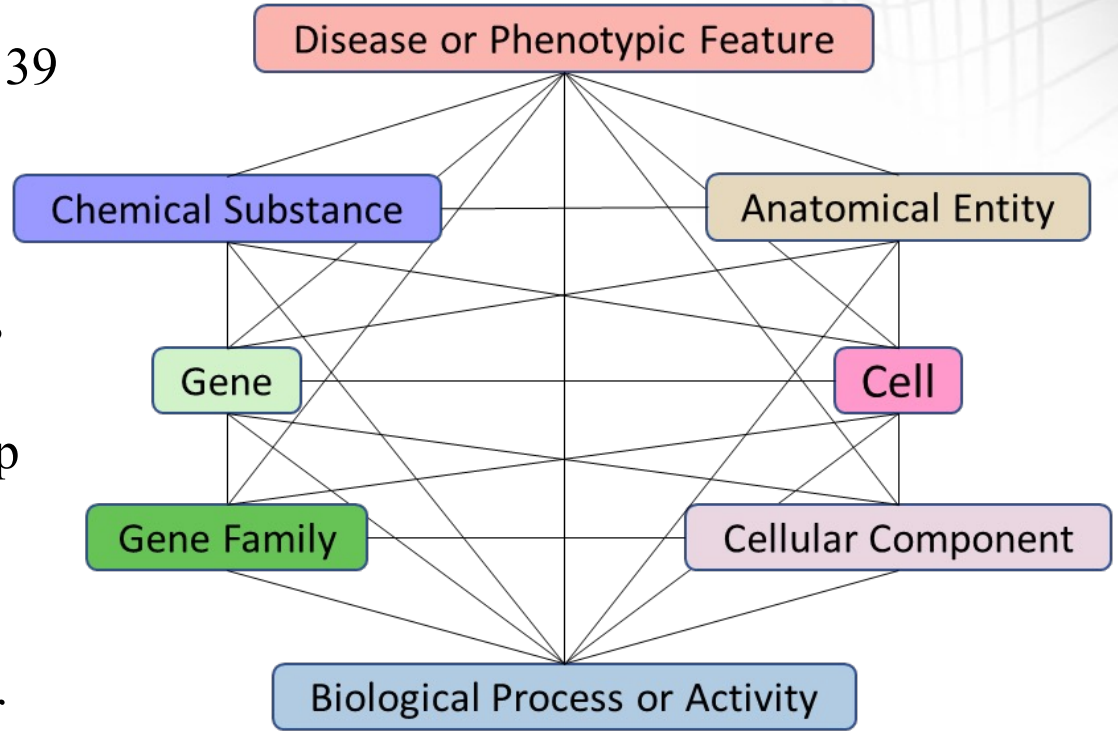
ROBOKOP

Reasoning Over Biomedical Objects linked in Knowledge Oriented Pathways

- ~9M nodes and ~130M edges drawn from 39 biomedical data sources and bio/chemical-ontologies

- Nodes represent biomedical entities; edges provide predicates, publications, and quantitative data that explain the relationship between connected nodes

- Leverages Biolink Model (https://biolink.github.io/biolink-model/) for semantic harmonization and standardization of node and edge types



Funders

U24ES035214

OT3TR002020

OT2TR003430

UL1TR002489

UL1TR002489-03S4

UNC ESHELMAN SCHOOL OF PHARMACY

Morton et al., ROBOKOP: An Abstraction Layer and User Interface for Knowledge Graphs to Support Question Answering. Bioinformatics. **2019** 35(24):5382-5384

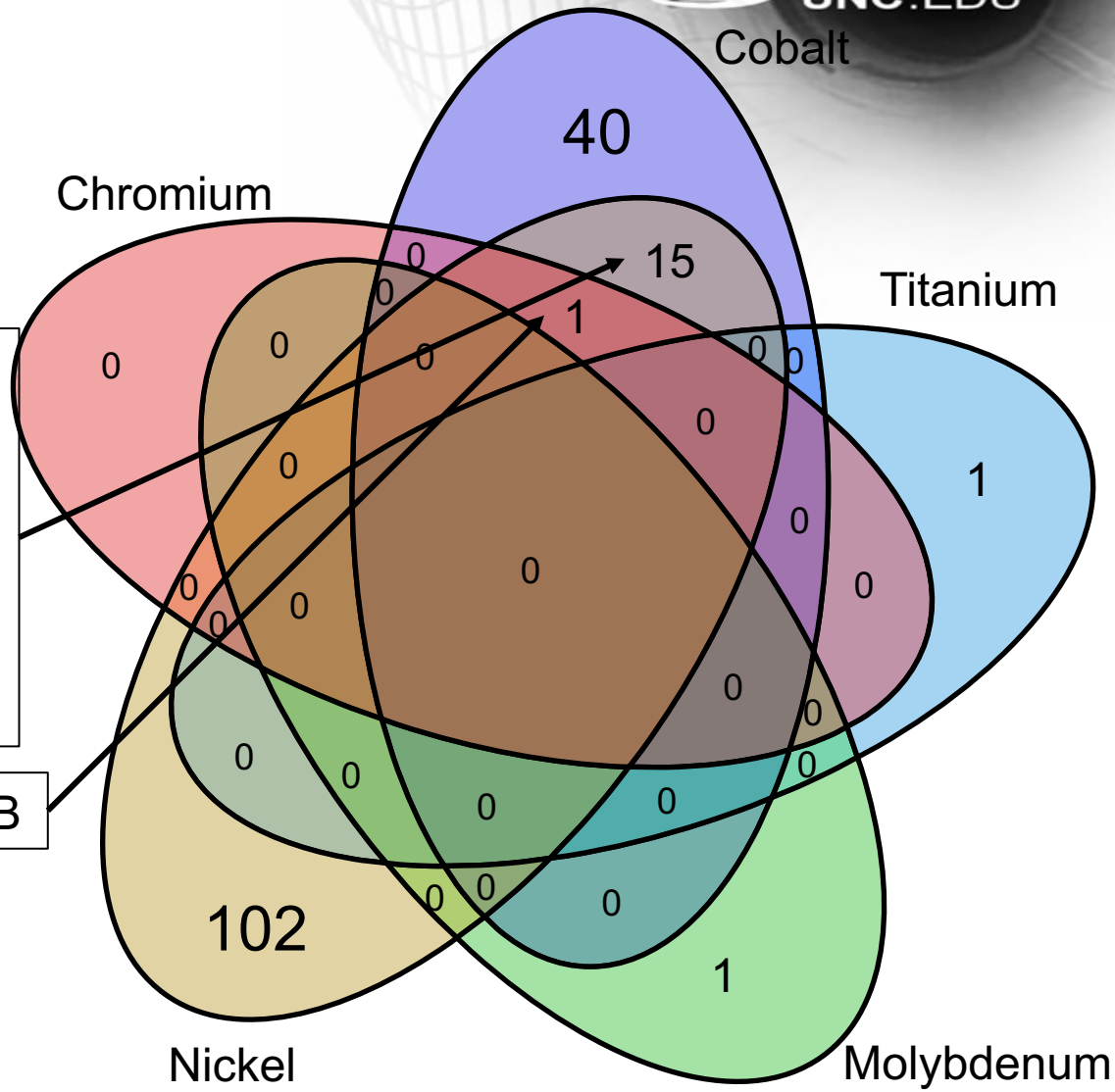
Chris Bizon, Steven Cox, James Balhoff, Yaphet Kebede, Patrick Wang, Kenneth Morton, Karamarie Fecho, and Alexander Tropsha *Journal of Chemical Information and Modeling* **2019** *59* (12), 4968-4973

Structure-based Knowledge of Metal-protein Interactions from MetalPDB

- Known metal-protein structures mined from MetalPDB*
- 56 cobalt, 1 chromium, 1 molybdenum, 118 nickel, and 1 titanium PDB entries

ARG1, B2M, BF, BFD, BRCA1, CA2, CDABP0092, CFB, HDCMA22P, JMJD3, JMJD5, KDM6B, KDM8, KIAA0346, RNF53

POLB



*Putignano, V., Rosato, A., Banci, L. & Andreini, C. MetalPDB in 2018: A database of metal sites in biological macromolecular structures. *Nucleic Acids Res.* **46**, D459–D464 (2018).

ROBOKOP was used to mine metal-gene interactions and generate hypothetical Adverse Outcome Pathways

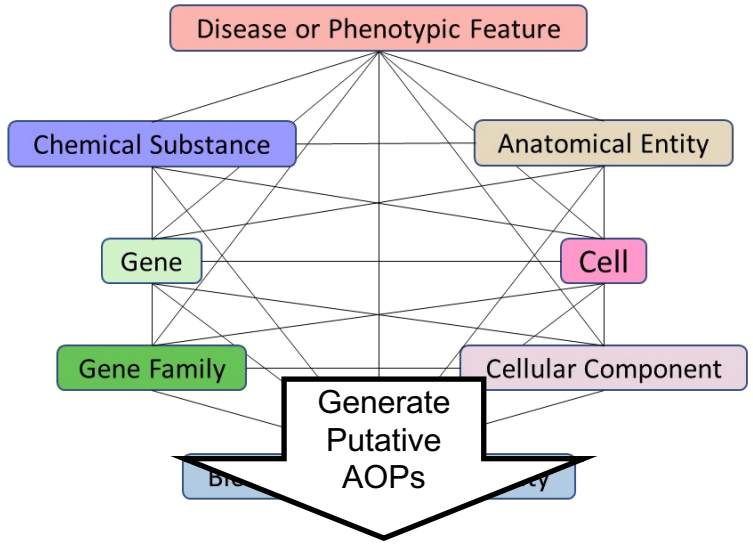
*Taken from [1]

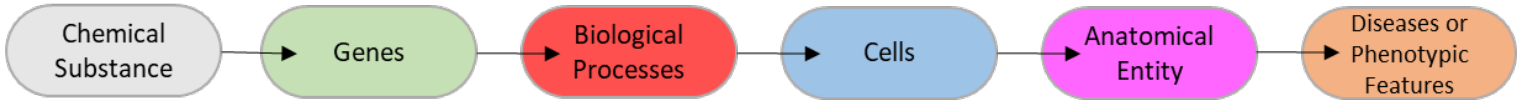
Metal Species	"chemical entity" ROBOKOP Terms	"disease or phenotypic feature" ROBOKOP Terms
Titanium	titanium atom	Innate Immune Response, Nlrp3 Inflammasome Complex, Periprosthetic Osteolysis
Cobalt	cobalt atom, , cobalt cation, cobalt(1+), cobalt(2+), cobalt(3+)	Innate Immune Response, Nlrp3 Inflammasome Complex, Periprosthetic Osteolysis, cardiomyopathy, heart failure , dyspnea, palpitations, chest tightness, Neurotoxicity Syndromes , Demyelinating Polyneuropathy, Cognitive Impairment, Tinnitus, Fibromyalgia, Chronic Fatigue, Hypothyroidism, Prostate Cancer, Melanoma (Disease)
Nickel	nickel atom, nickel ion, nickel cation, nickel(1+), nickel(2+), nickel(3+)	Innate Immune Response, Nlrp3 Inflammasome Complex, Periprosthetic Osteolysis, Prostate Cancer, Metal Allergy, Allergic Contact Dermatitis, Localized Skin Lesion, Squamous Cell Carcinoma, Anaplastic Large Cell Lymphoma
Chromium	chromium atom,chromium ion, chromium cation, chromium(2+), chromium(3+), chromium(4+), chromium(5+), chromium(6+)	Melanoma (Disease), Anaplastic Large Cell Lymphoma
Molybdenum	molybdenum atom, molybdenum cation, molybdenum(4+), molybdenum(5+), molybdenum(6+)	Anaplastic Large Cell Lymphoma

estions
OKOP
]

Robokop

Reasoning **O**ver **B**iomedical **O**bjects linked in **K**nowledge **O**riented **P**athways

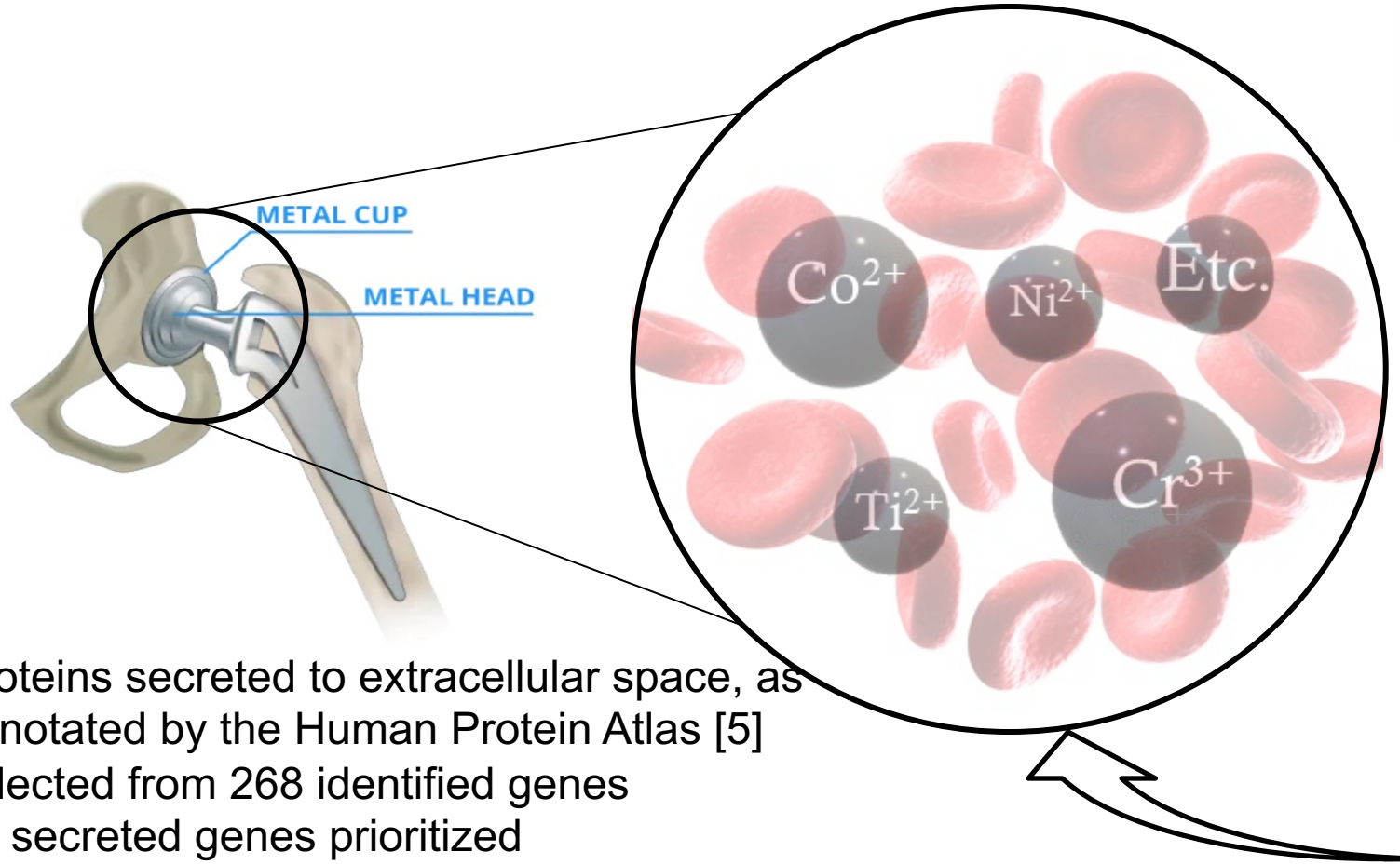


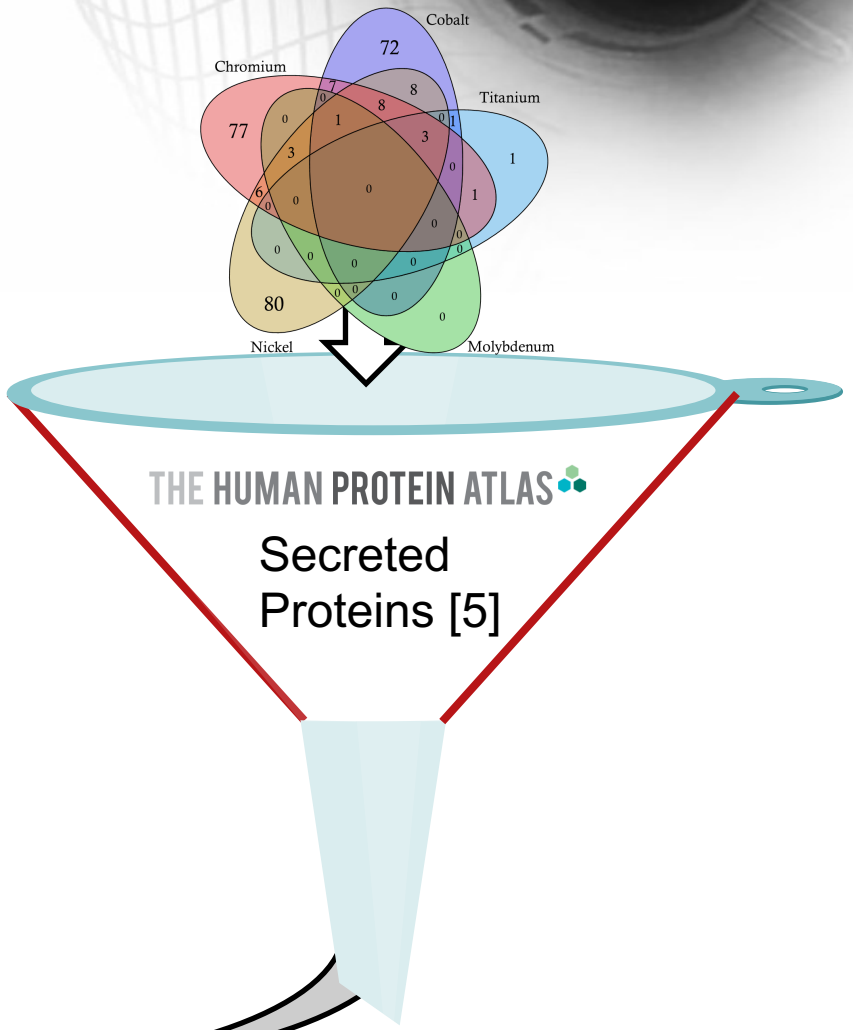


[1] U.S. Food and Drug Administration Center for Devices and Radiological Health. *Biological Responses to Metal Implants*. www.fda.gov/media/131150/download (2019).

[4] Morton, K. *et al.* ROBOKOP: an abstraction layer and user interface for knowledge graphs to support question answering. *Bioinformatics* **35**, 5382–5384 (2019).

Hypothetical AOPs prioritized if they contain a secreted protein

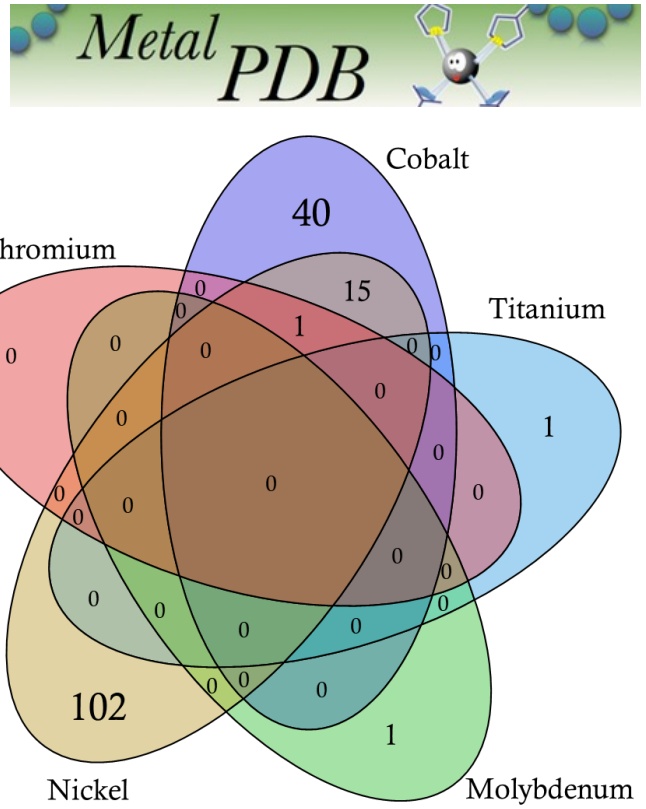


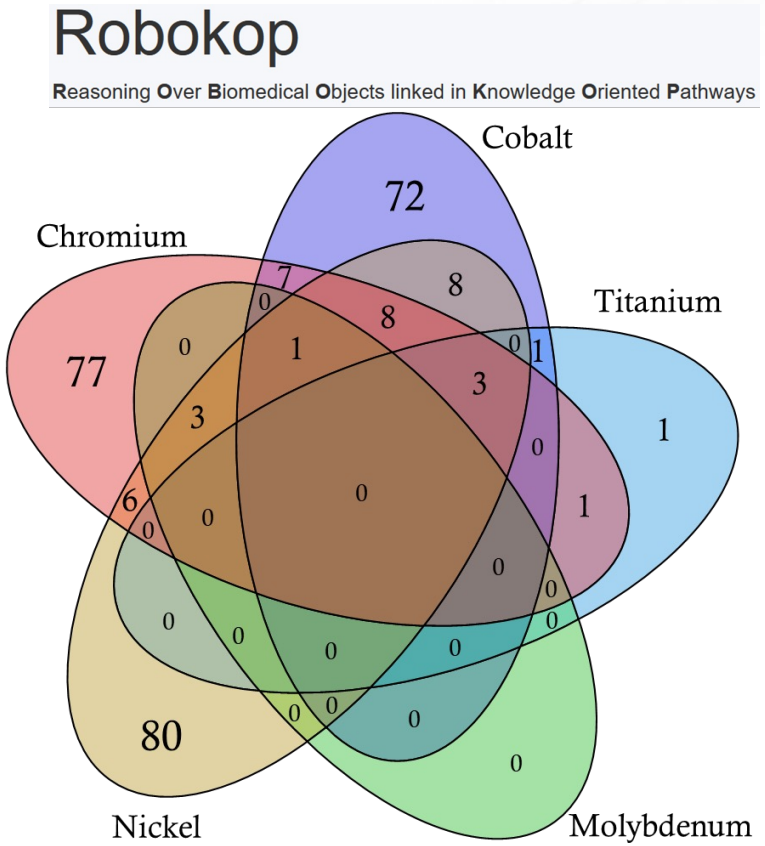


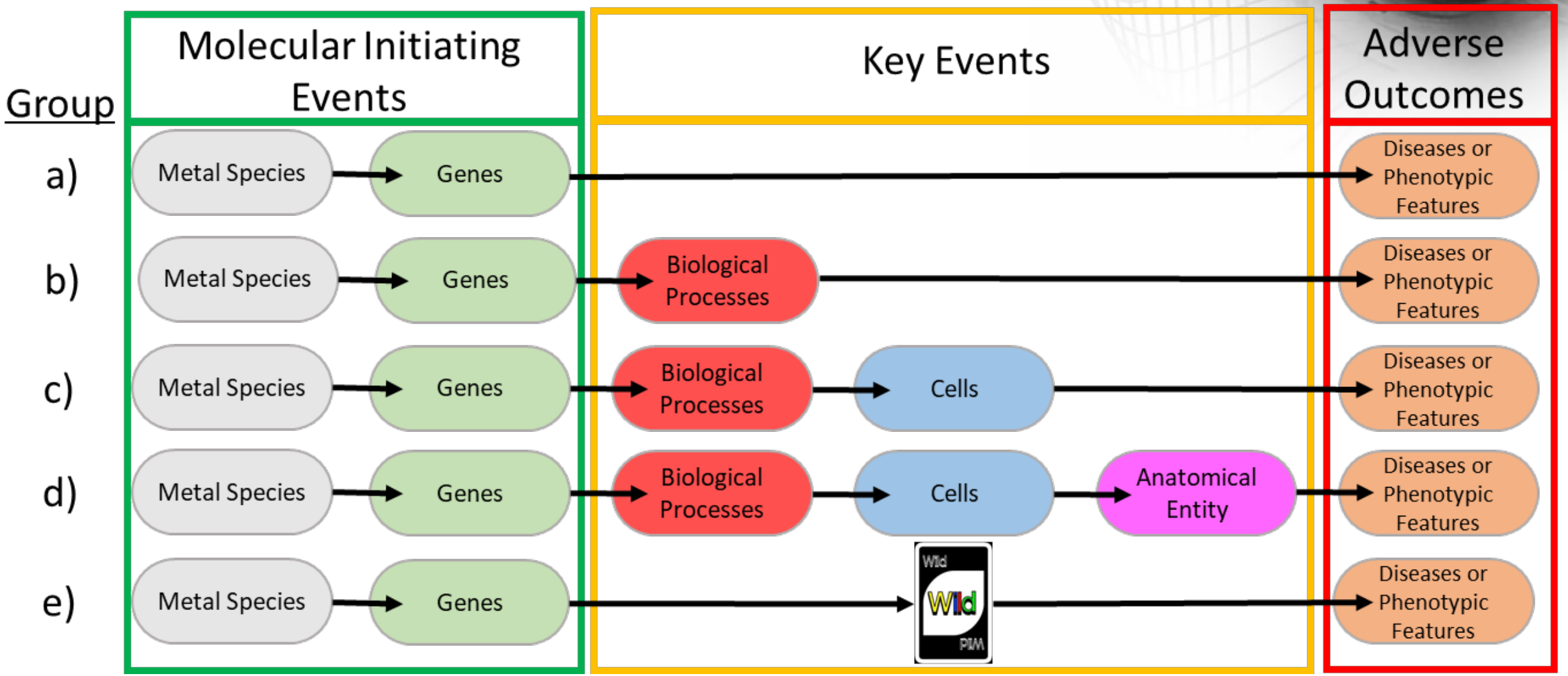
- Proteins secreted to extracellular space, as annotated by the Human Protein Atlas [5] selected from 268 identified genes
- 72 secreted genes prioritized

[5] M, U. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, (2015).

177 metal-interacting proteins identified by metalpdb, 268 identified by ROBOKOP

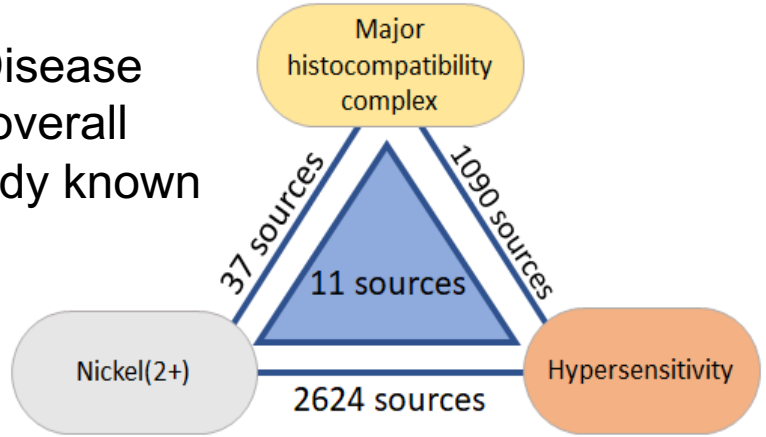


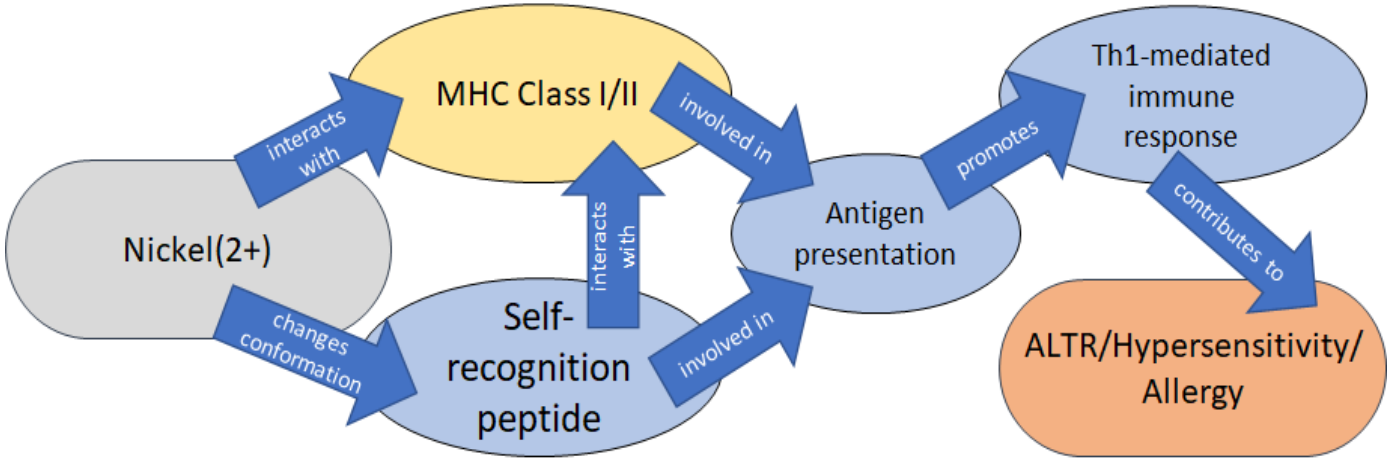




Edwards, S. W., *et al.* (2016). Adverse Outcome Pathways-Organizing Toxicological Information to Improve Decision Making. *The Journal of pharmacology and experimental therapeutics*, *356*(1), 170–181. https://doi.org/10.1124/jpet.115.228239
U.S. National Toxicology Program. Adverse Outcome Pathways. https://ntp.niehs.nih.gov/whatwestudy/niceatm/comptox/ct-aop/aop.html

(i) Metal-Protein-Disease connections and overall hypothesis was already known





reaction

Fisher, A. A. Allergic dermatitis persumably due to metallic foreign bodies containing nickel or cobalt. *Cutis* **19**, (1977).

Cramers, M. & Lucht, U. Metal Sensitivity in Patients Treated for Tibial Fractures with Plates of Stainless Steel. *http://dx.doi.org/10.3109/17453677708988763* **48**, 245–249 (2009).

Basketter, D. A., Briatico-Vangosa, G., Kaestner, W., Lally, C. & Bontinck, W. J. Nickel, cobalt and chromium in consumer products: a role in allergic contact dermatitis? *Contact Dermatitis* **28**, 15–25 (1993).

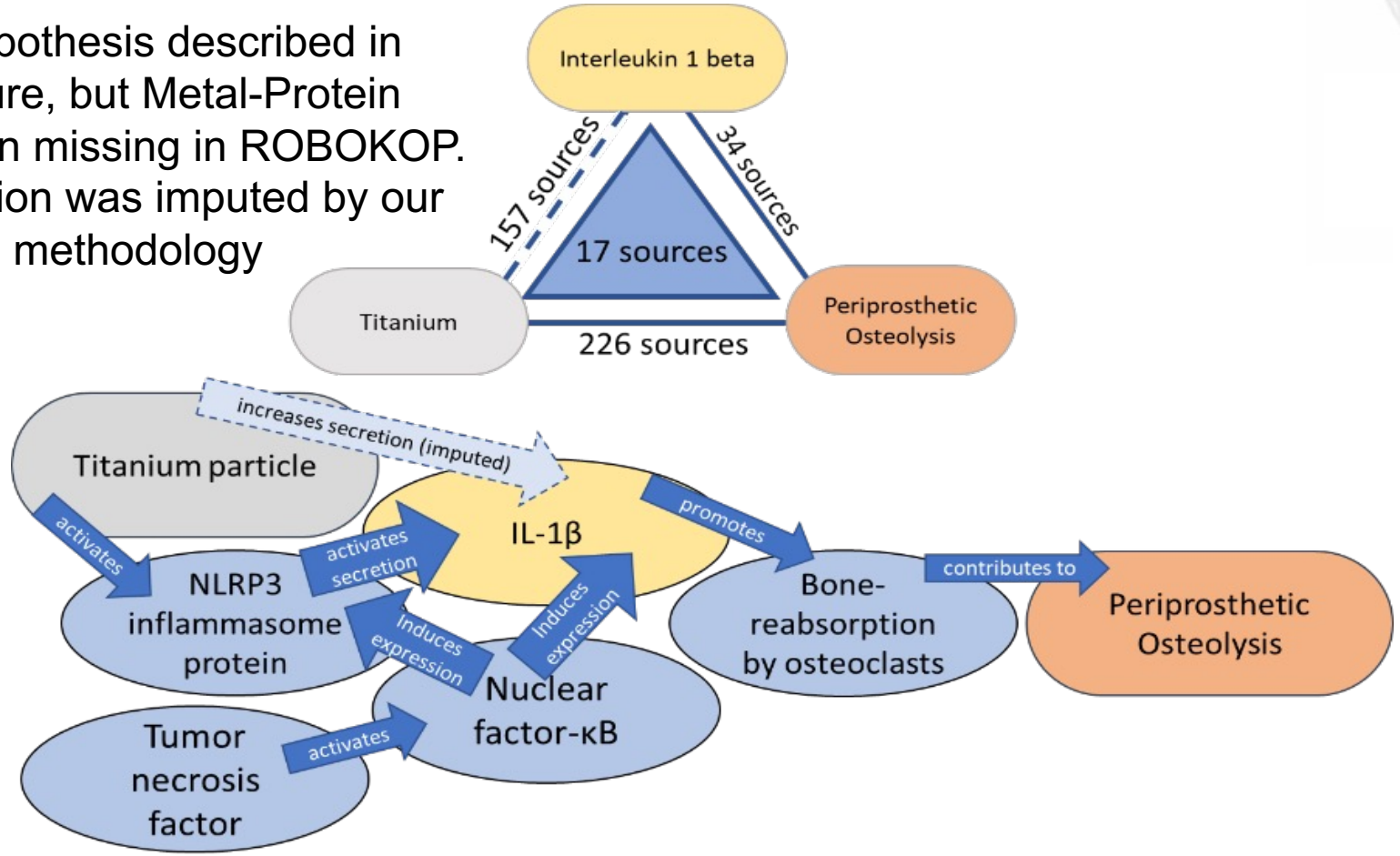
Hallab, N. J., Caicedo, M., Finnegan, A. & Jacobs, J. J. Th1 type lymphocyte reactivity to metals in patients with total hip arthroplasty. *J. Orthop. Surg. Res.* **3**, 6 (2008).

Wang, Y. & Dai, S. Structural basis of metal hypersensitivity. *Immunol. Res.* **55**, 83–90 (2013).

Driller, R. *et al.* Metal-triggered conformational reorientation of a self-peptide bound to a disease-associated HLA-B*27 subtype. *J. Biol. Chem.* **294**, 13269–13279 (2019).

Kilb, B. K. J. *et al.* Frank Stinchfield Award: Identification of the At-risk Genotype for Development of Pseudotumors Around Metal-on-metal THAs. in *Clinical Orthopaedics and Related Research* vol. 476 230–241 (Lippincott Williams and Wilkins, 2018).

(ii) Hypothesis described in literature, but Metal-Protein connection missing in ROBOKOP. Connection was imputed by our methodology



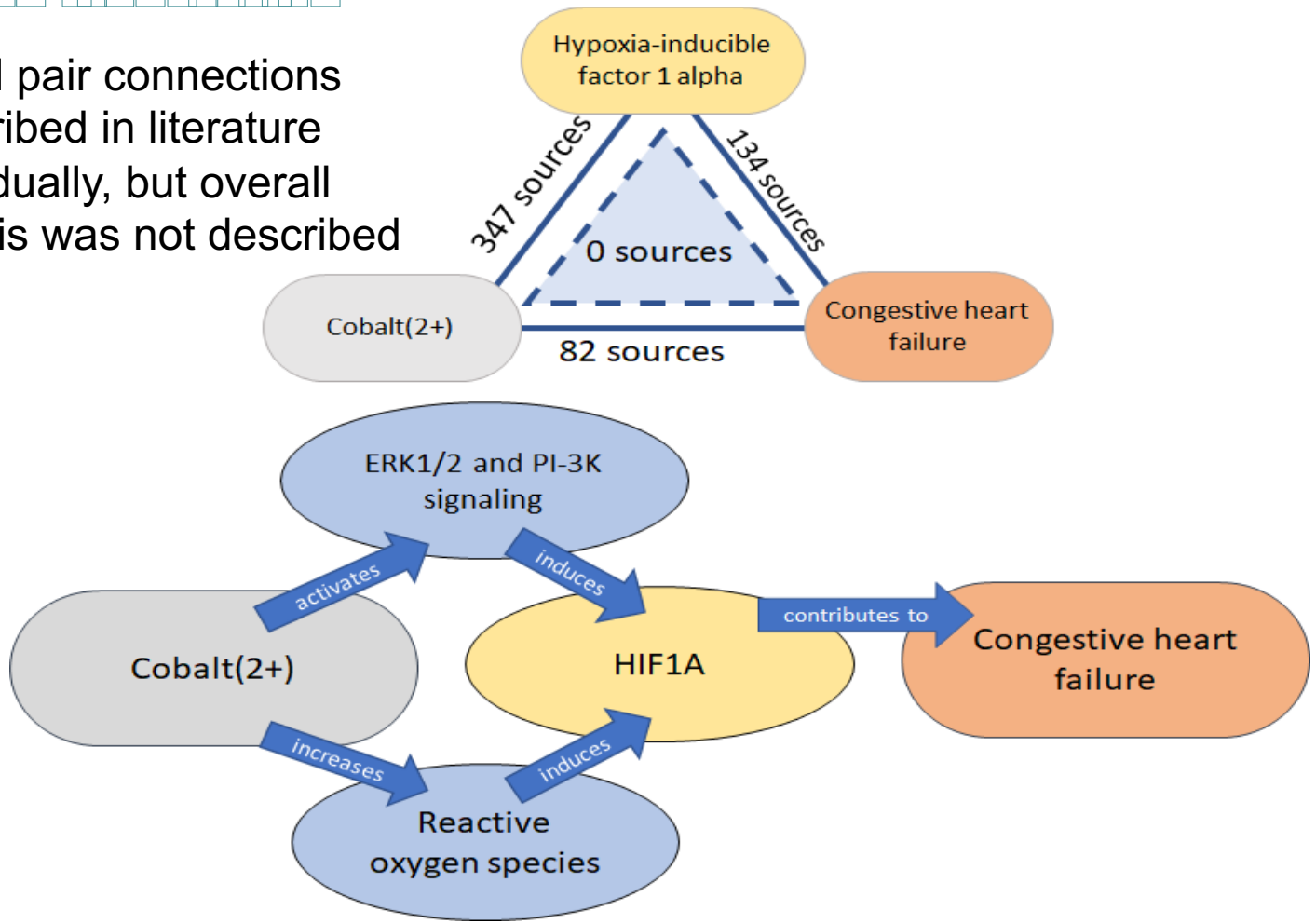
E, J. *et al.* Tumor necrosis factor primes and metal particles activate the NLRP3 inflammasome in human primary macrophages. *Acta Biomater.* **108**, 347–357 (2020).

Pierre, C. A. S. *et al.* Periprosthetic Osteolysis: Characterizing the innate immune response to titanium wear-particles. *J. Orthop. Res.* **28**, 1418 (2010).

FS, S., S, H. & SL, C. Mechanism of NLRP3 inflammasome activation. *Ann. N. Y. Acad. Sci.* **1319**, 82–95 (2014).

Y, H., H, H. & G, N. Mechanism and Regulation of NLRP3 Inflammasome Activation. *Trends Biochem. Sci.* **41**, 1012–1021 (2016).

(iii) All pair connections described in literature individually, but overall hypothesis was not described



A, T. *et al.* Cobalt induces hypoxia-inducible factor-1alpha (HIF-1alpha) in HeLa cells by an iron-independent, but ROS-, PI-3K- and MAPK-dependent mechanism. *Free Radic. Res.* **40**, 847–856 (2006).

L, S. *et al.* Cobalt-alloy implant debris induce HIF-1α hypoxia associated responses: a mechanism for metal-specific orthopedic implant failure. *PLoS One* **8**, (2013).

N, S. *et al.* MAPK signaling up-regulates the activity of hypoxia-inducible factors by its effects on p300. *J. Biol. Chem.* **278**, 14013–14019 (2003).

GL, S. Targeting HIF-1 for cancer therapy. *Nat. Rev. Cancer* **3**, 721–732 (2003).

Semenza, G. L. HIF-1 and tumor progression: pathophysiology and therapeutics. *Trends Mol. Med.* **8**, S62–S67 (2002).

Lee, J.-W., Bae, S.-H., Jeong, J.-W., Kim, S.-H. & Kim, K.-W. Hypoxia-inducible factor (HIF-1)α: its protein stability and biological functions. *Exp. Mol. Med. 2004 361* **36**, 1–12 (2004).

Hölscher, M. *et al.* Unfavourable consequences of chronic cardiac HIF-1α stabilization. *Cardiovasc. Res.* **94**, 77–86 (2012).

Summary

- 177 structure-based and 268 knowledge-based metal-interacting proteins uncovered
- 2170 unique metal-protein-disease triples mined from AOPs generated via ROBOKOP
- 72 secreted proteins prioritized for further review
- Cases studies demonstrated the utility of our workflow for accelerating hypothesis generation and literature review.
- Hypothetical AOPs can be used as the motivation for research into biomarkers of metal implant-related toxicity.

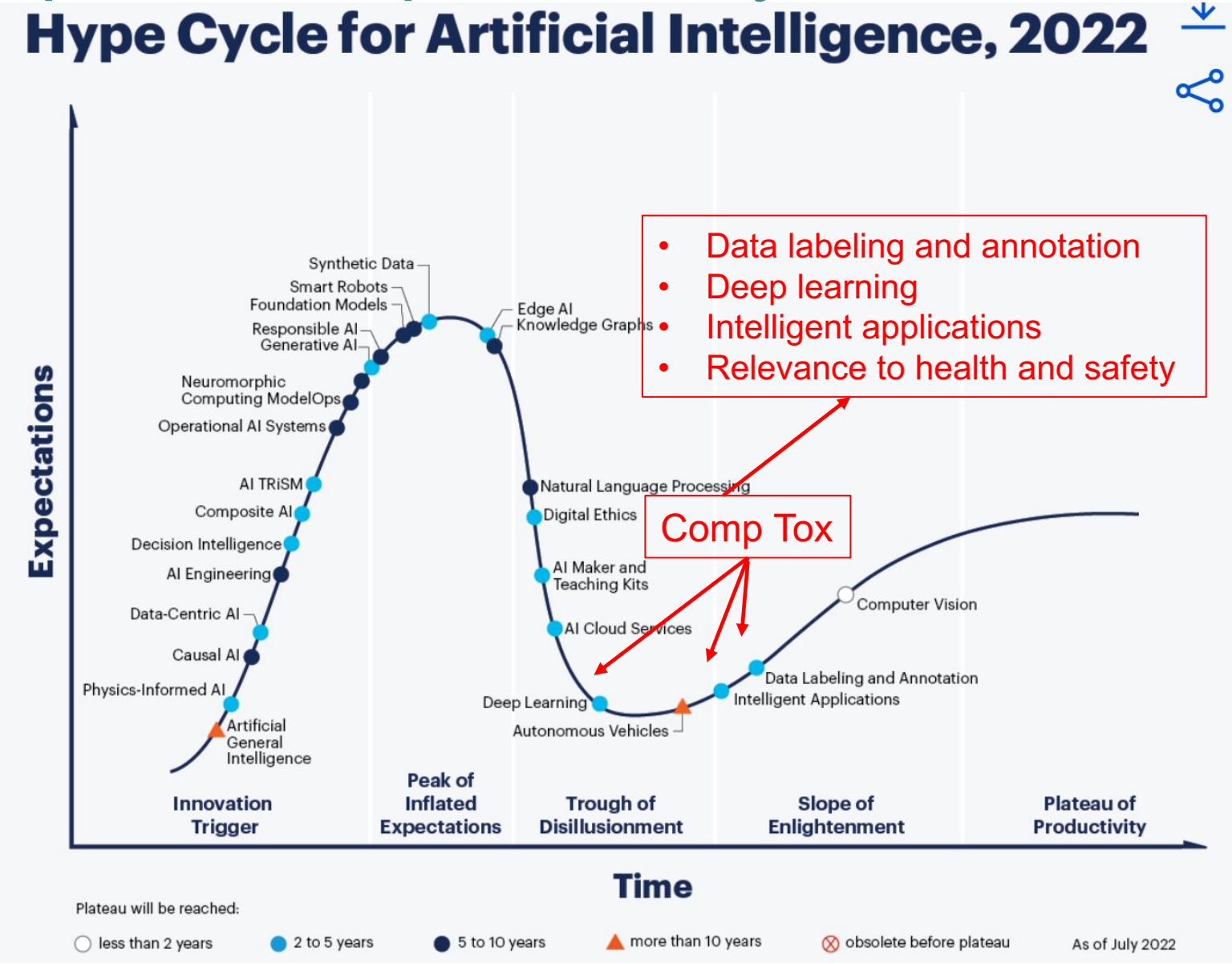
Questions that should be Addressed to Support the Acceptability of a QSAR model

- Compliance with OECD principles of model validation
- Descriptor values can be reproduced?
  - Values provided or their calculation described
- Model definition can be confirmed?
  - Modeling algorithm described explicitly
- External validation set independent of model?
  - External set has no overlaps with training set compounds
  - Descriptors filtered and selected based on training set only
  - Models tuned on training set only
- Model is available and can be tested independently?

Conclusions and Outlook

- Rapid accumulation of large biomolecular datasets :
  - Strong need for both chemical and biological data curation
- Novel approaches towards integration of inherent chemical properties with additional data streams
  - improve the outcome of structure – in vitro – in vivo extrapolation
- Interpretation of significant chemical and biological descriptors emerging from externally validated models
  - inform the selection or design of effective and safe chemicals
- Critical challenge: via integration of short-term assays and computational data modeling, achieve accuracy of *in vivo* toxicity acceptable for regulatory decision making
  - Minimize animal-based toxicity assessment
- Tool and data sharing
  - Commercial and public web portals

AI and CompTox: are we reaching the "plateau of productivity"*?



- Data labeling and annotation
- Deep learning
- Intelligent applications
- Relevance to health and safety

Comp Tox

https://www.gartner.com/en/articles/what-s-new-in-artificial-intelligence-from-the-2022-gartner-hype-cycle

UNC ESHELMAN SCHOOL OF PHARMACY

Molecular Modeling Lab

MML UNC.EDU

Acknowledgements

Principal Investigator
Alexander Tropsha
Research Professor
Eugene Muratov
Postdoctoral Fellows
Travis Maxfield
Haoxi Li
Graduate students
Jon-Michael Beasley
Joshua Hochuli
Matthew Hart
Holli-Joi Martin
Michael Brocidiacono
Kelvin Idanwekhai
James Wellnitz
Kathryn Kirchoff
Nyssa Tucker
Enes Kelestemur
Benjamin Strickland
Undergraduate students
Marielle Rath
David Major

Collaborators
Chris Bizon (RENCI/UNC)
Kara Fecho (RENCI/Copperline Professional Solutions)
Daniel Korn (EveryCure)
Kara Schatz (NCSU)
Rada Chirkova (NCSU)
Diego Rua (FDA)
Nicole Kleinstreuer (NIEHS)

MAJOR FUNDING

NIH
U24ES035214
OT3TR002020
OT2TR003430
UL1TR002489
1OT2TR003441
R01GM140154
R01FD007481UU
19AI171292

NSF
ITE-2333810
DMS-2324394

RTI
RTI 0282103.100.007