Interagency Coordinating Committee on the Validation of Alternative Methods

Moving Beyond Animal Data as the Gold Standard

Nicole Kleinstreuer

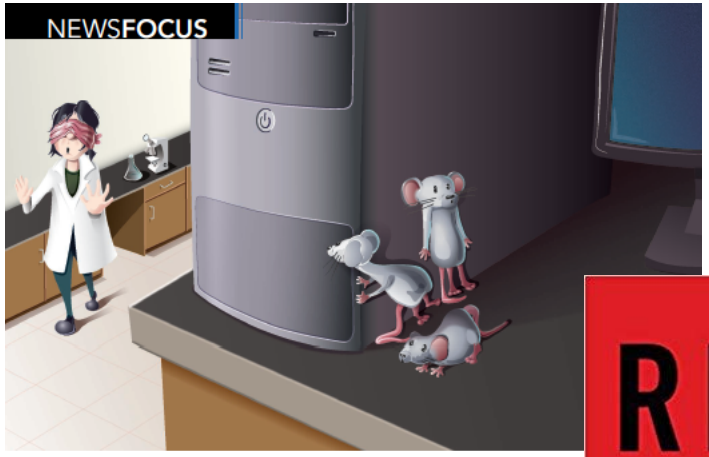
Deputy Director, NICEATM

SACATM

5-6 September, 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 Institute • National Institute of Standards and Technology • Occupational Safety and Health Administration

The Reproducibility Crisis



When Mice Mislead

Tackling a long-standing disconnect between animal and human studies, some charge tha researchers need stricter safeguards and better statistics to ensure their science is s

THREE MICE HAD VANISHED. AND ULRICH Dirnagl had a hunch about where they'd ended up: in the metaphorical dustbin housing animals—and there are lots of them—that line up at an experiment's starting line but are discarded before the finish. The paper that Dirnagl, director of the Center for Stroke Research at Charité University Medicine Berlin, was reviewing described how a new drug protected a rodent's brain after a stroke. The authors used 20 mice, half of which got the therapy. But mysteriously, only seven of the 10 treated animals appeared in a graph analyzing the results.

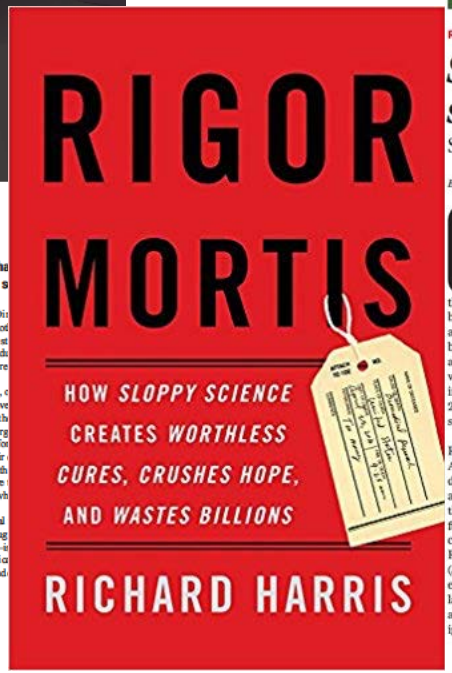
"I wrote to the editor and said, 'I cannot judge this paper, I need to know where the three mice went,'" Dirnagl recalls. For 6 months, radio silence. Then, the editor responded. He'd heard from the authors, he told Dirnagl. The three mice, suffering from massive strokes, had died, and the authors had simply left them out of the paper. Extra analysis of their stroke drug, however, revealed that those mice had an important message to bear: The therapy harmed the brain rather than helping it.

"This isn't fraud," says Dirnagl, who often works with mice. Dropping animals from a research study for any number of reasons, he explains, is an entrenched, accepted part of the culture. "You look at your data, there are no rules. … People exclude animals at their whim, they just do it and they don't report it." That bad habit, he believes, is one of several that plague animal studies.

For years, researchers, pharmaceutical companies, drug regulators, and even the general public have lamented how rarely therapies that cure animals do much of anything for humans. Much attention has focused on whether mice with different diseases accurately reflect what happens in sick people. But Dir
others suggest there's anot
problem. Many animal st
done, they say, and if condi
rigor they'd be a much more
of human biology.

It's hard to generalize, c
studies cut across a massive
tracking everything from th
molecules in a healthy org
of a new drug poised for
And many who stake their
studies conduct them with
weighing how to structure
and chasing the science wh
subjects take it.

That said, even animal
a big effect on human drug
work Dirnagl reviewed—i
fewer standards than clinic
There, volunteers are rand



REPRODUCIBILITY

Sloppy reporting on animal studies proves hard to change

Scientists appear to ignore guidelines adopted 7 years ago

By **Martin Enserink**

Closely read any paper on an animal experiment, and you're likely to have many questions. What strain of mice was used, and what were their sex and age? Were animals randomly assigned to control and treatment groups? Was the researcher who examined outcomes blinded to what group they were in? The absence of such details partly explains why between 51% and 89% of animal studies aren't reproducible. It may also help explain why so many treatments reported to work in animals have flopped in humans (*Science*, 22 November 2013, p. 922). Yet it's proving surprisingly hard to solve the problem.

In 2010, the U.K. National Centre for the Replacement, Refinement & Reduction of Animals in Research (NC3Rs) in London developed a checklist of items that any paper about in vivo research ought to include. More than 1000 scientific journals and two dozen funding agencies have endorsed the so-called ARRIVE guidelines—short for Animal Research: Reporting of In Vivo Experiments. (*Science* has not officially endorsed them, but encourages authors to comply.) But 7 years later, studies suggest that many scientists are either unaware of the guidelines or are ignoring them.

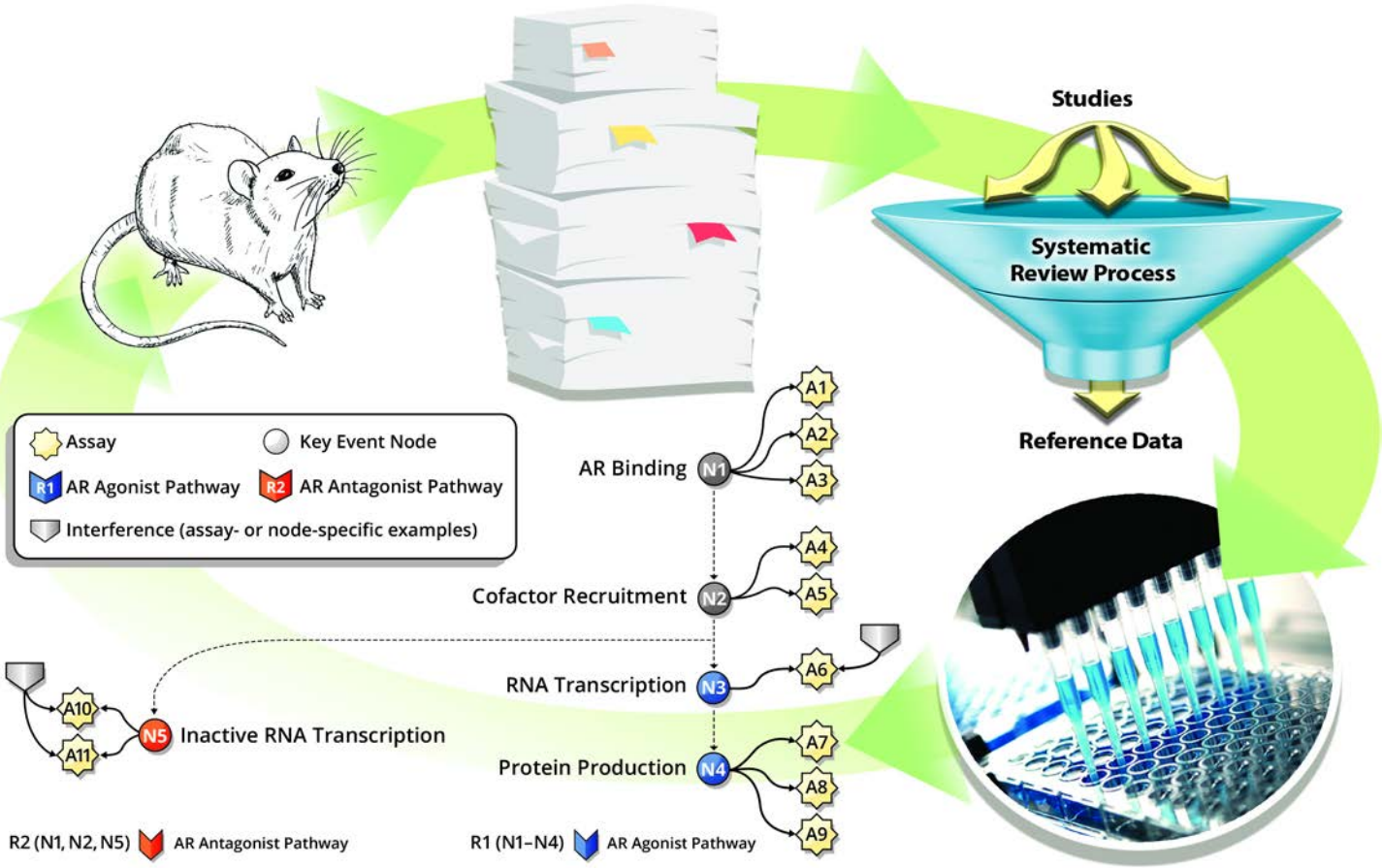
"We just don't seem to make much progress," says Merel Ritskes-Hoitinga of Radboud University Medical Center in Nijmegen, the Netherlands, who co-organized a 25 September roundtable in Edinburgh where scientists met with journal editors and funders such as the United Kingdom's Medical Research Council and the Wellcome Trust to discuss ways of speeding up implementation of the guidelines. One problem may be that ensuring compliance can take a lot of work, both for authors and journals.

The 38 items in the checklist provide a "gold standard," says Malcolm Macleod, a neurologist at the University of Edinburgh who has studied the problems in animal experimentation. The list covers a wide range of issues, from a paper's title and study design to how the animals were cared for, results, and conflicts of interest. But a 2014 survey showed almost no improvement in reporting in journals of Nature Publishing Group (NPG) and PLOS during the first 2 years after the guidelines were introduced, even though both publishers had endorsed ARRIVE. That study's last author, Sandra Amor of VU University Medical Center in Amsterdam, says that an as-yet-unpublished analysis shows that things weren't much better in the 2012–15 period.

Macleod and colleagues have tested one

Validation Workflow

Importance of Curated Reference Data



Kleinstreuer et al. 2018 RTX in press

Addressing Data Quality

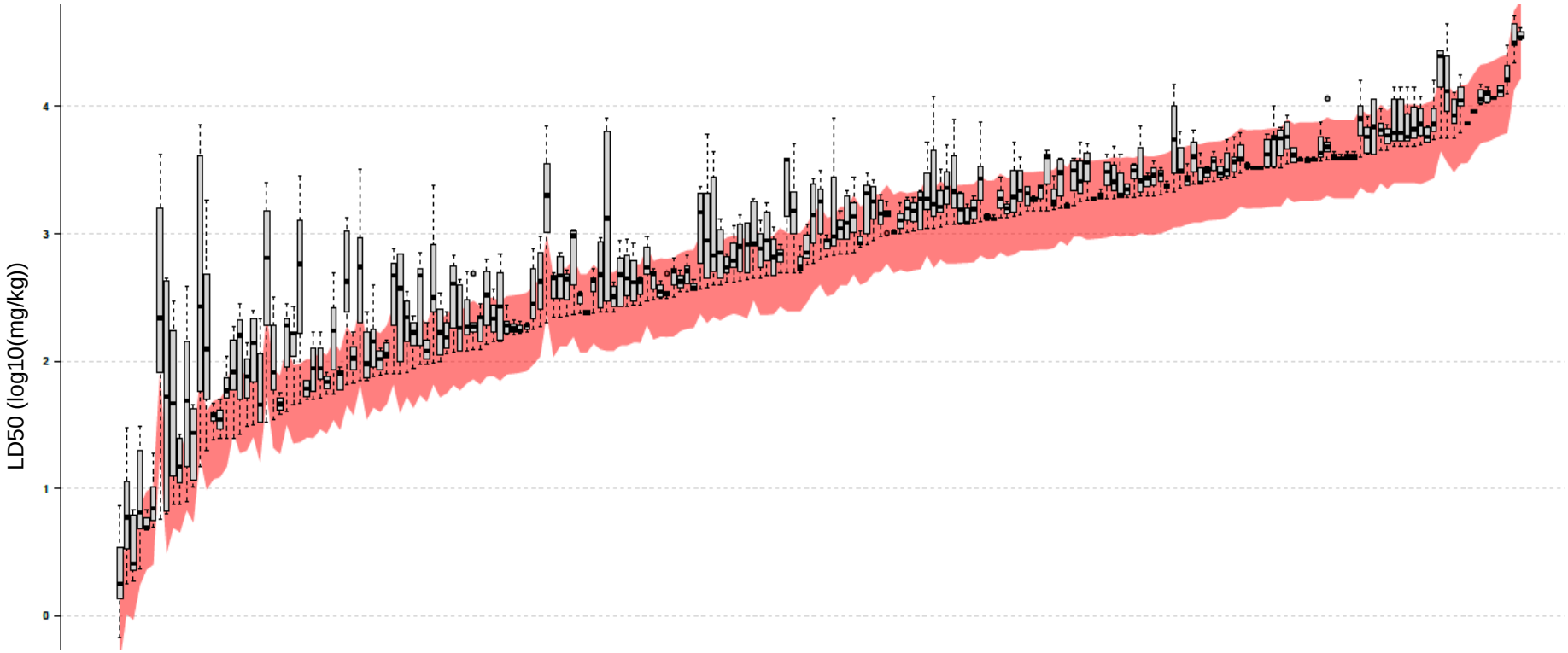
Ex: Rat oral acute toxicity LD50 Database

- Identify transcription errors (e.g. 20005000 mg/kg, >10 mg/kg, confidence intervals as values)
- Manual curation of highly variable chemicals; identify source data

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

Defining a Confidence Range

Bootstrapping of the standard deviations for repeat test chemicals identified a 95% confidence interval for LD50 values of ± 0.31 $\log_{10}$(mg/kg)



Variation in Classification

Ex: ECHA Ocular Data

CASRN	ECHA Data
100-41-4	Not Irritating
100-41-4	Category 2A
100-74-3	**Category 1**
100-74-3	**Category 2A**
102-06-7	Category 1
102-06-7	Category 2
10213-78-2	**Category 1**
10213-78-2	**Category 2A**
103-50-4	Not Irritating
103-50-4	Category 2B
10361-93-0	**Category 1**
10361-93-0	**Category 2A**

Reproducibility of Animal Data

Binary Hazard Classification

- Uterotrophic: ~74%
- Hershberger: ~72%

- Skin Sensitization: ~78%

- Acute Systemic: ~81%
- Skin Irritation: ~76%

- Eye Irritation: ~84%

Kleinstreuer et al. 2016; Browne et al. 2018; Kleinstreuer et al. 2018a; Dumont et al. 2016; Hoffmann et al. 2018; Kleinstreuer et al. 2018b; Karmaus et al. in prep; Leuchtefeld et al. 2018

Reproducibility of Animal Data

Ocular Potency Categorization

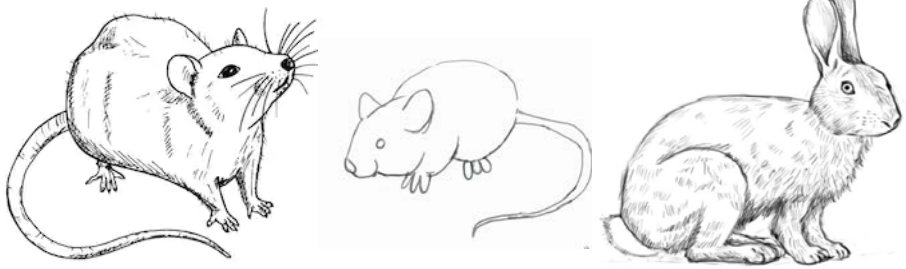
Conditional probability of Draize evaluations given a previous test result

491 substances with at least two Draize studies and extractable eye irritation category in REACH registrations 2008-2014

Prior Type	1	2A	2B	Non	Total
1	73%	16.1%	0.4%	10.4%	46
2A	4.2%	32.9%	3.5%	59.4%	138
2B	0.2%	4%	15.5%	80.2%	86
Non	1.1%	3.5%	1.5%	93.9%	400

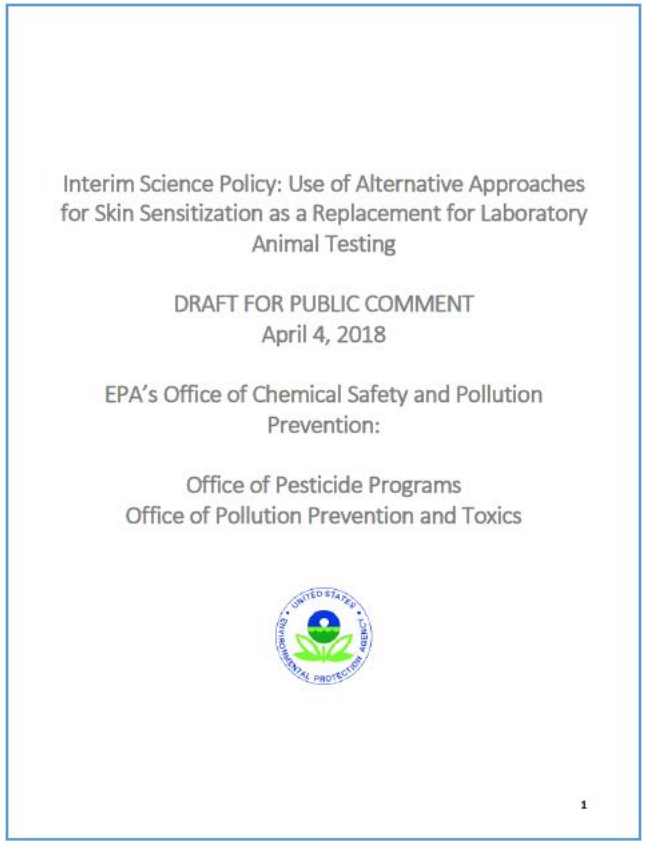
Leuchtefeld et al. 2017

How to Benchmark Alternative Models



How to Benchmark Alternative Models

Animal data reproducibility as threshold for performance



Interim Science Policy: Use of Alternative Approaches for Skin Sensitization as a Replacement for Laboratory Animal Testing

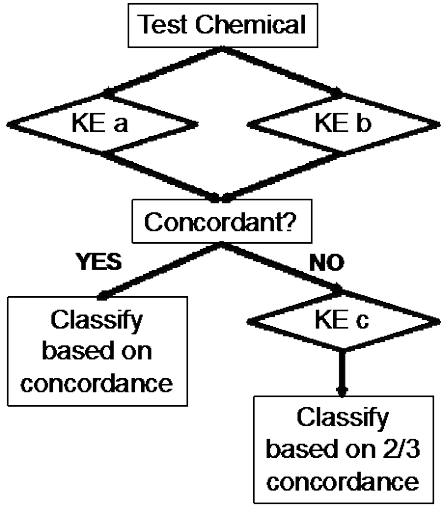
DRAFT FOR PUBLIC COMMENT
April 4, 2018

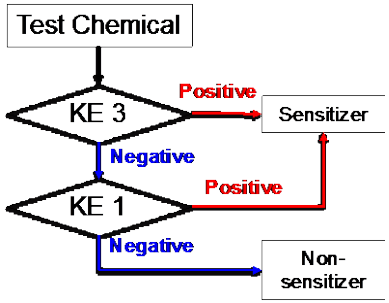
EPA's Office of Chemical Safety and Pollution Prevention:

Office of Pesticide Programs
Office of Pollution Prevention and Toxics

1

Defined Approaches (AOP WoE and KE 1/3 STS) accepted by EPA based on comparison to LLNA (mouse) data





Development of Predictive Models for Acute Oral Toxicity

- Use large database of rat oral LD50 values to train (and test) QSAR models to predict acute oral systemic toxicity
- 32 groups from the US, Europe, and Asia responded with 135 models for LD50, EPA and GHS categories, and binary nontoxic vs all others and very toxic vs all others.
- Models were qualitatively and quantitatively assessed and combined into consensus models.
- Consensus model performance compared with animal test reproducibility for binary, categorical, and quantitative models

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

Predictive Models for Acute Toxicity:

Performance vs Animal Data

	Rat Oral LD50: Reproducibility			Consensus Model Performance (Tr/Ts Avg)		
	Sensitivity	Specificity	BA	Sensitivity	Specificity	BA
VT	63%	99%	81%	77%	95%	86%
NT	96%	82%	89%	82%	92%	87%
EPA	74%	91%	82%	62%	94%	78%
GHS	66%	92%	79%	54%	92%	73%

	R2	RMSE	R2	RMSE
LD50	0.8	0.42	0.74	0.42

Communicating Variability to Stakeholders

Review

A Section 508–conformant HTML version of this article is available at http://dx.doi.org/10.1289/ehp.1510183.

A Curated Database of Rodent Uterotrophic Bioactivity

Nicole C. Kleinstreuer,[1] Patricia C. Ceger,[1] David G. Allen,[1] Judy Strickland,[1] Xiaoqing [illegible] Jonathan T. Hamm,[1] and Warren M. Casey[2]

[1]Integrated Laboratory Systems, in support of the National Toxicology Program Interagency Center for Toxicological Methods (NICEATM), Research Triangle Park, North Carolina, USA; [2]NICEATM, Division o Program, National Institute of Environmental Health Sciences, National Institutes of Health, Department Research Triangle Park, North Carolina, USA

Background: Novel *in vitro* methods are being developed to identify chemicals that may interfere with estrogen receptor (ER) signaling, but the results are difficult to put into biological context because of reliance on reference chemicals established using results from other *in vitro* assays and because of the lack of high-quality *in vitro* reference data. The Organisation for Economic Co-operation and Development (OECD)-validated rodent uterotrophic bioassay is considered the "gold standard" for identifying potential ER agonists.

Objectives: We performed a comprehensive literature review to identify and evaluate data from uterotrophic studies and to analyze study variability.

Methods: We reviewed 670 articles with results from 2,615 uterotrophic bioassays using 235 unique chemicals. Study descriptors, such as species/strain, route of administration, dosing regimen, lowest effect level, and test outcome, were captured in a database of uterotrophic results. Studies were assessed for adherence to six criteria that were based on uterotrophic regulatory test guidelines. Studies meeting all six criteria (458 bioassays on 118 unique chemicals) were considered guideline-like (GL) and were subsequently analyzed.

Results: The immature rat model was used for 76% of the GL studies. Active outcomes were more prevalent across rat models (74% active) than across mouse models (36% active). Of the 70 chemicals with at least two GL studies, 18 (26%) had discordant outcomes and were classified as both active and inactive. Many discordant results were attributable to differences in study design (e.g., injection vs. oral dosing).

Conclusions: This uterotrophic database provides a valuable resource for understanding *in vivo* outcome variability and for evaluating the performance of *in vitro* assays that measure estrogenic activity.

Citation: Kleinstreuer NC, Ceger PC, Allen DG, Strickland J, Chang X, Hamm JT, Casey WM. 2016. A curated database of rodent uterotrophic bioactivity. Environ Health Perspect 124:556–562; http://dx.doi.org/10.1289/ehp.1510183

Introduction

Understanding the impact of endocrine bioactive chemicals on human health and the environment is a high priority for U.S. and international agencies. The large number of untested chemicals in commerce (> 80,000) necessitates the use of high-throughput screening (HTS) programs such as the U.S. Environmental Protection Agency (EPA) ToxCast™ initiative and the Tox21 U.S. federal partnership to quickly identify potential endocrine disruptors and to help characterize any hazards they may pose (Dix et al. 2007; Judson et al. 2010; Kavlock et al. 2012; Tice et al. 2013; U.S. EPA 2011a, 2012). Furthermore, there is growing societal pressure to avoid animal testing and to develop alternative approaches that replace, reduce, or refine the use of animals in toxicity testing [Hartung 2009; Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Authorization Act of 2000].

To determine the usefulness and limitations of a novel alternative method for identifying endocrine activity and to show that it is fit for its intended purpose, the method must be evaluated against a set of chemicals that have demonstrated activity and well-defined properties (potency and efficacy) against the target nuclear receptor and the subsequent biological pathway. At the present time, reference chemicals used to validate *in vitro* assays aimed at detecting potential endocrine disruptors (estrogen, androgen, and thyroid receptors) are selected based only on their activity in other *in vitro* assays, a circular validation paradigm that arose because of the lack of sufficient *in vivo* data [ICCVAM et al. 2011; Organisaton for Economic Co-operation and Development (OECD) 2012]. To facilitate work that will better elucidate and characterize the relationship between the *in vitro* and *in vivo* estrogen bioactivity of chemicals, the National Toxicology Program Interagency Center for Evaluation of Alternative Toxicological Methods (NICEATM) developed a curated database of high-quality *in vivo* rodent uterotrophic bioassay data extracted from published studies (http://ntp.niehs.nih.gov/pubhealth/evalatm/tox21-support/endocrine-disruptors/edhts.html).

The uterotrophic bioassay [Test Guideline (TG) 440] was validated by the OECD as a short-term screening test to evaluate the

CRITICAL REVIEWS IN TOXICOLOGY, 2018
VOL. 48, NO. 5, 359–374
https://doi.org/10.1080/10408444.2018.1429386

REVIEW ARTICLE

Non-animal methods to predict skin sensitization (II): an assessme approaches**

Nicole C. Kleinstreuer[a], Sebastian Hoffmann[b], Nathalie Alépée[c], David Allen[d], Takao Ashikaga Elodie Clouet[f], Magalie Cluzel[g], Bertrand Desprez[h], Nichola Gellatly[i], Carsten Göbel[j], Petra S. Martina Klaric[k], Jochen Kühnl[l], Silvia Martinozzi-Teissier[c], Karsten Mewes[m], Masaaki Miyazawa Erwin van Vliet[n], Qingda Zang[d] and Dirk Petersohn[m]

[a]NIH/NIEHS/DNTP/NICEATM, Research Triangle Park, NC, USA; [b]SEH Consulting + Services, Paderborn, Germany; [c]L'Or Aulnay-sous-Bois, France; [d]ILS, Research Triangle Park, NC, USA; [e]Shiseido, Yokohama-shi, Kanagawa, Japan; [f]Pierre Fa [g]LVMH, St Jean de Braye, France; [h]Cosmetics Europe, Brussels, Belgium; [i]Unilever, London, UK; [j]Coty, Darmstadt, Germ Services Company NV, Strombeek-Bever, Belgium; [k]Beiersdorf AG, Hamburg, Germany; [m]Henkel AG & Co. KGaA, Düss Corporation, Haga, Tochigi, Japan; [n]Services & Consultations on Alternative Methods (SeCAM), Magliaso, Switzerland

ABSTRACT

Skin sensitization is a toxicity endpoint of widespread concern, for which the mechanistic understanding and concurrent necessity for non-animal testing approaches have evolved to a critical juncture, with many available options for predicting sensitization without using animals. Cosmetics Europe and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods collaborated to analyze the performance of multiple non-animal data integration approaches for the skin sensitization safety assessment of cosmetics ingredients. The Cosmetics Europe Skin Tolerance Task Force (STTF) collected and generated data on 128 substances in multiple *in vitro* and *in chemico* skin sensitization assays selected based on a systematic assessment by the STTF. These assays, together with certain *in silico* predictions, are key components of various non-animal testing strategies that have been submitted to the Organization for Economic Cooperation and Development as case studies for skin sensitization. Curated murine local lymph node assay (LLNA) and human skin sensitization data were used to evaluate the performance of six defined approaches, comprising eight non-animal testing strategies, for both hazard and potency characterization. Defined approaches examined included consensus methods, artificial neural networks, support vector machine models, Bayesian networks, and decision trees, most of which were reproduced using open source software tools. Multiple non-animal testing strategies incorporating *in vitro*, *in chemico*, and *in silico* inputs demonstrated equivalent or superior performance to the LLNA when compared to both animal and human data for skin sensitization.

Table of contents

Introduction ... 360
Methods ... 360
Selection of defined approaches ... 360
Qualitative evaluation criteria ... 361
Database ... 361
Predictive performance assessment ... 361
Results ... 362
Qualitative evaluation ... 362
BASF "2 out of 3" ... 363
RIVM STS ... 363
DuPont IATA-SS ... 363
L'Oreal Stacking Meta-model ... 364
ICCVAM SVM ... 364
JRC CCT ...
Givaudan ITS ...
Shiseido ANN-EC3 ...
Procter & Gamble BN-ITS 3 ...
Kao STS ...
Kao ITS ...
Unilever SARA ...
Summary of qualitative evaluatio
LLNA and human reference data
Quantitative performance assess
BASF "2 out of 3" ...
Kao STS ...
Kao ITS ...
ICCVAM SVM ...

CONTACT Nicole C. Kleinstreuer nicole.kleinstreuer@nih.gov NIH/NIEHS/DNTP/NICEATM, P.O. Box 12233, Mail Stop K2 27709, USA
*Present Address: Japanese Center for the Validation of Alternative Methods (JaCVAM), National Institute of Health Sciences (N Tokyo, Japan.
**This manuscript is part of a set of papers; the companion manuscript can be found here: https://doi.org/10.1080/10408444.201

Predictive Models for Acute Oral Systemic Toxicity: A Workshop to Bridge the Gap from Research to Regulation

Authors: Kleinstreuer NC, Karmaus A, Mansouri K, Allen D, Fitzpatrick J, Patlewicz G

Acknowledgments: The authors thank the ICCVAM ATWG members and the Predictive Models for Acute Oral Systemic Toxicity Workshop Organizing Committee: D. Asturiol, S. Bell, L. Burgoon, D. Cronce, J. Gearhart, J. Gordon, S. Marty, L. Milchak, E. Odenkirchen, P. Pradeep, L. Scarano, and J. Strickland

Highlights

- Towards implementation of the ICCVAM Strategic Roadmap, a global modeling project was organized to build predictive in silico models for acute oral systemic toxicity.
- An international workshop was held in April 2018 at the NIH to discuss the results of the modeling project, with a diverse group of scientists and stakeholders participating in 2 days of presentations and breakout group discussions.
- Relative strengths and weaknesses of the models for different regulatory purposes were discussed, and recommendations and next steps are presented.

Disclaimer: *The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency or the National Institutes of Health. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.*

Abstract

In early 2018, the Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM) published the "Strategic Roadmap for Establishing New Approaches to Evaluate the Safety of Chemicals and Medical Products in the United States" (ICCVAM 2018). Cross-agency federal workgroups have been established to implement this roadmap for various toxicological testing endpoints, with an initial focus on acute toxicity testing. The ICCVAM acute toxicity workgroup (ATWG) helped organize a global collaboration to build predictive in silico models for acute oral systemic toxicity, based on a large dataset of rodent studies and targeted towards regulatory needs identified across federal agencies. Thirty-two international groups across government, industry, and academia participated in the project, culminating in a workshop in April 2018 held at the National Institutes of Health (NIH). At the workshop, computational modelers and regulatory decision makers met to discuss the feasibility of using predictive

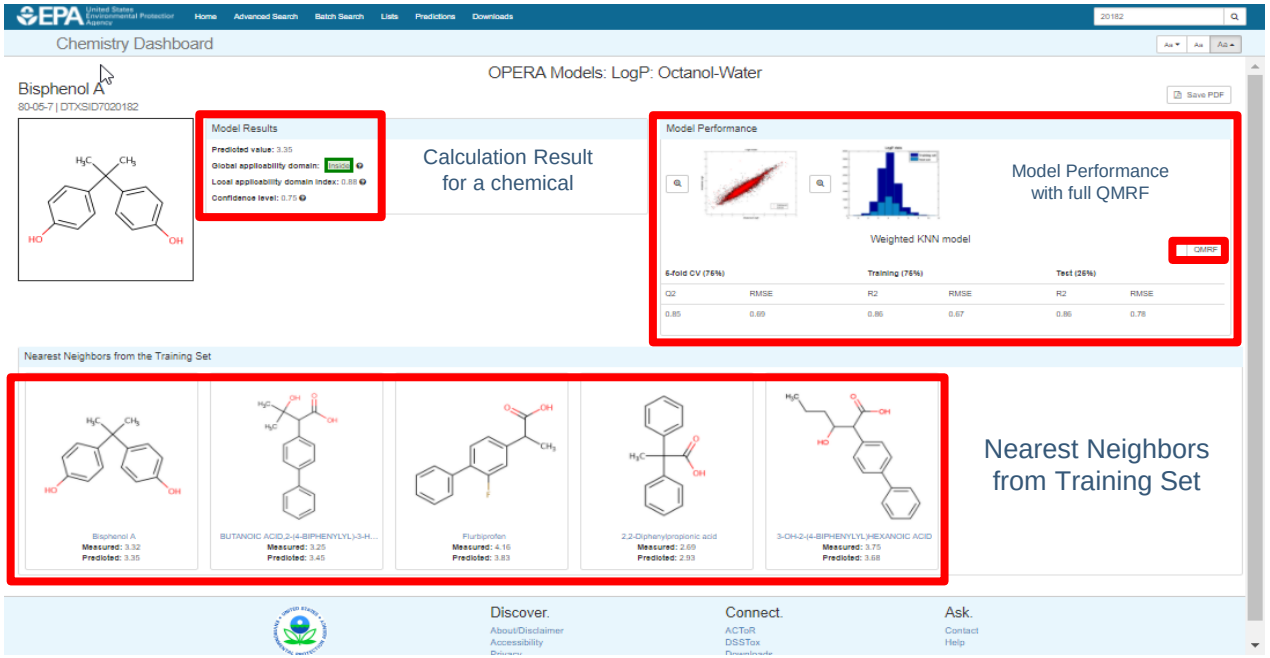
Recent Workshop: Modelers + Regulators

Predictive Models for Acute Oral Systemic Toxicity

William H. Natcher Conference Center
National Institutes of Health, Bethesda, Maryland
April 11 – 12, 2018

Attendees in-person: 89; webcast: 215

Model Accessibility and Transparency

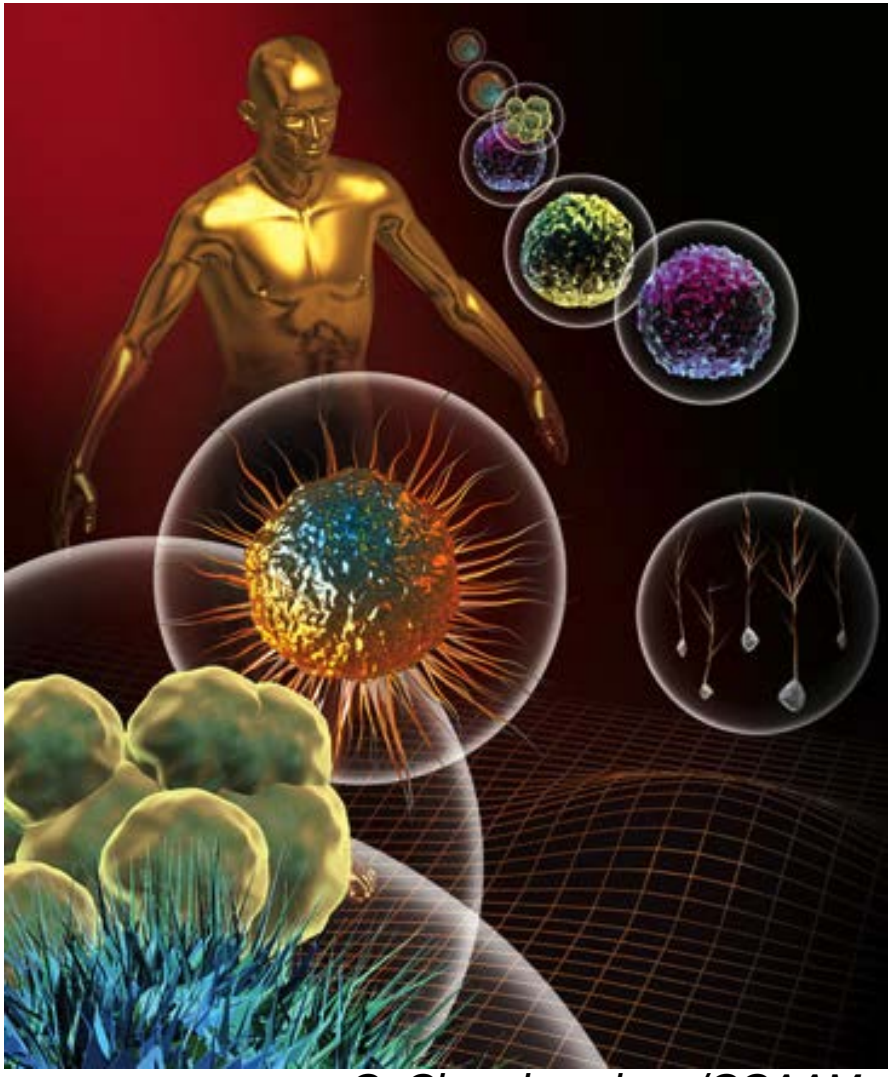


OPERA
OPEn (q)saR App

Mansouri et al. OPERA models (https://link.springer.com/article/10.1186/s13321-018-0263-1)

https://github.com/kmansouri/OPERA

How to Benchmark Alternative Models

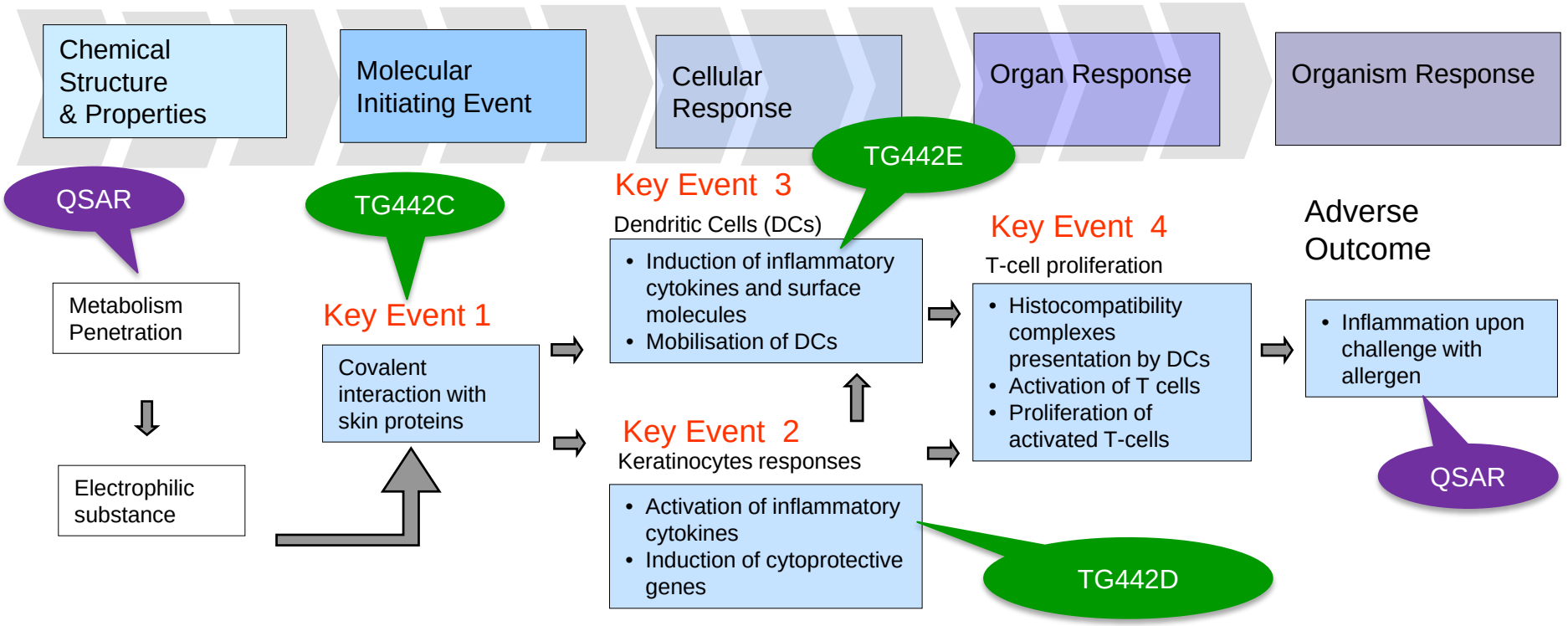


C. Chandrasekera/CCAAM

Human data and human biology as the gold standard

How to Benchmark Alternative Models

Example: Skin Sensitization



Defined Approaches (DAs) combine *in vitro* and *in silico* data using simple decision trees or machine learning algorithms to predict skin sensitization.

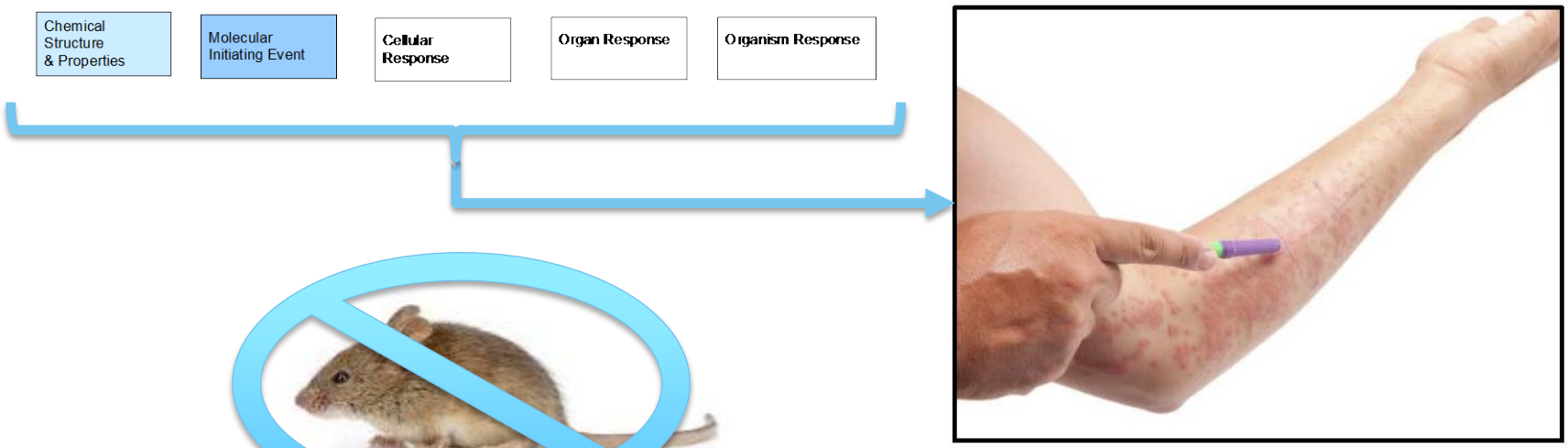
How to Benchmark Alternative Models

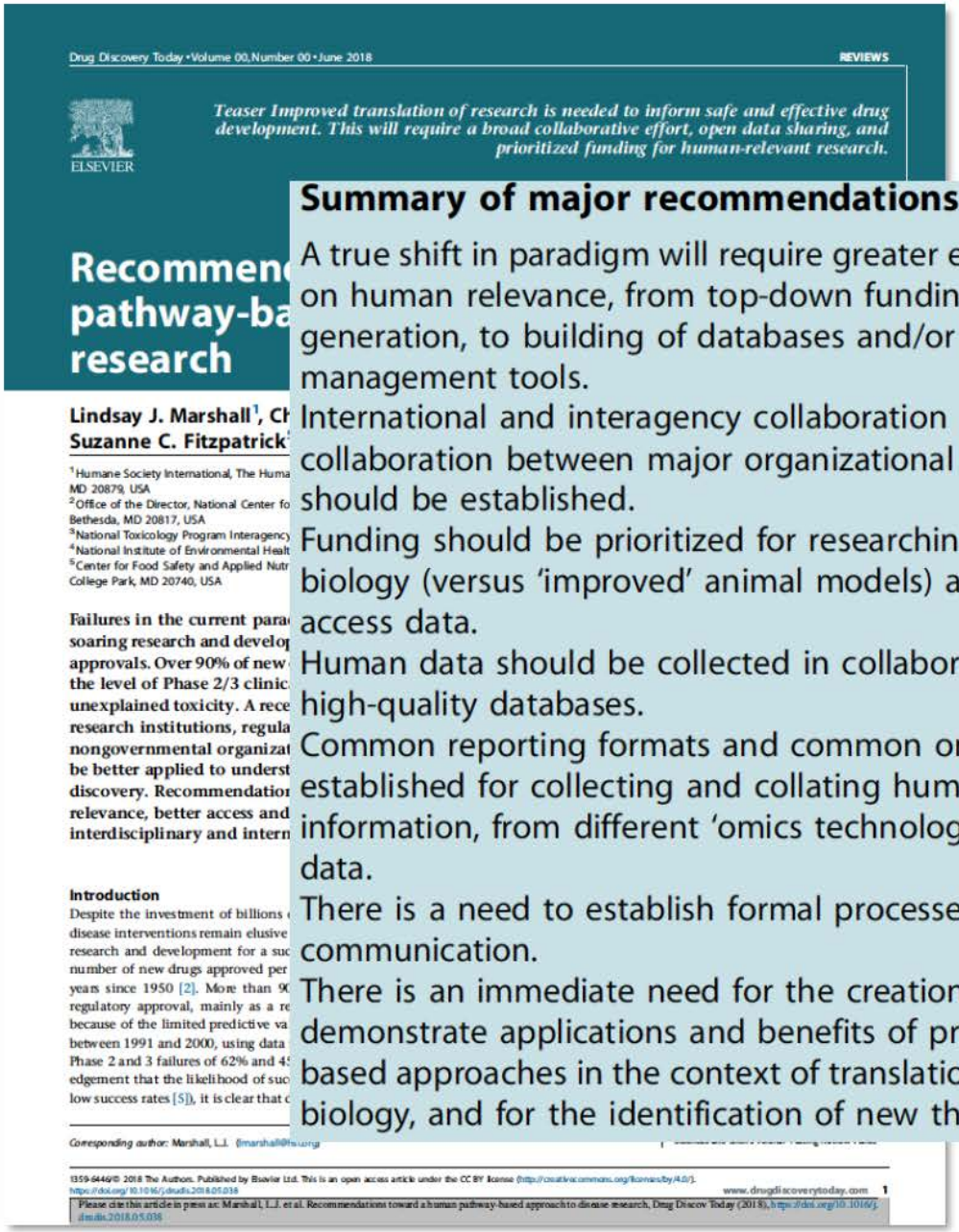
Example: Skin Sensitization

All non-animal defined approaches evaluated perform as well or **better** than the mouse at predicting human skin sensitization:

Hazard: 74% (mouse) vs. 75-85% (DAs)

3-class Potency: 59% (mouse) vs. 55-69% (DAs)





Drug Discovery Today • Volume 00, Number 00 • June 2018 REVIEWS

ELSEVIER

Teaser Improved translation of research is needed to inform safe and effective drug development. This will require a broad collaborative effort, open data sharing, and prioritized funding for human-relevant research.

Recommen
pathway-ba
research

Lindsay J. Marshall[1], Ch
Suzanne C. Fitzpatrick[
[1]Humane Society International, The Huma
MD 20879, USA
[2]Office of the Director, National Center fo
Bethesda, MD 20817, USA
[3]National Toxicology Program Interagency
[4]National Institute of Environmental Healt
[5]Center for Food Safety and Applied Nutr
College Park, MD 20740, USA

**Failures in the current para
soaring research and develop
approvals. Over 90% of new
the level of Phase 2/3 clinic
unexplained toxicity. A rece
research institutions, regula
nongovernmental organizat
be better applied to underst
discovery. Recommendatio
relevance, better access and
interdisciplinary and intern**

Introduction

Despite the investment of billions
disease interventions remain elusive
research and development for a suc
number of new drugs approved per
years since 1950 [2]. More than 9
regulatory approval, mainly as a re
because of the limited predictive va
between 1991 and 2000, using data
Phase 2 and 3 failures of 62% and 4
edgement that the likelihood of suc
low success rates [5]), it is clear that

Corresponding author: Marshall, L.J. (lmarshall@hsi.org

https://doi.org/10.1016/j.drudis.2018.05.038

Please cite this article in press as: Marshall, L.J. et al. Recommendations toward a human pathway-based approach to disease research, Drug Discov Today (2018), https://doi.org/10.1016/j.drudis.2018.05.038

Summary of major recommendations

A true shift in paradigm will require greater emphasis to be placed on human relevance, from top-down funding decisions to data generation, to building of databases and/or knowledge management tools.

International and interagency collaboration is critical: formal collaboration between major organizational and funding bodies should be established.

Funding should be prioritized for researching human-based biology (versus 'improved' animal models) and promoting open access data.

Human data should be collected in collaborative, open-access high-quality databases.

Common reporting formats and common ontologies should be established for collecting and collating human biology information, from different 'omics technologies to human clinical data.

There is a need to establish formal processes for cross-sector communication.

There is an immediate need for the creation of case studies to demonstrate applications and benefits of predictive, mechanism-based approaches in the context of translation and human disease biology, and for the identification of new therapeutics.

Example: Eye Irritation

OECD/OCDE **492**

Adopted:
25 June 2018

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Reconstructed human Cornea-like Epithelium (RhCE) test method for identifying chemicals not requiring classification and labelling for eye irritation or serious eye damage

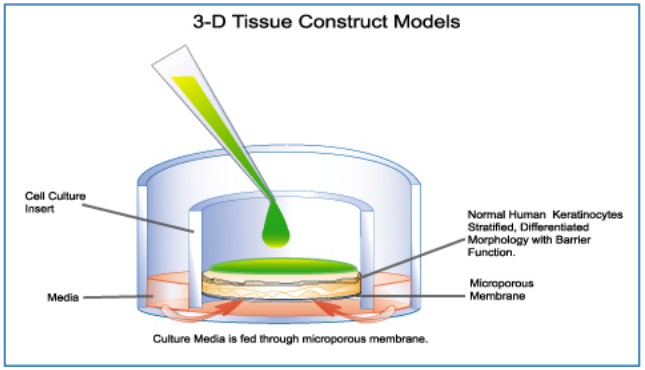
INTRODUCTION

1. *Serious eye damage* refers to the production of tissue damage in the eye, or serious physical decay of vision, following application of a test chemical to the anterior surface of the eye, which is not fully reversible within 21 days of application, as defined by the United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS) (1). Also according to UN GHS, *eye irritation* refers to the production of changes in the eye following the application of a test chemical to the anterior surface of the eye, which are fully reversible within 21 days of application. Test chemicals inducing serious eye damage are classified as UN GHS Category 1, while those inducing eye irritation are classified as UN GHS Category 2. Test chemicals not classified for eye irritation or serious eye damage are defined as those that do not meet the requirements for classification as UN GHS Category 1 or 2 (2A or 2B) i.e., they are referred to as UN GHS No Category.

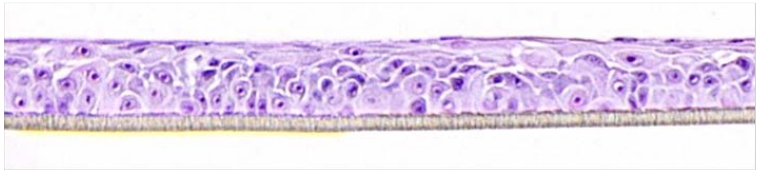
2. The assessment of serious eye damage/eye irritation has typically involved the use of laboratory animals (OECD Test Guideline (TG) 405; adopted in 1981 and revised in 1987, 2002, 2012 and 2017) (2). The choice of the most appropriate test method and the use of this Test Guideline should be seen in the context of the OECD Guidance Document on an Integrated Approaches on Testing and Assessment (IATA) for Serious Eye Damage and Eye irritation (3).

3. This Test Guideline describes an *in vitro* procedure allowing the identification of chemicals (substances and mixtures) not requiring classification and labelling for eye irritation or serious eye damage in accordance with UN GHS. It makes use of reconstructed human cornea-like epithelium (RhCE) which closely mimics the histological, morphological, biochemical and physiological properties of the human corneal epithelium. Four other *in vitro* test methods have been validated, considered scientifically valid and

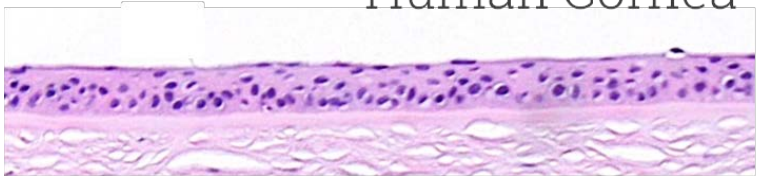
In accordance with the Decision of the Council on a Delegation of Authority to amend Annex I of the Decision of the Council on the Mutual Acceptance of Data in the Assessment of Chemicals [C(2018)49], this Guideline was amended by the OECD's Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology by written procedure on 25 June 2018. guideline was adopted by the OECD Council by written procedure on 25 June 2018.



EpiOcular



Human Cornea



Eye Irritation Classification:

OECD TG 492 Proficiency Chemicals vs. Sigma SDS

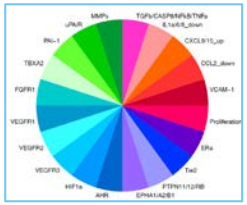
Chemical Name	CASRN	OECD TG 492 (in vivo data)	SDS (in vivo data)
Methythioglycolate	2365-48-2	Category 1	Category 2A
2,5-Dimethyl-2,5-hexanediol	110-03-2	Category 1	Not classified
1-Ethyl-3-methylimidazolium ethylsulphate	342573-75-5	Not Classified	Category 1
Diethyl toluamide	134-62-3	Category 2B	Category 2A
Camphene	79-92-5	Category 2B	Category 2A

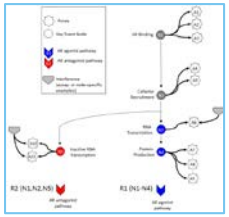
Mechanistic Mapping of HTS Assays

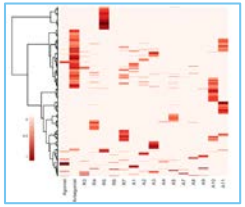
Example: Developmental Toxicity

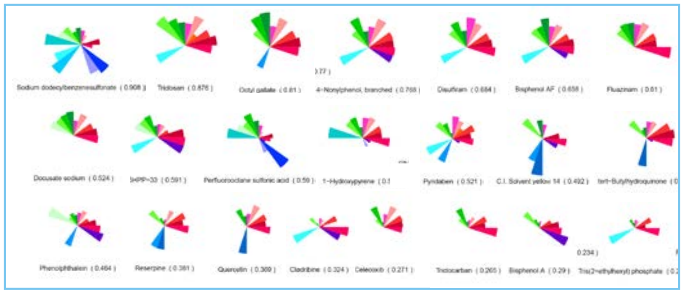
Human Teratogenic Mechanisms

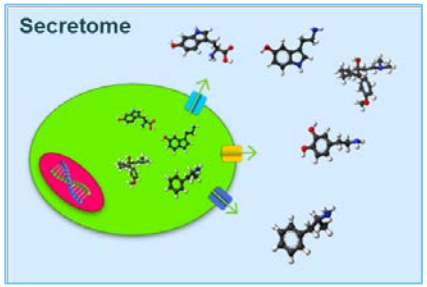
- Endocrine disruption
- Oxidative stress
- Vascular disruption
- Folate antagonism
- Neural crest cell disruption
- Specific receptor- or enzyme-mediated











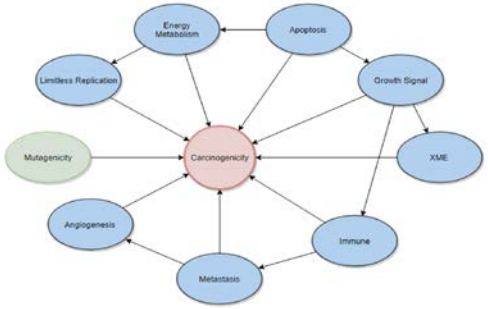
Van Gelder et al. 2010; Knudsen and Kleinstreuer 2011

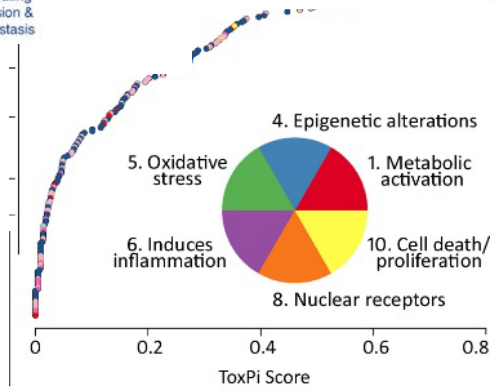
Mechanistic Mapping of HTS Assays

Example: Carcinogenicity

Hallmarks of Cancer & Characteristics of Carcinogens

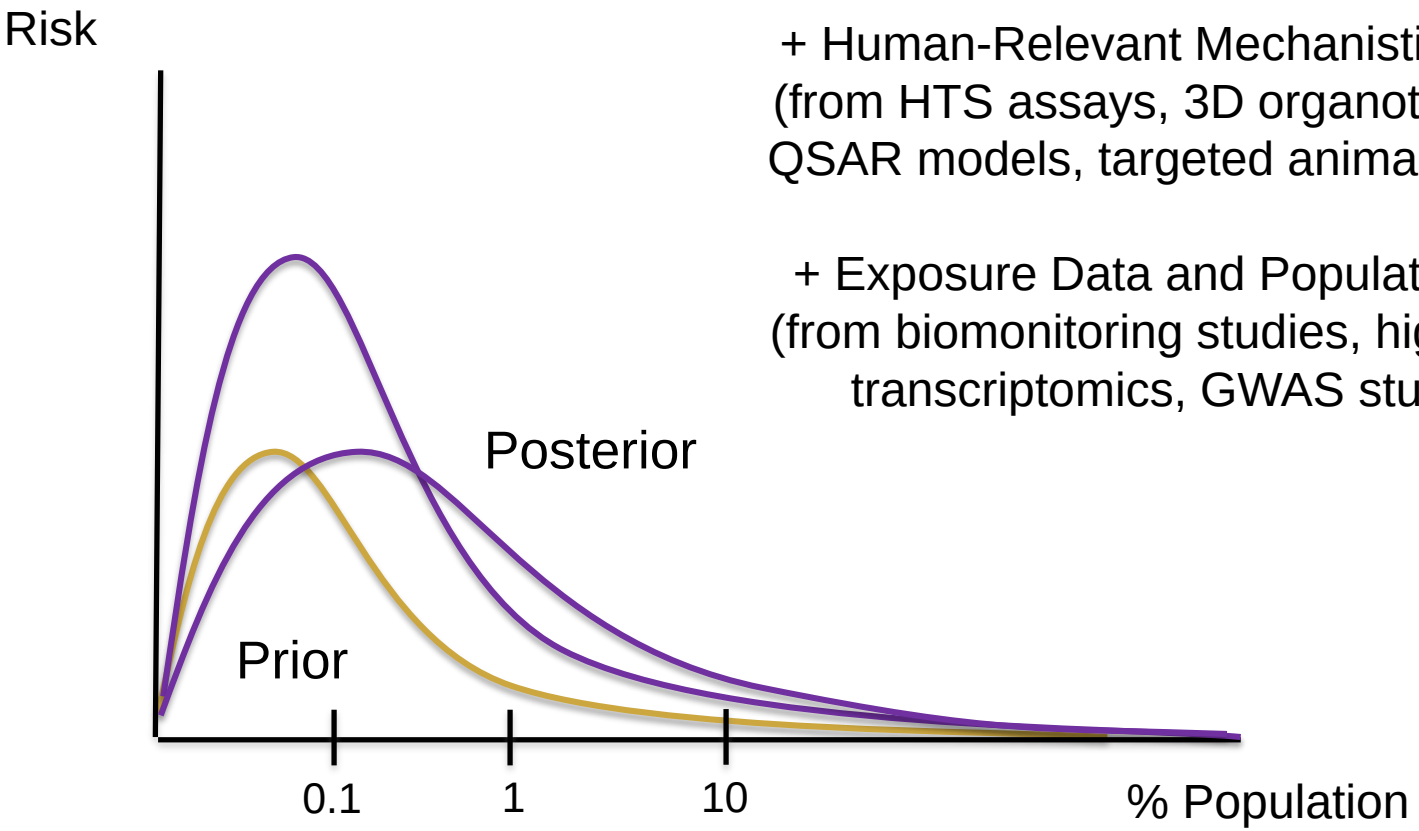
- Inflammation
- Oxidative stress
- Genotoxicity/instablitiy
- Angiogenesis
- Immortalization/proliferation
- Immunosuppression
- Invasion/metastasis
- Specific receptor- or enzyme-mediated





Hanahan & Weingberg 2011; Smith et al. 2016; Guyton et al. 2018; Chiu et al. 2018

Addressing Risk Probabilistically



+ Human-Relevant Mechanistic Information (from HTS assays, 3D organotypic systems, QSAR models, targeted animal studies, etc.)

+ Exposure Data and Population Genetics (from biomonitoring studies, high-throughput transcriptomics, GWAS studies, etc.)

Challenges

- Scientific
  - Considering population/genetic variability
  - Incorporating metabolic competence
  - Developing complex systems models
  - Reporting and collection of reference data
- Non-scientific
  - Increasing awareness, education, and training
  - Cross-sector communication
  - Funding for human-centric research and education