Machine Learning Models: Regulatory Application, Acceptance, and Implementation

Nicole C. Kleinstreuer

NICEATM Deputy Director

ICCVAM CoP
23rd January, 2018

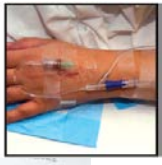
Agency for Toxic Substances and Disease Registry • Consumer Product Safety Commission • Department of Agriculture
Department of Defense • Department of Energy • Department of the Interior • Department of Transportation
Environmental Protection Agency • Food and Drug Administration • National Institute for Occupational Safety and Health
National Institutes of Health • National Cancer Institute • National Institute of Environmental Health Sciences
National Institute of Standards and Technology • National Library of Medicine • Occupational Safety and Health Administration

Outline

- Regulatory Needs & Challenges
- Current and Future Applications:
  - QSAR models (EPA, FDA)
  - TSCA Prioritization
  - Endocrine: CERAPP/CoMPARA
  - Skin Sensitization: Defined Approaches
  - Acute Oral Toxicity: Predictive Models

Environmental Chemical Disease Contributions

- Pesticides
  - Cancer, neurodegenerative diseases, thyroid
- Consumer products
  - Neurological, developmental, systemic
- Air pollutants
  - Childhood ADHD, autism, allergic asthma
- Drinking water contaminants
  - Systemic effects, cancer, neurological
- Endocrine Disruptors
  - Developmental impairment, decreased fertility, cancer



…….and many others…….

https://www.niehs.nih.gov/health/materials/index.cfm

Chemicals >> Data

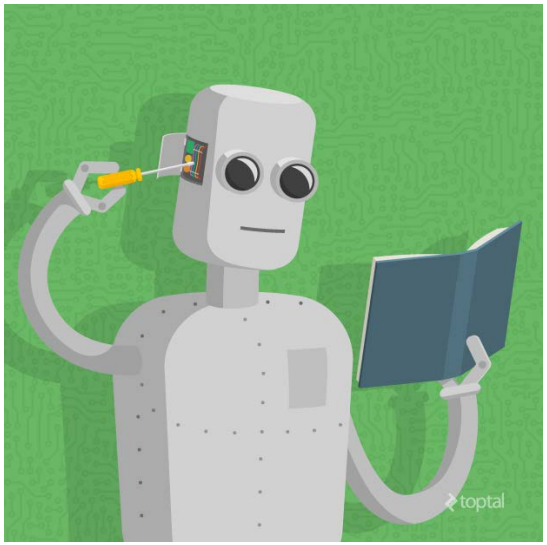
- 80+ million substances synthetized
- 140,000 chemicals in commerce (plus mixtures, natural products and metabolites)
- Less than 10% tested

Curated Legacy Data

e.g. REACH, ToxRefDB, ICE

High-Throughput Screening

e.g. ToxCast, Tox21



Omics techologies

e.g. transcriptomics, metabolomics, exposomics

High-Content Imaging

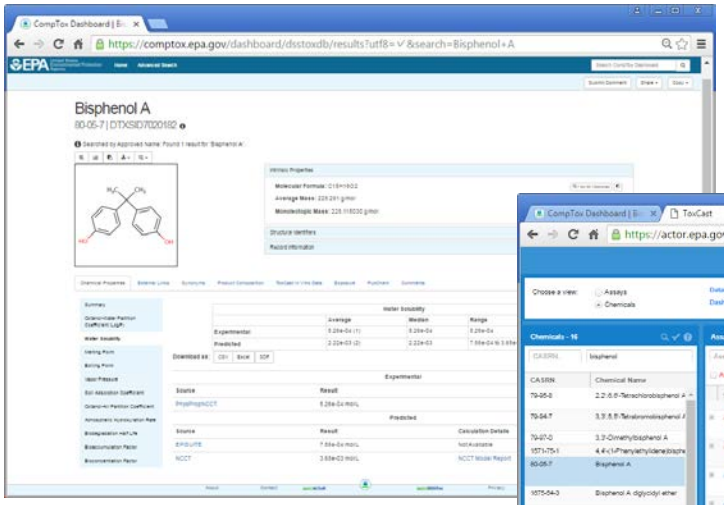
e.g. EuToxRisk

Current Regulatory Use of ML: Structure Based Models

- EPA/OPPT: Predictive Methods to Assess Hazard under TSCA
  - EcoSAR, OncoLogic, EPISuite
- FDA/CDER, CDRH: Genotoxicity and Carcinogenicity
  - Bacterial mutagenicity (expert-rule based & statistical)

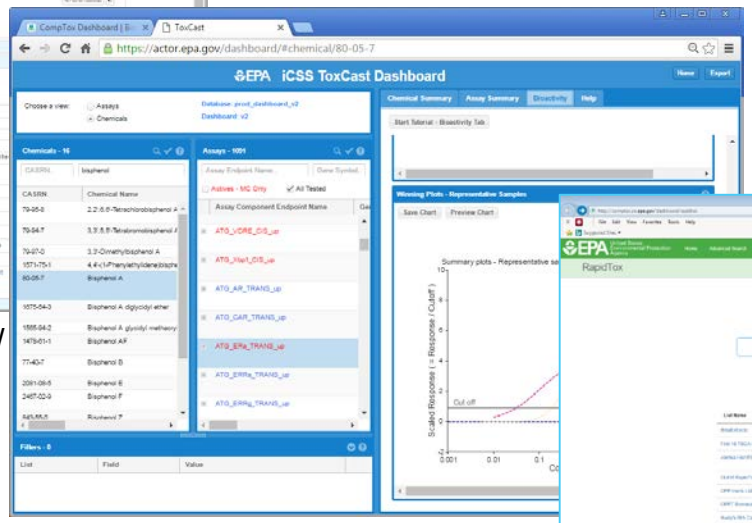
EPA/NCCT Decision Support Tools Deliver Data and Models

Comptox Chemistry Dashboard



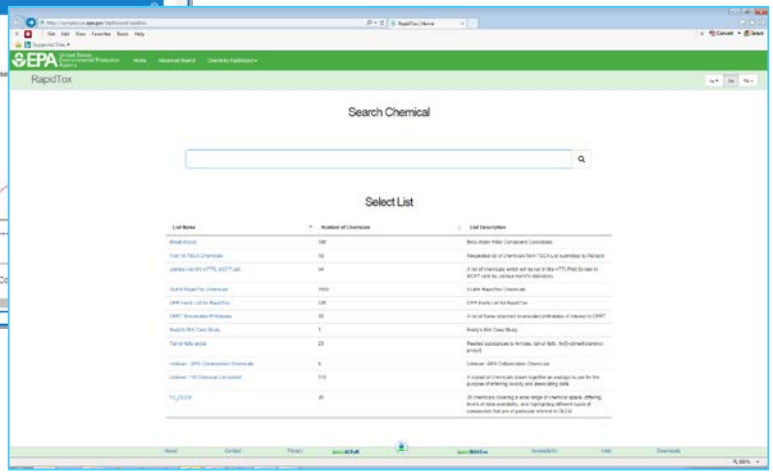
https://comptox.epa.gov/dashboard/

ToxCast Dashboard



https://actor.epa.gov/dashboard/

RapidTox Dashboard

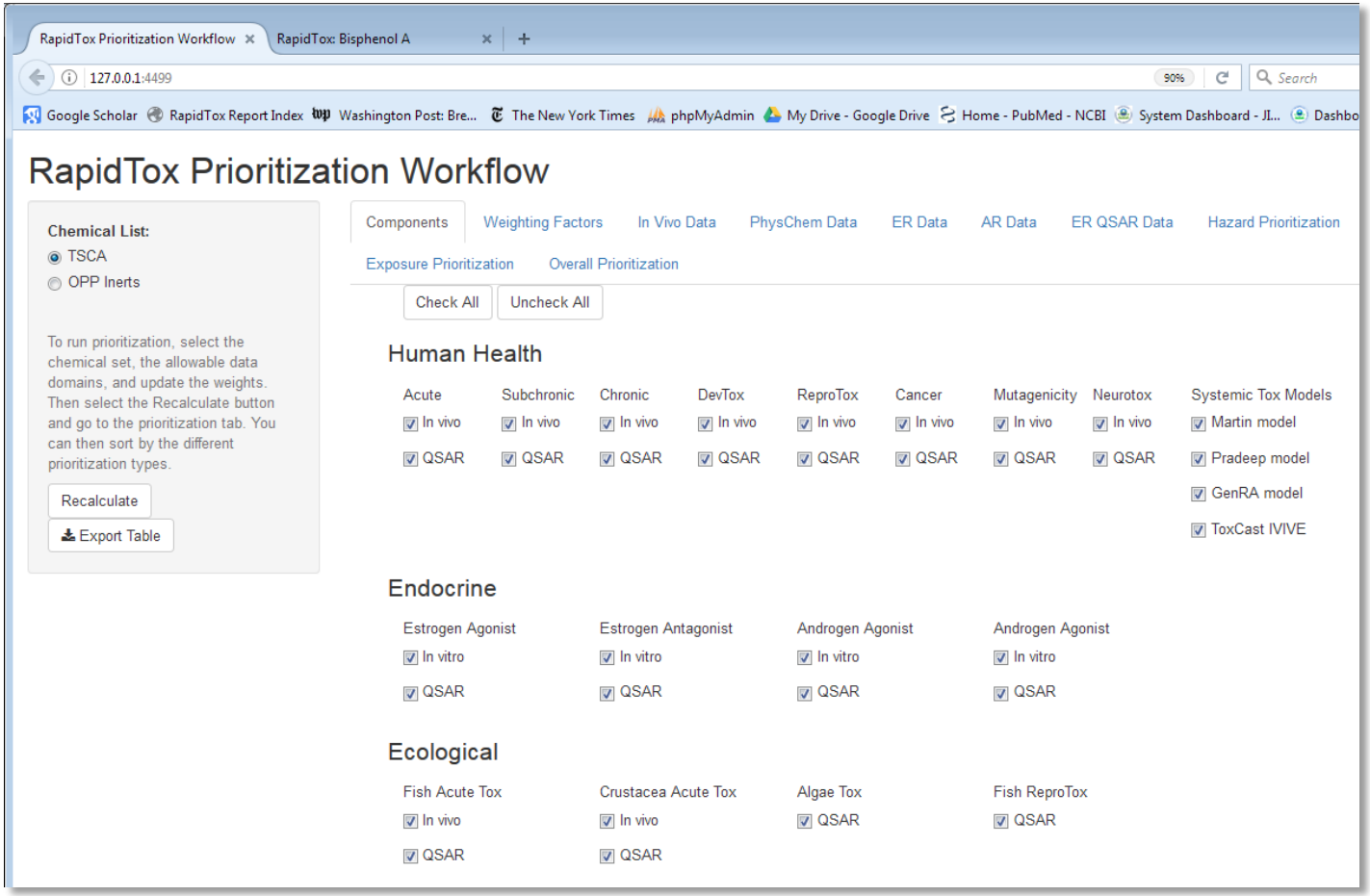


Internal Beta

Data Delivery Tools → ← Workflow Management Tool →

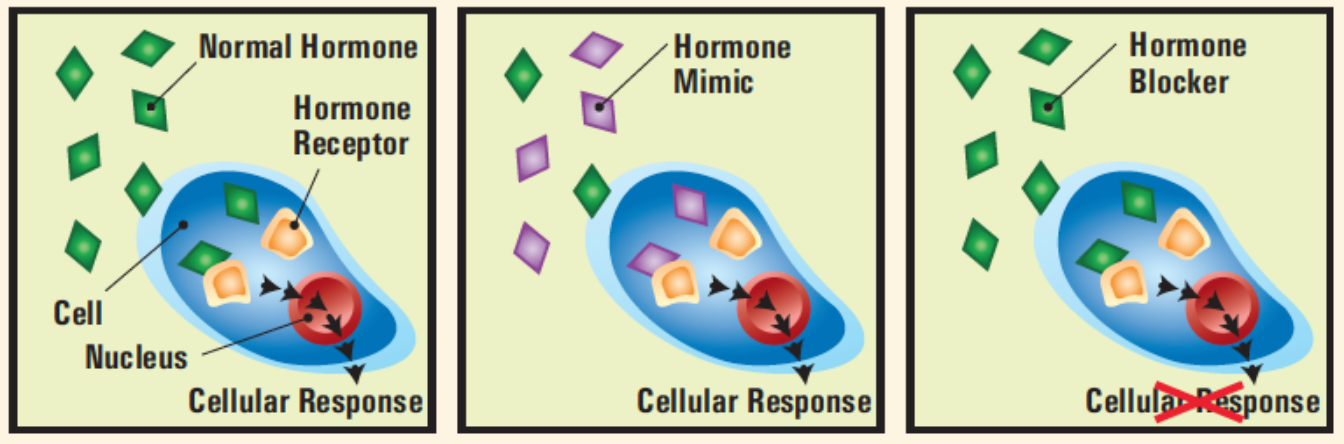
Rusty Thomas, EPA/NCCT

RapidTox: Prioritization Workflow



Rusty Thomas, EPA/NCCT

Environmental Endocrine Disruptors

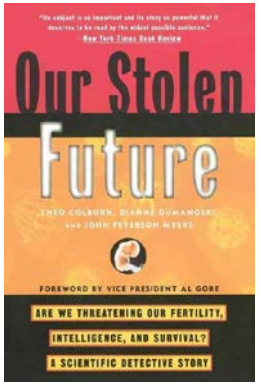


Reviews & Commentaries • Colborn et al.

Developmental Effects of Endocrine-Disrupting Chemicals in Wildlife and Humans

Theo Colborn,[1] Frederick S. vom Saal,[2] and Ana M. Soto[3]
[1]W. Alton Jones Foundation and World Wildlife Fund, Washington, DC, 20037 USA; [2]Division of Biological Sciences and John M. Dalton Research Center, University of Missouri, Columbia, MO 65211 USA; [3]Department of Anatomy and Cellular Biology, Tufts University, Boston, MA 02111 USA

1993, Environmental Health Perspectives



Legislative Mandates:

1996 Federal Food, Drug and Cosmetic Act
1996 Safe Drinking Water Act Amendments

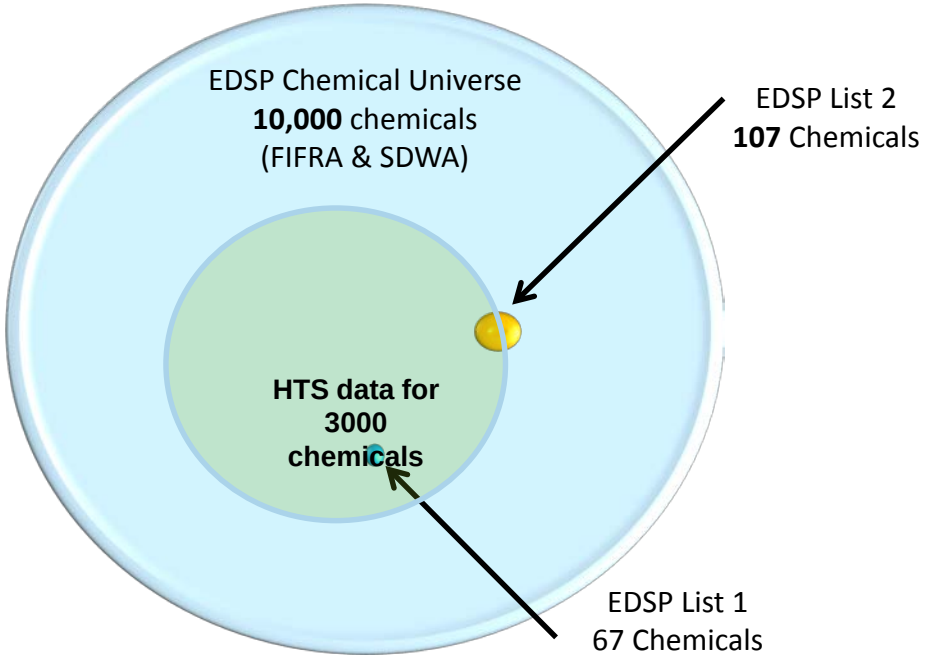
U.S. EPA EDSP

Evolution of the Endocrine Disruptor Screening Program

EDSP Tier 1 Testing: for the purposes of prioritization and screening, identify chemicals with the potential to disrupt estrogen, androgen, or thyroid hormone receptor signaling.

Mismatch between resources needed for EDSP Tier 1 testing and the number of chemicals to be tested

- ~$1M per chemical for Tier 1
- 11 low-throughput & animal tests

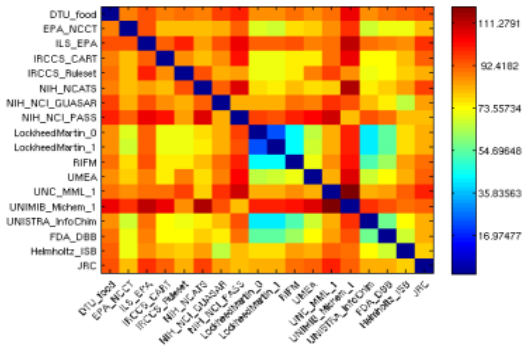


New Approach: **EDSP + Tox21 = EDSP21**

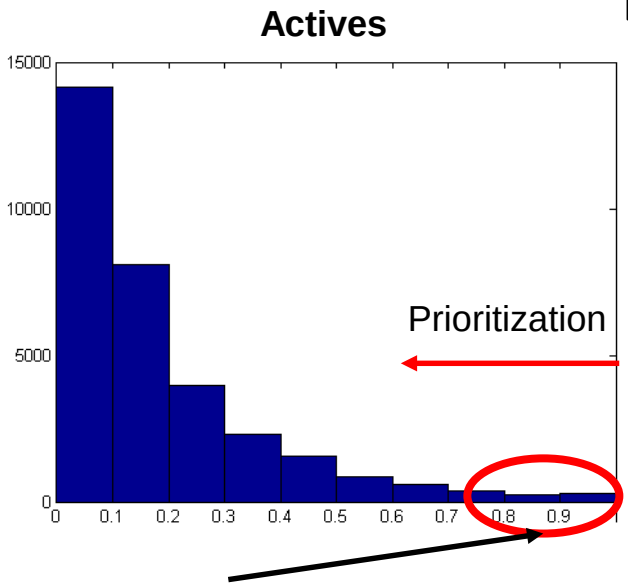
- Pathway-based predictive models (HTS in vitro assays)
- Validate to replace selected Tier 1 screening assays
- Train QSAR models to prioritize chemical universe

ER/AR QSAR Models

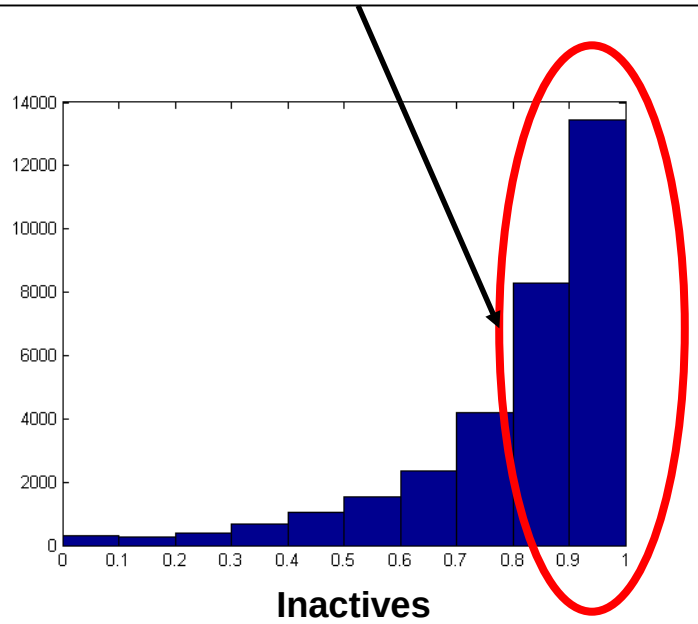
- Training set (ToxCast): 1677 chemical structures
- CERAPP: Global collaborative project for ER
  - 17 international groups participated
  - Individual and consensus models
  - *Mansouri et al. 2016 EHP*
- Prediction Set (EDSP):
  - 32,464 chemical structures
  - 5-10% predicted to be ER-active: Prioritize for further testing
- COMPARA: Global collaborative project for AR
  - 34 international groups participating
  - *Mansouri et al. 2018 in prep*



Correlation matrix of the CERAPP continuous ER models predictions



Most models predict most chemicals as inactive



757 chemicals have >75% positive concordance


ehp ENVIRONMENTAL HEALTH PERSPECTIVES

Articles | Collections | Authors | EHP 中文版 | Careers and Funding | E-Mail Alerts | About EHP

RESEARCH ARTICLE — VOLUME 124 | ISSUE 7 | JULY 2016

Environ Health Perspect; DOI:10.1289/ehp.1510267

CERAPP: Collaborative Estrogen Receptor Activity Prediction Project

Kamel Mansouri,[1,2] Ahmed Abdelaziz,[3] Aleksandra Rybacka,[4] Alessandra Roncaglioni,[5] Alexander Tropsha,[6] Alexandre Varnek,[7] Alexey Zakharov,[8] Andrew Worth,[9] Ann M. Richard,[1] Christopher M. Grulke,[1] Daniela Trisciuzzi,[10] Denis Fourches,[6]

CERAPP:
Only a small fraction of chemicals are prioritized for further testing

Mansouri et al. (2016) EHP 124:1023–1033
DOI:10.1289/ehp.1510267

Kamel Mansouri, Scitovation

Regulatory Use

US Government Information

One stop source for US Government Information

HOME CONSUMER DEFENSE & INTERNATIONAL RELATIONS EDUCATION & EMPLOYMENT FAMILY, HOME, & COMMUNITY HEALTH MONEY PUBLIC SAFETY & LAW REFERENCE & SCIENCE & TECHNOLOGY ABOUT

EDSP Prioritization: Collaborative Estrogen Receptor Activity Prediction Project (CERAPP) (SOT)

Humans are potentially exposed to tens of thousands of man-made chemicals in the environment. It is well known that some environmental chemicals mimic natural hormones and

EDSP dashboard:
http://actor.epa.gov/edsp21/

EPA US Environmental Protection Agency

Español | 中文: 繁體版 | 中文: 简体版 | Tiếng Việt | 한국어

Learn the Issues | Science & Technology | Laws & Regulations | About EPA | Search EPA.gov

Related Topics: Safer Chemicals Research | Contact Us | Share

Safer Chemicals Research Update June 2016

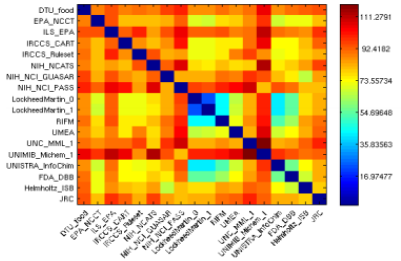
US EPA's Office of Research and Development provides quarterly updates, highlights, events and news about its chemical safety research. This is the June 2016 edition.

You will need Adobe Reader to view some of the files on this page. See EPA's About PDF page to learn more.

- June 2016 CSS Pathways News Anticipating Impacts of Chemicals (PDF) (13 pp, 1 MB)

Consensus Modeling: Powering Prediction Through Collaboration

Predictive computational models can efficiently help us prioritize thousands of chemicals for additional testing and evaluation. CSS scientists Kamel Mansouri and Richard Judson, from the U.S. EPA's National Center for Computational Toxicology (NCCT), led a large-scale modeling project called the Collaborative Estrogen Receptor Activity Prediction Project (CERAPP). CERAPP demonstrated the efficacy of using computational models with high-throughput screening (HTS) data to predict potential estrogen receptor (ER) activity of over 32,000 chemicals. This international collaborative effort (17 research groups from the United States and Europe) used both quantitative structure-activity relationship models and docking approaches to evaluate binding, agonist and antagonist activity of chemicals. A total of 48 models were developed. Each model was evaluated and weighed for its predictive accuracy using ToxCast and



Correlation matrix of the CERAPP continuous ER models predictions

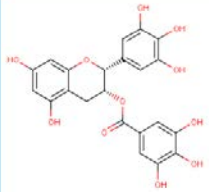
EPA United States Environmental Protection Agency

EDSP21 Dashboard

Endocrine Disruption Screening Program for the 21st Century

Chemical Summary | Public Information | Bioactivity Summary | Bioactivity | High-Throughput Exposure | Assay Definitions

Chemical Structure and Data

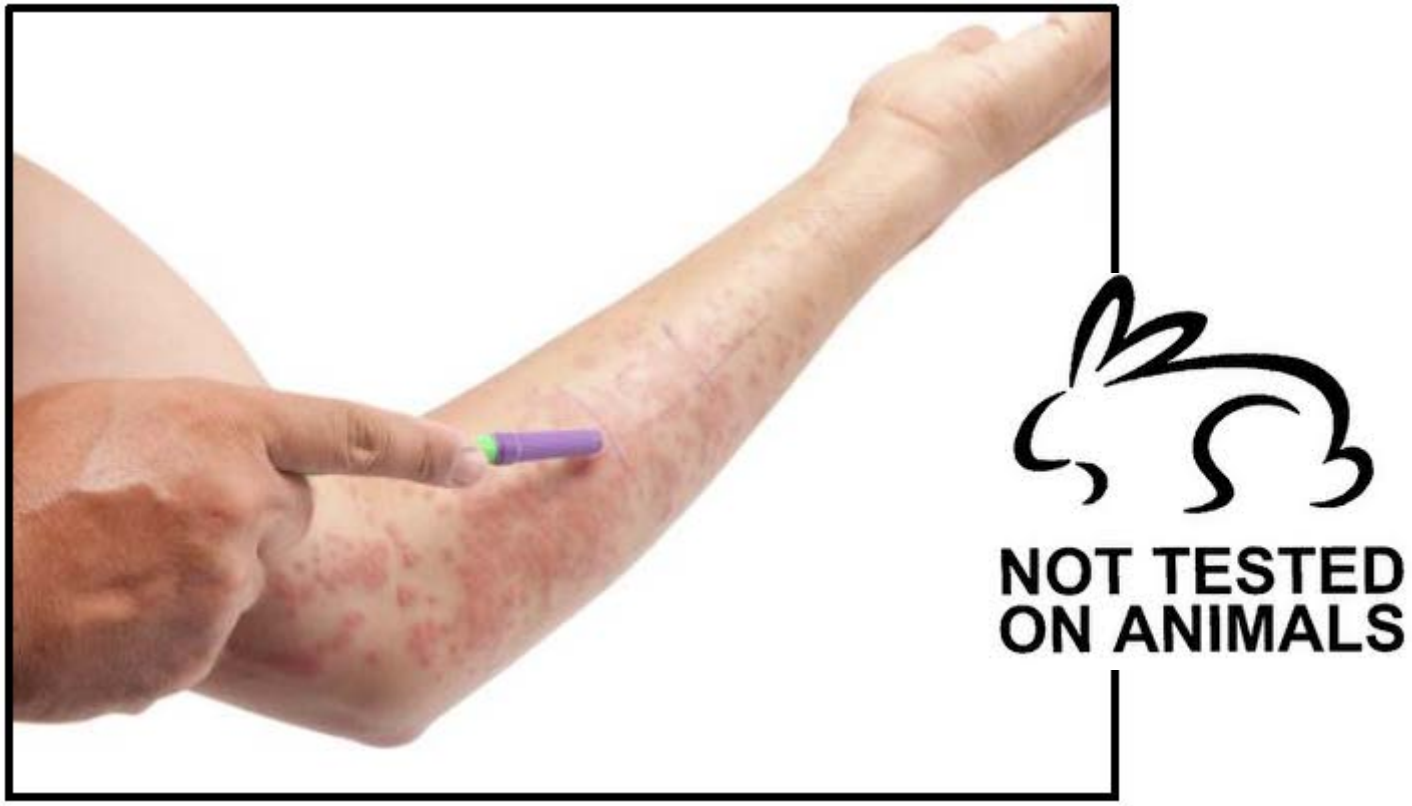


DSSTOX GSID	29889
CASRN	989-51-5
CASRN Type	Single Compound
Name	(-)-Epigallocatechin gallate
SMILES	OC1=CC(O)=C2C[C@@H](OC(=O)C3=CC(O)=C(O)C(O)=C3)[C@H](OC
InChI	InChI=1S/C22H18O11/c23-10-5-12(24)11-7-18(33-22(31)9-3-15(27)20(30)
InChI Key	WMBWREPUVVBILR-WIYYLYMNSA-N
Molecular Wt.	458.37
Chemical Formula	C22H18O11
Cytotoxicity Limit (uM)	0
Chemical Type	Organic
Chiral/Stereo	
dbl/Stereo	

Kamel Mansouri, Scitovation

Skin Sensitization

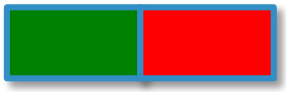
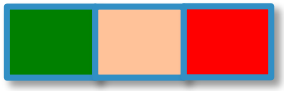
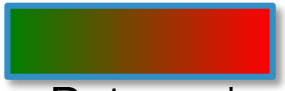
"Allergic Contact Dermatitis"



Accounts for 10-15% of all occupational disease *(Anderson et al. 2010)*

Major testing requirement for cosmetics, pesticides, industrial chemicals, etc.

U.S. Agency Requirements/Considerations

		Reference Animal Method	Classification Criteria
EPA	Pesticides Industrial chem	 LLNA	NS S  Hazard
CPSC	Household Products	 LLNA	NS S SS  Potency
FDA	Dermatological Products	 GPMT*	 Potency*

*human data preferred

Accuracy Against Human Clinical Data (~150 chems)

LLNA

Hazard	Potency
72%-82%	54% - 60%

GPMT / Buehler

Hazard	Potency
~72%	~60%

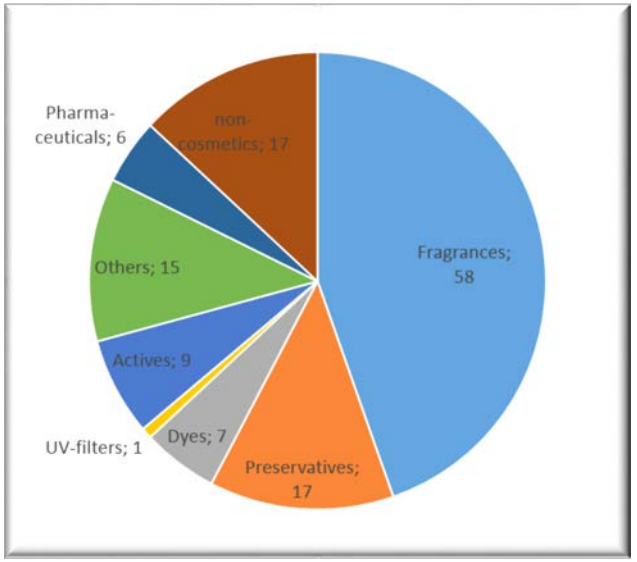
Reproducibility of Multiple Tests (~100 chems)

Hazard	Potency
~78%	~62%

ICCVAM. 1999. NIH Publication No. 99-4494
ICCVAM. 2010. NIH Publication No. 11-7709
Urbisch et al. 2015. Reg Tox Pharm 71:337-351.
Dumont et al. 2016. Tox In Vitro 34: 220-228
Hoffmann et al. 2018 Crit Rev Tox in press

Global Skin Sensitization Project

- Objective: analysis of available non-animal defined approaches (DAs)
- Collaboration with Cosmetics Europe
  - 128 substance dataset
  - LLNA (mouse) and human data
  - Curation/generation of
    - *in vitro* cell-based data that maps to AOP
    - *in silico* computer predictions, chemical structural features & properties



Spectrum of 128 substances

- Analyze non-animal DAs in an open source and transparent way
- Evaluate performance against the LLNA and human hazard/potency categories

Kleinstreuer et al. 2018 Crit Rev Tox in press

Research article
Journal of Applied Toxicology
Received: 13 October 2016, Revised: 26 October 2016, Accepted: 1 November 2016 Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI 10.1002/jat.3424

Prediction of skin sensitization potency using machine learning approaches

Qingda Zang[a], Michael Paris[a], David M. Lehmann[b], Shannon Bell[a], Nicole Kleinstreuer[c] and Warren Casey[c]

ABSTRACT: The replacement of ... agencies that use data from such ... out using animal data have been ... classified into potency categories ... node assay (LLNA) and human ...

Research article
Journal of Applied Toxicology
Received: 16 February 2016, Revised: 21 June 2016, Accepted: 21 June 2016 Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI 10.1002/jat.3366

Multivariate models for prediction of human skin sensitization hazard

Judy Strickland[a]*, Qingda Zang[a], Michael Paris[a], David M. Lehmann[b], David Allen[a], Neepa Choksi[a], Joanna Matheson[d], Abigail Jacobs[e], Warren Casey[c] and Nicole Kleinstreuer[c]

ABSTRACT: One of the Interagency Coordinating Committee on the V... the development and evaluation of non-animal approaches to ident... events necessary to produce skin sensitization suggests that no single ... imal tests. ICCVAM is evaluating an integrated approach to testing ar...

Research article
Journal of Applied Toxicology
Received: 9 October 2015, Revised: 10 November 2015, Accepted: 2 December 2015 Published online in Wiley Online Library: 6 February 2016
(wileyonlinelibrary.com) DOI 10.1002/jat.3281

Integrated decision strategies for skin sensitization hazard

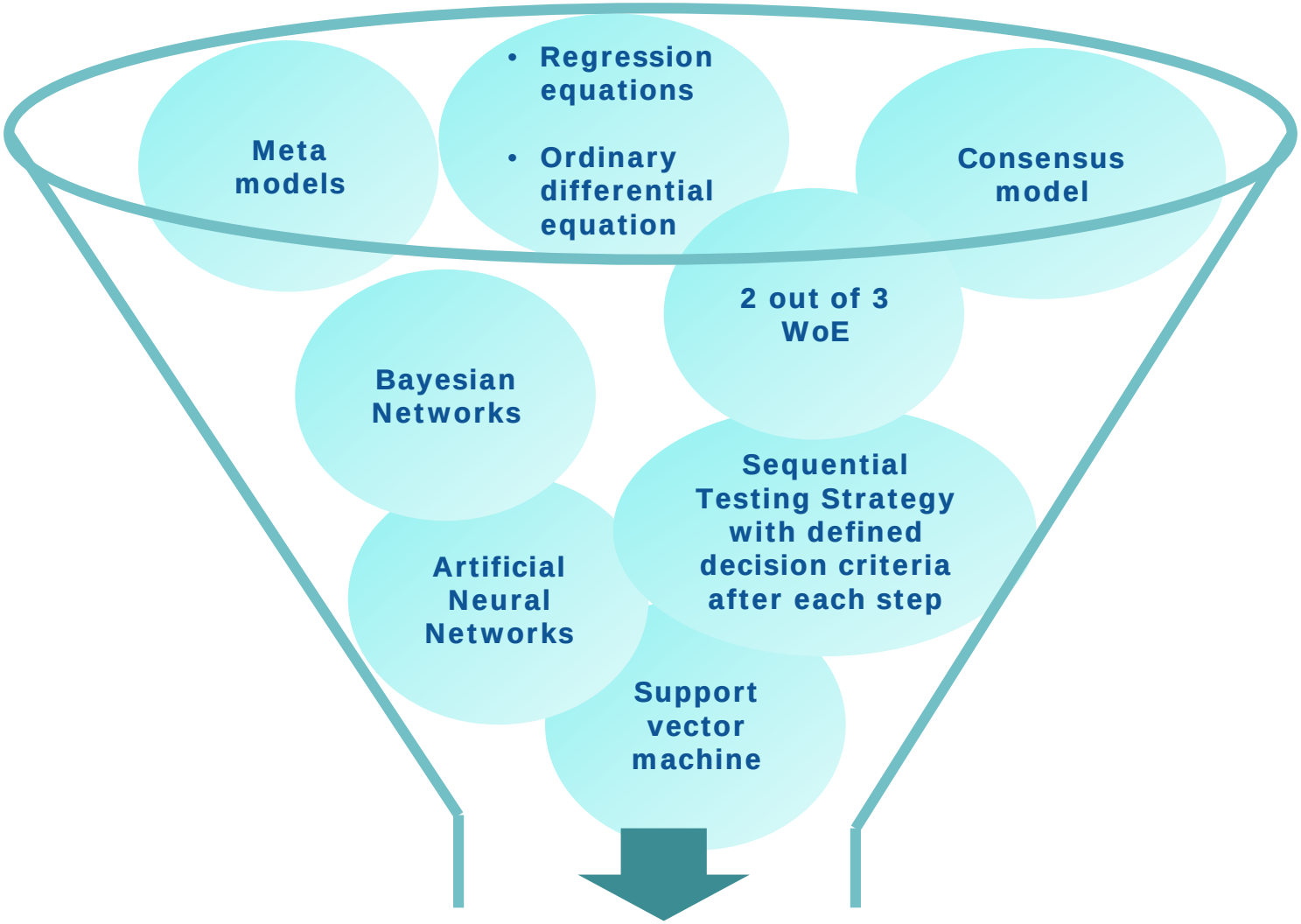
Judy Strickland[a], Qingda Zang[a], Nicole Kleinstreuer[a], Michael Paris[a], David M. Lehmann[b], Neepa Choksi[a], Joanna Matheson[c], Abigail Jacobs[d], Anna Lowit[e], David Allen[a] and Warren Casey[f]*

ABSTRACT: One of the top priorities of the Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM) is the identification and evaluation of non-animal alternatives for skin sensitization testing. Although skin sensitization is a complex process, the key biological events of the process have been well characterized in an adverse outcome pathway (AOP) proposed by the Organisation for Economic Co-operation and Development (OECD). Accordingly, ICCVAM is working to develop

NTP National Toxicology Program
U.S. Department of Health and Human Services

ATSDR AGENCY FOR TOXIC SUBSTANCES AND DISEASE REGISTRY

Different Modeling Approaches



Silvia Casati, JRC

Types of Information Sources

Adopted in *chemico, in vitro* methods				2 out of 3 WoE – BASF STS sequential strategy – Kao
Adopted in *chemico, in vitro* method(s)	*In silico* prediction(s)			ITS battery system -Kao
Adopted in *chemico, in vitro* method(s)	*In silico* prediction(s)	Phys-chem properties		SSWG SVM model - ICCVAM Non-testing pipeline approach -G. Patlewicz Decision Strategy – L'Oréal Bayesian Network – P&G
Adopted in *chemico, in vitro* method(s)	Non-standard Method(s)	Phys-chem properties		Artificial Neural Network model - Shiseido
Adopted in *chemico, in vitro* method(s)	Non-standard Method(s)	*In silico* prediction(s)		STS- RIVM
Adopted in *chemico, in vitro* method(s)	Non-standard Method(s)	*In silico* prediction(s)	Phys-chem properties	Sensitizer potency prediction- Givaudan
Modified adopted test methods				SARA model for risk assessment - Unilever
In silico prediction(s)				Consensus model - JRC

Silvia Casati, JRC

Defined Approach Evaluation

Most non-animal defined approaches evaluated so far perform **better** than the LLNA at predicting human skin sensitization hazard and potency.

(And when compared to the LLNA, are equivalent in performance to the LLNA at predicting itself.)

International Harmonization

- OECD proposal (SPSF) submitted November 2016
  - Co-led by U.S., EU, and Canada
  - Create an international performance based test guideline for non-animal defined approaches to skin sensitization testing
  - Achieve widespread replacement of mouse test
- National coordinators (WNT) voted unanimously to include the project in OECD workplan, April 2017
- Special session of the WNT met in December 2017 to review progress and discuss next steps
  - Achieved consensus on evaluation framework for consideration and assessment of DAs

OECD BETTER POLICIES FOR BETTER LIVES

Rat oral acute toxicity: LD50 Database

- Multiple existing resources containing rat oral acute toxicity LD50 data were mined and merged

Data source	Number of LD50 values	Number of unique chemicals
ECHA ChemProp	5,533	2,136
NLM HSDB	3,981	2,205
JRC AcutoxBase	637	138
NLM ChemIDplus	13,072	12,977
NICEATM PAI	364	293
OECD eChemPortal	10,119	2,290

Total:
34,511 LD50 values
16,307 chemicals

↓ Identify unique data in mg/kg

21,210 LD50 values
15,698 chemicals

- LD50 data comprised point estimates as well as limit tests

Agnes Karmaus, ILS/NICEATM

Development of Predictive Models for Acute Oral Toxicity

- International modeling community invited to build models to predict acute oral systemic toxicity
- ICCVAM agencies informed model endpoints
- Training and test data derived from large dataset compiled by NICEATM and EPA/NCCT
  - 11,992 QSAR-ready structures (75% training, 25% test)
  - Quantitative & qualitative evaluation
  - Models will be integrated to yield consensus predictions

https://ntp.niehs.nih.gov/go/tox-models

Endpoints to be Modeled

Participants are asked to develop models for any/all of the following endpoints identified based on regulatory needs provided by ICCVAM agencies:

1. Very toxic (< 50 mg/kg vs. all others)
2. Nontoxic (>2000 mg/kg vs. all others)
3. LD50 point estimate
4. EPA hazard categories (n=4)
5. GHS hazard categories (n=5)*

*GHS categories 5 and “not classified” are combined into one category

Modeling Approach

- Modelers are encouraged to consider different modeling approaches
  - Machine learning, global/local, hybrid/consensus models, etc.
- Models could include any variety of data inputs:
  - Chemical features/structure classes, physiochemical properties, product use categories, production volumes, in vitro data (measured or predicted), etc.

Evaluation Criteria

The OECD QSAR validation principles to be considered as guidance:

1. **A defined endpoint**
2. **An unambiguous algorithm**
3. **A defined domain of applicability**
4. **Appropriate measures of goodness-of–fit, robustness and predictivity**
5. **Mechanistic interpretation, if possible**

Timeline

- **November 17, 2017:** Release of Training Data to the public**.**
- **December 15, 2017:** Release of Prediction Data to the public.
- **February 9, 2018:** Deadline for submission of model results and documentation to NICEATM.
- **March 9, 2018:** Organizing Committee finalizes selection of models to be invited for platform presentations and notifications are sent to presenters.
- **April 11-12, 2018:** Predictive Models for Acute Oral Systemic Toxicity Workshop, NIH Natcher Conference Center, Bethesda, MA.

https://ntp.niehs.nih.gov/go/tox-models

Summary

- Toxicology data can be synthesized and modeled effectively using machine learning approaches.
  - Also: exposure, use case, systematic review, etc.
- Machine learning models (i.e. QSARs) have already achieved limited acceptance in the regulatory space.
- Additional education, training, and communication will facilitate more widespread adoption.

https://ntp.niehs.nih.gov/pubhealth/evalatm/natl-strategy/

Acknowledgments

- Gino Scarano
- Jennifer Goode
- Paul Browne
- Agnes Karmaus
- Grace Patlewicz
- Rusty Thomas
- Kamel Mansouri
- Richard Judson
- Sebastian Hoffmann
- Silvia Casati
- ICCVAM partners
- Cosmetics Europe STTF
- ILS/NICEATM
- EURL ECVAM/JRC
- Health Canada
- ICATM partners

Extra Slides

Previous CoP webinars

- **Incorporating Chemical Information: Resources, Limitations, and Characterizing the Domain of Applicability for 21st Century Toxicity Testing (January 24, 2017)**
- https://ntp.niehs.nih.gov/go/commprac-2017
- **Fundamentals of Using Quantitative Structure-Activity Relationship Models and Read-across Techniques in Predictive Toxicology (January 26, 2016)**
- https://ntp.niehs.nih.gov/go/commprac-2016

Big biological data in toxicology

Name	General Information	Data description
PubChem	Over 50 million compounds, over 700,000 bioassays, over 13 billion data points	Toxicity, genomics and literature data
ChEMBL	Over 600,000 compounds, 3.3 million bioassay readout data	Literature toxicity data
ACToR	The toxicity results from 100 various data resources	Both in vitro and in vivo toxicity data
ToxNET	Over 50,000 environmental compounds from 16 different resources	Both in vitro and in vivo toxicity data
SEURAT	Over 5,500 cosmetic-type compounds in the current COSMOS database web portal	Animal toxicity data
REACH	816,048 studies for 9,800 substances and 3,600 study types	Data submitted in EU chemical legislation, made machine-readable by Luechtefeld et al. 2016a (this issue)
CTD	Over 13,000 compounds, over 32,000 genes, over 6000 diseases	Compound, gene and disease relationships
CEBS	About 10,000 toxicity bioassays from various sources	Gene expression data
DrugMatrix	About 600 drug molecules and 10,000 genes	Gene expression data
Cmap	About 1,300 compounds and 7,000 genes	Gene expression data

Zhu et al. 2016 ALTEX

Impact of Variability on Hazard Classification