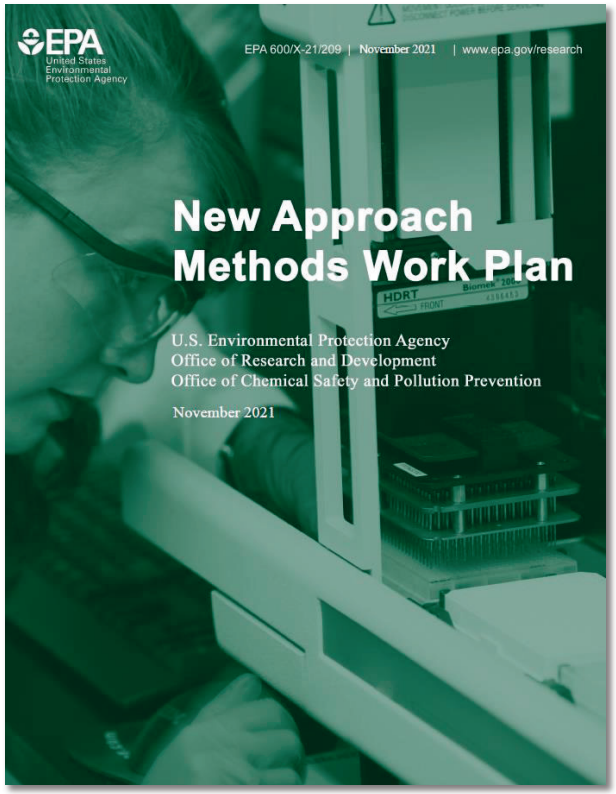
Updated NAM Work Plan Identified Objectives, Strategies and Deliverables for Applying NAMs



- Five objectives for reducing animal testing and research while ensuring that Agency decisions remain fully protective of human health and the environment
  - Evaluate Regulatory Flexibility
  - Develop Baselines and Metrics
  - Establish Scientific Confidence and Demonstrate Application
  - Develop NAMs to Address Information Gaps
  - Engage and Communicate with Stakeholders
- Updated NAM Work Plan released in December 2021
  - Expansion of the species covered in the work plan to include all vertebrate animals to be consistent with TSCA.
  - Modified deliverable timelines that reflect the expansion of covered species and incorporate feedback received over the preceding years.
  - New case studies for building confidence and demonstrating application of NAMs.
  - A pilot study to develop NAMs training courses and materials.

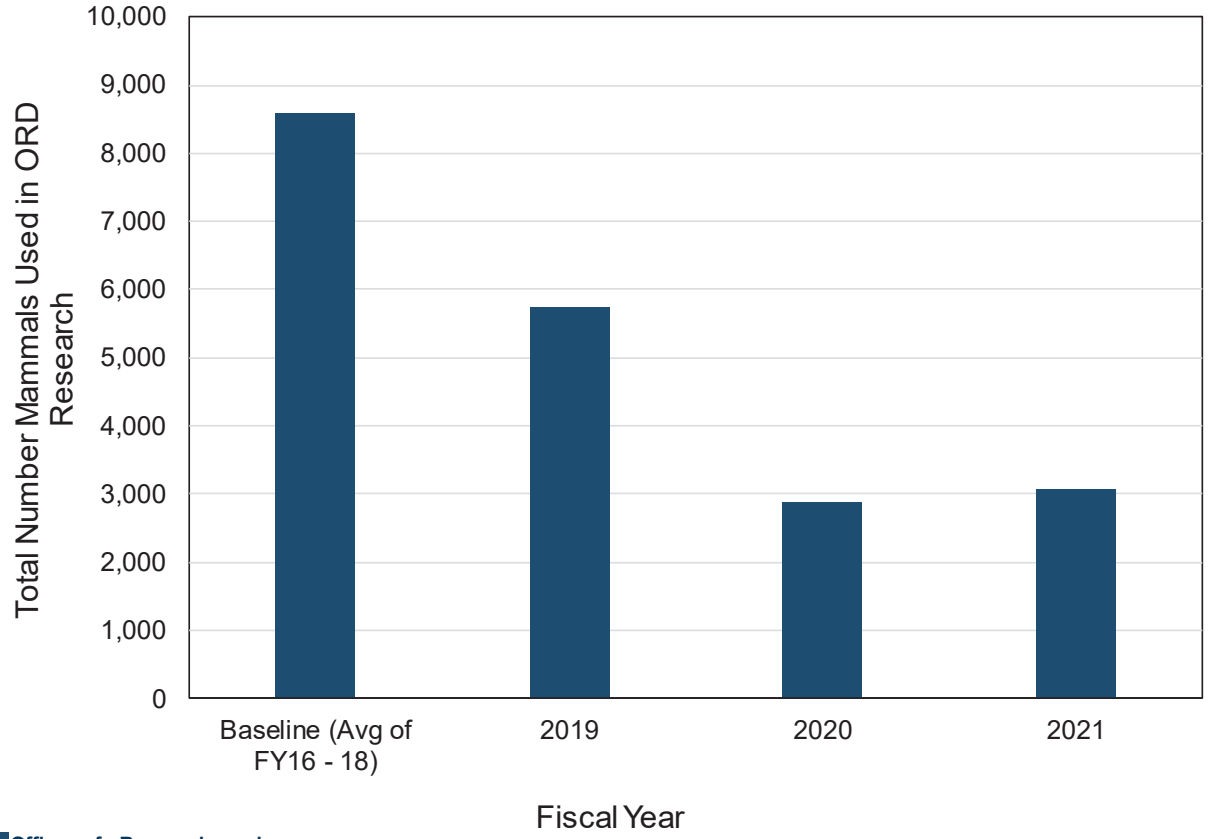
EPA United States Environmental Protection Agency

Status of NAM Work Plan Deliverables

Milestones/Deliverables	Proposed Dates
Evaluate Regulatory Flexibility for Accommodating the Use of NAMs	
EPA report on a review of existing statutes, programmatic regulations, policies, and guidance that relate to vertebrate animal testing and the implementation and use of appropriate NAMs for regulatory purposes.	2022
Develop Baselines and Metrics for Assessing Progress	
Progress and summary metrics on reducing vertebrate animal testing requests and use.	Annually starting in Q4 2022
Establish Scientific Confidence in NAMs and Demonstrate Application to Regulatory Decisions	
U.S. National Academies of Sciences, Engineering, and Medicine study that evaluates the variability and relevance of existing mammalian toxicity tests and reviews frameworks for validation and establishing scientific confidence in testing methods. The study is funded by the EPA, but the timing is determined by the National Academies.	2023
A scientific confidence framework to evaluate the quality, reliability, and relevance of NAMs.	Q4 2024
An initial set of reporting templates which may be used by EPA and stakeholders that capture the range of specific NAMs used for Agency decisions.	Q4 2024
Case studies for evaluating application to risk assessment and demonstrating protection of human health and the environment.	Ongoing

FY19 – FY21 Animal Use Metrics for ORD

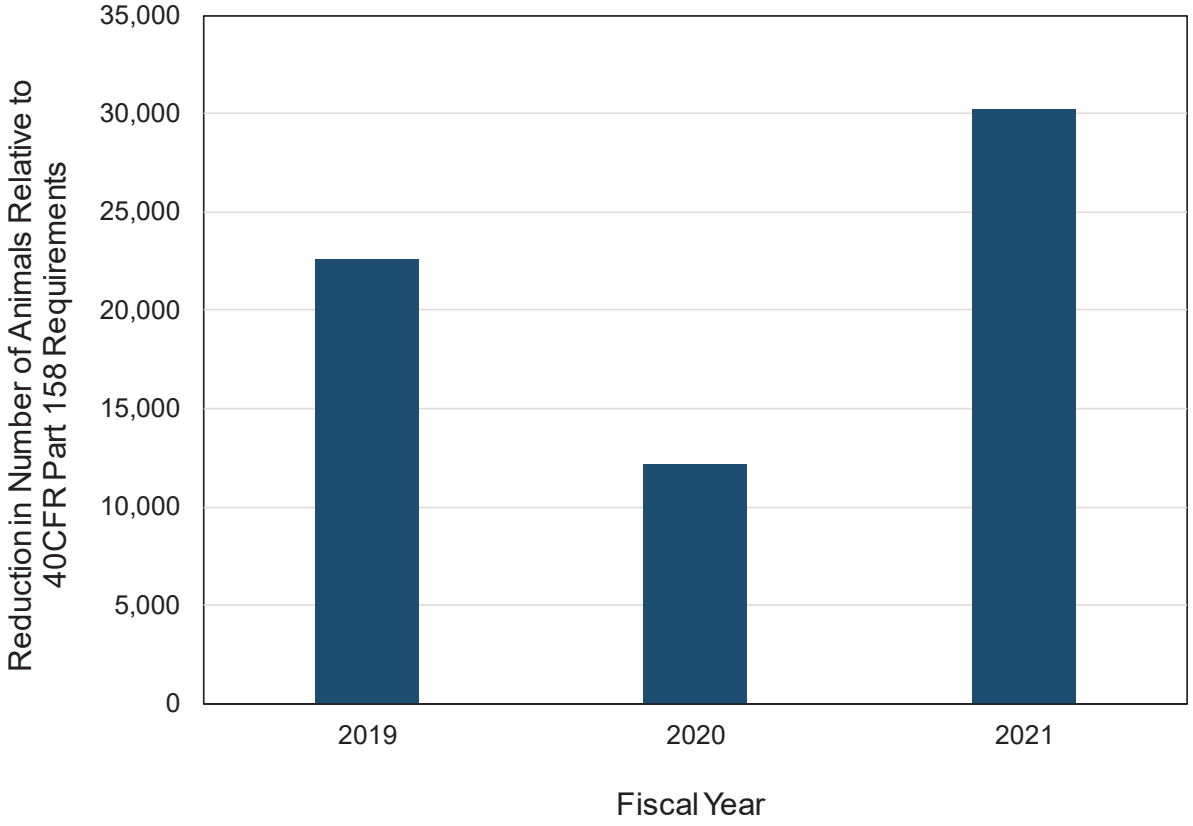
Milestone/Deliverable: Progress and summary metrics on reducing vertebrate animal testing requests and use. (FY22+).



- The numbers in FY19 – 21 include those mammals used in contract research activities.
- Baseline numbers (FY16 – 18) do not include mammals used in contract research activities due to a lack of tracking at that time.
- The numbers in FY19 are likely reduced due to impacts of the ORD reorganization and lab remodeling.
- The numbers in FY20 – 21 are likely reduced due of the impact of the pandemic on research activities.

FY19 – FY21 Animal Reduction Metrics for OPP

Milestone/Deliverable: Progress and summary metrics on reducing vertebrate animal testing requests and use. (FY22+).



- The reduction in the number of animals were due to Hazard and Science Policy Council (HASPOC), Chemistry and Acute Toxicology Science Advisory Council (CATSAC), and Acute Dermal waivers.
- Acute dermal waivers granted specifically under the updated dermal waiver polices (2016/2020).
- The total number waivers granted from FY19 – 21 were:
  - HASPOC - 163
  - CATSAC - 54
  - Acute Dermal - 123
- The number of NAM-related endpoint data submissions from FY19 – 21 were:
  - Eye Irritation - 57
  - Skin Irritation - 42
  - Skin Sensitization - 15

EPA United States Environmental Protection Agency

Status of NAM Work Plan Deliverables

Milestones/Deliverables	Proposed Dates
Develop NAMs to Address Scientific Challenges and Fill Important Information Gaps	
EPA Strategic Research Action Plans outlining research products to develop and apply NAMs.	Q1 2023 ✓
Encourage development of NAMs through mechanisms such as the STAR program and facilitate partnerships with organizations focused on establishing scientific confidence in alternative methods.	Ongoing ✓
Engage and Communicate with Stakeholders	
EPA website to house information about NAM efforts and progress being upon release of the work plan.	2020 ✓
Public webinars and, where appropriate, peer-review on deliverables from this work plan.	Ongoing ✓
Complete NAMs pilot training program in the fourth quarter (Q4) of 2023 and provide regular scientific exchanges and progress updates through Agency sponsored and partner organized events.	Q4 2023 and Ongoing ✓

EPA Research Planning

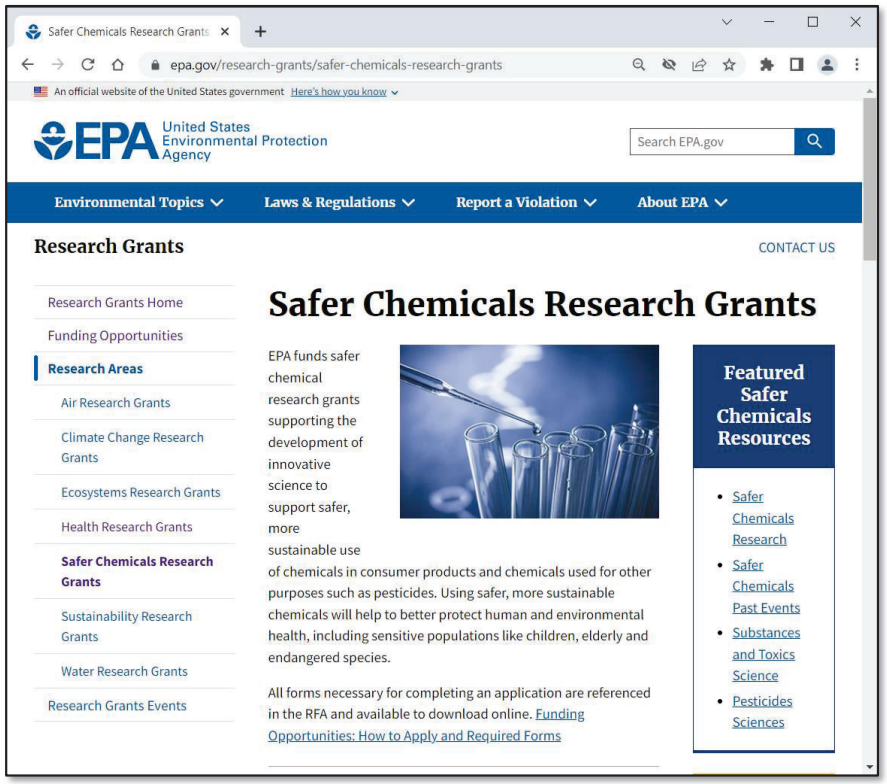
Milestone/Deliverable: EPA Strategic Research Action Plans outlining research products to develop and apply NAMs. (2023).

https://www.epa.gov/research/strategic-research-planning

- FY23 – 26 Strategic Research Action Plans (StRAPs) released outlining the next four years of ORD research activities
- More than 100 research products directly related to research on NAM development and application
  - Human Health Toxicity-related NAMs
  - Ecological Toxicity-related NAMs
  - Toxicokinetic-related NAMs
  - Case Studies
  - OPPT New Chemicals Research Program
  - Communication and Training
- Many other research products indirectly supporting NAM development and application (e.g., development of databases and tools).

EPA STAR Grants

Milestone/Deliverable: Encourage development of NAMs through mechanisms such as the STAR program and facilitate partnerships with organizations focused on establishing scientific confidence in alternative methods. (Ongoing).

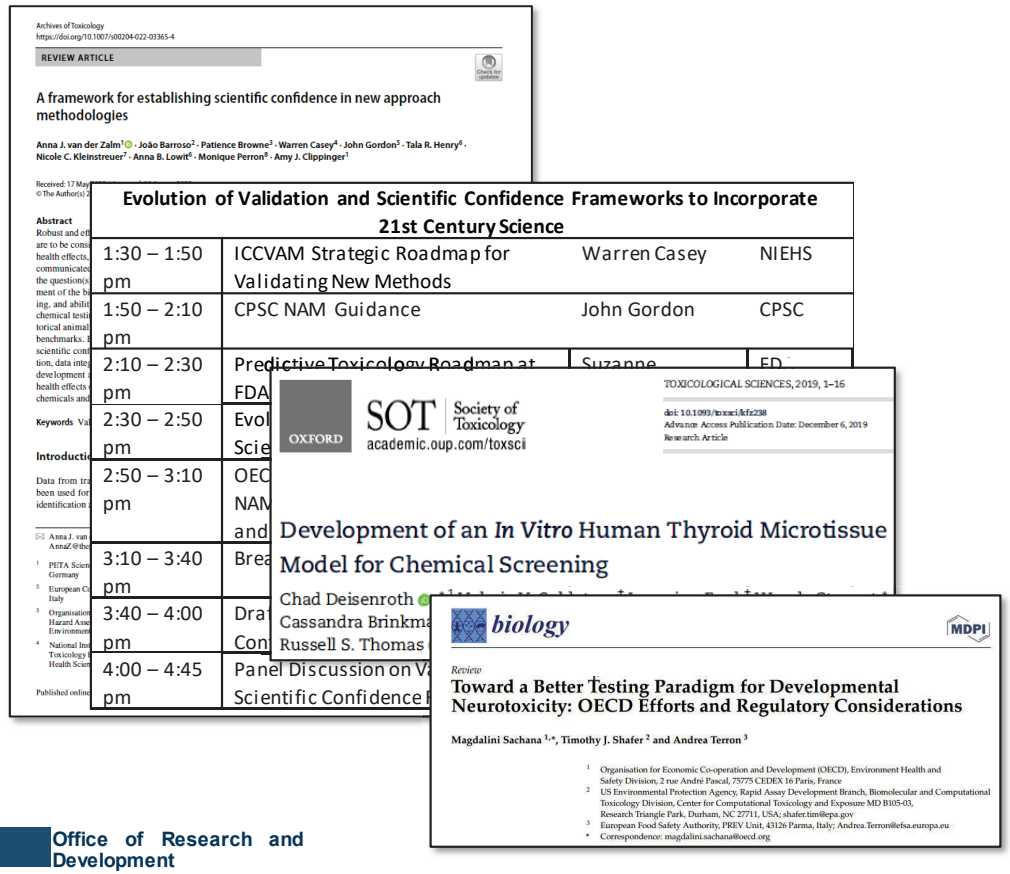


https://www.epa.gov/research-grants/star

- EPA STAR grants on *Advancing Actionable Alternatives to Vertebrate Animal Testing for Chemical Safety Testing (2019-22/24)*
  - Awarded ~$4.2 million to 5 universities
  - Vanderbilt University, University of California Riverside, Louisiana State University, Oregon State University, Johns Hopkins University
- EPA STAR grants on *Advancing Toxicokinetics for Efficient and Robust Chemical Evaluations (2020 – 2023)*
  - Awarded ~$4 million to 5 institutions
  - Purdue University, Woods Hole Oceanographic Institution, Vanderbilt University, Texas A&M, and University of Nevada Reno
- EPA STAR grants on *Development of Innovative Approaches to Assess the Toxicity of Chemical Mixtures (2023-26)* – Coming Soon!

Partnerships with External Organizations Focused on Scientific Confidence

Milestone/Deliverable: Encourage development of NAMs through mechanisms such as the STAR program and facilitate partnerships with organizations focused on establishing scientific confidence in alternative methods. (Ongoing).



Archives of Toxicology
https://doi.org/10.1007/s00204-022-03365-4

REVIEW ARTICLE

A framework for establishing scientific confidence in new approach methodologies

Anna J. van der Zalm[1] · João Barroso[2] · Patience Browne[3] · Warren Casey[4] · John Gordon[5] · Tala R. Henry[6] · Nicole C. Kleinstreuer[7] · Anna B. Lowit[8] · Monique Perron[8] · Amy J. Clippinger[1]

Evolution of Validation and Scientific Confidence Frameworks to Incorporate 21st Century Science

1:30 – 1:50 pm	ICCVAM Strategic Roadmap for Validating New Methods	Warren Casey	NIEHS
1:50 – 2:10 pm	CPSC NAM Guidance	John Gordon	CPSC
2:10 – 2:30 pm	Predictive Toxicology Roadmap at FDA	Suzanne	FD
2:30 – 2:50 pm	Evol Scie		
2:50 – 3:10 pm	OEC NAM and		
3:10 – 3:40 pm	Brea		
3:40 – 4:00 pm	Drat Con		
4:00 – 4:45 pm	Panel Discussion on V Scientific Confidence F		

OXFORD SOT Society of Toxicology
academic.oup.com/toxsci

TOXICOLOGICAL SCIENCES, 2019, 1–16
doi: 10.1093/toxsci/kfz238
Advance Access Publication Date: December 6, 2019
Research Article

Development of an *In Vitro* Human Thyroid Microtissue Model for Chemical Screening

Chad Deisenroth
Cassandra Brinkma
Russell S. Thomas

biology MDPI

Review

Toward a Better Testing Paradigm for Developmental Neurotoxicity: OECD Efforts and Regulatory Considerations

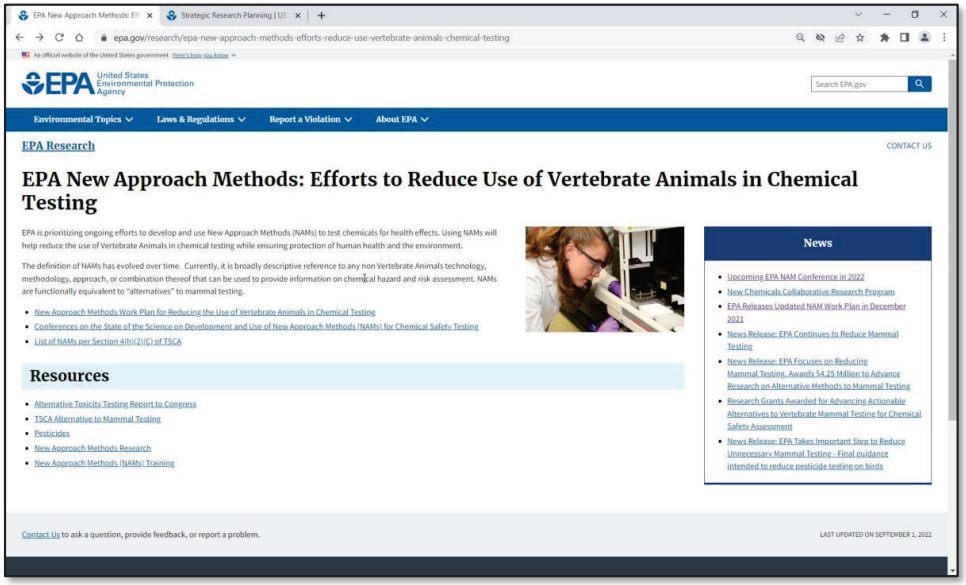
Magdalini Sachana [1,*], Timothy J. Shafer [2] and Andrea Terron [3]

[1] Organisation for Economic Co-operation and Development (OECD), Environment Health and Safety Division, 2 rue André Pascal, 75775 CEDEX 16 Paris, France
[2] US Environmental Protection Agency, Rapid Assay Development Branch, Biomolecular and Computational Toxicology Division, Center for Computational Toxicology and Exposure MD B105-03, Research Triangle Park, Durham, NC 27711, USA; shafer.tim@epa.gov
[3] European Food Safety Authority, PREV Unit, 43126 Parma, Italy; Andrea.Terron@efsa.europa.eu
* Correspondence: magdalini.sachana@oecd.org

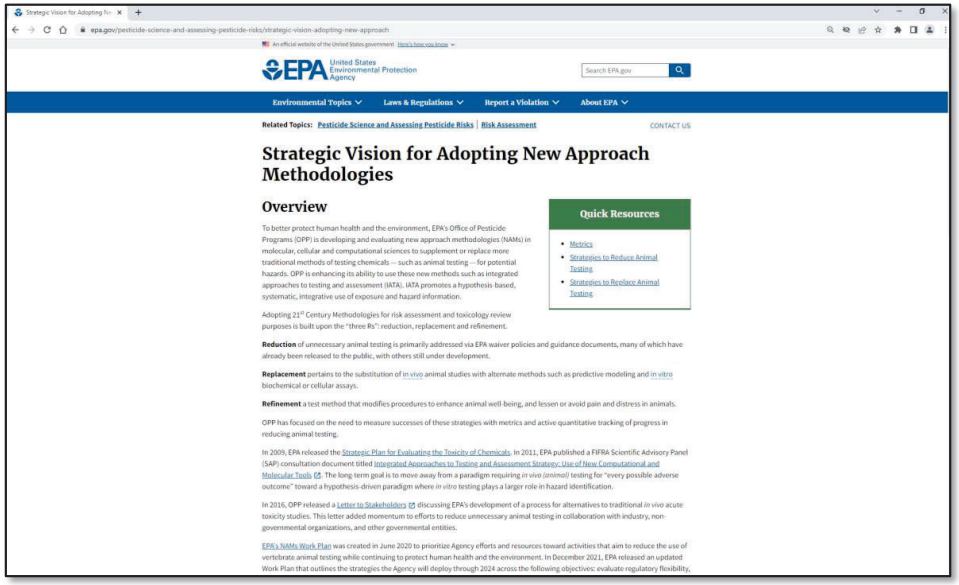
- EPA partnered with 5 national and international organizations to develop a framework for establishing scientific confidence in NAMs (Zalm *et al.*, Arch Toxicol., 2022).
- Session in this EPA NAM Conference to discuss experiences with validation and establishing scientific confidence.
- Partnering with 4 external organizations on an inter-laboratory prevalidation study of a human thyroid microtissue assay.
- Partnering with 5 external organizations on the development and validation of 17 assays for developmental neurotoxicity.

EPA NAM Websites

Milestone/Deliverable: EPA website to house information about NAM efforts and progress being upon release of the work plan. (2020).



https://www.epa.gov/nam



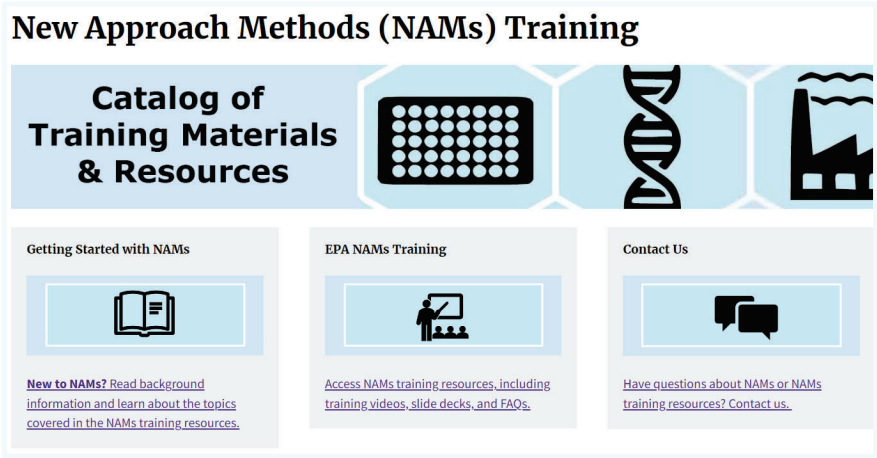
https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/strategic-vision-adopting-new-approach

EPA NAM Pilot Training Program and Regular Scientific Exchanges and Progress Updates

Milestone/Deliverable: Complete NAMs pilot training program in the fourth quarter (Q4) of 2023 and provide regular scientific exchanges and progress updates through Agency sponsored and partner organized events. (2023 and Ongoing).

Progress

- Public NAMs training website released to serve as a resource for training materials and recordings for EPA tools and databases that contribute to NAMs research (May 2022)
- Interactive training on ECOTOX Knowledgebase (May 2022, 350+ attendees)
- New NAMs Update email bulletin established to share progress and updates
- Two-way communication via NAM@epa.gov
- Upcoming:
  - October 18, 2022: Interactive training on CompTox Chemicals Dashboard (1100+ registrants)
  - Spring 2023: Interactive training on Generalized Read-Across (GenRA)



https://www.epa.gov/chemical-research/new-approach-methods-nams-training

EPA United States Environmental Protection Agency

Goals For The Meeting

- First and foremost… Enjoy the meeting, seeing colleagues again, the great science that is going to be presented, and the subsequent discussions.
- Upcoming NAM Work Plan deliverables are focused on variability and relevance of current animal models and development of an Agency-wide scientific confidence framework for NAMs. We would like to stimulate a deeper discussion in the community on –
  - Generalizable conclusions from the studies evaluating the variability and inter-species concordance of laboratory mammalian toxicity studies and implications for NAMs.
  - Conservation of mode-of-action between the animal toxicity testing models and humans in a risk assessment context and opportunities for NAMs.
  - Concordance between laboratory mammalian models and humans in the adverse effects following chemical exposure and implications for NAMs.
  - Key components in a fit-for-purpose validation paradigm or scientific confidence framework for NAMs.

Variability of Chronic Rodent Bioassays

Christoph Helma

October 12, 2022

Content

Rodent Carcinogenicity

E Gottmann, S Kramer, B Pfahringer and C Helma

Data quality in predictive toxicology: reproducibility of rodent carcinogenicity experiments
Environ Health Perspect 109:509–514 (2001)
https://doi.org/10.1289/ehp.01109509

Lowest observed adverse effect level (LOAEL)

C Helma, D Vorgrimmler, D Gebele, M Gütlein, B Engeli, J Zarn, B Schilter and E Lo Piparo

Modeling Chronic Toxicity: A Comparison of Experimental Variability With (Q)SAR/Read-Across Predictions
Front Pharmacol 9 (2018)
https://doi.org/10.3389/fphar.2018.00413

Carcinogenicity Data

- Carcinogenic Potency Database(CPDB, Gold 1997)
- 1,289 unique compounds
- 2 Subsets
  - *National Toxicology Program (NTP)*
  - *General literature*
- 121 common compounds in both subsets

Carcinogenicity Classification

•**57%** concordant classifications (69/121 compounds, 39 carcinogens, 30 non-carcinogens)

Rats

62% concordant classifications

Mice

49% concordant classifications

Multi species carcinogens

58% concordant classifications

Multi organ carcinogens:

52% concordant classifications

•poor reproducibility of sex, species and organ specific effects

Carcinogenicity TD50's

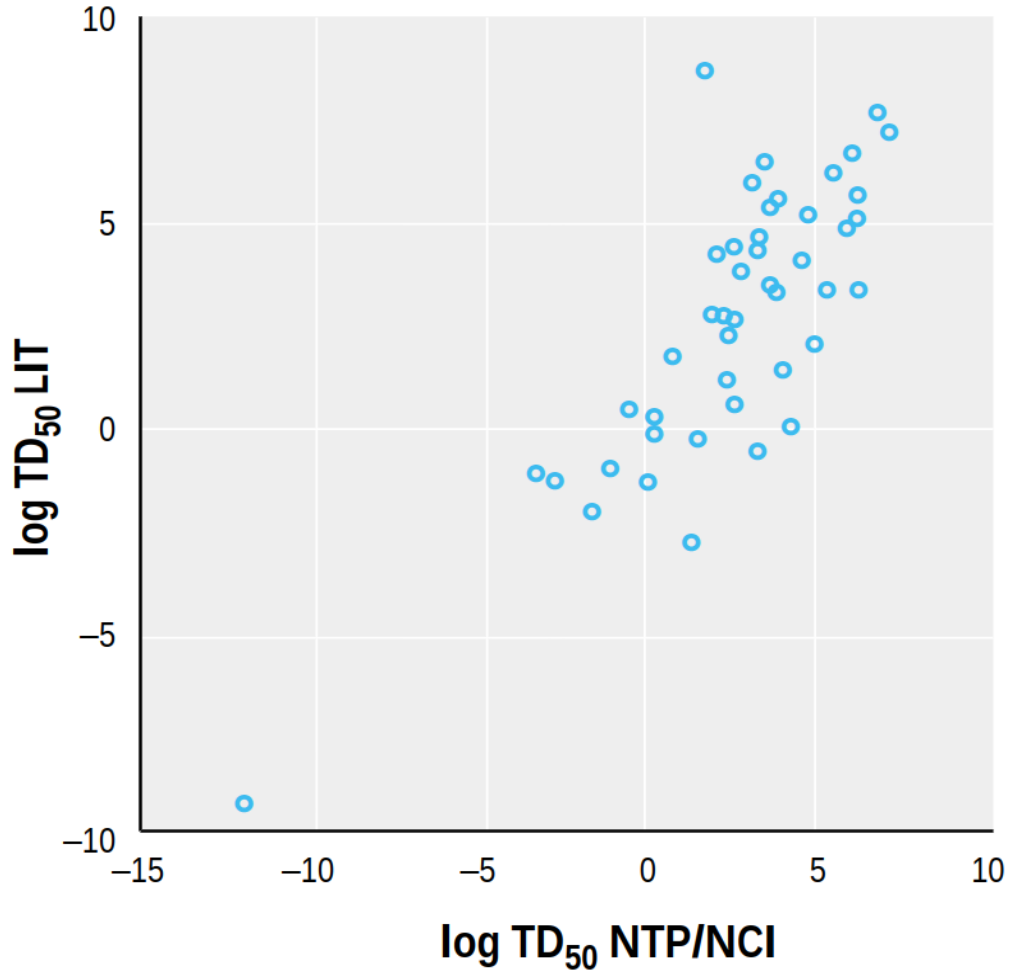


Figure 2. Correlation of carcinogenicity TD_{50} values from the NTP/NCI and the literature (LIT) part of the CPDB ($r^2 = 0.63$).

Carcinogenicity caveats

•low sample size

•no standardized protocols for literature data

Gold et al. (1987)

- 38 compounds from the literature
- 93% reproducibility for rats
- 76% for mice
- 34 studies were published by the same authors (!)

LOAEL Data

Chronic (>180 days) lowest observed effect levels (LOAEL) for rats (Rattus norvegicus) after oral (gavage,

Nestlé Database

567 LOAEL values for 445 unique chemical structures from the literature (Mazzatorta et al., 2008)

Swiss Food Safety and Veterinary Office (FSVO) Database

493 rat LOAEL values for 381 unique chemical structures from pesticide evaluations (Zarn et al., 2011, 201

- European Food Safety Authority (EFSA) (EFSA, 2014)
- Joint FAO/WHO Meeting on Pesticide Residues (JMPR) (WHO, 2011)
- US EPA (US EPA, 2011)

Combined dataset

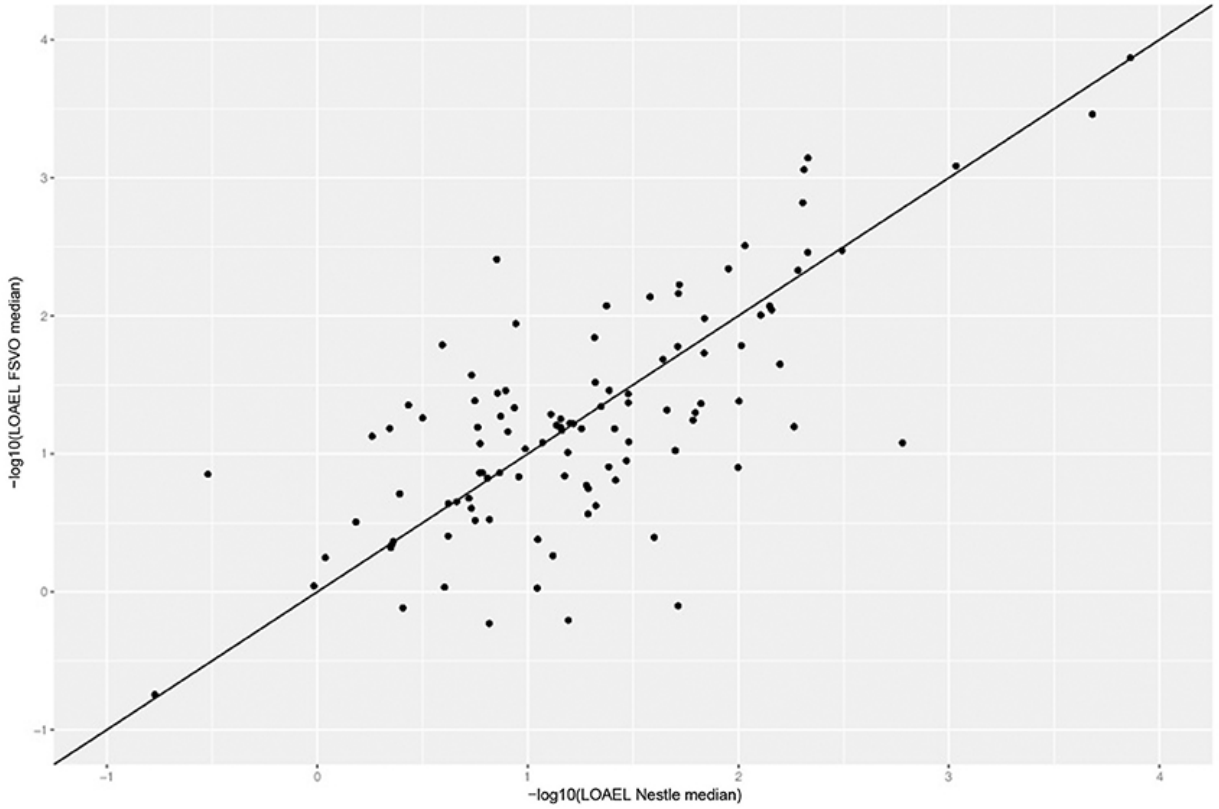
- compounds that occur in both databases
- 375 LOAEL values for 155 unique chemical structures

LOAEL Variability

Both datasets contain substances with multiple measurements

All datasets have almost the same experimental variability (standard deviations: 0.56 mg/kg_bw/day (Nestlé), 0.57 mg/kg_bw/day (FSVO), 0.56 mg/kg_bw/day (combined))

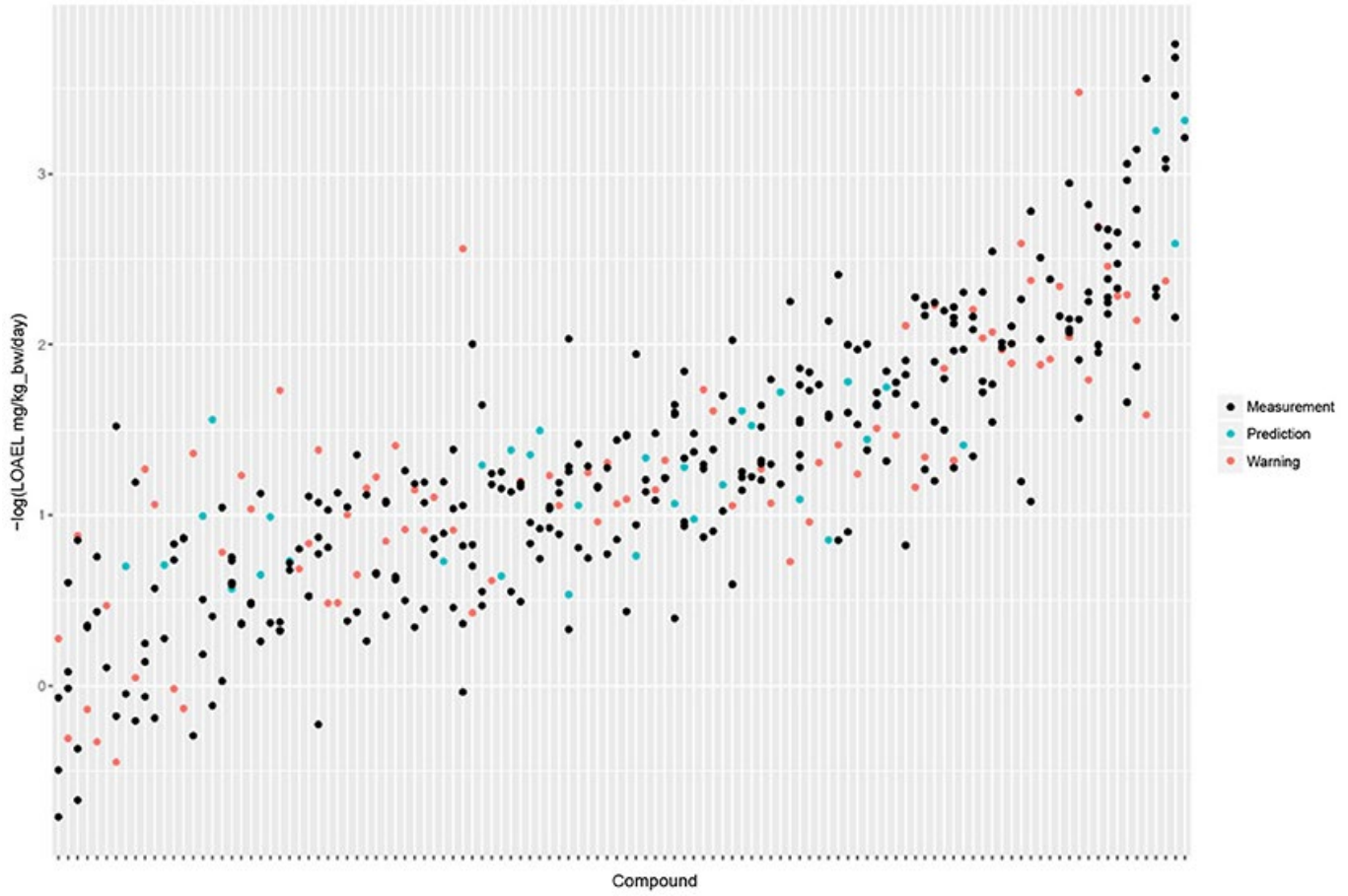
LOAEL Correlation



r^2: 0.52, RMSE: 0.59, p-value < 2.2e-16

As both databases **contain duplicates medians were used for the correlation plot and** statistics

LOAEL Experiments vs Predictions



Conclusions

- Carcinogenicity classifications seem to be poorly reproducible (57% concordant classifications for repeated experiments)
- Experimental LOAEL values have a variablity of approximately 1.5 log units (orders of magnitude)
- Variability in chronic *in vivo* bioassays might be caused by
  - *biological complexity*
  - *long term experimental conditions*
  - *evaluation complexity*
  - *statistical limitations (low number of animals/treatment)*
- Good *in-silico* models have the same accuracy as biological experiments (*in-vivo* and *in-vitro*) for **compounds in their applicability domain**

https://in-silico.ch/presentations/epa-nam-2022/

Using Big Data to Evaluate the Concordance of Toxicity of Pharmaceuticals between Animals and Humans

EPA NAM Conference 2022

Thomas Steger-Hartmann
Bayer AG, Pharmaceuticals

The Issue

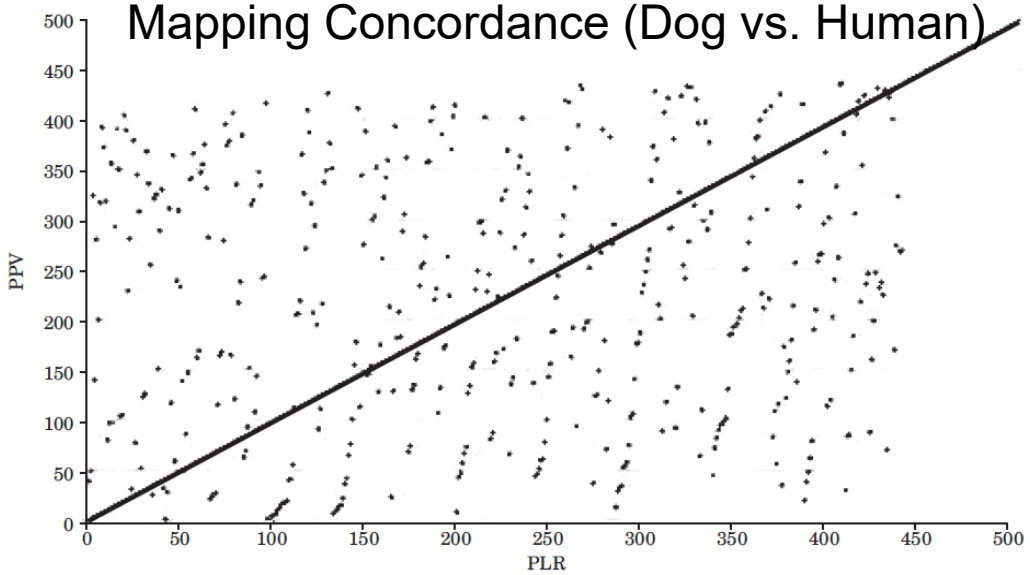
Why are we interested in the concordance between animal studies and human outcome?

Despite the development of NAMs, animal studies will remain to deliver pivotal contributions to human safety assessment in the next decade.

This holds particularly true for the pharmaceutical sector.

Animals Do not Predict at All

Why should we still use them?

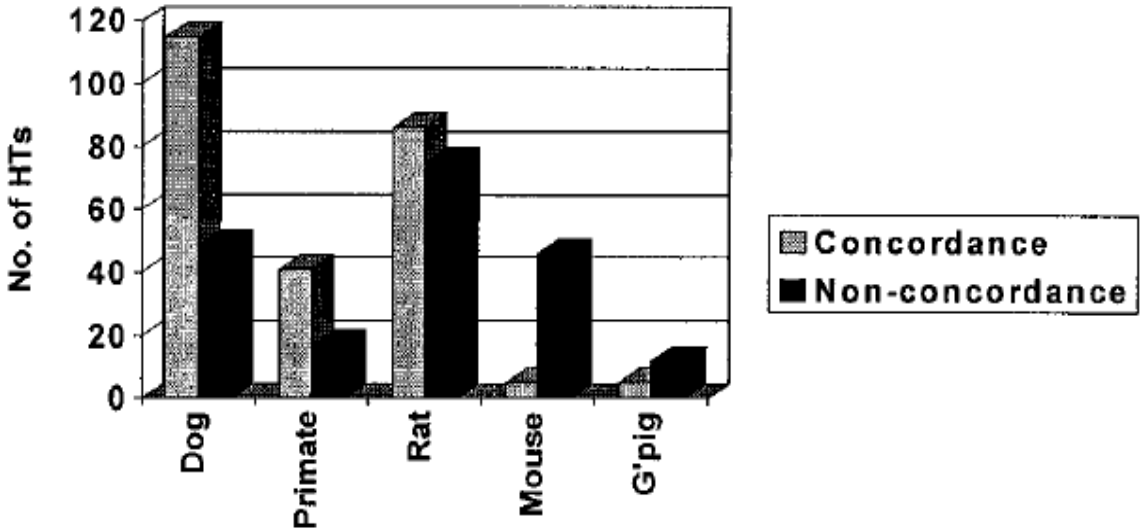


"Positive Predictive Values (PPV) and Positive Predictive Likelihood Ratios (PLRs) for all 436 results ordered according to their value, with the highest ranking first and the lowest last.
If a perfect correlation exists, all points should lie on the line, (…). However, the significant scatter of the data points demonstrates that little correlation exists between PPV and PLR." (Bailey at al. ATLA 41, 335-350, 2013).

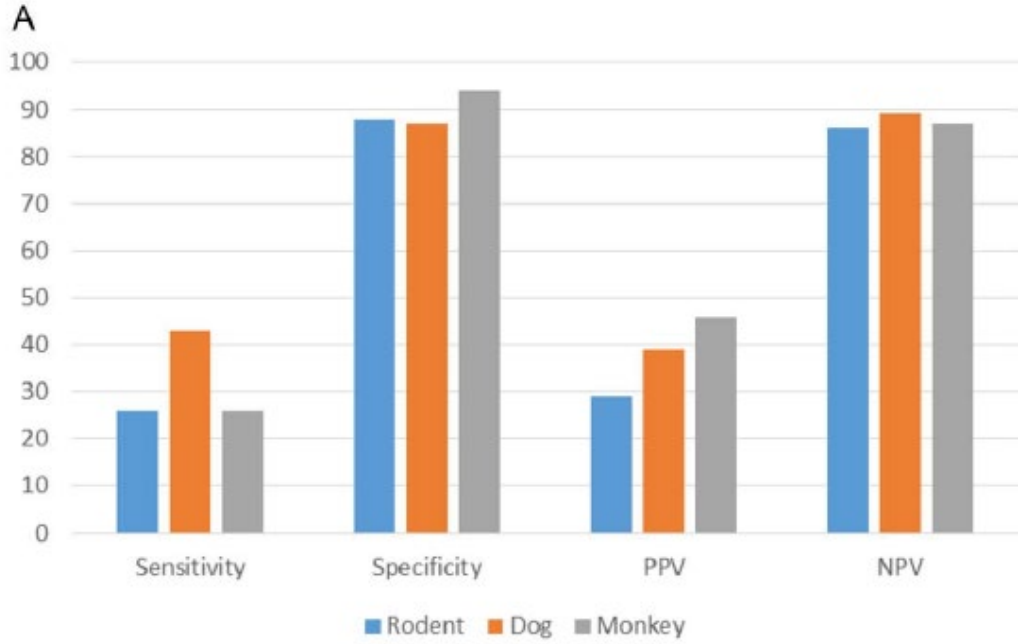
"…results from tests on animals … are highly inconsistent predictors of toxic responses in humans and are little better than what would result merely by chance.." Bailey et al. ATLA 42, 181–199, 2014

Olson et al. (2000) and Later Studies

Animals do Predict Human Outcome



Concordance rates of preclinical results for human toxicities (absolute values); n=150 compounds (Phase I-III)
(Olson et al. Regulatory Toxicology and Pharmacology 32, 56–67, 2000)



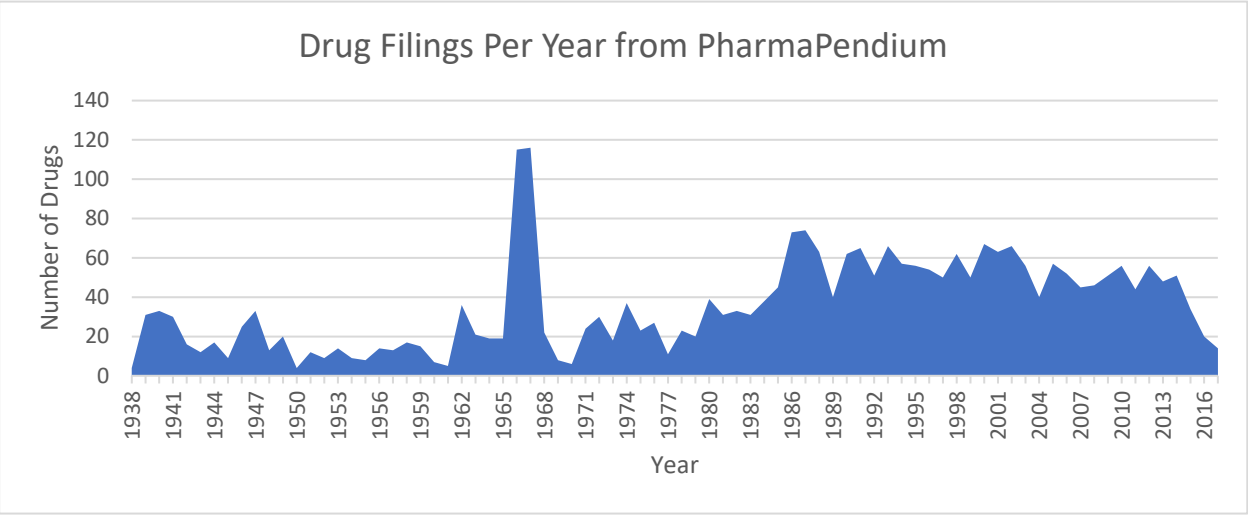
Concordance parameters by test species evaluated. A. sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV); n=182 compounds (First-in-man)
(Monticello et al. Toxicology and Applied Pharmacology 334, 100–109, 2017)

- There is evidence that preclinical species predict human toxicities to a certain extent
- Analyzed data sets were still rather small
- ***Can we drill even deeper in terms of species and findings?***

Methodology of a Systematic Analysis

A Big Data Approach using PharmaPendium

- Key Facts on PharmaPendium
  - 1,637,449 preclinical observation & adverse event reports
  - 3,920 drugs and drug formulations
  - spans a period of drug approvals of more than 70 years
  - No post-marketing data

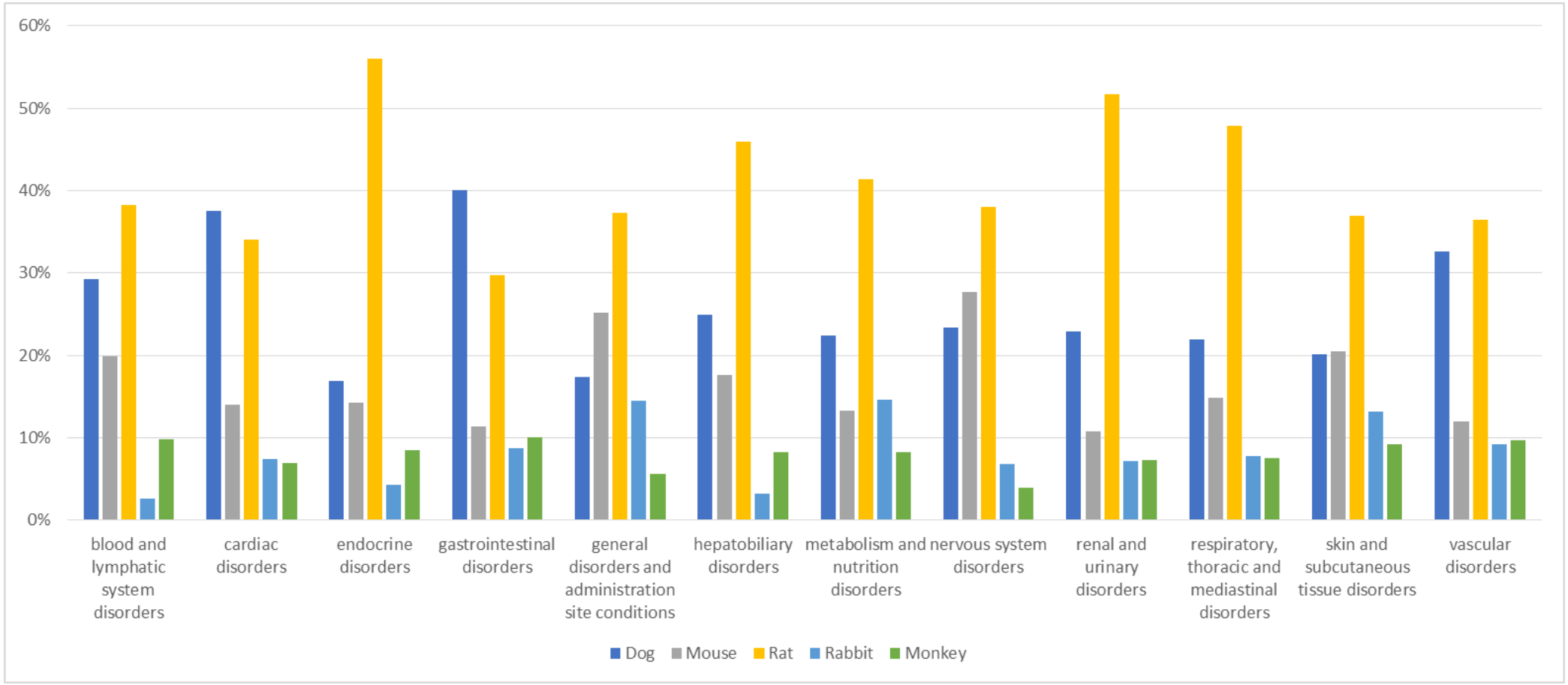


Species	Count of Observations
Human	1,361,367
Rat	155,807
Dog	51,175
Mouse	49,388
Rabbit	20,836
Cynomolgus monkey	14,662
Monkey (unspecified)	6,760
Rhesus monkey	2,743
Pig	2,059
Guinea pig	1,326

- Curation in PharmaPendium: preclinical observations & adverse events are coded to MedDRA preferred terms by the PharmaPendium curators

Results of Analysis

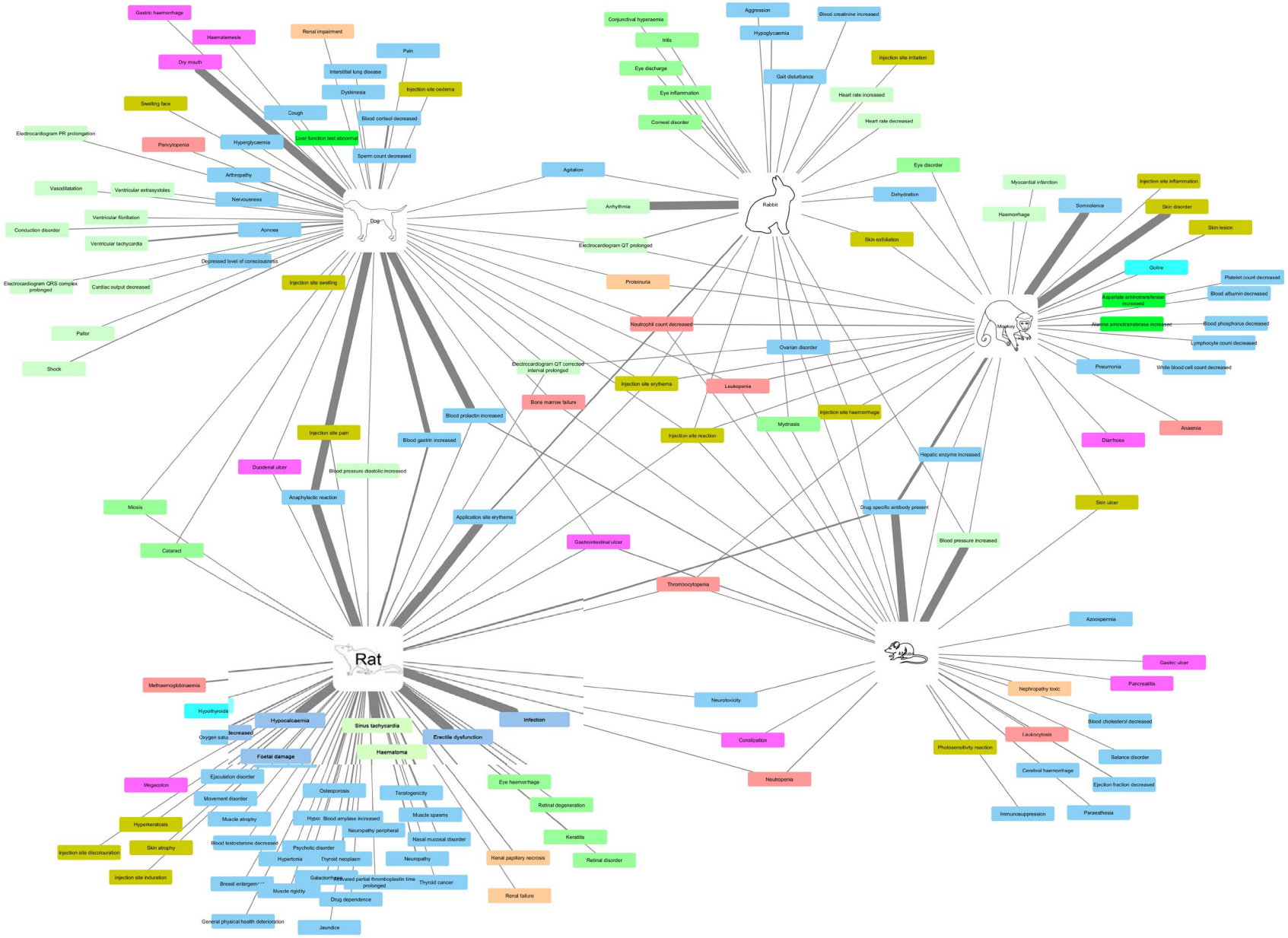
True positives per organ class and species adjusted for the frequency of species use



// Highest rates of TPs (normalized for frequency of animal use) are found for rat and dog

Results of Analysis

Line thickness is proportional to positive likelihood ratio (LR+)



- Blood and lymphatic disorders
- Renal and urinary disorders
- Gastrointestinal disorders
- Cardiac and vascular disorders
- Hepatobiliary disorders
- Skin and subcutaneous disorders
- Eye disorders
- Other

Conclusions from PharmaPendium Analyses

- Certain animal findings are confirmed as being highly predictive, such as cardiac disorders
- Negative predictivity is generally low
- Predictivity of observations is highly species-specific, but also influenced by frequency of animal use for specific endpoints

Regulatory Toxicology and Pharmacology 96 (2018) 94–105

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

A big data approach to the concordance of the toxicity of pharmaceuticals in animals and humans

Matthew Clark[a,*], Thomas Steger-Hartmann[b]

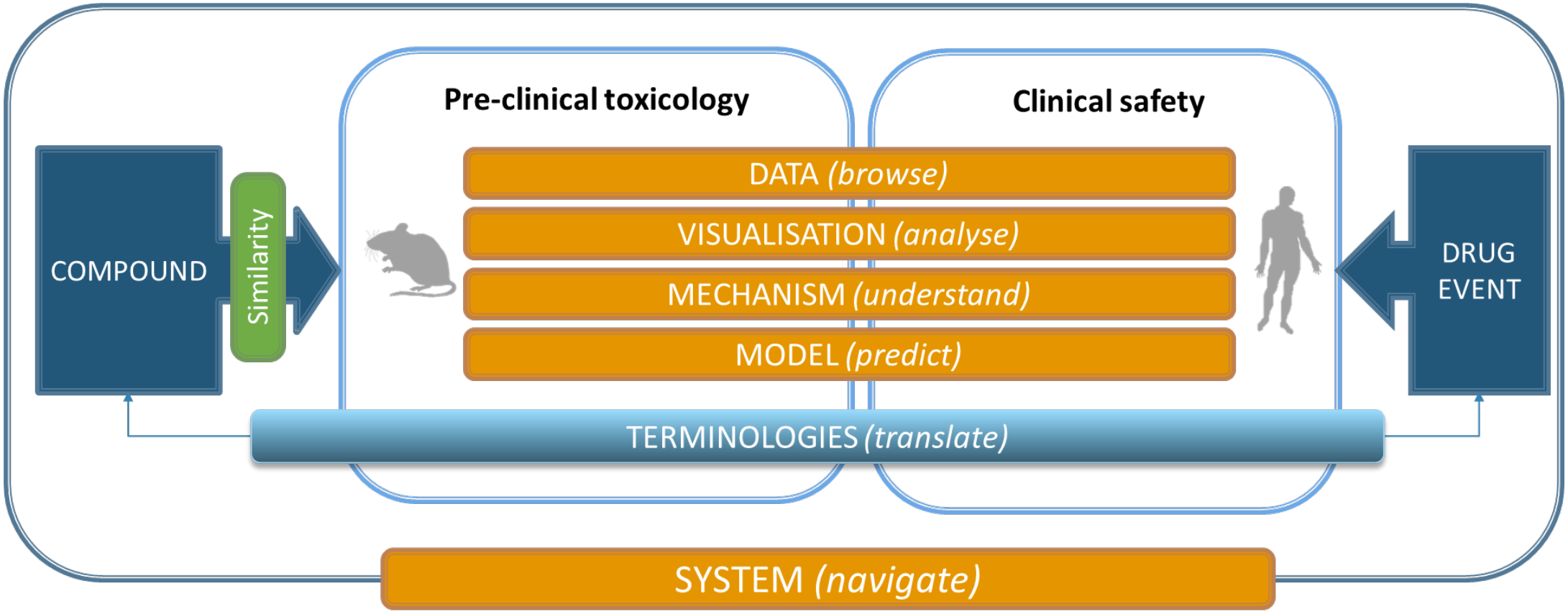
[a] Elsevier R&D Solutions, 1600 JFK Blvd, Philadelphia, PA, 19103, USA
[b] Investigational Toxicology, Bayer AG, 13353, Berlin, Germany

- ***Statistical analyses are influenced by size of data, data subset (early clinical phases vs. marketed compounds vs. PV data) and subjective terminology assignment***

Can we increase data size and overcome terminology issues?

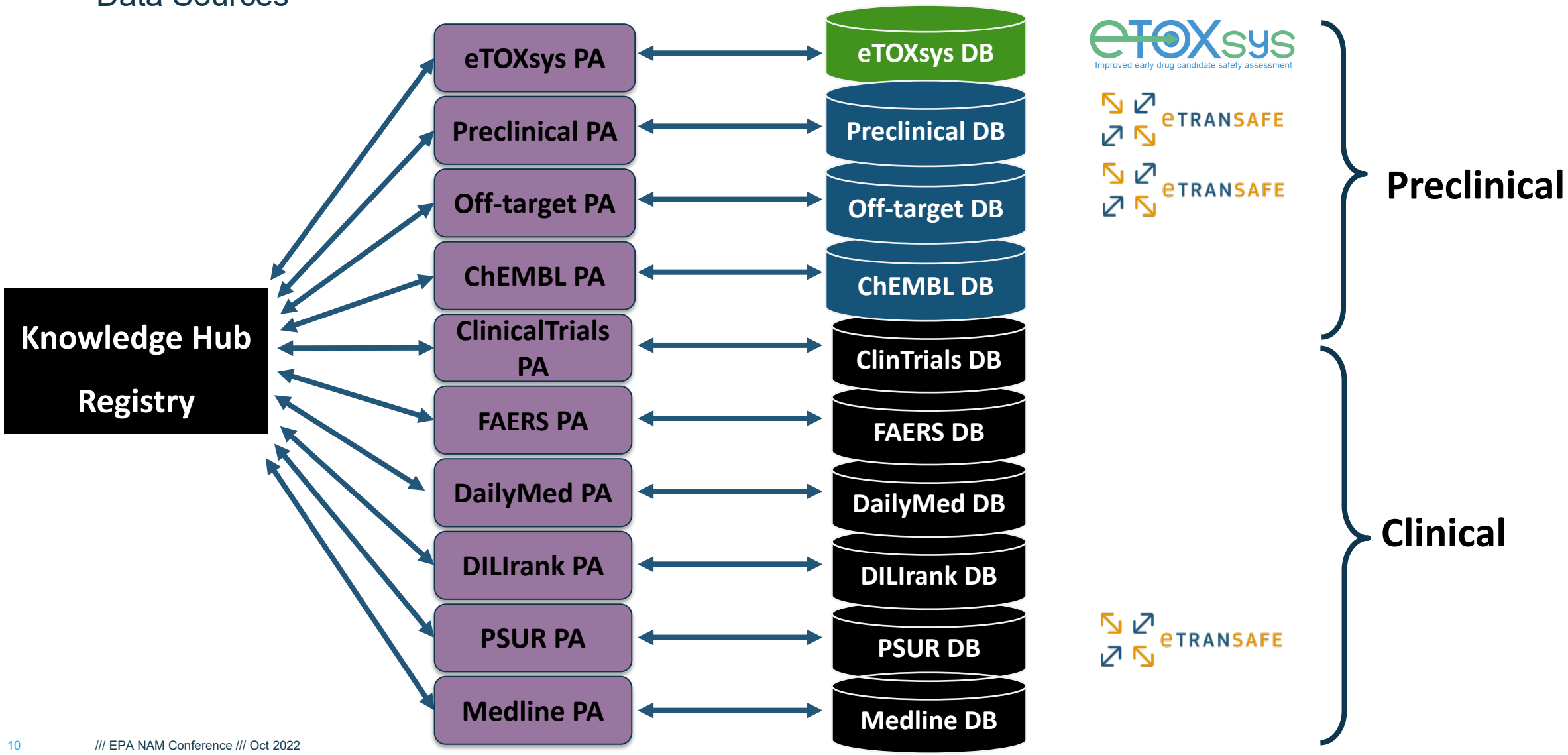
ToxHub – A Translational System for Safety Assessment

Functionalities



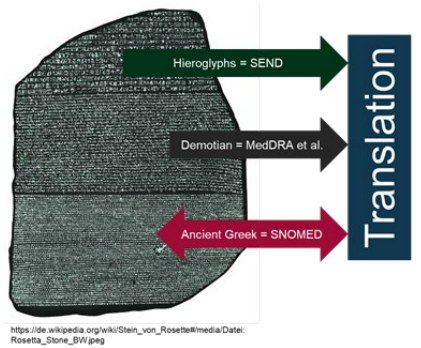
ToxHub – A Translational System for Safety Assessment

Data Sources



Translational Analysis of Safety Data

E.g., matching for term “steatosis” (another form of DILI)

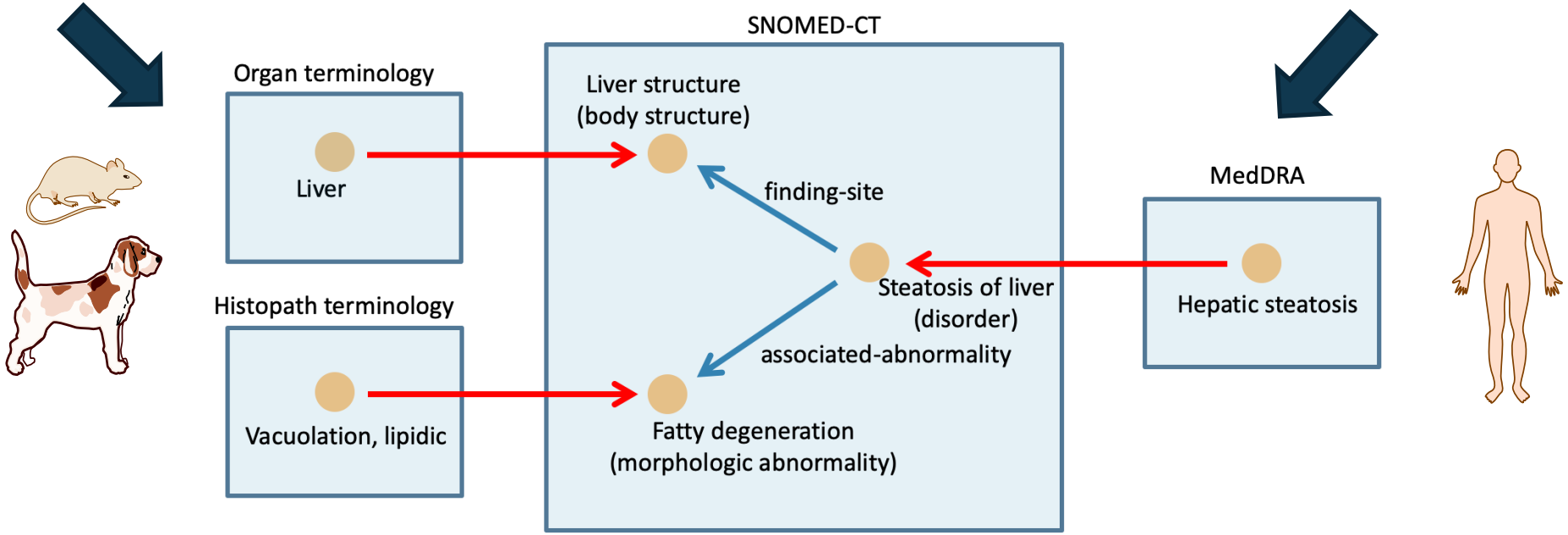


Histopathology diagnosis of steatohepatitis :

- Liver – parenchymal cells - hepatocytes
- Fat accumulation
- Increased Intracellular lipid content
- Vacuolation, lipidic
- Fat necrosis
- Treatment-related

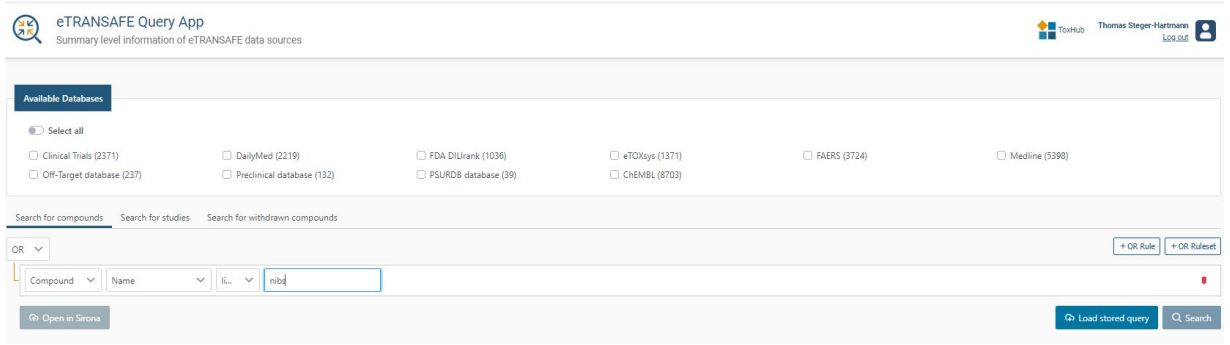
Possible signs and symptoms of steatohepatitis :

- Abdominal swelling (ascites)
- Enlarged blood vessels just beneath the skin's surface
- Enlarged breasts in men
- Enlarged spleen
- Red palms
- Yellowing of the skin and eyes (jaundice)



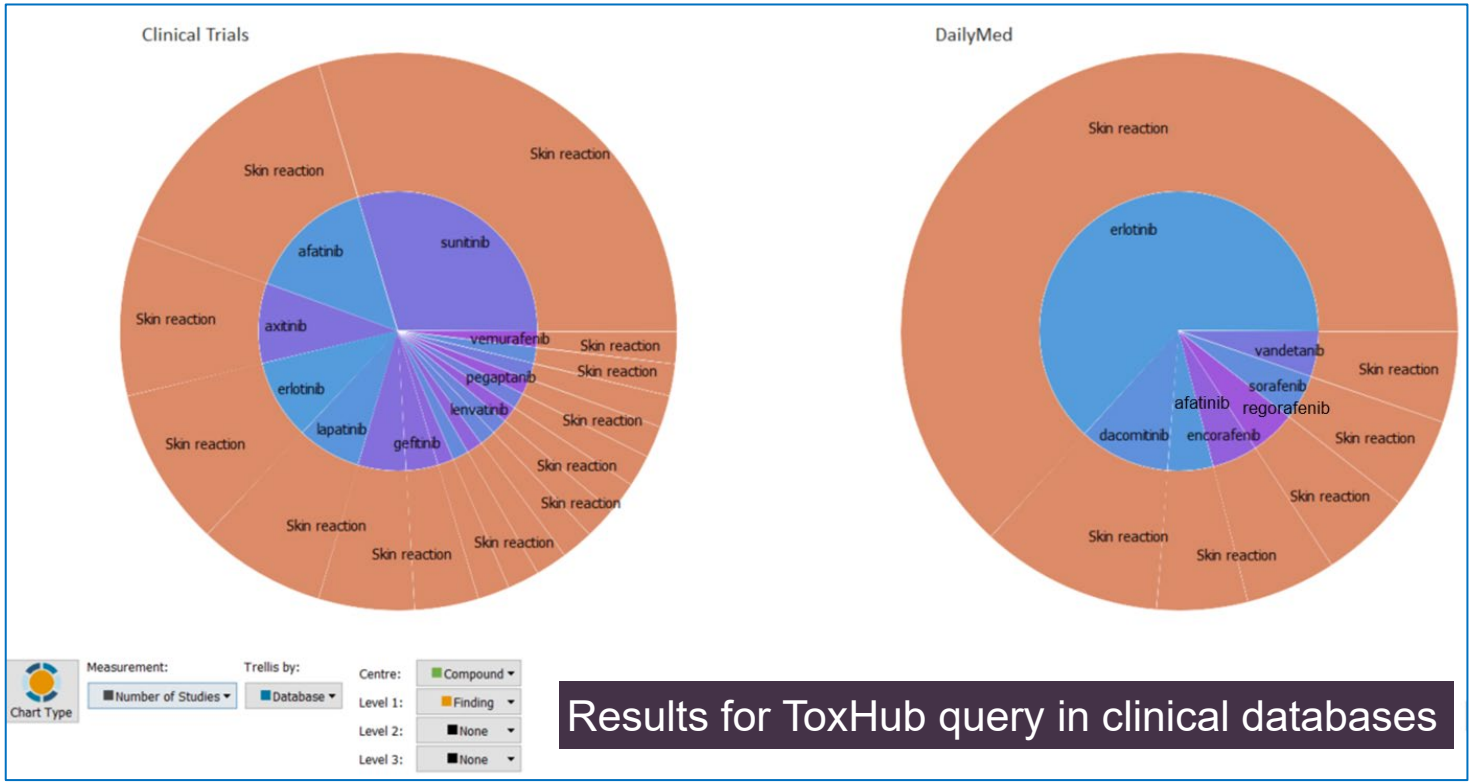
ToxHub – A Use Case

Investigating the translational value of animal data – kinase inhibitors as an example



Questions:

- Is it possible to identify differences or commonalities between the clinical safety profiles of kinase inhibitors with regard to skin toxicity?
- Can these profiles be correlated with the preclinical findings?
- ***Can conclusions be drawn with regard to translational predictivity of preclinical findings, relevance of species selection?***



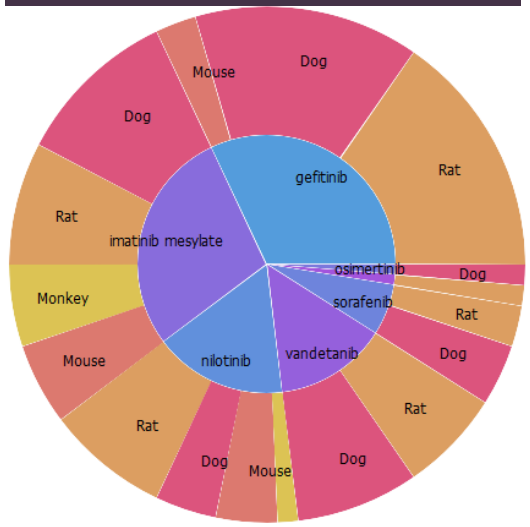
Results for ToxHub query in clinical databases

Kinase Inhibitors (“nibs”)

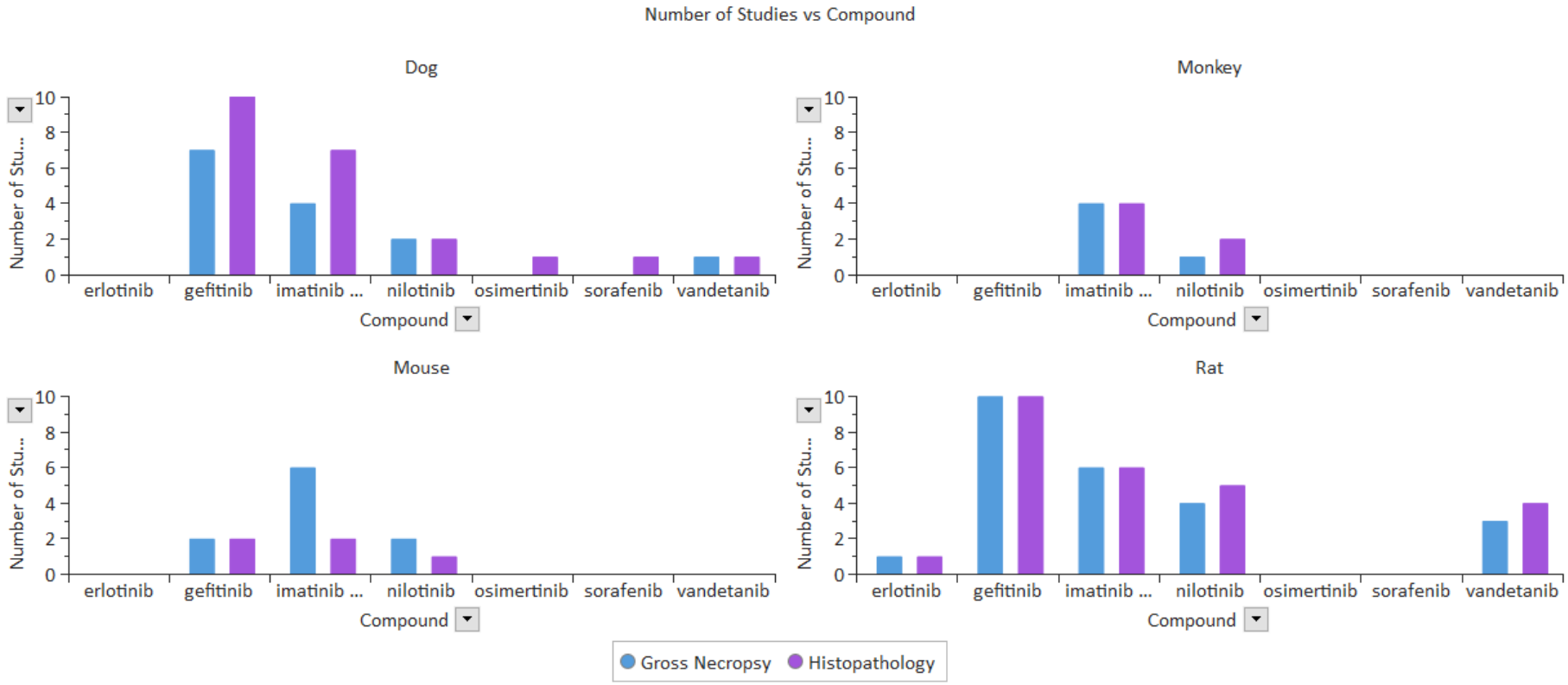
Results for coverage in ToxHub’s preclinical databases

- For 7 compounds there are skin findings in the preclinical databases (erlotinib, gefitinib, imatinib, nilotinib, osimertinib, sorafenib, vandetanib)

General species coverage

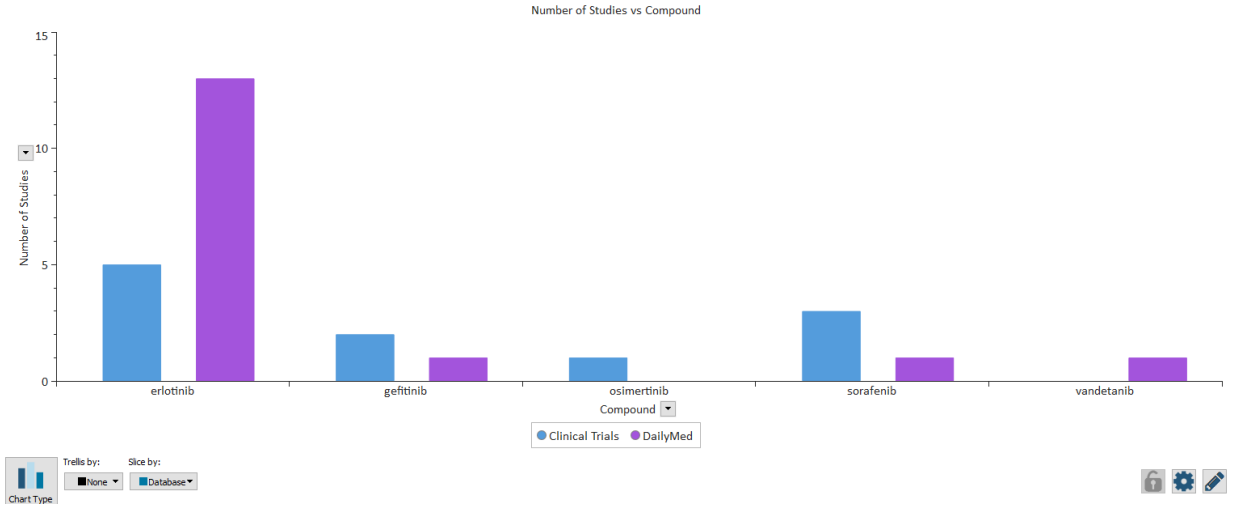


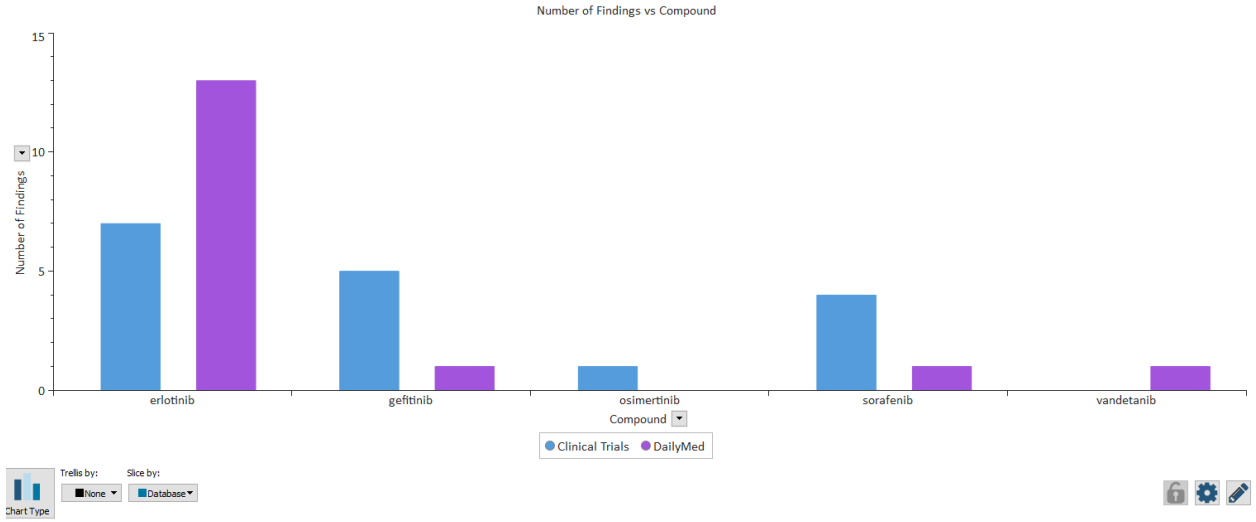
Species coverage for skin findings



Kinase Inhibitors (“nibs”)

Skin findings in clinical databases (ClinTrials & DailyMed) for overlapping compounds





- Imatinib and nilotinib have no entries for skin findings in the two clinical databases.

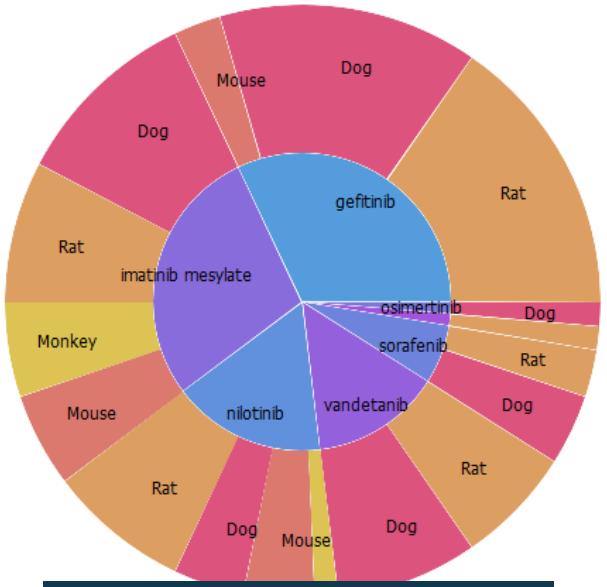
→ the preclinical skin findings for these two compounds were evidently not predictive for the clinical outcome.

Kinase Inhibitors (“nibs”)

Conclusions

- Where data for more then one species are available, the rat seems to be the more sensitive (gefitinib, imatinib, vandetanib) whereas the monkey is evidently less sensitive.
- Regarding translational predictivity (animal → human), it is obvious that adding a further species to the rat for the purpose of assessing skin reactions does not add any value. Particularly, the NHP does not seem to be more predictive than rats.
- The translation of observed preclinical skin findings into adverse in clinical trial is particularly questionable for non-(V)EGFR tyr kinases (imatinib, nilotinib).
- ***The higher translational value of the rat regarding skin findings over other species confirms previous analyses***

(Clark & Steger-Hartmann, 2018, https://doi.org/10.1016/j.yrtph.2018.04.018)



Species coverage for nibs

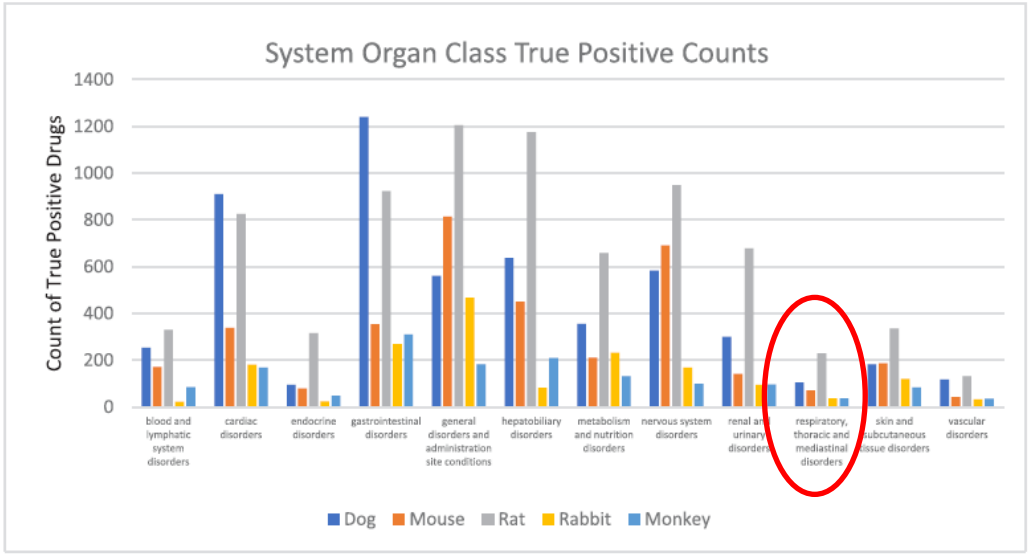


Fig. 2. Count of true positive drugs by system organ class and species.

Summary

Access to Big Data and application of advanced data science technologies will improve our understanding of the translational value of animal studies and may in the future contribute to a re-design of preclinical programs.

This will complement NAMs' strive to reduce animal use.

Acknowledgements

- **Members of my Bayer team:** Annika Kreuchwig
- **eTRANSAFE:** Francois Pognan, Ferran Sanz, Manuel Pastor, Gavin Nicholson (Optibrium) and many other participants

Parts of this work have received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement No. 777365 (eTransafe), resources of which are composed of financial contributions from European Union's Horizon 2020 research and innovation programme as well as EFPIA companies in kind contribution

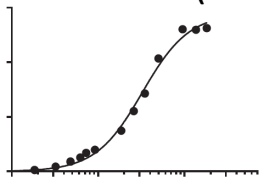
Disclosure Statement

- Member of several science advisory boards (public and private sector): ILSI, ILSI Europe, Cosmetics Europe LRSS, MSU Center for Research on Ingredient Safety, A*STAR Food and Chemical Safety Programme Singapore, Owlstone Medical, PCPC Expert Group on Carcinogenicity, Plastics Europe Brigid (microplastics) project
- Member/chair of several national and international scientific advisory committees: UK COT, UK COMEAP, JMPR, JECFA, WHO TobReg, ISO TC126 WG10 Intense Smoking Regime
- I have no financial interests in the subject matter of the session

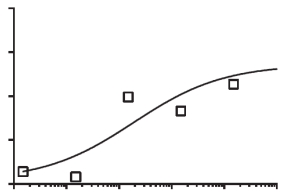
AOP/MOA

Exposure

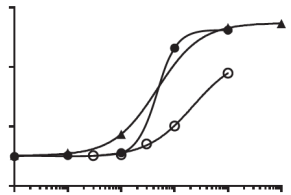
Dose-MIE (KE1)



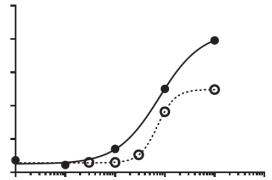
KE1-KE2



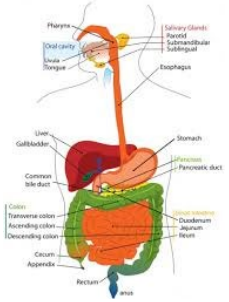
KE2-KE3



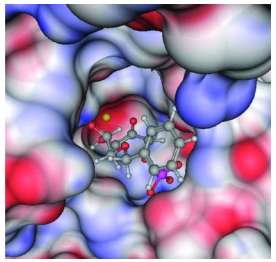
KE3-AO



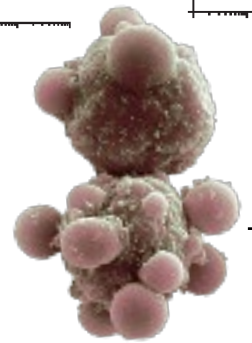
Population



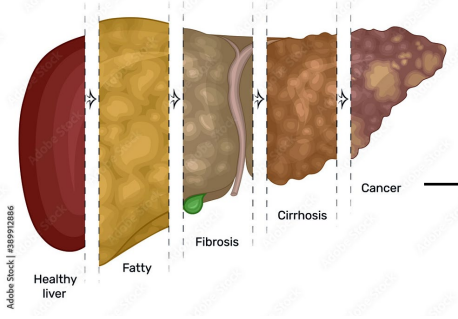
ADME/TK



Molecular interaction



Cellular Effects



Organ Effects

Individual
(Adverse Outcome)

Adverse Outcome Pathway

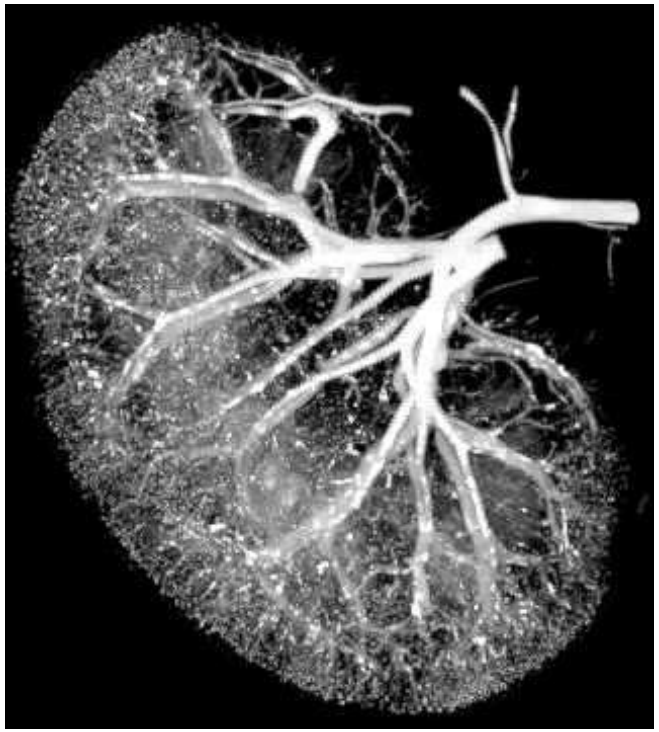
Mode of Action

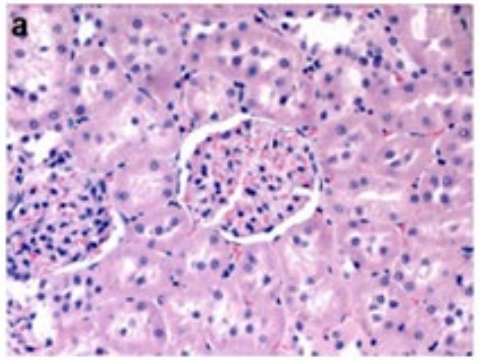
Key events (based on Bradford Hill considerations)

Level of confidence

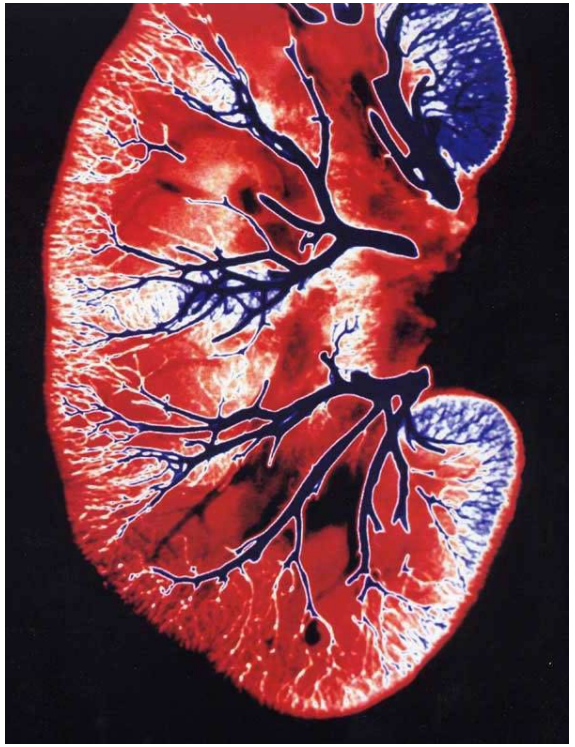
Anatomy, e.g. kidney

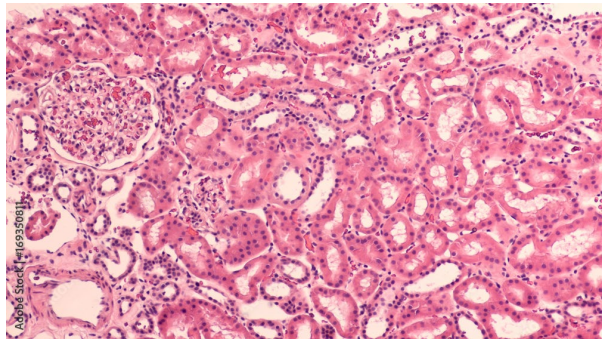
Rat



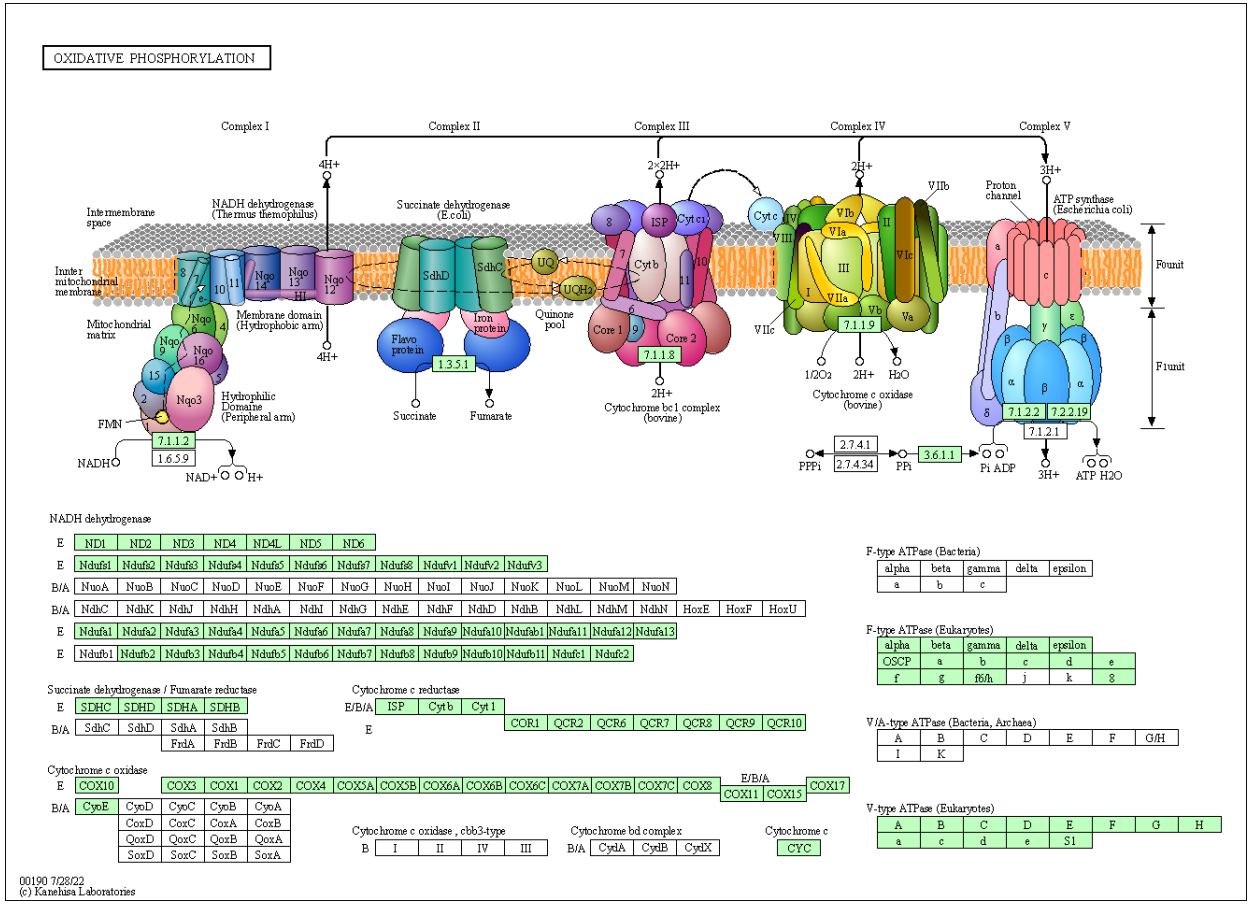


Human





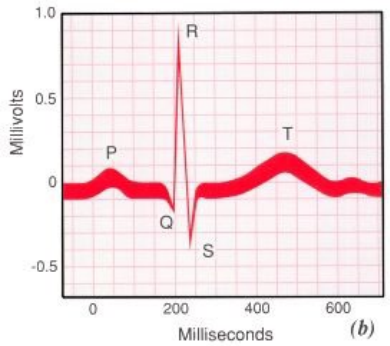
Biochemical pathways e.g. oxidative phosphorylation

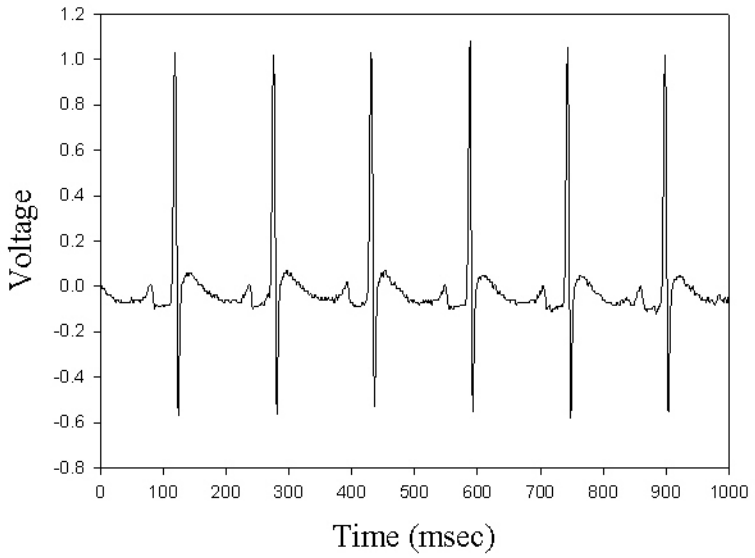


Mouse/Rat/Human

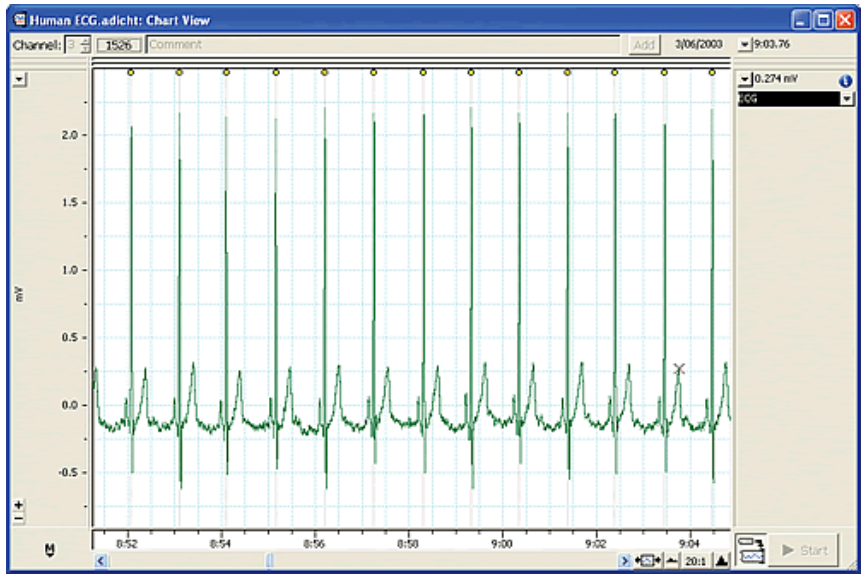
From https://www.kegg.jp/pathway

Physiology, e.g. cardiac function (ECG)





Rat



Human

Carbofuran: MOA for neurotoxicity

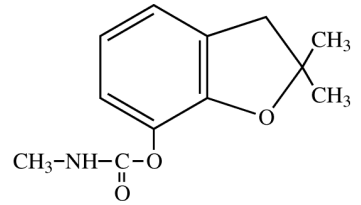
Is the weight of evidence sufficient to establish a mode of action (MOA) in animals?

YES

AChE inhibited in vitro at in vivo concentrations

Inhibition of neuronal AChE

Carbofuran

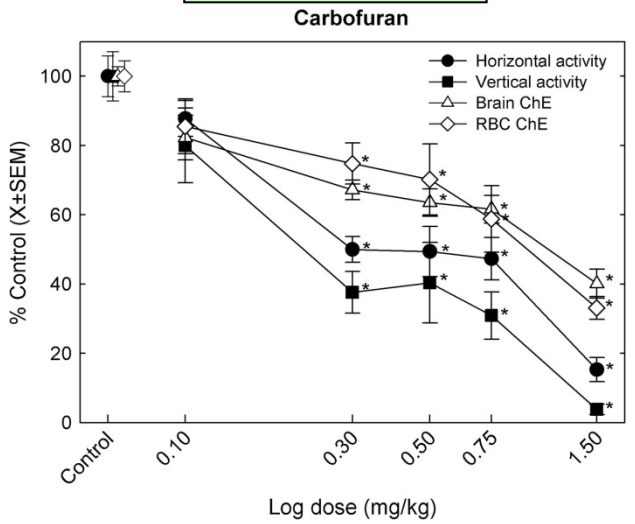


ACh levels increased In brain and periphery

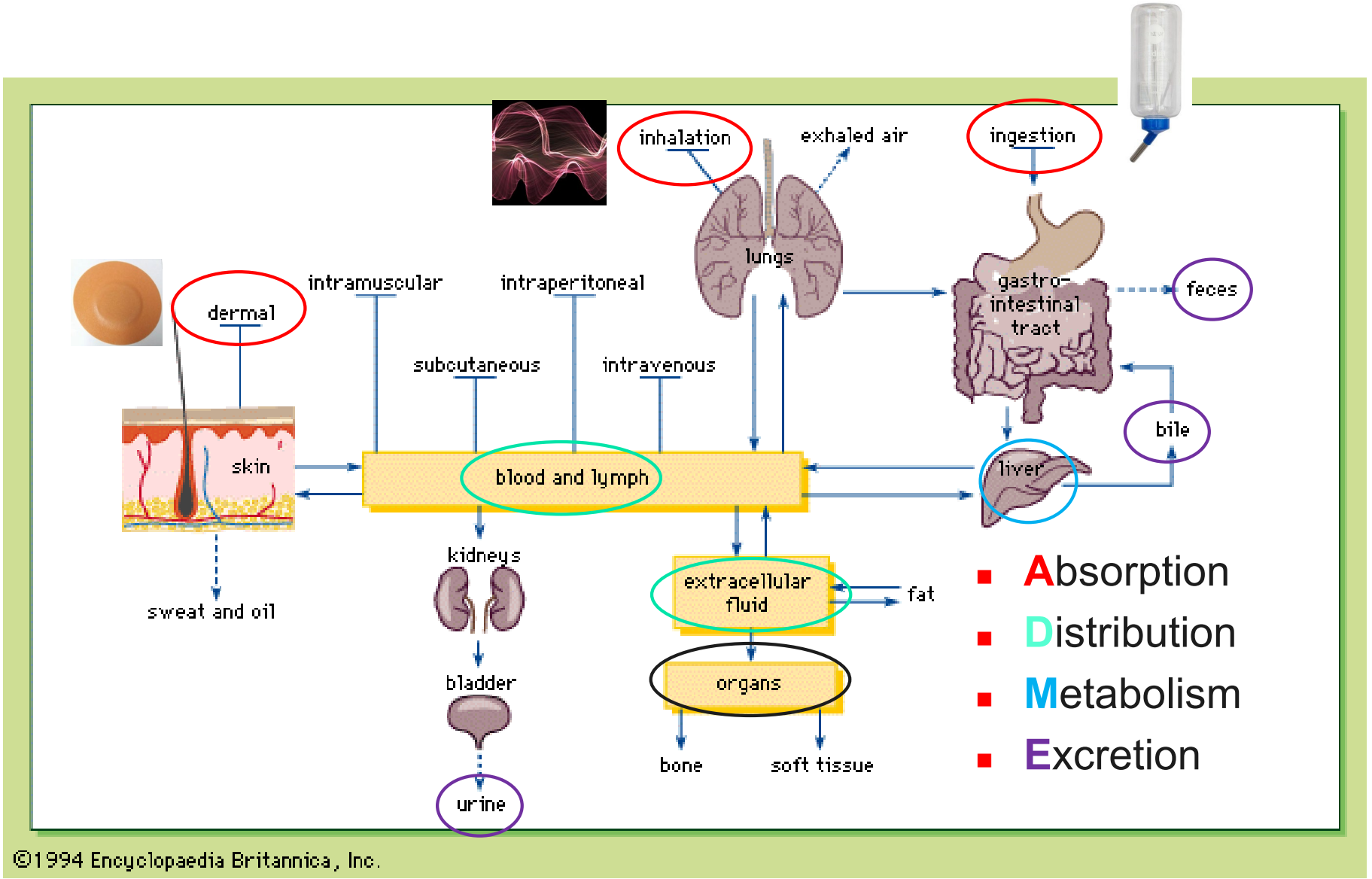
Increased ACh in synaptic cleft

Overstimulation of ACh receptors

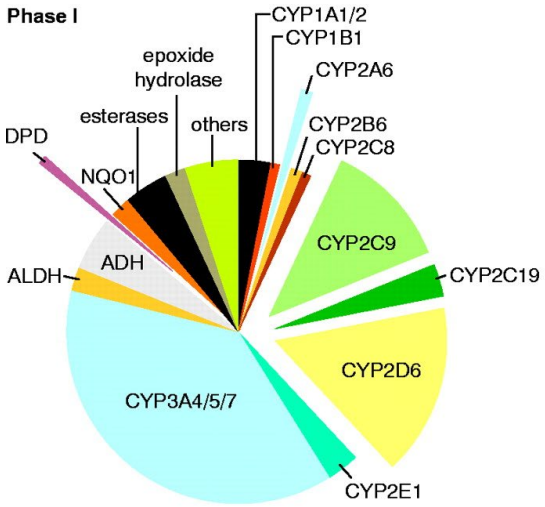
Clinical signs

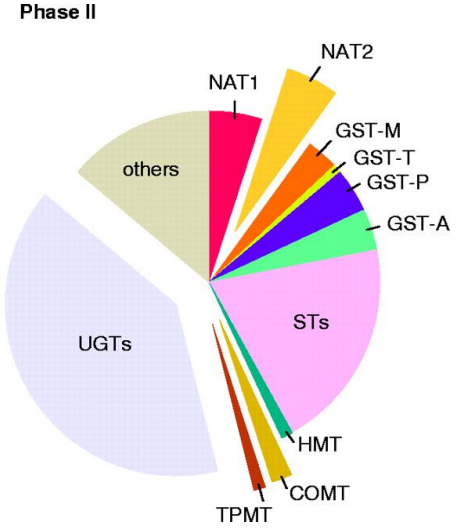


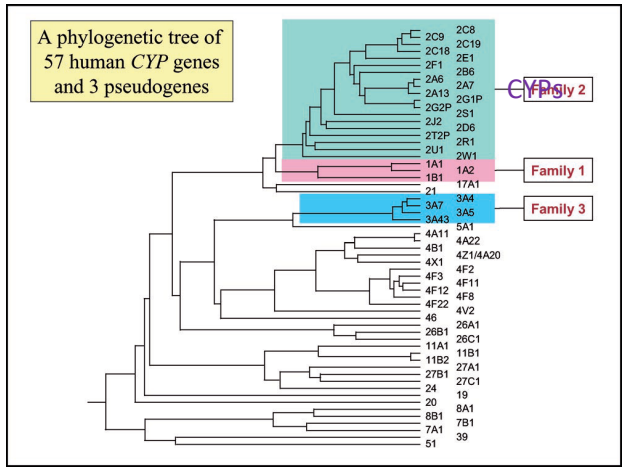
Absorption, distribution, metabolism, excretion (ADME) determine exposure



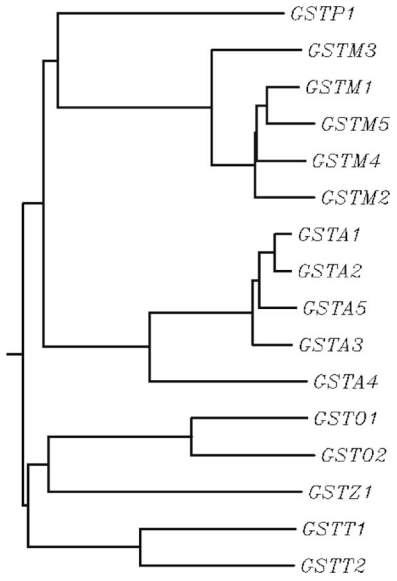
Xenobiotic disposition



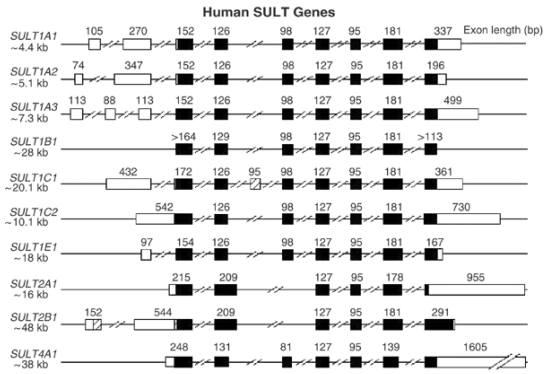




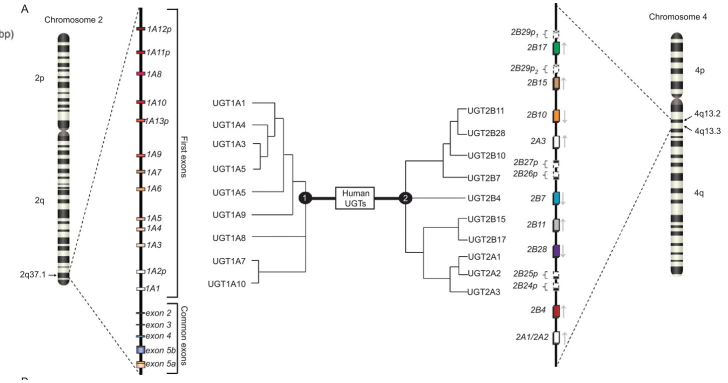
GSTs



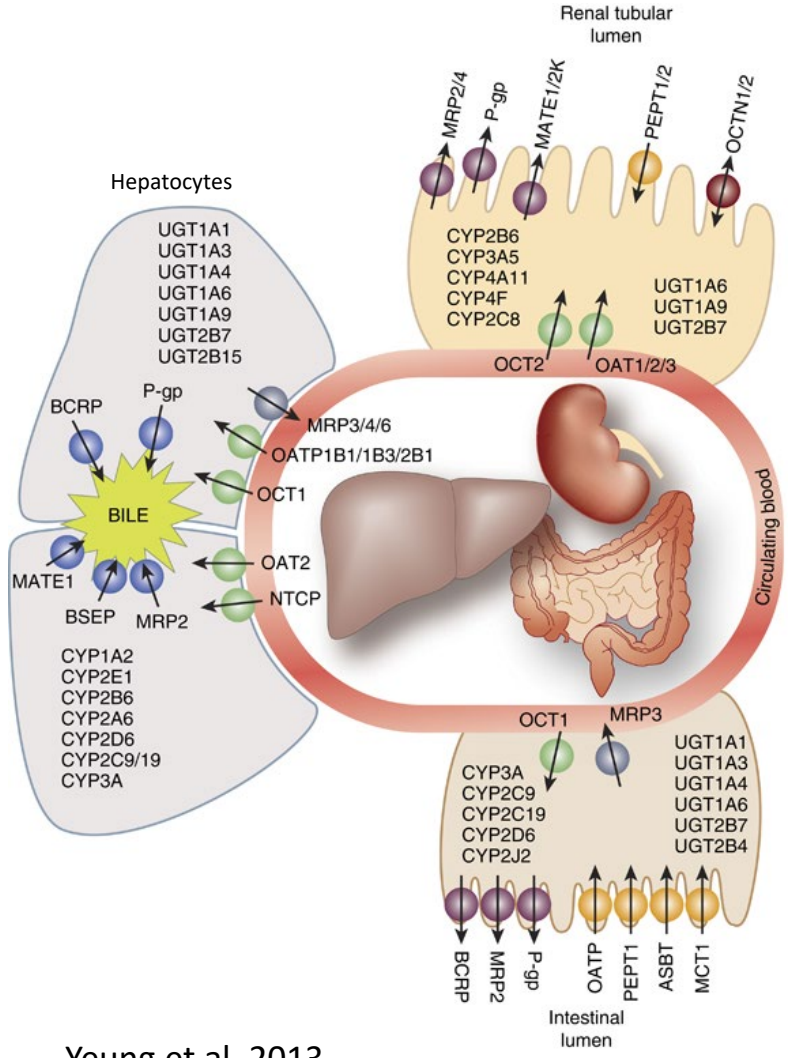
SULTs



UGTs

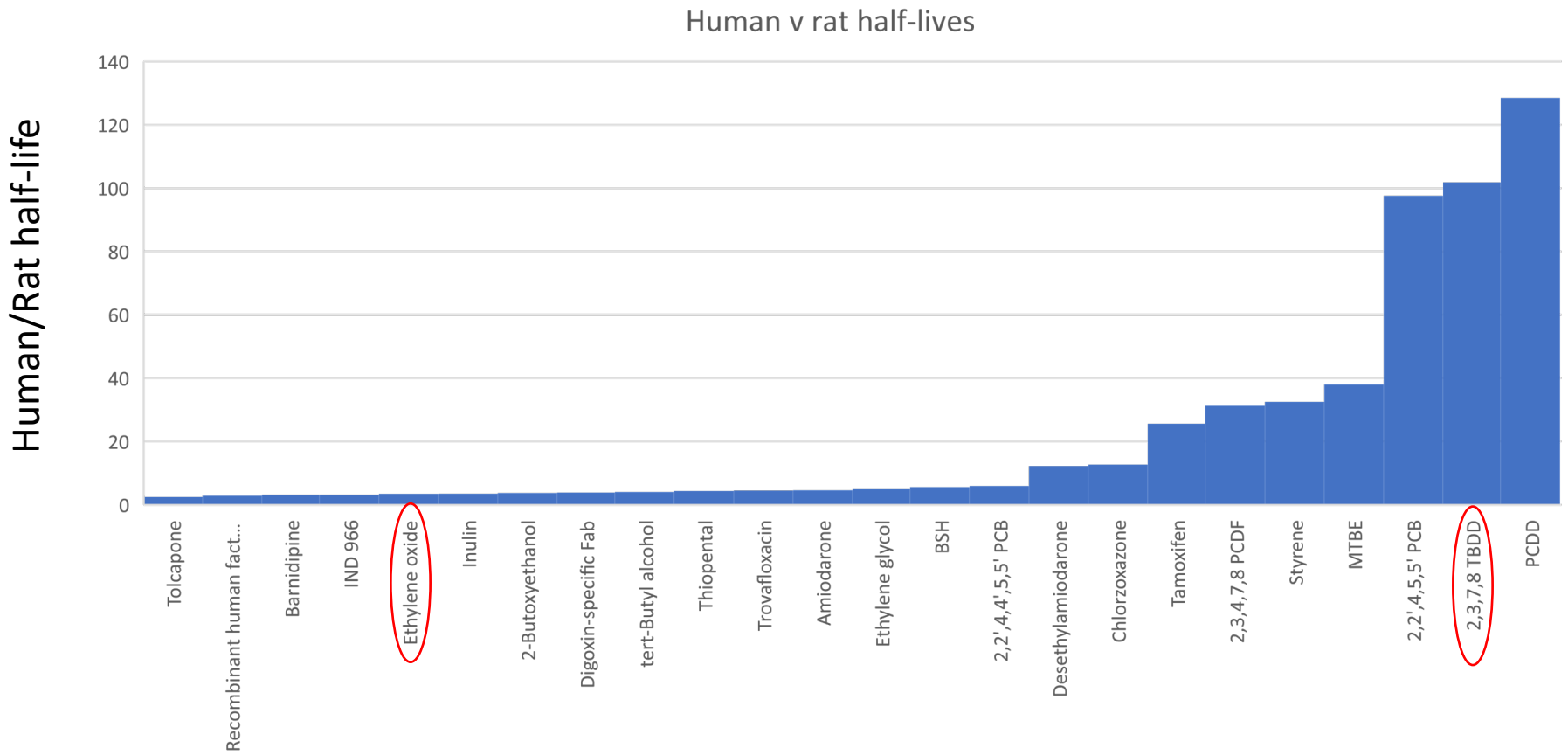


- Specificity
- Maximum rate (V_{max})
- Affinity (K_m)



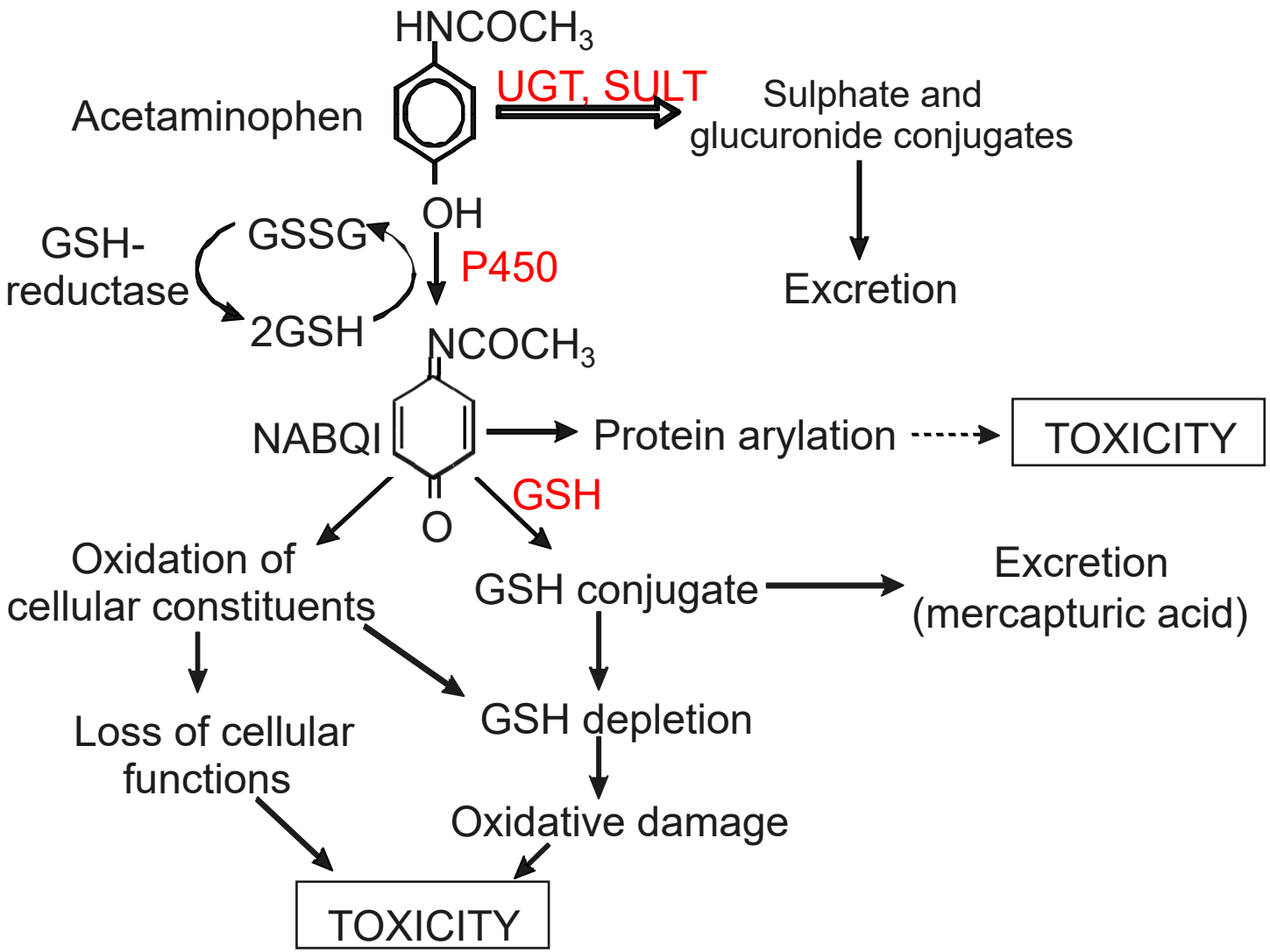
Yeung et al, 2013

Species comparison of plasma half-lives



Data from Sarver et al, 1997

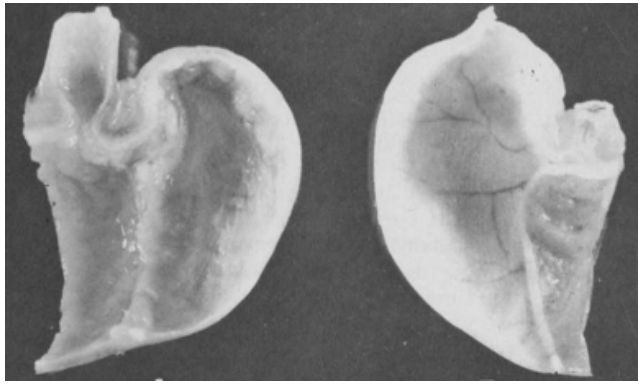
Mode of action for acetaminophen hepatotoxicity



Not all MOAs observed in rodent studies are relevant to humans

- Forestomach tumours induced in mice and rats by butylated hydroxyanisole (local irritation)
- Bladder tumours induced in rats by sodium saccharin (local irritation)
- Mammary tumours induced in female rats by atrazine (suppression of LH surge)
- Thyroid tumors in rats induced by phenobarbital (induction of UGT)
- Renal toxicity in male rats induced by D-limonene (α2u-golbulin)
- Developmental effects of sulfoxaflor in rats (nAChR agonism)

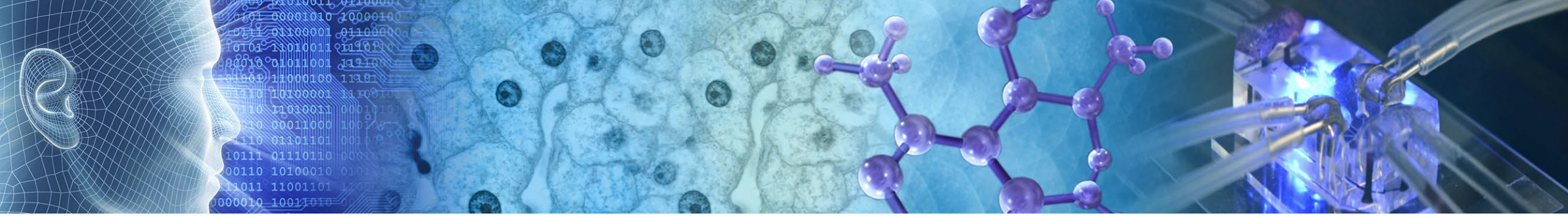
BHA in rat (left)



Forestomach changes

Conclusions

- There is considerable conservation of biochemistry, signalling, anatomy and physiology between rodents and humans
  - Many shared AOPs/MOAs
  - Some quantitative differences in dose-response and response-response
- Some AOPs/MOAs are rodent specific
  - Many were identified early as focus was on disproving human relevance
  - Relatively well understood
- Qualitative similarities in toxicokinetics, but many important quantitative differences
  - Often conservation when TK plays a key role in MOA (e.g. metabolic activation, active uptake)



Variability and Relevance of Animal Studies for Acute Toxicity, Skin Sensitization, and Mechanistic Responses

Nicole C. Kleinstreuer
NICEATM Director (Acting)

EPA NAMs Meeting
October 12-13, 2022

Why Does Variability Matter?

- Data from traditional mammalian guideline toxicology studies are used by regulatory agencies to make decisions about chemical classification and labeling and inform risk assessments
- In vivo guideline studies have been the reference upon which alternative method performance is often assessed
  - Do we reproduce the same outcome (sufficiently sensitive alternatives)?
  - Affects our confidence and context for interpreting results
- Better characterizing the in vivo guideline study reproducibility could provide additional insight to set appropriate expectations for alternatives

Evaluating Reproducibility

Assessing Impact on Categorical Endpoints

- Many guideline studies are interpreted by hazard category classification
- Variability cannot be assessed quantitatively (e.g., by standard deviation)
- Instead, reproducibility is evaluated to determine how often the same category is identified across replicate studies

Chemical X

Study 1: category 3

Study 2: category 2

Study 3: category 2

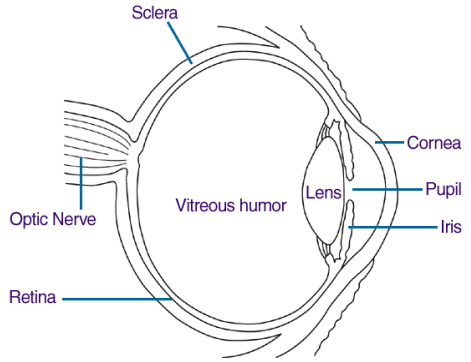
Study 4: category 1

Prior type	1	2	3	4	Total Studies
1	25%	50%	25%	-	1
2	25%	50%	25%	-	2
3	25%	50%	25%	-	1
4	-	-	-	-	0

Rabbit Eye Test Scoring

Cornea, iris, and conjunctiva are subjectively evaluated and scored

- Corneal opacity (CO)
  - 1 = Scattered or diffuse area – details of iris visible
  - 2 = Easily discernible translucent areas – details of iris slightly obscured
  - 3 = Opalescent areas, no details of iris visible, size of pupil barely discernable
  - 4 = Opaque – iris not visible
- Iris
  - 1 = Folds above normal, congestion, swelling, circumcorneal injection (any one or all of there, or combination of any thereof), iris still reacting to light
  - 2 = No reaction to light, hemorrhage, gross destruction (any one or all of these
- Conjunctival redness (CR)
  - 1 = Vessels definitely injected above normal
  - 2 = More diffuse, deeper crimson red, individual vessels not easily discernable
  - 3 = Diffuse, beefy red
- Conjunctival chemosis (CC)
  - 1 = Any swelling above normal (includes nictitating membrane)
  - 2 = Obvious swelling with partial eversion of the lids
  - 3 = Swelling with lids about half closed
  - 4 = Swelling with lids half to completely closed



OECD/OCDE

405

Adopted:
2 October 2012

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Acute Eye Irritation/Corrosion

United States
Environmental Protection
Agency

Prevention, Pesticides
and Toxic Substances
(7101)

EPA 712–C–98–195
August 1998

EPA

Health Effects Test Guidelines
OPPTS 870.2400
Acute Eye Irritation

Eye Irritation Hazard Classification

EPA Classification

- **Category I**: Corrosive (irreversible destruction of ocular tissue) or corneal involvement or irritation persisting for more than 21 days.
- **Category II**: Corneal involvement or irritation clearing in 8-21 days.
- **Category III**: Corneal involvement or irritation clearing in 7 days or less.
- **Category IV**: Minimal effects clearing in less than 24 hours.

- Maximum score in any animal used for classification
- Positive: CO or IR ≥ 1 or CC or CR ≥ 2

GHS Classification

- **Category 1**: Effects on the cornea, iris or conjunctiva that are not expected to reverse or that have not fully reversed within 21 days.
- **Category 2A**: Effects on the cornea, iris or conjunctiva that fully reverse within 21 days.
- **Category 2B**: Effects on the cornea, iris or conjunctiva that fully reverse within 7 days.

Category	*In Vivo* Effect
1	≥ 1 animal with CO = 4 at any time OR ≥ 2 animals with mean* CO ≥ 3 or IR ≥ 1.5 OR ≥1 animal at day 21 with CO or IR ≥ 1 or CC or CR≥ 2
2A	≥ 2 animals with mean* CO or IR ≥ 1 or CC or CR≥ 2 which reverses within 21 days.
2B	≥ 2 animals with mean* CO or IR ≥ 1 or CC or CR≥ 2 which reverses within 7 days.

*Mean values calculated over days 1-3

Reproducibility of Categorical Outcomes

Rabbit Draize Eye Test

GHS Classification

- **Category 1**: Effects on the cornea, iris or conjunctiva that are not expected to reverse or that have not fully reversed within 21 days.
- **Category 2A**: Effects on the cornea, iris or conjunctiva that fully reverse within 21 days.

- **Category 2B**: Effects on the cornea, iris or conjunctiva that fully reverse within 7 days.

Prior type	1	2A	2B	NC	Total Studies
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
NC	1.1%	3.5%	1.5%	93.9%	400

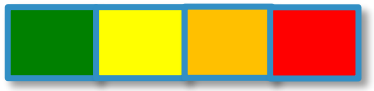
- ECHA database evaluation
- GHS hazard categories
- 491 substances with at least 2 Draize eye studies

Luechtefeld *et al.*, 2016. ALTEX 33(2)

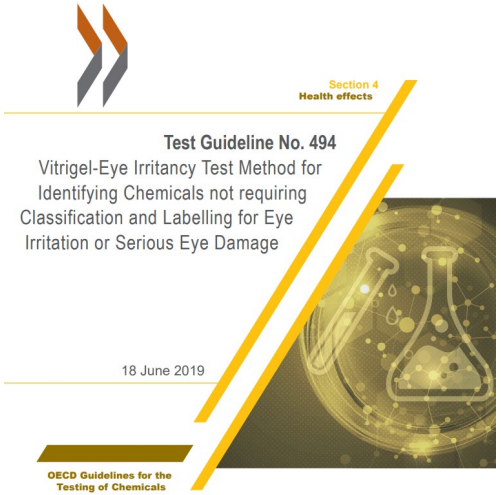
OECD Guidelines for in vitro/ex vivo eye irritation testing – assessed based on comparison to the rabbit test…

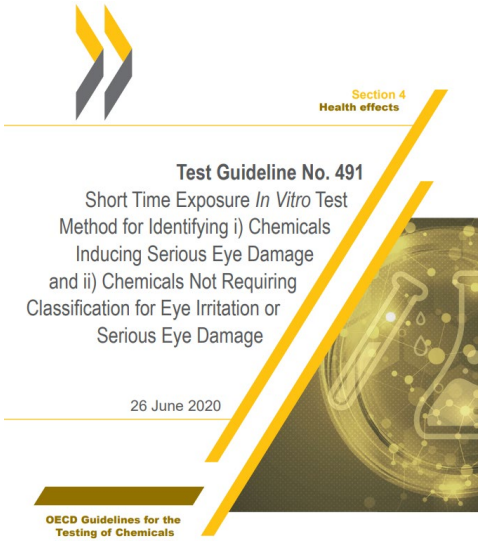
✓ ? ? ✓

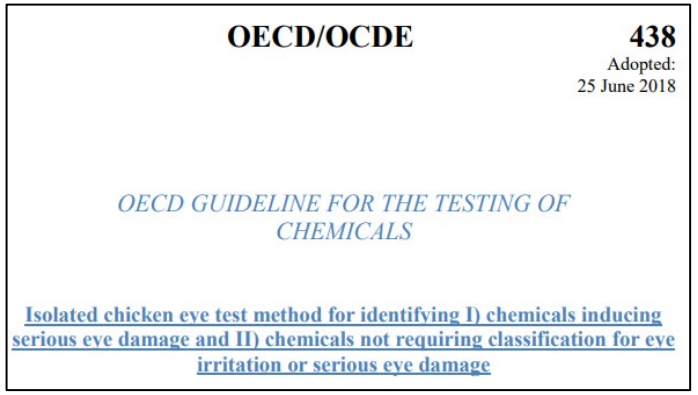
IV III II I



Hazard



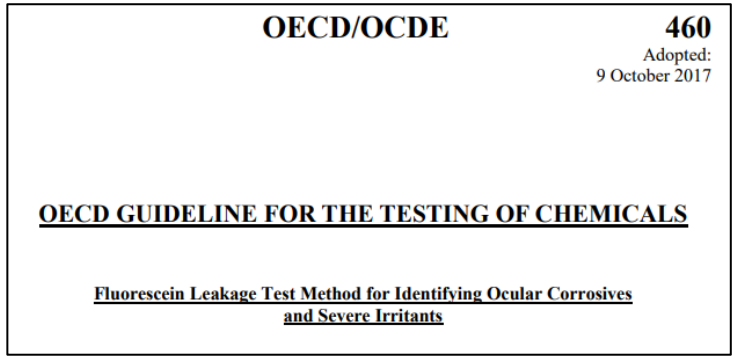




OECD/OCDE **438**
Adopted:
25 June 2018

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

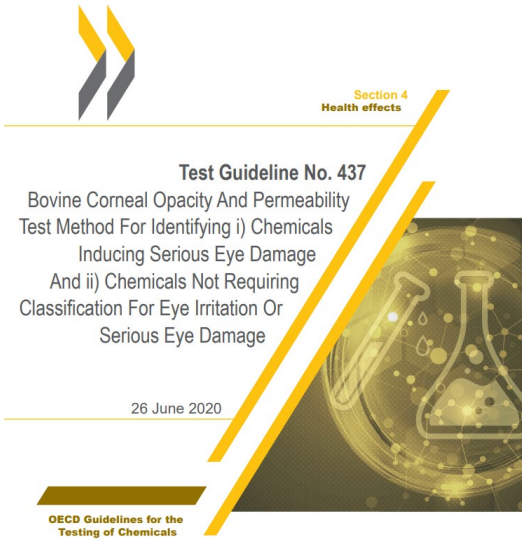
Isolated chicken eye test method for identifying I) chemicals inducing serious eye damage and II) chemicals not requiring classification for eye irritation or serious eye damage

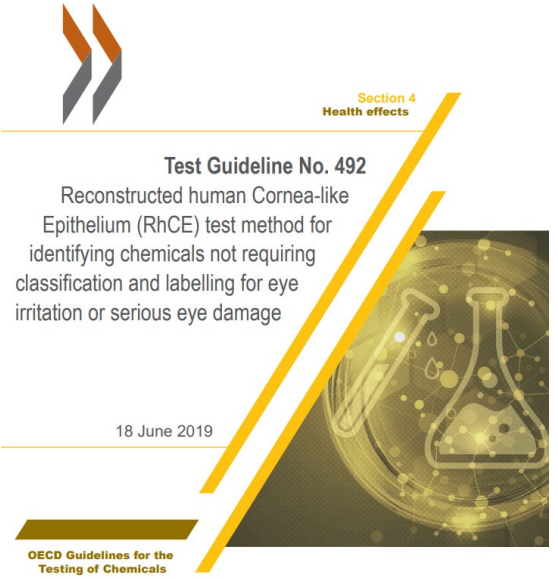


OECD/OCDE **460**
Adopted:
9 October 2017

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Fluorescein Leakage Test Method for Identifying Ocular Corrosives and Severe Irritants



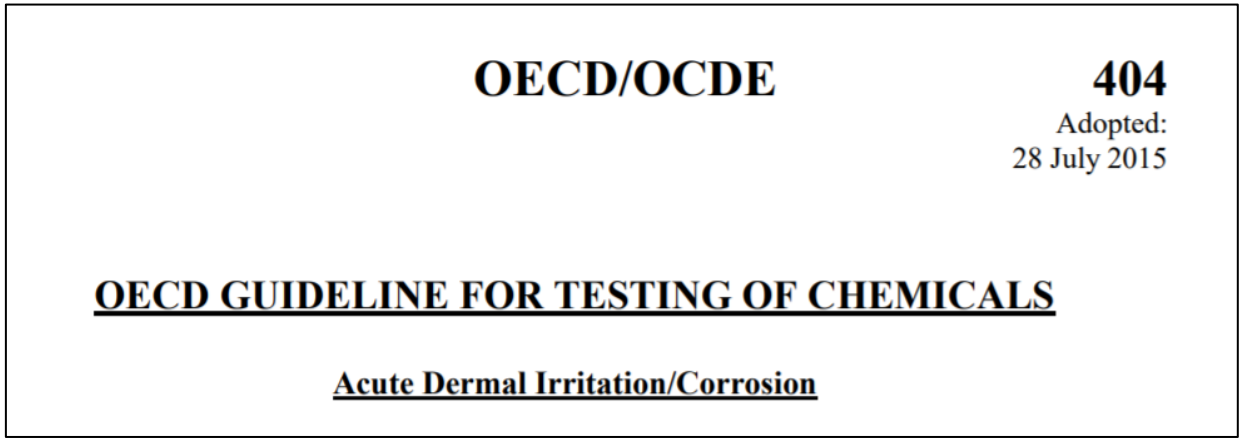


In Vivo Skin Irritation

- EPA Skin Irritation guidelines:
  - Intact skin, fur removed by clipping or shaving.
  - At least 3 animals unless corrosive.
  - 4 hour exposure (recommended).
  - Semiocclusive coverage (recommended).
  - Scoring at 1, 24, 28 and 72 hours after substance removal. Continued monitoring for up to 14 days.
  - Scoring via Draize scale (0-4 for erythema and edema).
  - PDII = average erythema score + average edema score (4 time points: 30-60 min, 24h, 48h and 72h after substance removal)

Erythema and Eschar Formation:	Score
No erythema	0
Very slight erythema (barely perceptible)	1
Well-defined erythema	2
Moderate to severe erythema	3
Severe erythema (beet redness) to slight eschar formation (injuries in depth)	4

Edema Formation:	Score
No edema	0
Very slight edema (barely perceptible)	1
Slight edema (edges of area well defined by definite raising)	2
Moderate edema (raised approximately 1 mm)	3
Severe edema (raised more than 1 mm and extending beyond area of exposure	4



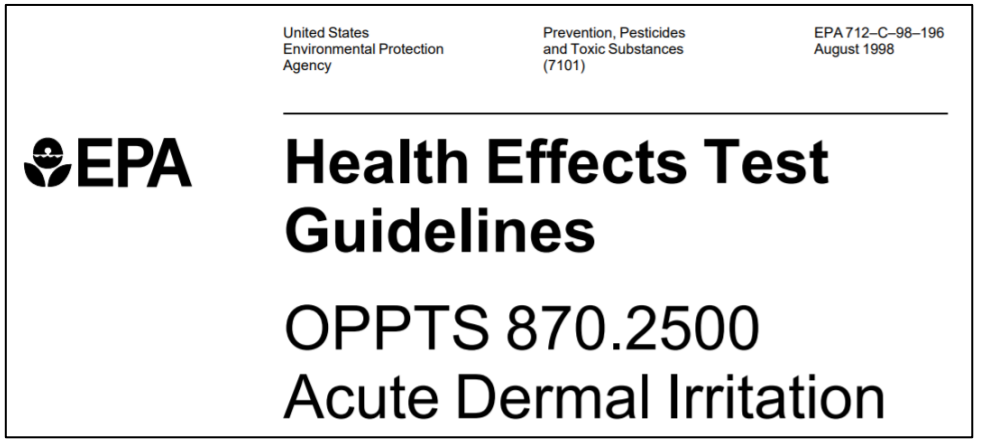
OECD/OCDE

404

Adopted:
28 July 2015

OECD GUIDELINE FOR TESTING OF CHEMICALS

Acute Dermal Irritation/Corrosion



United States Environmental Protection Agency

Prevention, Pesticides and Toxic Substances (7101)

EPA 712–C–98–196 August 1998

EPA

Health Effects Test Guidelines

OPPTS 870.2500
Acute Dermal Irritation

Reproducibility of Categorical Outcomes

Acute Dermal Skin Irritation/Corrosion

	Irritant		Non-irritant	
EPA	**Category I**	**Category II**	**Category III**	**Category IV**
PDII	Corrosive	>5.0	2.1-5.0	0-2.0
Signal Word	DANGER	WARNING	CAUTION	CAUTION
PPE Required	Coveralls worn over long-sleeved shirt and long pants	Coveralls worn over short-sleeved shirt and short pants	Long-sleeved shirt and long pants	Long-sleeved shirt and long pants
	Socks; chemical-resistant footwear	Socks; chemical-resistant footwear	Socks; shoes	Socks; shoes
	Waterproof or chemical-resistant gloves	Waterproof or chemical-resistant gloves	Waterproof or chemical-resistant gloves	No minimum

- ECHA database evaluation
- EPA hazard categories
- 425 substances with at least two studies

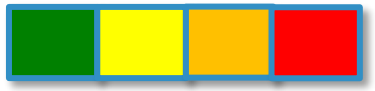
Prior type	I (Corrosive)	II	III	IV	Total Studies
I (Corrosive)	86.3%	4.2%	7.1%	2.5%	207
II	14.1%	44.9%	20.5%	20.5%	35
III	6.9%	5.2%	53.6%	34.3%	133
IV	0.9%	2.0%	9.1%	88.0%	690

Rooney *et al.*, 2021. Reg Tox Pharm 122:104920

OECD Guidelines for in vitro skin irritation testing – assessed based on comparison to the rabbit test…

✓ ? ? ✓

IV III II I



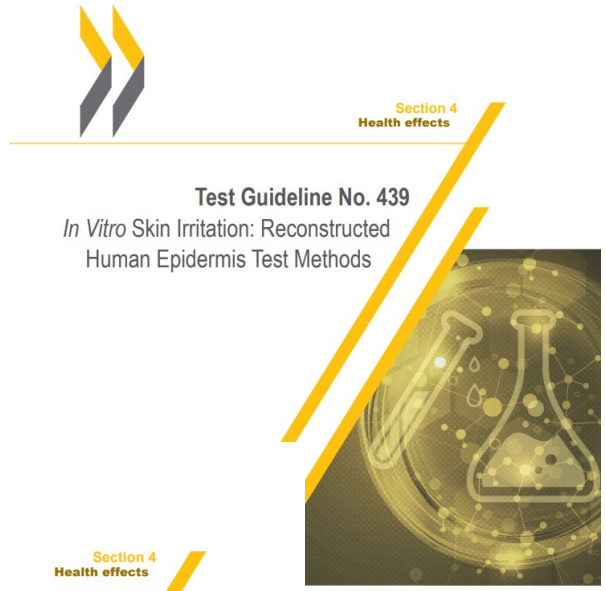
Hazard

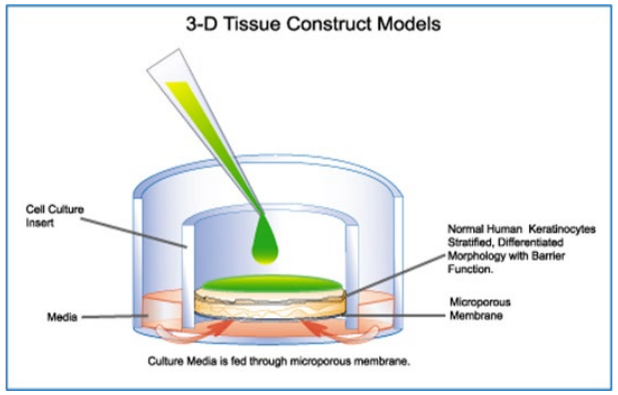
OECD/OCDE **430**

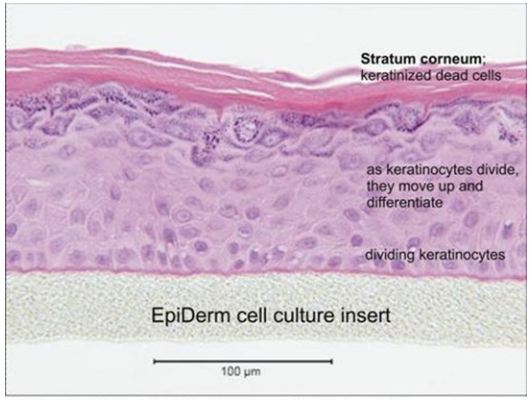
Adopted:
28 July 2015

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

***In Vitro* Skin Corrosion: Transcutaneous Electrical Resistance Test Method (TER)**







EpiDerm™ cell culture insert. Quelle & Rechte: MatTek Corporation

Acute Oral Toxicity Categories

EPA Categories

Hazard

I (≤ 50 mg/kg)

II (>50 ≤ 500 mg/kg)

III (>500 ≤ 5000 mg/kg)

IV (>5000 mg/kg)

GHS Categories

Packing Group

I (≤ 5 mg/kg)

II (>5 ≤ 50 mg/kg)

III (>50 ≤ 300 mg/kg)

IV (>300 ≤ 2000 mg/kg)

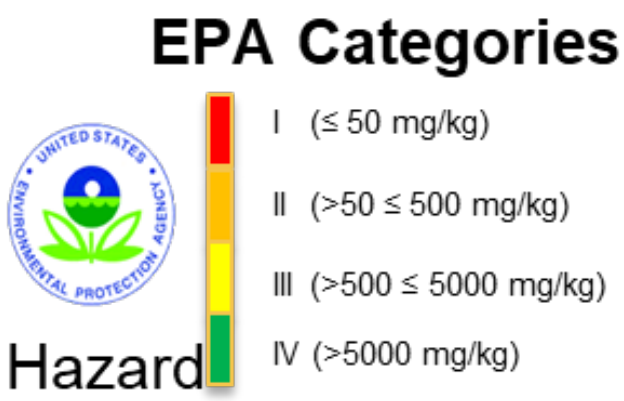
NC (> 2000 mg/kg)

OSHA Hazard

EPA Category	Signal Word	Statement
I (LD_{50} ≤ 50 mg/kg)	Danger/Poison	Fatal if swallowed.
II (50>LD_{50} ≥ 500 mg/kg)	Warning	May be fatal if swallowed.
III (500>LD_{50} ≥ 5000 mg/kg)	Caution	Harmful if swallowed.
IV (LD_{50} > 5000 mg/kg)	Caution (optional)	No statement is required. May use Category III statement

Reproducibility of Categorical Outcomes

Rat Acute Oral Toxicity



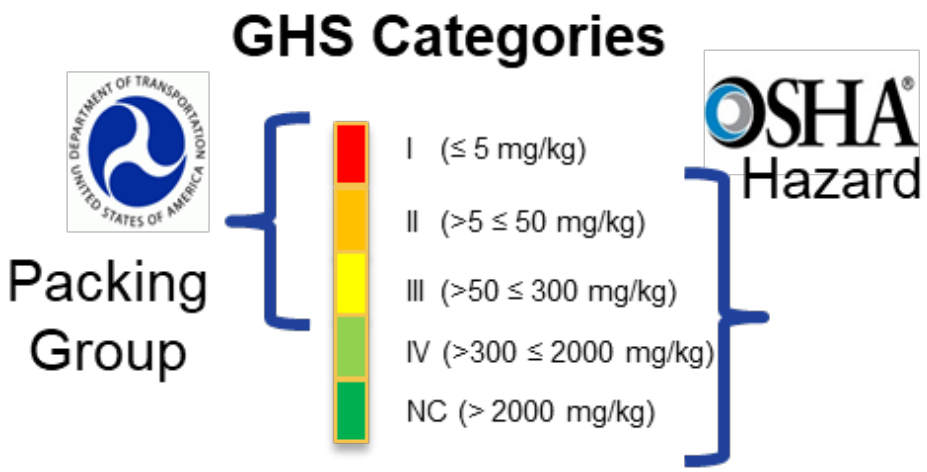
Prior type	I	II	III	IV	Total Studies
I	57.9%	34.5%	6.2%	1.3%	446
II	5.7%	66.5%	27.5%	0.4%	1694
III	0.5%	11%	79.8%	8.7%	4646
IV	0.1%	0.6%	44.7%	54.6%	788

- Comprehensive compilation of data from multiple global resources
- Data heavily curated manually
- Includes limit tests and point estimate data

Karmaus *et al.*, 2022. Tox Sci 188(1)

Reproducibility of Categorical Outcomes

Rat Acute Oral Toxicity

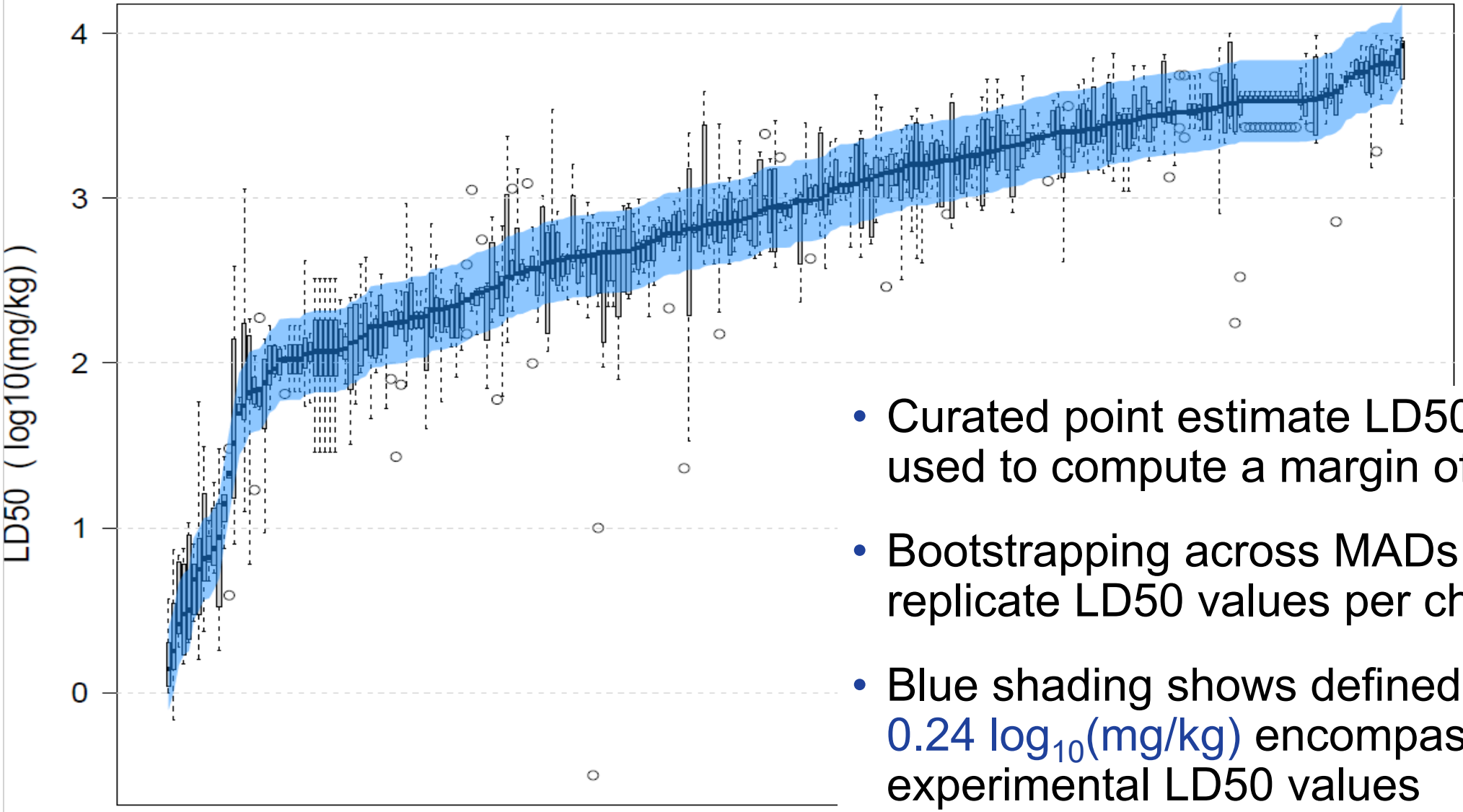


Prior type	1	2	3	4	5	Total Studies
1	53.3%	34.9%	1.5%	5.1%	5.1%	104
2	7.7%	48.9%	33.2%	8.9%	1.3%	342
3	0.2%	7.1%	61.9%	28.9%	1.9%	1166
4	0.1%	1%	11%	66.1%	21.8%	3095
5	0%	0.2%	1%	23.8%	75%	2867

- Comprehensive compilation of data from multiple global resources
- Data heavily curated manually
- Includes limit tests and point estimate data

Karmaus *et al.*, 2022. Tox Sci 188(1)

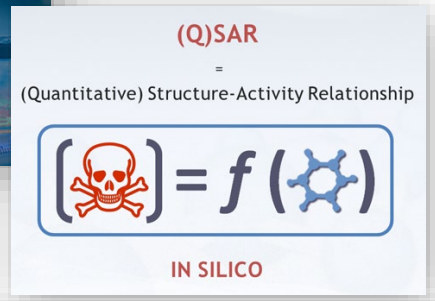
Defining a Margin of Uncertainty

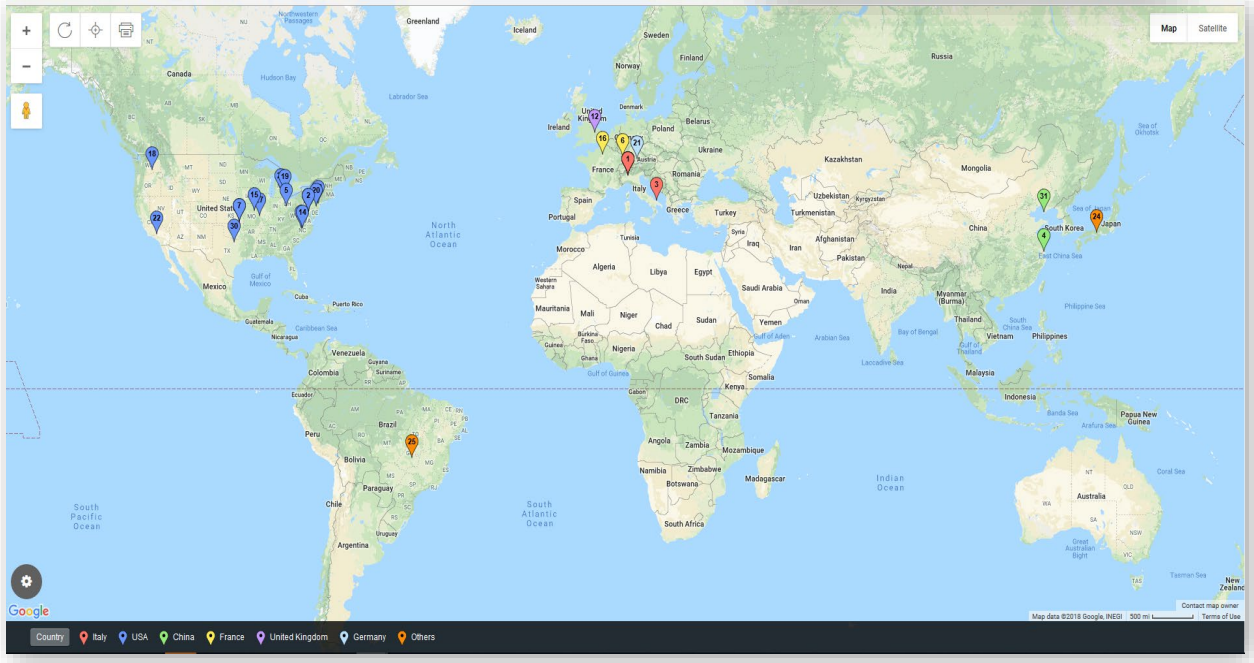


- Curated point estimate LD50 values were used to compute a margin of uncertainty
- Bootstrapping across MADs derived from replicate LD50 values per chemical
- Blue shading shows defined range 0.24 $\log_{10}$(mg/kg) encompasses most experimental LD50 values

Karmaus *et al.*, 2022. Tox Sci 188(1)

Global Crowdsourcing Predictive Models





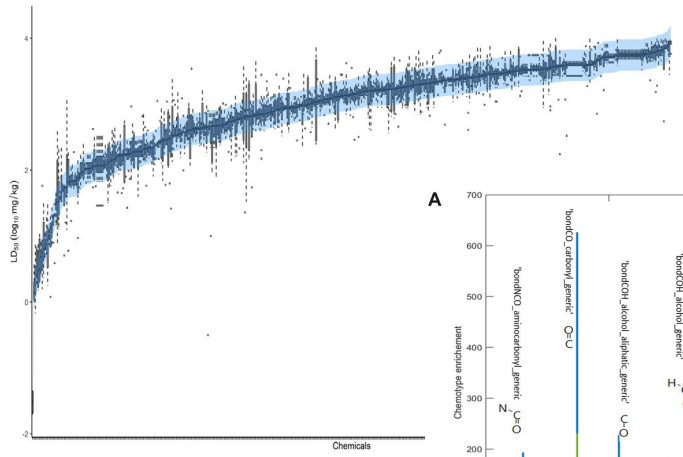
- 35 Groups: academia, industry, govt
- Curate reference data to train & test models: >10k chemicals
- Use molecular structure and chemical properties to predict toxicity
- Combine best models together into “ensemble” approaches
- Accessible via open access AI/ML modeling suite

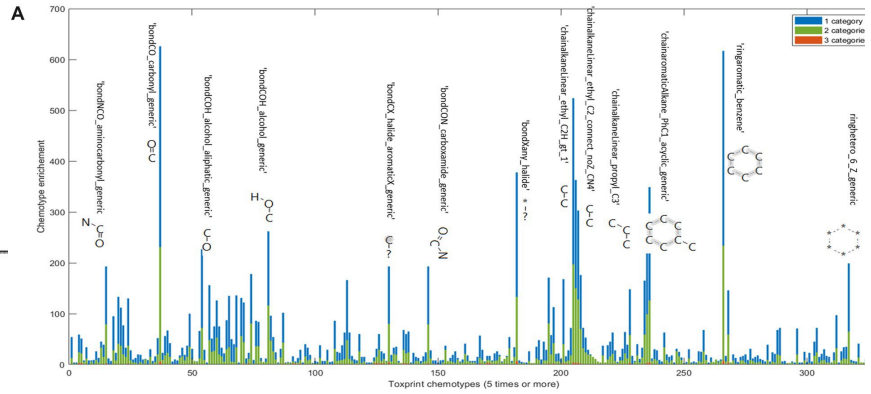
OPERA
OPEn (q)saR App

https://github.com/NIEHS/OPERA

Kleinstreuer et al. Comp Tox (2018); Mansouri et al. J Cheminform (2018), Env Health Persp (2020, 2022)

Applying Variability to Model Evaluation and Predictions





Analyzing sources of variability in acute oral toxicity data & applying 95% confidence interval to predictions

	0	5	50	300	500	2000	5000 mg/kg
VT	0	0	1	1	1	1	1
NT	1	1	1	1	1	0	0
EPA	0	0	1	1	0	0	0
GHS	0	0	1	0	0	0	0
LD50	0	0	1	1	1	0	0
WoE	1	1	5	4	3	1	1

160 (-0.3) 316 (+0.3) 613

Collaborative Acute Toxicity Modeling Suite (CATMoS) Performance

	Very Toxic		Non-Toxic		EPA		GHS	
	Train	Eval	Train	Eval	Train	Eval	Train	Eval
Sensitivity	0.87	0.70	0.88	0.67	0.81	0.62	0.80	0.58
Specificity	0.99	0.97	0.97	0.90	0.92	0.86	0.95	0.90
Balanced Accuracy	0.93	0.84	0.92	0.78	0.87	0.74	0.88	0.74
***In vivo* Balanced Accuracy**	0.81		0.89		0.82		0.79	

	LD50 values		LD50 values
	Train	Eval	***In Vivo***
R2	0.85	0.65	0.80
RMSE	0.30	0.49	0.42

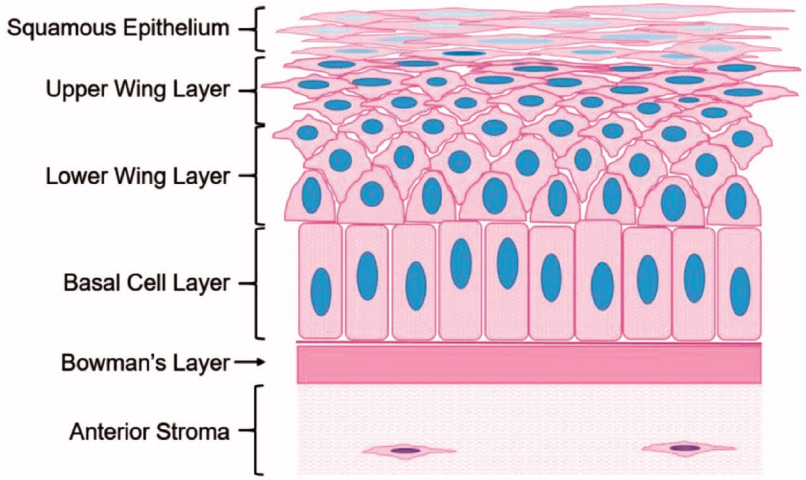
CATMoS QSAR predictions perform just as well as replicate *in vivo* data at predicting oral acute toxicity outcome

Karmaus et al. Toxicol Sci. 2022; Mansouri et al. EHP 2021

Using mechanistic information and human relevance

Clippinger et al. 2021 Cut Ocu Tox

(a)



Superficial Conjunctival or Corneal Epithelium (Figure 2c)
3D Reconstructed Human Cornea-like Epithelial Tissue
EYEIRR-IS
Vitrigel-Eye Irritancy
Bovine Corneal Opacity and Permeability
Isolated Chicken Eye
Isolated Rabbit Eye
Porcine Cornea Opacity Reversibility Assay
Ex Vivo Eye Irritation Test (EVEIT)
Fluorescein Leakage
Short Time Exposure
Neutral Red Release
Cytosensor Microphysiometer
Ocular Irritection
OptiSafe

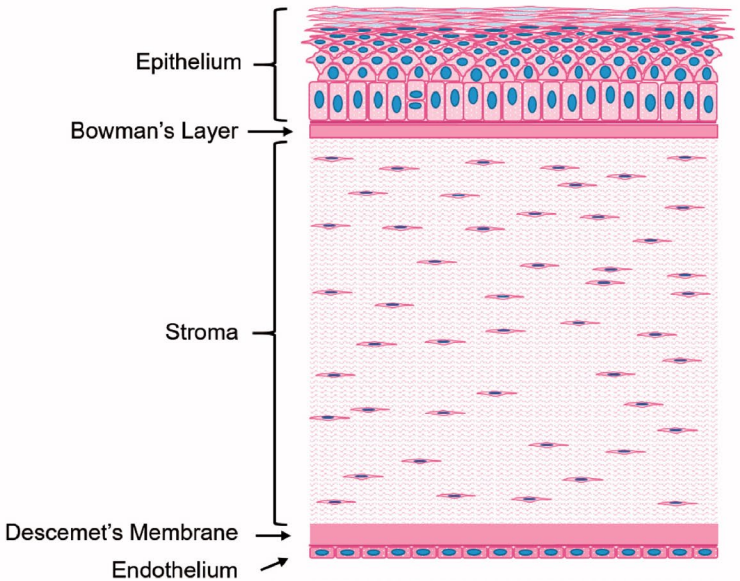
Wing Cell Layer of the Epithelium (Figure 2d)
3D Reconstructed Human Cornea-like Epithelial Tissue
EYEIRR-IS
Vitrigel-Eye Irritancy
Bovine Corneal Opacity and Permeability
Isolated Chicken Eye
Isolated Rabbit Eye
Porcine Cornea Opacity Reversibility Assay
Ex Vivo Eye Irritation Test (EVEIT)
Ocular Irritection
OptiSafe

Lower Wing Cell and Basal Cell Layers of the Epithelium (Figure 2e)
3D Reconstructed Human Cornea-like Epithelial Tissue
EYEIRR-IS
Bovine Corneal Opacity and Permeability
Porcine Cornea Opacity Reversibility Assay
Isolated Chicken Eye
Isolated Rabbit Eye
Ex Vivo Eye Irritation Test (EVEIT)
Ocular Irritection
OptiSafe

Consider strengths and limitations of all available methods with respect to:

- their relevance to human ocular anatomy
- the mechanisms of eye irritation/corrosion in humans

(b)



Corneal Stroma (Figure 2f)
Bovine Corneal Opacity and Permeability
Isolated Chicken Eye
Isolated Rabbit Eye
Ex Vivo Eye Irritation Test (EVEIT)
Ocular Irritection
OptiSafe

Corneal Endothelium (Figure 2g)
Bovine Corneal Opacity and Permeability
Isolated Chicken Eye
Isolated Rabbit Eye
Ex Vivo Eye Irritation Test (EVEIT)

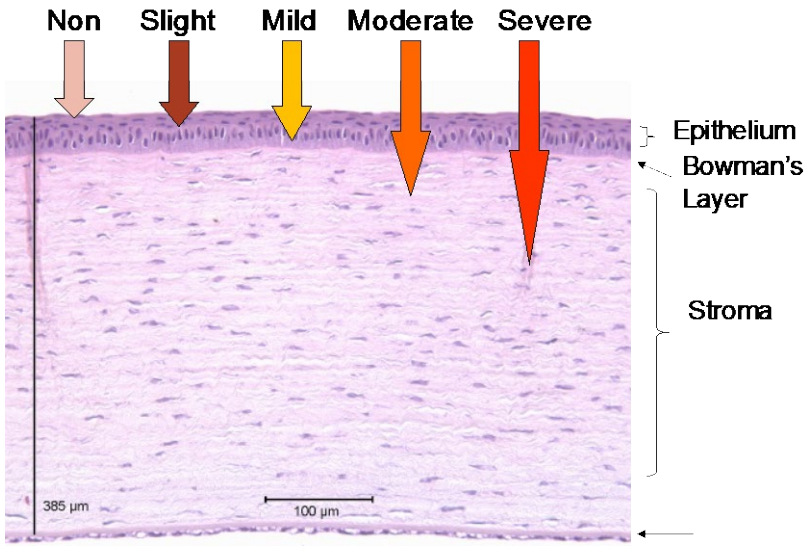
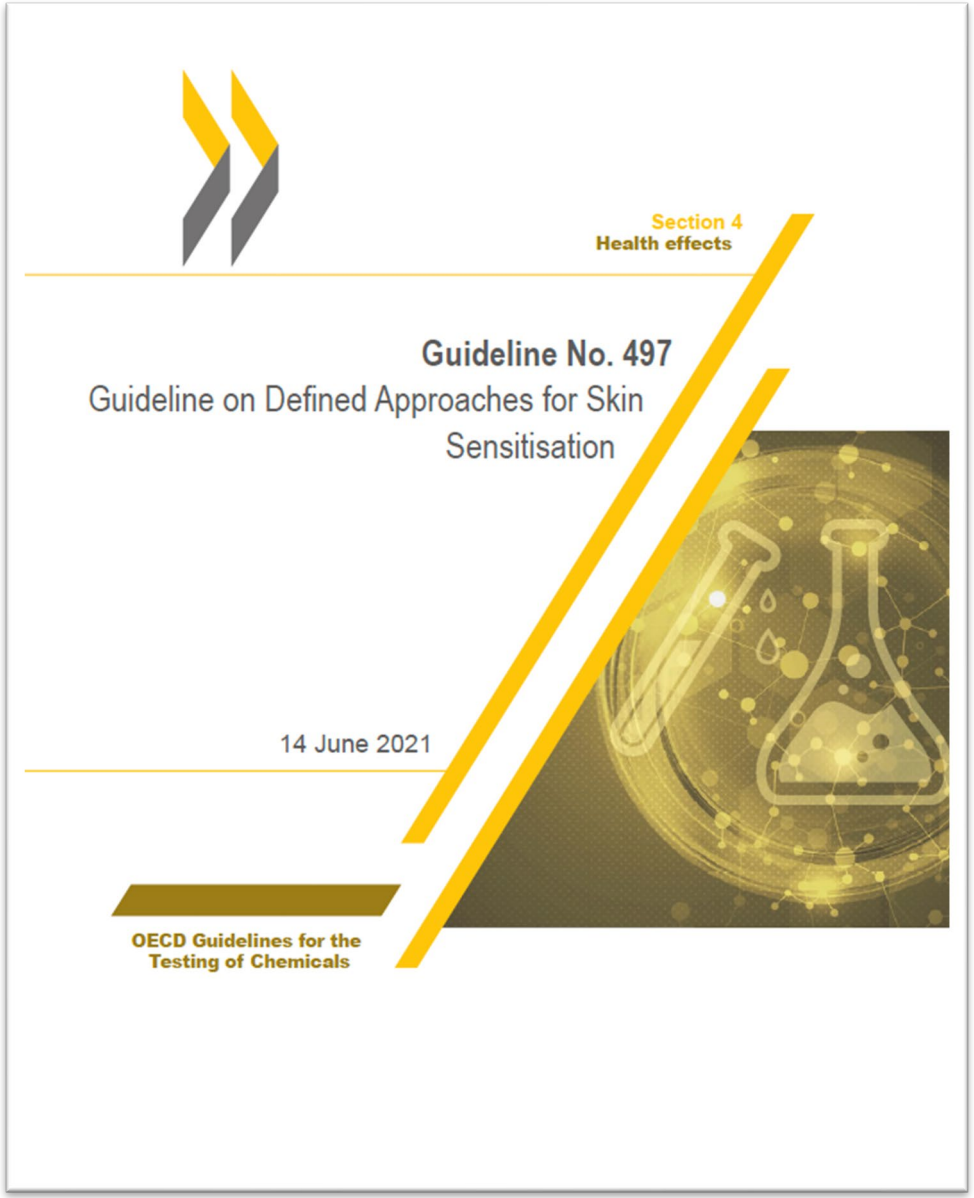
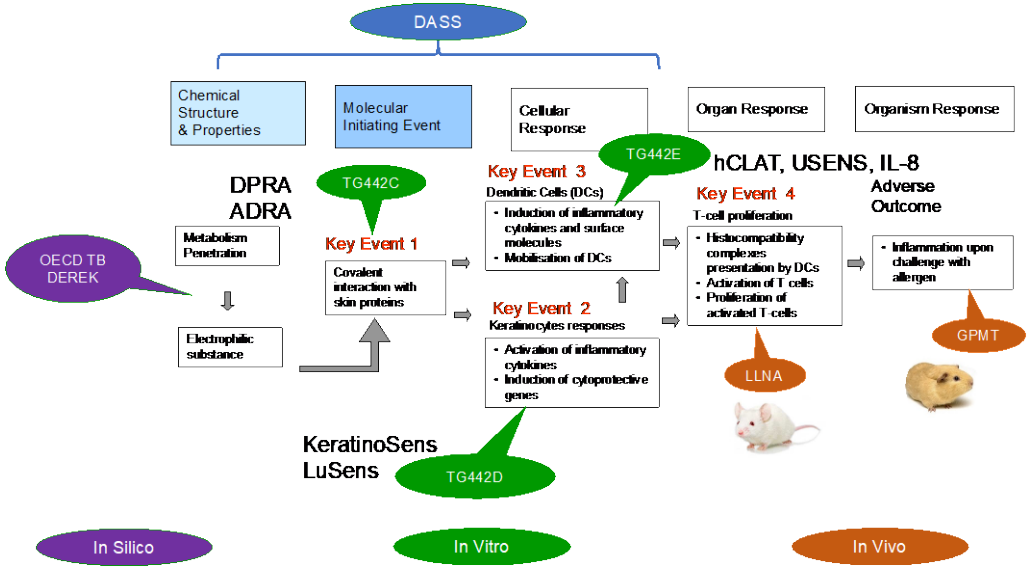


Image modified from Scott, et al., 2010

Defined Approaches for Skin Sensitization Guideline





DA/Method	Information Sources	Capability (Hazard and/or Potency)	Hazard Performance vs. LLNA N~168	Hazard Performance vs. Human N~63	GHS Potency Performance vs. LLNA (Accuracy)	GHS Potency Performance vs. Human (Accuracy)
2o3 DA	DPRA, KeratinoSens™, h-CLAT	Hazard	84% BA, 82% Sens, 85% Spec	88% BA, 89% Sens, 88% Spec	-	-
ITSv1 DA	DPRA, h-CLAT, DEREK Nexus v6.1.0	Hazard, Potency (GHS)	81% BA, 92% Sens, 70% Spec	69% BA, 93% Sens, 44% Spec	70% NC, 71% 1B, 74% 1A	44% NC, 77% 1B, 65% 1A
ITSv2 DA	DPRA, h-CLAT, OECD QSAR Toolbox v4.5	Hazard, Potency (GHS)	80% BA, 93% Sens, 67% Spec	69% BA, 94% Sens, 44% Spec	67% NC, 72% 1B, 72% 1A	44% NC, 80% 1B, 67% 1A
LLNA (provided for comparison)	*in vivo*	Hazard, Potency	-	58% BA, 94% Sens, 22% Spec	-	25% NC, 74% 1B, 56% 1A

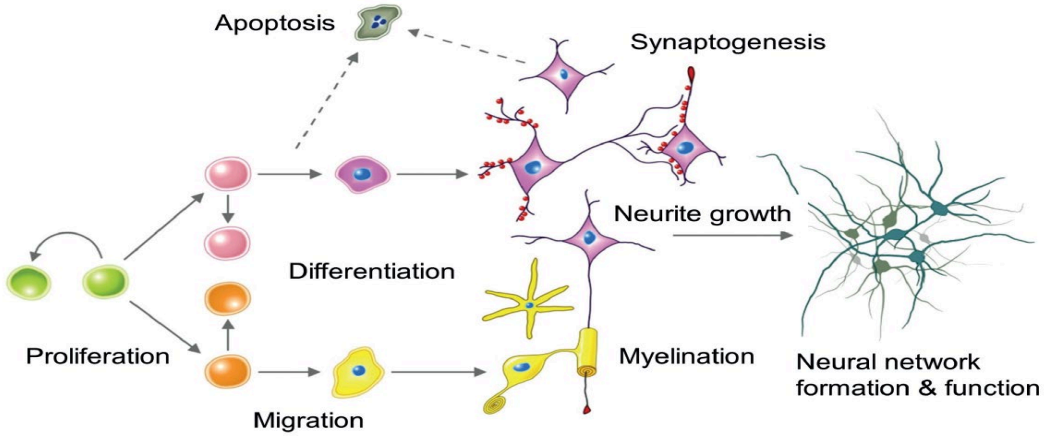
Test Readiness Criteria of NAMs for DNT

Human Relevance Consideration

Phase I	Max. score	UKN2 cMINC	Phase II	Max. score	UKN2 cMINC	Phase III (optional)	max. score	UKN2 cMINC
1 Test system	10	9	8 Testing strategy	4	3	*13 Screening hits*	*4*	*4*
2 Exposure scheme	3	3	9 Robustness	4	3	Score 0 = D		
3 Documentation/SOP	5	5	10 Test benchmarks	4	4	Score 1 = C		
4 Main endpoints	4	4	11 Prediction model	4	3	Score 2 = B		
5 Cytotoxicity	5	5	12 Applicability domain	3	1	Score 3 - 4 = A		
6 Test method controls	4	4						
7 Data evaluation	4	4						
Sum	**35**	**34**	**Sum**	**19**	**14**	**Sum**		**4**

The scores of the different phases are evaluated and result in the ranks of readiness

Phase I		Phase II			
Score	Grading	Score	Grading	Explanation of grading	
< 7	D	< 4	D	D	Not ready at all
8 - 17	C	5 - 9	C	C	Substantial improvements required to be ready
18 - 28	B	10 - 14	B	B	Improvements required to be ready
29 - 35	A	15 - 19	A	A	Test method is close to ready or ready



Criteria	Description
1 Test system	
1a What is modelled	Is there a clear rationale given for what target organ/tissue relevant for human poisoning/pathology the test systems should reflect
1b Relevance	Is the chosen test system known to be a key component in pathogenesis, or why is it thought to reflect a key component, mechanism or tissue
1c System uncertainties and human correlate (HC)	(i) Is there a discussion on where the test system differs from the mimicked human tissue, and which gaps of analogy need to be considered? (ii) Do toxicant-altered genes (or other biomarkers) correspond to changes in mimicked human tissue (after poisoning or in relevant pathologies)

Is the target organ/tissue relevant for human poisoning/pathology?

Are correlation/differences to human tissue discussed?

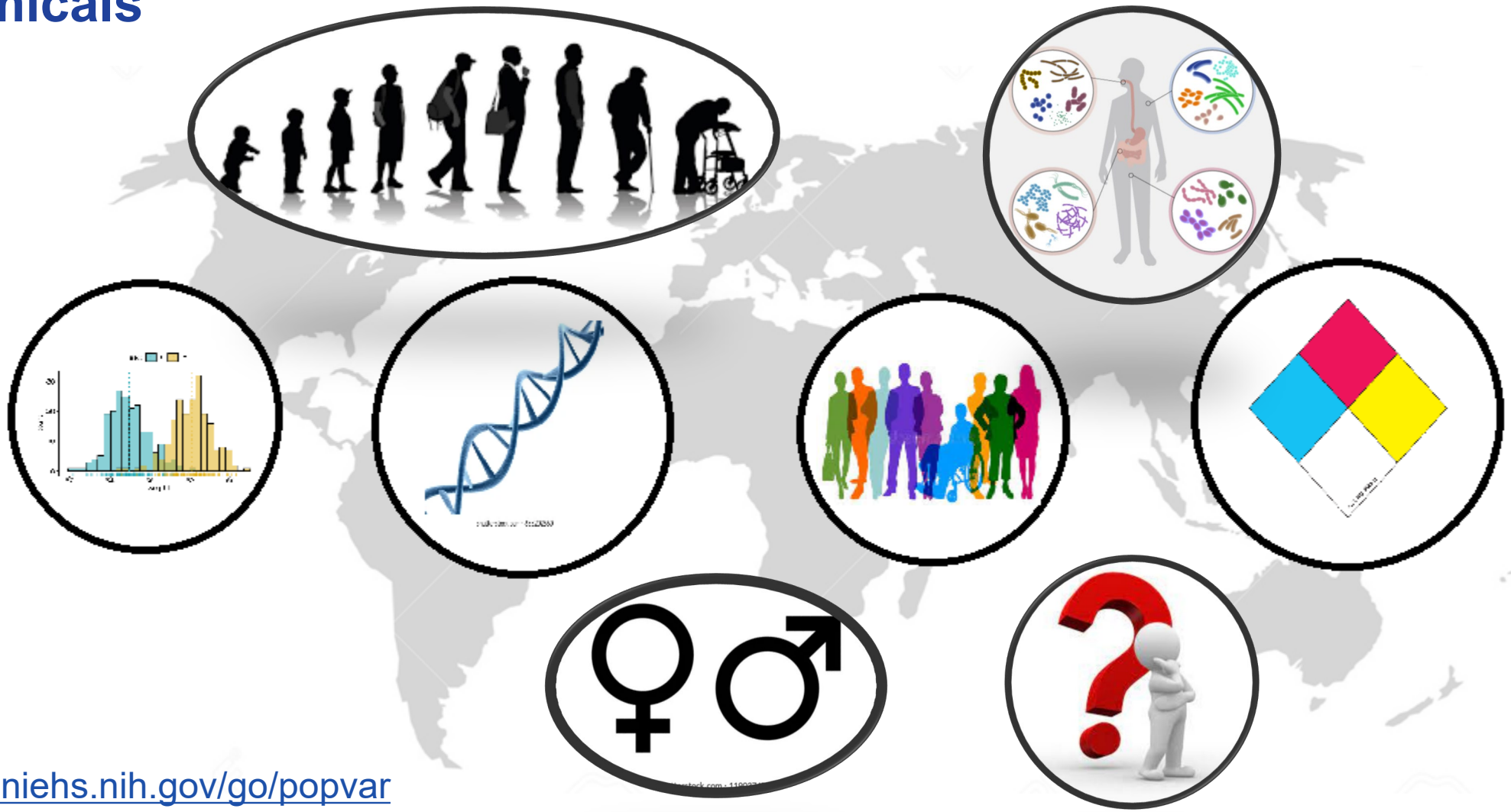
*OECD IATA Case Study Published Sept. 2022

Summary

- In vivo data have been used to derive thresholds for hazard categorization, precautionary labeling, and perform quantitative risk assessments
- Establishing confidence in NAMs should include considerations of variability in in vivo test methods
- In vivo variability should also be considered to determine if concordance with NAMs is an appropriate comparison
- Mechanistic relevance to humans should also be carefully considered to adequately determine confidence.

Workshop: Oct 26 – 27, 2022

Identify opportunities and needs for NAMs to provide relevant information on population variability and susceptibility to environmental chemicals



https://ntp.niehs.nih.gov/go/popvar

Acknowledgments

The NICEATM Group

- ICCVAM Agencies
- EPA Partners
- OECD Secretariat/WGs
- NGO Collaborators

Integrated Chemical Environment

Report for 2020-2021 is out now!

https://ntp.niehs.nih.gov/go/2021iccvamreport

Subscribe to NICEATM News email list
https://ntp.niehs.nih.gov/go/niceatm

Qualitative and Quantitative Variability of Repeat Dose Animal Toxicity Studies

Katie Paul Friedman, PhD

paul-friedman.katie@epa.gov

Toxicologist, Center for Computational Toxicology and Exposure,

Office of Research and Development, US EPA

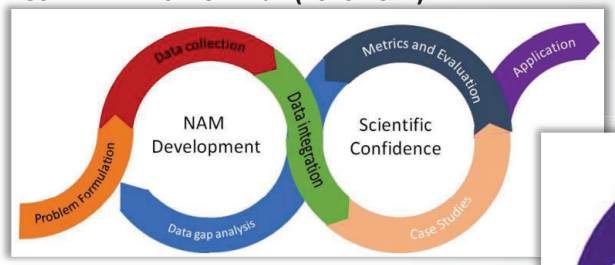
October 12, 2022
EPA NAM Conference 2022

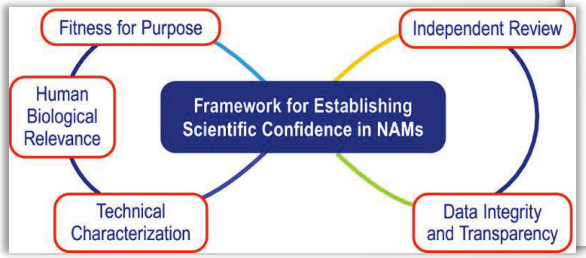
The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA

Variability of *in vivo* repeat dose data informs NAM performance expectations and a part of scientific confidence

- In Section 4(h) in the Lautenberg amendment to Toxic Substances Control Act:
  - *"…Administrator shall reduce and replace, to the extent practicable and scientifically justified…the use of vertebrate animals in the testing of chemical substances or mixtures…"*
  - New approach methods (NAMs) need to provide *"information of equivalent or better scientific quality and relevance…"* than the traditional animal models
- Multiple frameworks suggest scientific confidence may depend in part on fitness for purpose, biological relevance, and characterization of NAM performance, which in some cases relates to traditional animal study performance or reference data.

US EPA NAMs WorkPlan (2020-2021)





van der Zalm et al. (2022). 10.1007/s00204-022-03365-4

Parish et al. (2020). 10.1016/j.yrtph.2020.104592

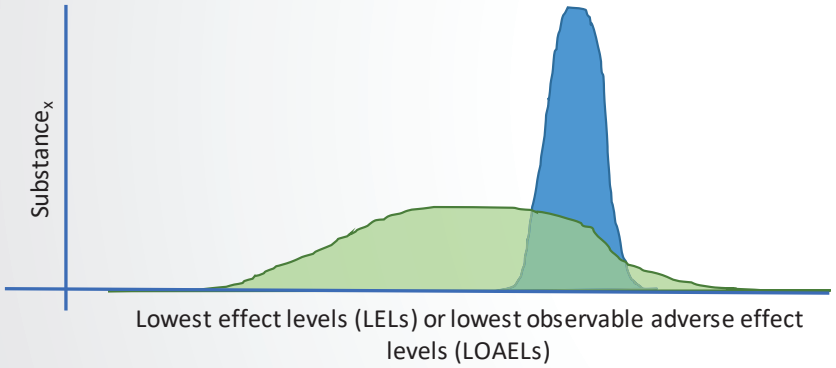
How do we define expectations of *in silico, in chemico,* and *in vitro* models for predicting repeat-dose toxicity?

***In silico, in chemico,* and *in vitro* models cannot predict *in vivo* systemic effect values from animal studies with greater accuracy than those animal models reproduce themselves.**

How can variability in traditional animal studies be expressed for use as reference or training data?

Quantitative: variance is a measure of how far values are spread from the average.

We need to know what the "spread" or variability of traditional effect levels might be to know the range of acceptable or "good" values from a NAM.



Qualitative: We need to know if a specific effect is always observed or not.

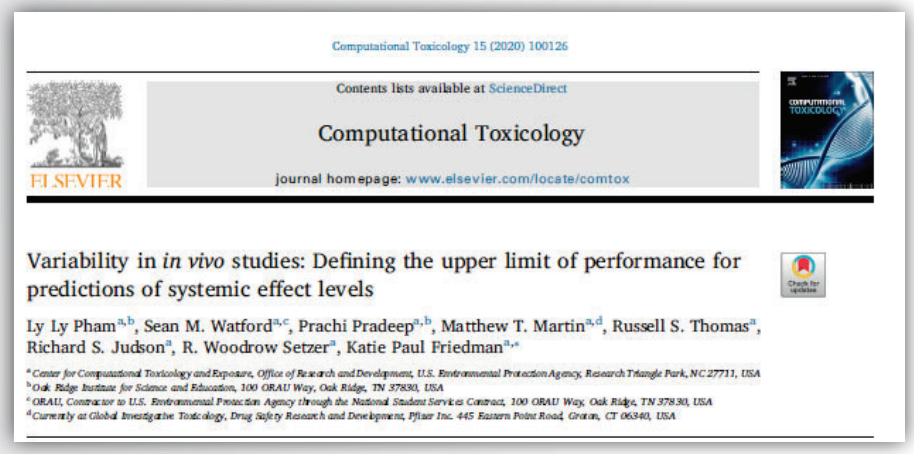
We need to know something about classification performance or about reference data for a phenotype.

		"Truth" (traditional toxicology)	
		Negative	Positive
Predicted (NAM)	Negative	True negative	False negative
	Positive	False positive	True positive

If we are going to learn from variable and uncertain data, we will propagate this variability and uncertainty to any NAMs developed.

If we are going to evaluate NAM performance based on comparison to *in vivo* data, we should account for variability and uncertainty in these reference data.

Part I: Benchmarks on quantitative reproducibility of systemic findings in repeat dose animal studies



Computational Toxicology 15 (2020) 100126

Contents lists available at ScienceDirect

Computational Toxicology

journal homepage: www.elsevier.com/locate/comtox

Variability in *in vivo* studies: Defining the upper limit of performance for predictions of systemic effect levels

Ly Ly Pham[a,b], Sean M. Watford[a,c], Prachi Pradeep[a,b], Matthew T. Martin[a,d], Russell S. Thomas[a], Richard S. Judson[a], R. Woodrow Setzer[a], Katie Paul Friedman[a,*]

[a] Center for Computational Toxicology and Exposure, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA
[b] Oak Ridge Institute for Science and Education, 100 ORAU Way, Oak Ridge, TN 37830, USA
[c] ORAU, Contractor to U.S. Environmental Protection Agency through the National Student Services Contract, 100 ORAU Way, Oak Ridge, TN 37830, USA
[d] Currently at Global Investigative Toxicology, Drug Safety Research and Development, Pfizer Inc. 445 Eastern Point Road, Groton, CT 06340, USA

Pham LL, Watford S, Pradeep P, Martin MT, Thomas RS, Judson RS, Setzer RW, Paul Friedman K. 2020. 10.1016/j.comtox.2020.100126

Primary Research Question	Statistical approaches
What is the range of possible effect values (mg/kg/day) in replicate studies for a given chemical?	• Residual root mean square error (RMSE) is an estimate of variance in the same units as the systemic effect values. • The RMSE can also be used to define a minimum prediction interval, or estimate range, for a model.
What is the maximal accuracy of a new model that attempts to predict effect values for a chemical?	• The mean square error (MSE) is used to approximate the unexplained variance (not explained by study descriptors). • This unexplained variance limits the R-squared on a new model.

Based on the study descriptors in ToxRefDB v2.0, we developed statistical models of the variance in quantitative systemic effect level values.

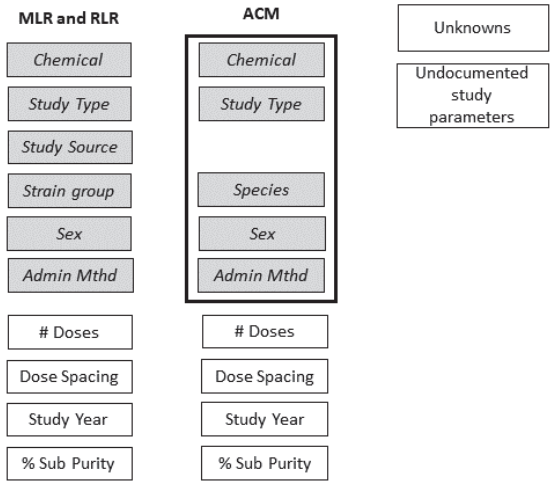
Total variance | **Approximated by mean square error**

Observed Variance (LEL or LOAELs)	=	Variance Explained by Study Parameters	+	Unexplained Variance

Chemical → Study → Observed Effect Level

Study	Effect	Treatment Related Effect* (mg/kg/day)	Critical Effect**
1	Body weight	5[+]	0
1	Liver	15[++]	1
1	kidney	20	0
1	heart	10	0

* Dose at which a treatment related effect was observed
** Expert driven designation
[+] Observed effect level used in LEL dataset
[++] Observed effect level used in LOAEL dataset



Using two approaches:

	Multilinear regression (MLR, RLR)	Augmented cell means (ACM)
Aggregation level	Chemical	Chemical-Study Type-Species-Sex-Admin Method combination
Replicate definition stringency	Not stringent	Stringent
N	Maximized; ↓ impact of outliers/database error rate	Small; may bias variance estimate
Study descriptors	Contribute independently to variance	Accounts for possible interactions among descriptors

Figure 2. Statistical model of the variance. *LEL = lowest effect level; LOAEL = lowest observable adverse effect level. The LEL is the lowest treatment-related effect observed for a given chemical in a study, and the LOAEL is defined by expert review as coinciding with the critical effect dose level from a given study. Multiple studies for a given chemical yield multiple LELs and LOAELs for computation of variance. MLR = multilinear regression; RLR = robust linear regression; ACM = augmented cell means; Adm. Method = administration method; % Sub Purity = % substance purity used in the study. The gray shaded study descriptor boxes are categorical variables, and the white study descriptor boxes are continuous variables. The box around five categorical study descriptors for the ACM indicates these were concatenated to a factor to define study replicates.*

Pham LL, Watford S, Pradeep P, Martin MT, Thomas RS, Judson RS, Setzer RW, Paul Friedman K. 2020. 10.1016/j.comtox.2020.100126

Variance results suggest that repeat dose studies for regulatory toxicology, as conducted and curated, may have inherent irreducible amount of unexplained variance.

- 28 different statistical models were constructed.
- RMSE is used to define a 95% minimum prediction interval (i.e., based on the standard deviation or spread of the residuals).
- The % explained variance (amount explained by study descriptors) likely approaches 55-73%.
- This means that the R^2 on some new, predictive model would approach 0.55 to 0.73 as an upper bound on accuracy.

	Total Variance (log_{10} mg/kg/day)^2	Unexplained Variance (MSE) (log_{10} mg/kg/day)^2	RMSE (log_{10} mg/kg/day)	% explained variance	Minimum prediction interval (log_{10}-mg/kg/day)
Range	0.744 - 1.013	0.2 - 0.395	0.448 - 0.629	54.9 - 73.3	± 0.878 - ± 1.23
Median (MAD)	0.825 (0.065)	0.301 (0.068)	0.549 0.061	66.1 4.89	± 1.07 (0.12)
Mean (SD)	0.838 (0.070)	0.300 (0.055)	0.545 (0.050)	65.3 (4.86)	± 1.07 (0.098)

Based on tables from Pham LL, Watford S, Pradeep P, Martin MT, Thomas RS, Judson RS, Setzer RW, Paul Friedman K. 2020. 10.1016/j.comtox.2020.100126

Table 3
Comparison of performance of the current model with previous publications.

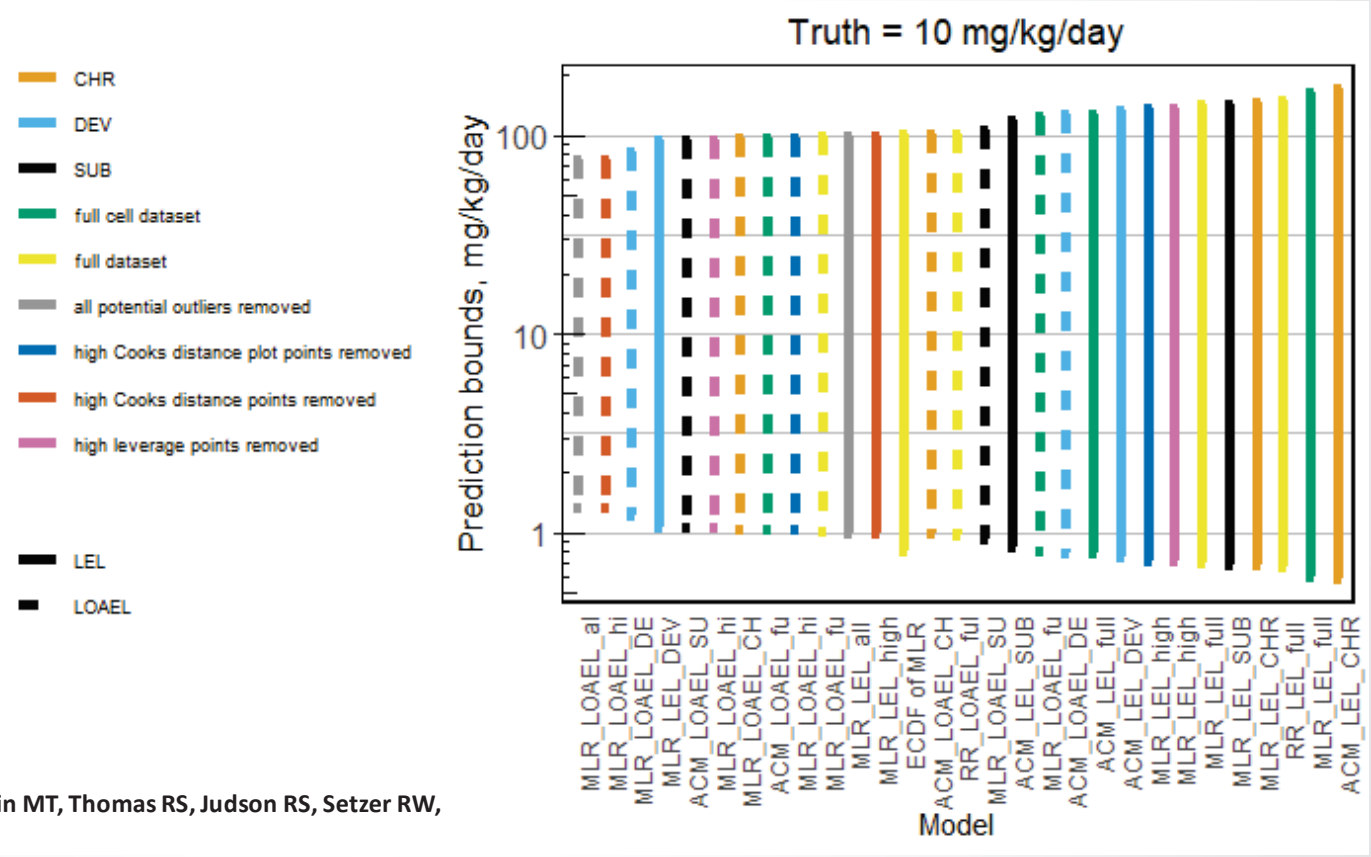
Study	Reference	Number of chemicals	RMSE (log_{10}-mg/kg/day)	R^2
Current	Current	3592	0.70	0.57
Mumtaz et al.	[16]	234	0.41	0.84
Hisaki et al.	[17,18]	421	0.53, 0.56, 0.51	–
Toropova et al.	[19]	218	0.51–0.63	0.61–0.67
Veselinovic et al.	[20]	341	0.46–0.76	0.49–0.70
Novotarskyi et al.	[22]	1,854	1.12 ± 0.08	0.31
Truong et al.	[24]	1247	0.69	0.43

- A multi-linear regression QSAR model of chronic oral rat LOAEL values for approximately 400 chemicals, demonstrated a RMSE of 0.73 log_{10}(mg/kg-day) which was similar to the size of the variability in the training data, ±0.64 log_{10}(mg/kg-day), suggested that the error in the model approached the error in the reference data from different laboratories (Mazzatorta et al. 2008; Helma et al. 2018).

Pradeep P, Paul Friedman K, Judson RS. (2020). 10.1016/j.comtox.2020.100139

Range of 95% minimum prediction intervals across the modeling approaches, effect levels, and study types is 58-284-fold

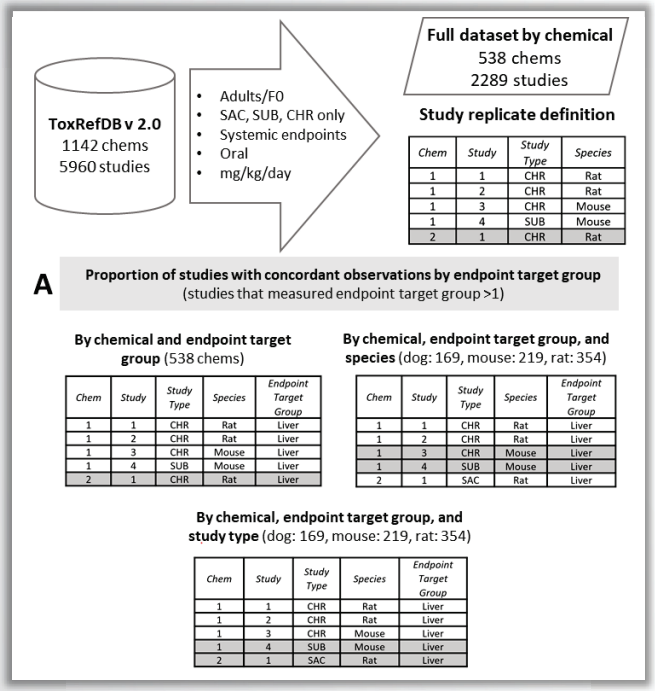
If attempting to use a NAM-based predictive model for prediction of a reference systemic effect level value of 10 mg/kg/day, it is likely that given the variability in reference data of this kind, that a model prediction of somewhere between 1 and 100 mg/kg/day would be the greatest amount of accuracy achievable.



Based on tables from Pham LL, Watford S, Pradeep P, Martin MT, Thomas RS, Judson RS, Setzer RW, Paul Friedman K. 2020. 10.1016/j.comtox.2020.100126

How reproducible are organ level effects in replicate studies and studies of different duration?

A. What is the reproducibility of systemic findings in repeat dose animal studies?



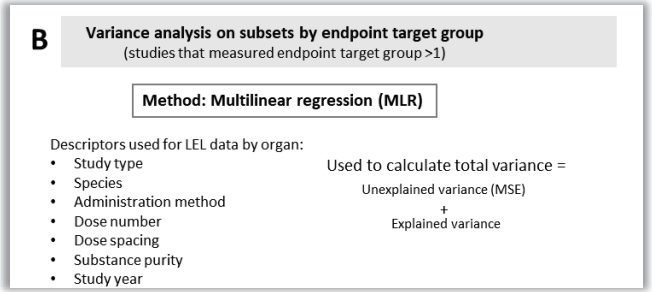
Chem	Study	Study Type	Species
1	1	CHR	Rat
1	2	CHR	Rat
1	3	CHR	Mouse
1	4	SUB	Mouse
2	1	CHR	Rat

Chem	Study	Study Type	Species	Endpoint Target Group
1	1	CHR	Rat	Liver
1	2	CHR	Rat	Liver
1	3	CHR	Mouse	Liver
1	4	SUB	Mouse	Liver
2	1	CHR	Rat	Liver

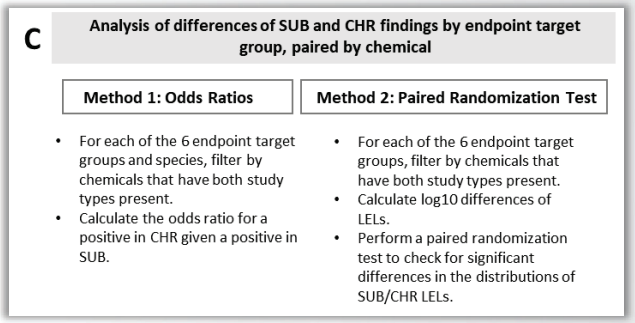
Chem	Study	Study Type	Species	Endpoint Target Group
1	1	CHR	Rat	Liver
1	2	CHR	Rat	Liver
1	3	CHR	Mouse	Liver
1	4	SUB	Mouse	Liver
2	1	SAC	Rat	Liver

Chem	Study	Study Type	Species	Endpoint Target Group
1	1	CHR	Rat	Liver
1	2	CHR	Rat	Liver
1	3	CHR	Mouse	Liver
1	4	SUB	Mouse	Liver
2	1	SAC	Rat	Liver

B. Are variance estimates reduced for organ-level effects only in repeat dose animal studies, using LELs, BMDs, etc.?



C. Understanding NAM alternatives are not necessarily 1:1 replacements, would estimates of subchronic and chronic effect levels be necessary?



A: How qualitatively reproducible are organ level findings in repeat dose studies?

Primary Research Question	Statistical approaches
How concordant are organ-level effects for multiple repeat dose study observations?	Calculate concordance of findings between replicate studies when grouped by chemical and organ; chemical, organ, and species; and chemical, organ, and study type

$$\%Concordance = \frac{chemical\ with\ positive\ finding\ in\ all\ studies + chemicals\ with\ negative\ finding\ in\ all\ studies}{total\ chemicals\ tested}$$

- Qualitative reproducibility of organ-level effect observations in repeat dose studies of adult animals was 33-88%, depending on grouping.
- Organs associated with more negative chemicals (stomach, thyroid, adrenal) had higher rates of concordance.
- Within-species concordance tended to be greater than within-study concordance.

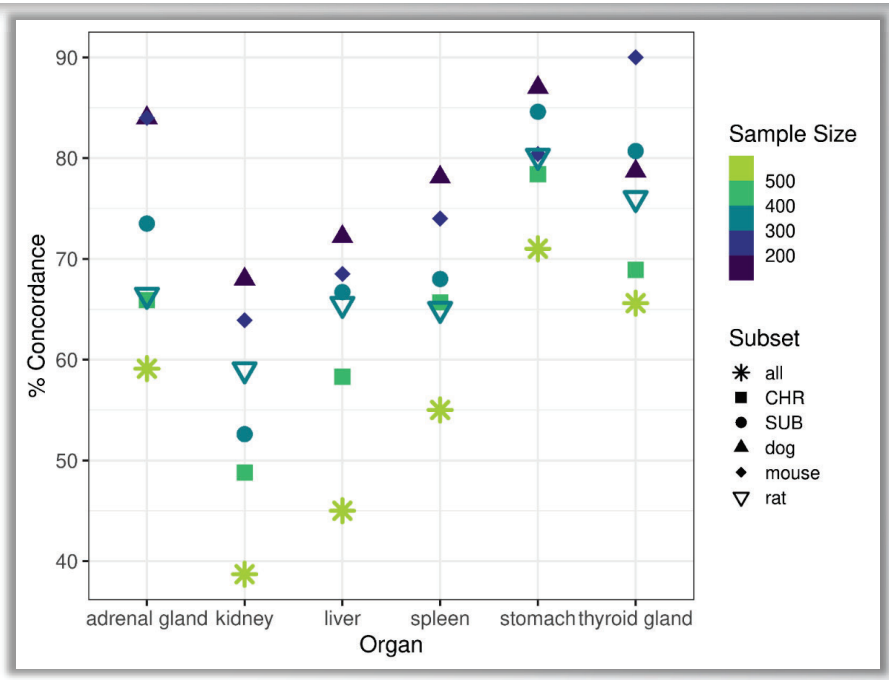


Figure 2, Paul Friedman et al. (in prep).

Indeed, previous literature reports suggest variable inter-species concordance of carcinogenic findings, within the range we observed across organs

Reference	Comparison	% Agreement	Description of N
Haseman and Lockhart, 1993 10.1289/ehp.9310150	Intraspecies species sex concordance in site-specific carcinogenesis	65	379 studies
Gottmann *et al.*, 2001 10.1289/ehp.01109509	Intraspecies concordance of carcinogens	62% for rats 49% for mice	44 substances with replicate studies 34 substances with replicate studies
Haseman and Lockhart, 1993 10.1289/ehp.9310150	Interspecies concordance of site-specific carcinogenesis (rats – mice)	36	379 studies
Gottmann *et al.*, 2001 10.1289/ehp.01109509	Interspecies concordance of carcinogens	57	121 substances
Huff *et al.*, 1991 10.1289/ehp.9193247	Interspecies concordance of rodent liver tumor incidence (rats – mice)	80	~60 studies with rats and mice (15% of 400 carcinogenesis studies)
Gold *et al.*, 1991 10.1289/ehp.9193233	Interspecies concordance of carcinogens (rats – mice)	71-76 for any site; 48-52 for same site	533 studies with rats and mice

Examining organ effect levels specifically failed to reduce estimates of variance (RMSE)

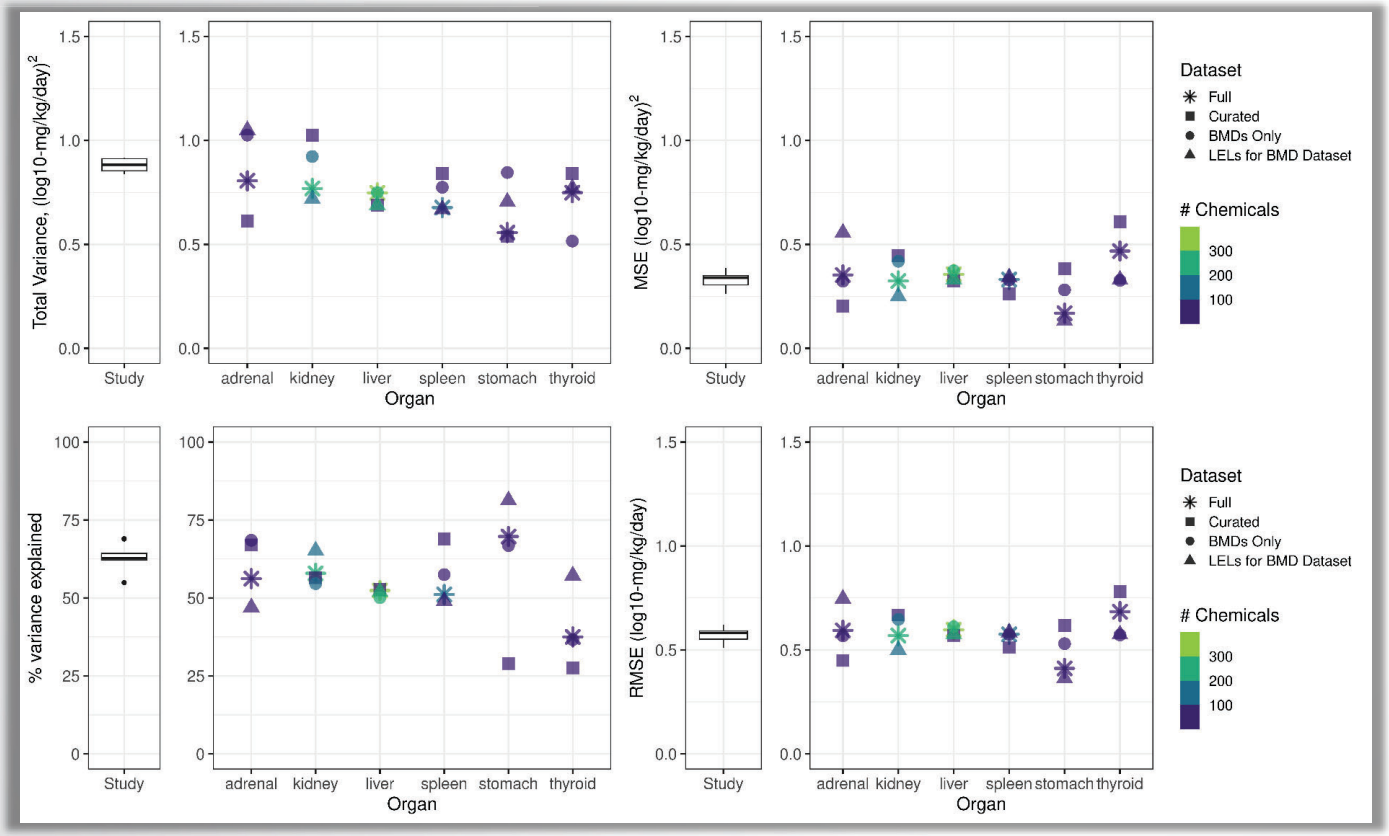


Figure 3, Paul Friedman et al. (in prep).

Primary Research Question	Statistical approaches
Can the estimate of variance for chemicals with replicate studies be reduced by estimating variance in specific organs?	Use multi-linear regression to approximate total variance, unexplained variance (MSE), RMSE, and % variance explained.

Predictions of an organ-level finding within ±1 log10-mg/kg/day may be an upper limit expectation on NAM performance.

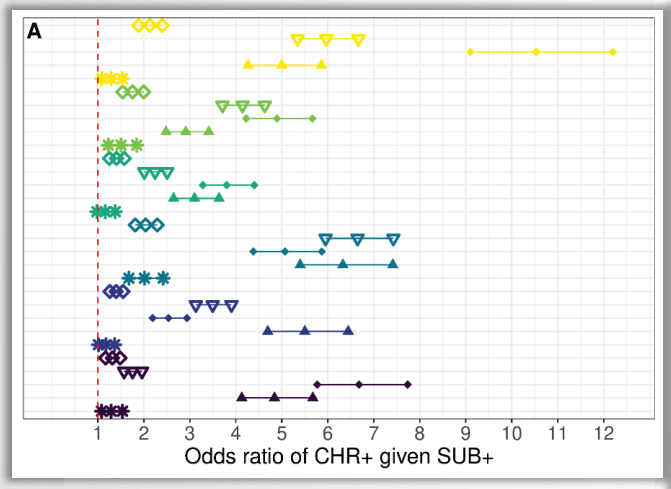
Qualitative reproducibility of organ-level findings between SUB and CHR studies may inform NAM strategy

- *In silico* NAMs for repeat dose toxicity could potentially be improved by combining SUB and CHR data for greater chemical coverage in training/testing.
  - Is it reasonable to expect similar organs will be affected by different study durations?

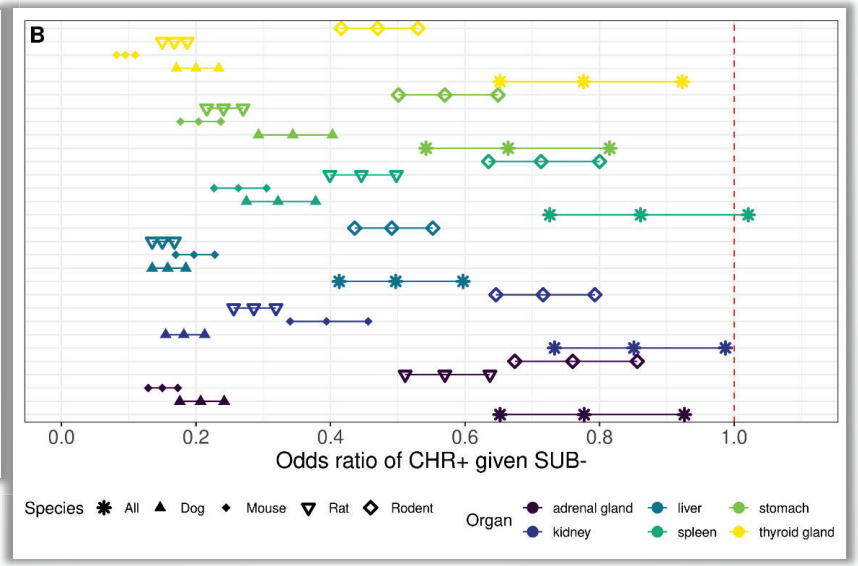
- Would a strategy focused on identification of a protective repeat dose point of departure using shorter-term studies or NAMs, without a chronic exposure study, miss organ-level effects?
  - Consider the contribution of cheminformatics and toxicoinformatics in identifying substances with longer serum half-life.
  - Excluding consideration of adversity of the findings in the organ.

Odds ratios for a positive in a tissue in a CHR given a negative in SUB are all less than 1, indicating this is an unlikely scenario.

Primary Research Question	Statistical approaches
What are the odds a chemical will produce any organ-level effect in a chronic (1-2 yr) study if the subchronic study was negative?	Calculate odds ratios for chemicals with subchronic and chronic study information



A positive in SUB tends to indicate a greater likelihood of a positive in CHR at that tissue, with some variability by species and tissue.



A negative in the SUB indicates a greater likelihood of negative in the CHR.

Possible indication: a repeat dose POD for a target organ at 90 days, particularly for liver and kidney where we have the largest datasets, is likely protective for a chronic finding.

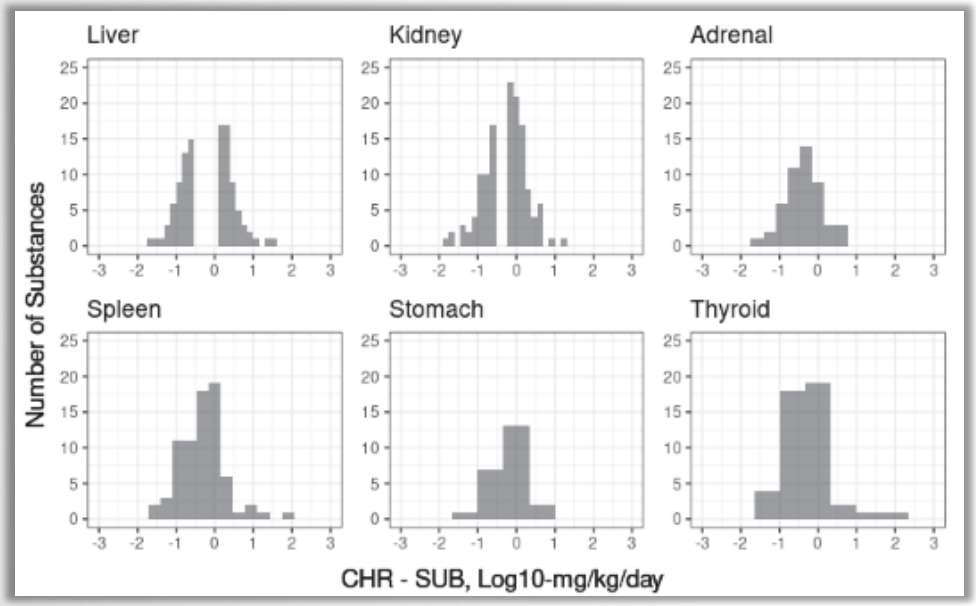
(without accounting for level of adversity)

Quantitative reproducibility of organ level findings between SUB and CHR studies may inform NAM strategy

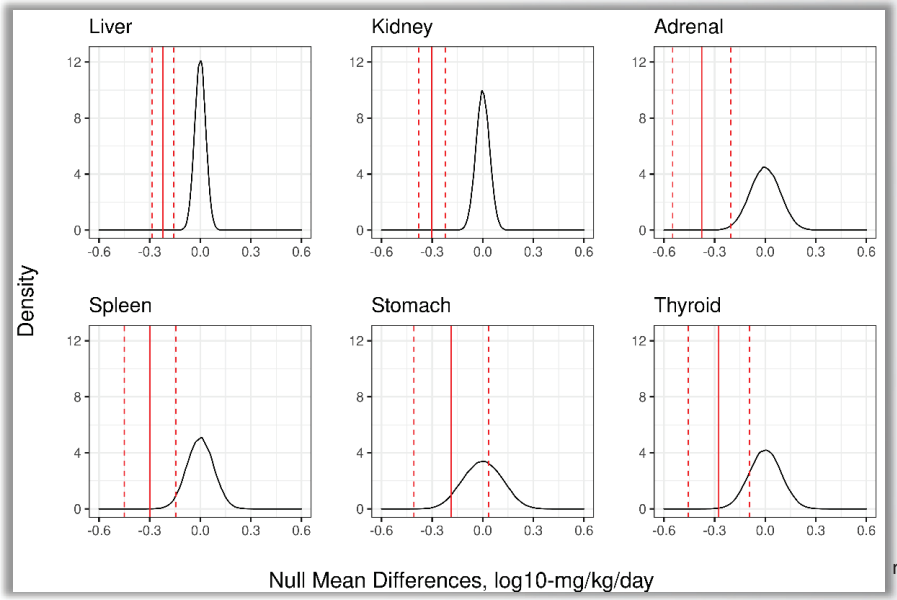
- What is a strategy for data-poor substances with no repeat-dose toxicity information?
  - Can reference or training data from subchronic and chronic studies be combined to develop *in silico* NAMs for repeat dose point of departure prediction?
  - Should a NAM-based repeat-dose point of departure estimate based on all data be adjusted for chronic exposure duration?

Generally, the chronic effect level values are 0.3 log10-mg/kg/day less than subchronic effect level values

Raw differences in CHR –SUB LELs



Sample mean differences ± CI compared to distribution of null mean differences



sample mean difference from the original data (log10(CHR/SUB))

2-sided 95% confidence interval ($p < 0.05$); *if the interval includes 0 then we cannot say that the true mean difference is different from 0*

Distribution of log10-transformed LEL differences following 100,000 randomization tests

- The mean differences in CHR – SUB min LEL values by organ approach estimates of variance in replicate repeat dose studies.
- *In silico* and *in vitro* NAMs for repeat dose point-of-departure estimation could combine SUB and CHR data in training.
- Current uncertainty or adjustment factors for SUB to CHR are protective.

Organ	Mean log10 difference, CHR SUB	Upper Bound, 95% CI	Lower Bound, 95% CI	p-value	N
Liver	-0.223	-0.159	-0.286	0	251
Kidney	-0.302	-0.223	-0.380	0	191
Adrenal	-0.377	-0.205	-0.548	0	49
Spleen	-0.298	-0.145	-0.450	1.00E-04	75
Stomach	-0.187	0.034	-0.408	0.0982	23
Thyroid	-0.275	-0.093	-0.458	0.0024	45

Conclusions: Primary takeaways from this work

- Part I: Variability in *in vivo* toxicity studies used in training or evaluation limits predictive accuracy of NAMs.
  - Maximal R-squared for a NAM-based predictive model of systemic effect levels may be 55 to 73%; i.e., as much as 1/3 of the variance in these data may not be explainable using study descriptors *at the study and the organ level.*
  - The estimate of variance (RMSE) in curated LELs and/or LOAELs approaches a 0.5 log10-mg/kg/day *at the study and the organ level.*
  - **Understanding that a prediction of an animal systemic effect level within ± 1 log10-mg/kg/day fold demonstrates a *very good* NAM is important for acceptance of NAMs for chemical safety assessment.**
- Part II: Qualitative and quantitative reproducibility of organ-level effect observations in repeat dose studies of adult animals
  - Qualitative concordance of organ-level effects was 33-88%, with highest concordance within species.
  - Quantitative variability in organ-level effects are similiar to estimates of variance at the study-level.
  - Subchronic and chronic *in vivo* observations can likely be combined for modeling to increase N.
  - It is unlikely that there are effects in organs like liver or kidney in a chronic study if these organs were unaffected in a subchronic study.
  - A repeat dose point of departure could be predicted by a NAM and adjusted to create a chronic-protective prediction.
- Construction of NAM-based effect level estimates that offer an equivalent level of public health protection as effect levels produced by methods using animals may provide a bridge to major reduction in the use of animals as well as identification of cases in which animals may provide scientific value.

Thank you for listening

Select References

Gold, L. S., et al. (1989). "Interspecies extrapolation in carcinogenesis: prediction between rats and mice." Environ Health Perspect **81: 211-219.**

Gottmann, E., et al., 2001. Data quality in predictive toxicology: Reproducibility of rodent carcinogenicity experiments. Environmental Health Perspectives. 109, 509-514.

Haseman, J. K. (2000). "Using the NTP database to assess the value of rodent carcinogenicity studies for determining human cancer risk." Drug Metab Rev **32(2): 169-186.**

Mazzatorta, P., et al., 2008. Modeling Oral Rat Chronic Toxicity. Journal of Chemical Information and Modeling. 48, 1949-1954.

Monticello, T. M., et al. (2017). "Current nonclinical testing paradigm enables safe entry to First-In-Human clinical trials: The IQ consortium nonclinical to clinical translational database." Toxicol Appl Pharmacol **334: 100-109.**

Pham L, Watford S, Pradeep P, Martin M, Judson R, Thomas RT, Setzer RW, Paul Friedman K. (2020). Variability in *in vivo* studies: Defining the upper limit of performance for predictions of systemic effect levels. Computational Toxicology. DOI: https://doi.org/10.1016/j.comtox.2020.100126

Pradeep P, Paul Friedman K, Judson RS. (2020). Structure-based QSAR models to predict repeat dose toxicity points of departure. Computational Toxicology. DOI: https://doi.org/10.1016/j.comtox.2020.100139

Toropov, A. A., et al., 2015. CORAL: model for no observed adverse effect level (NOAEL). Molecular diversity. 19, 563-75.

Toropova, A. P., et al., 2017. The application of new HARD-descriptor available from the CORAL software to building up NOAEL models. Food and Chemical Toxicology.

Toropova, A. P., et al., 2015. QSAR as a random event: a case of NOAEL. Environ Sci Pollut Res Int. 22, 8264-71.

Veselinović, J. B., et al., 2016. The Monte Carlo technique as a tool to predict LOAEL. European Journal of Medicinal Chemistry. 116, 71-75.

Wang, B. and G. Gray (2015). "Concordance of Noncarcinogenic Endpoints in Rodent Chemical Bioassays." Risk Anal **35(6): 1154-1166.**

Watford, S., et al., 2019. ToxRefDB version 2.0: Improved utility for predictive and retrospective toxicology analyses. Reprod Toxicol. 89, 145-158.

Thanks especially to Richard Judson, Woody Setzer, Ly Ly Pham, Prachi Pradeep, MJ Foster, Sean Watford, and Rusty Thomas

Office of Research and Development
Center for Computational Toxicology & Exposure (CCTE)
Bioinformatic and Computational Toxicology Division (BCTD)
Computational Toxicology and Bioinformatics Branch (CTBB)

Select Inter-Species Endpoint Comparison using National Toxicology Program (NTP) bioassay studies

Chad Blystone, PhD, DABT

Division of Translational Toxicology, NIEHS

10/12/2022

Background

- National Toxicology program has generated close to 600 carcinogenicity bioassays

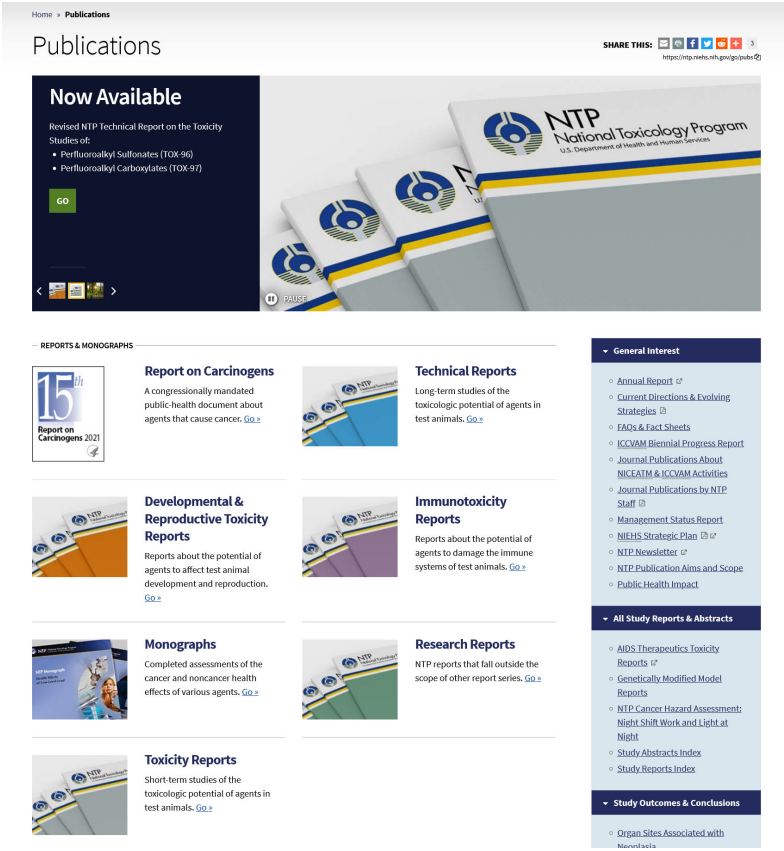
- Typically using two species:
  - Rat stock: F344/N, Osborne Mendel, Wistar Han, Hsd:SD Sprague Dawley (Current)
  - Mouse strain: B6C3F1

- Large variety of chemicals and routes of exposures evaluated within these studies
  - Gavage
  - Drinking water
  - Dermal
  - Feed studies
  - Inhalation

Data available online

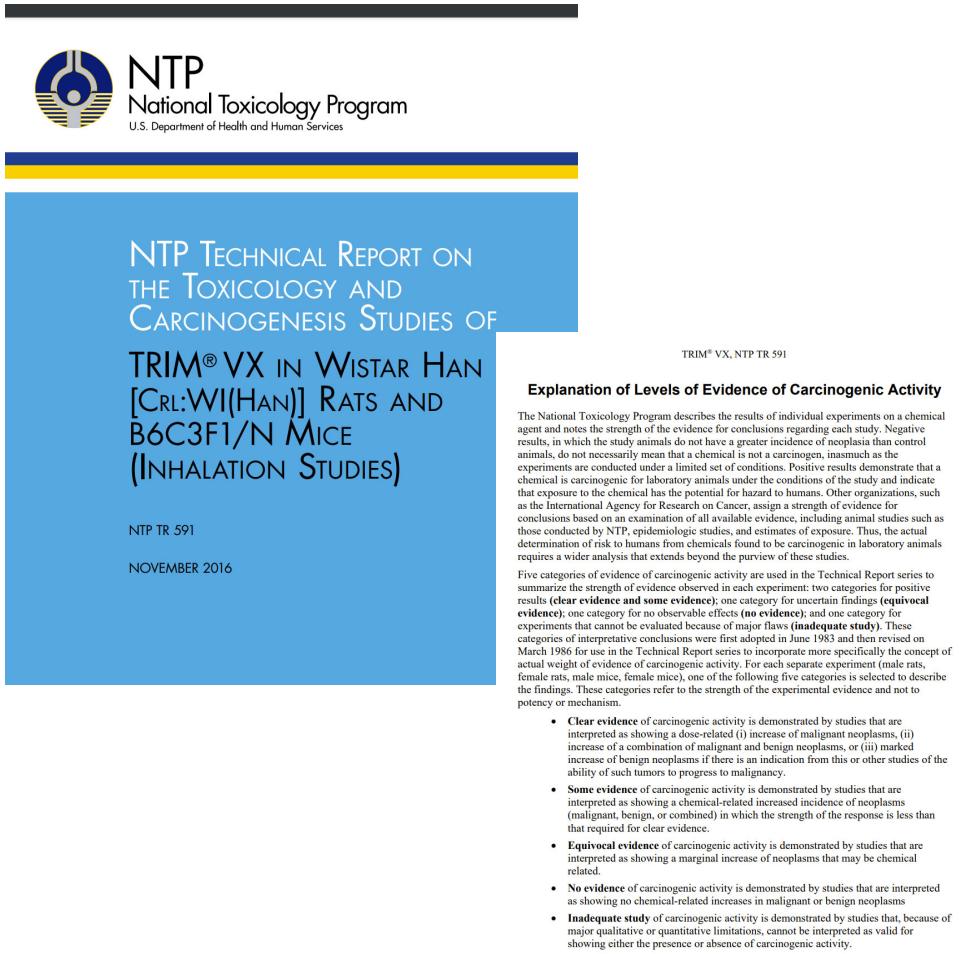
- Publications are available: https://ntp.niehs.nih.gov/publications/index.html

- Organ sites with neoplasia: https://cebs.niehs.nih.gov/organsites/



Level of evidence categorization

- Level of evidence call is made for each sex and species
- Categorizes confidence of carcinogenic response, based on increased neoplasms (benign or malignant) within a tissue
- Can result in highlighting rare non-statistically significant findings; downgrading statistically significant noisy background neoplasms



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

NTP TECHNICAL REPORT ON THE TOXICOLOGY AND CARCINOGENESIS STUDIES OF TRIM® VX IN WISTAR HAN [CRL:WI(HAN)] RATS AND B6C3F1/N MICE (INHALATION STUDIES)

NTP TR 591

NOVEMBER 2016

TRIM® VX, NTP TR 591

Explanation of Levels of Evidence of Carcinogenic Activity

The National Toxicology Program describes the results of individual experiments on a chemical agent and notes the strength of the evidence for conclusions regarding each study. Negative results, in which the study animals do not have a greater incidence of neoplasia than control animals, do not necessarily mean that a chemical is not a carcinogen, inasmuch as the experiments are conducted under a limited set of conditions. Positive results demonstrate that a chemical is carcinogenic for laboratory animals under the conditions of the study and indicate that exposure to the chemical has the potential for hazard to humans. Other organizations, such as the International Agency for Research on Cancer, assign a strength of evidence for conclusions based on an examination of all available evidence, including animal studies such as those conducted by NTP, epidemiologic studies, and estimates of exposure. Thus, the actual determination of risk to humans from chemicals found to be carcinogenic in laboratory animals requires a wider analysis that extends beyond the purview of these studies.

Five categories of evidence of carcinogenic activity are used in the Technical Report series to summarize the strength of evidence observed in each experiment: two categories for positive results **(clear evidence and some evidence)**; one category for uncertain findings **(equivocal evidence)**; one category for no observable effects **(no evidence)**; and one category for experiments that cannot be evaluated because of major flaws **(inadequate study)**. These categories of interpretative conclusions were first adopted in June 1983 and then revised on March 1986 for use in the Technical Report series to incorporate more specifically the concept of actual weight of evidence of carcinogenic activity. For each separate experiment (male rats, female rats, male mice, female mice), one of the following five categories is selected to describe the findings. These categories refer to the strength of the experimental evidence and not to potency or mechanism.

- **Clear evidence** of carcinogenic activity is demonstrated by studies that are interpreted as showing a dose-related (i) increase of malignant neoplasms, (ii) increase of a combination of malignant and benign neoplasms, or (iii) marked increase of benign neoplasms if there is an indication from this or other studies of the ability of such tumors to progress to malignancy.
- **Some evidence** of carcinogenic activity is demonstrated by studies that are interpreted as showing a chemical-related increased incidence of neoplasms (malignant, benign, or combined) in which the strength of the response is less than that required for clear evidence.
- **Equivocal evidence** of carcinogenic activity is demonstrated by studies that are interpreted as showing a marginal increase of neoplasms that may be chemical related.
- **No evidence** of carcinogenic activity is demonstrated by studies that are interpreted as showing no chemical-related increases in malignant or benign neoplasms
- **Inadequate study** of carcinogenic activity is demonstrated by studies that, because of major qualitative or quantitative limitations, cannot be interpreted as valid for showing either the presence or absence of carcinogenic activity.

Concordance of species calls vs endpoints

- Concordance between positive calls (clear, some, positive) between species and sexes will be higher than endpoint concordance
  - TR-494: p-Chloro-α,α,α-trifluorotoluene (PCTFT)

Positive Findings	Male	Female
Rat	Thyroid gland Lung	Thyroid gland Adrenal gland Uterus
Mouse	Liver	Liver Harderian gland

Assumptions/Caveats in species endpoint comparisons

- Study design similar
  - Dose selection rational similar across sexes and species
    - Caveat – Dose selection constrained in some instances
  - Exposure paradigm similar: young adult animals exposed for two years
    - Caveat – recent incorporation of in utero/lactational exposure in rats complicates direct comparisons to mice (adult only exposure)

- Evaluations of outcomes not necessarily interpreted independently
  - For example, a strong response in male rats may influence interpretation of moderate to weak response in male mice

Species endpoint concordance

- Endpoint can be defined from molecular target to apical endpoint
  - Focused on apical carcinogenic outcome within a tissue here
- Genetics will highly influence response within a species
  - "Species comparisons" can be highly skewed depending on animal model used
- Degree of concordance in this talk based on neoplastic response. Non-neoplastic response in separate sex/species could be on continuum, but not evaluated here.

GI Tract (Small and Large Intestine)

Test Article	Male Rat	Female Rat	Male Mouse	Female Mouse
1-Amino-2,4-dibromoanthraquinone	x	x		
1-Bromopropane	x	x		
2,2-bis(Bromomethyl)-1,3-propanediol	x			
2,3-Dibromo-1-propanol	x	x		
3,3'-Dimethoxybenzidine dihydrochloride	x	x		
3,3'-Dimethylbenzidine dihydrochloride	x	x		
4,4'-Thiodianiline	x			
Aloe vera whole leaf extract (native)		x		
Asbestos, chrysotile(IR)	x			
Bromodichloromethane	x	x		
C.I. Acid Red 114		x		
C.I. Direct Blue 15	x	x		
Captan			x	x
Methylene blue trihydrate			x	
o-Nitroanisole	x	x		
o-Nitrotoluene			x	x
Phenazopyridine hydrochloride	x	x		
Sodium dichromate dihydrate (VI)			x	x
Tribromomethane	x	x		
Bromochloroacetic acid	x	x		

- Typical neoplasms: adenoma, carcinoma
- 20 chemicals with positive calls in the intestine (small and large); 18 tested in rats and mice
- Sex concordance 14/20 = 70%
- Species concordance 0/18 = 0%

Urinary Bladder

Test Article	Male Rat	Female Rat	Male Mouse	Female Mouse
1,3-Dichloropropene (Telone II)				x
11-Aminoundecanoic acid	x			
1-Amino-2,4-dibromoanthraquinone	x	x		
2,2-bis(Bromomethyl)-1,3-propanediol	x			
4-Amino-2-nitrophenol	x			
4-Chloro-o-phenylenediamine	x	x		
Allyl isothiocyanate	x			
Anthraquinone	x	x		
C.I. Disperse Blue 1	x	x		
Chloroprene	x	x		
Glycidol				x
m-Cresidine	x	x		
Melamine	x			
Nitrilotriacetic acid trisodium monohydrate		x		
N-Nitrosodiphenylamine	x	x		
o-Anisidine hydrochloride	x	x	x	x
o-Nitroanisole	x	x		
o-Toluidine hydrochloride		x		
p-Benzoquinone dioxime		x		
p-Cresidine	x	x	x	x
Pulegone		x		
Salicylazosulfapyridine	x	x		

- Typical Neoplasm: Transitional cell/epithelial papilloma or carcinoma
- 21 chemicals with positive calls in urinary bladder
- Sex concordance 11/21 = 52%
- Species concordance 2/21 = 10%

Thyroid Gland Follicular Cell

Test Article	Male Rat	Female Rat	Male Mouse	Female Mouse
1,5-Naphthalenediamine			x	
2,2-bis(Bromomethyl)-1,3-propanediol	x	x		
2,3,7,8-Tetrachlorodibenzo-p-dioxin	x			x
2,4-Diaminoanisole sulfate	x	x	x	x
2-Methylimidazole	x	x	x	
3,3',4,4'-Tetrachloroazobenzene	x			
3-Amino-4-ethoxyacetanilide			x	
4,4'-Methylenebis(N,N-dimethyl)benzenamine	x	x		
4,4'-Methylenedianiline dihydrochloride	x	x	x	x
4,4'-Oxydianiline	x	x		x
4,4'-Thiodianiline	x	x	x	x
Acrylamide	x	x		
Anthraquinone			x	x
C.I. Basic Red 9 Monohydrochloride	x	x		
C.I. Pigment Red 3			x	
Chlorinated paraffins: C12, 60% chlorine		x		x
Chloroprene	x	x		
Cumene			x	
Ethylene thiourea (ETU)	x	x	x	x
Ginkgo biloba extract	x	x	x	
Glycidamide	x	x		
Glycidol	x	x		
HC Blue 1			x	
Iodinated glycerol	x			
Isobutene	x			
Isobutyl nitrite			x	
Malonaldehyde, sodium salt	x	x		
Mercuric chloride	x			
Metal Working Fluids: CIMSTAR 3800				x
N,N'-Diethylthiourea	x	x		
N,N-Dimethyl-p-toluidine	x			
o-Anisidine hydrochloride	x			
Oxazepam				x
Pentabromodiphenyl Ether Mixture [DE-71 (Technical Grade)]	x			
Primidone (primaclone)			x	
tert-Butyl alcohol				x
Trimethylthiourea		x		
Tris(2-Chloroethyl) Phosphate	x	x		
Water disinfection byproducts (Sodium chlorate)	x			

- Typical neoplasms: adenomas or adenocarcinomas
- 32 chemicals with positive calls, 31 tested in mouse and rat
- Sex concordance 13/32 = 41%
- Species concordance 8/31 = 26%

Mammary Gland

Test Article	Male Rat	Female Rat	Male Mouse	Female Mouse
1,2,3-Trichloropropane		x		
1,2-Dibromo-3-chloropropane		x		
1,2-Dibromoethane		x		x
1,2-Dichloroethane		x		x
1,3-Butadiene				x
2,2-bis(Bromomethyl)-1,3-propanediol	x	x		x
2,3-Dibromo-1-propanol		x		
2,4- & 2,6-Toluene diisocyanate		x		
2,4-Diaminotoluene (2,4-toluene diamine)		x		
2,4-Dinitrotoluene		x		
3,3'-Dimethoxybenzidine dihydrochloride		x		
3,3'-Dimethylbenzidine dihydrochloride		x		
5-Nitroacenaphthene		x		
Acronycine		x		
Acrylamide		x		x
Benzene		x		
C.I. Acid Red 114		x		
C.I. Basic Red 9 Monohydrochloride		x		
Chloroprene		x		x
Cytembena		x		
Endocrine disruptor (Genistein)		x		
Ethylene oxide				x
Furosemide				x
Glycidamide		x		x
Glycidol	x	x		x
Glycidol				
Hydrazobenzene		x		
Indium phosphide		x		
Isophosphamide		x		
Isoprene	x	x		
Methylene chloride	x	x		
Methyleugenol	x			
Nithiazide		x		
Nitrofurazone		x		
Nitromethane		x		
Ochratoxin A		x		
o-Nitrotoluene	x	x		
o-Toluidine hydrochloride		x		
Phenesterin		x		
Procarbazine hydrochloride	x	x		
Reserpine				x
Sulfallate		x		x
Urethane				x
Water disinfection byproducts (Bromochloroacetic acid)		x		
Water disinfection byproducts (Bromodichloroacetic Acid)		x		

- Typical neoplasms: adenomas, adenocarcinomas, fibroadenomas
- 44 chemicals with positive calls in mammary gland; 38 tested in two species
- Sex concordance 6/44 = 14%
- Species concordance 11/38 = 29%

Lung

Test Article	Male Rat	Female Rat	Male Mouse	Female Mouse
1,2-Dibromo-3-chloropropane			x	X
1,2-Dibromoethane		X	x	X
1,2-Dichloroethane			x	X
1,2-Epoxybutane	x			
1,3-Butadiene			x	X
1,3-Dichloropropene (Telone II)				X
1,5-Naphthalenediamine				X
1-Amino-2,4-dibromoanthraquinone			x	X
1-Bromopropane				X
2,2-bis(Bromomethyl)-1,3-propanediol	x		x	X
2,3-Dibromo-1-propanol			x	X
2,4,5-Trimethylaniline	x			
3,3',4,4'-Tetrachloroazobenzene	x	x	x	
3,3'-Dimethylbenzidine dihydrochloride	x	x		
4-Methylimidazole			x	X
4-Vinyl-1-cyclohexene diepoxide				X
5-Nitroacenaphthene	x	x		
8-Methoxypsoralen	x			
Acrylamide			x	x
Acrylonitrile				x
Antimony Trioxide	x	x	x	x
AZT transplacental carcinogenesis study			x	
Benzene			x	x
Benzofuran			x	x
beta-Picoline		x		x
bis(2-Chloro-1-methylethyl) ether			x	x
Bromoethane (ethyl bromide)	x			
C.I. Acid Red 114	x	x		
Chlorendic acid	x			
Chloroprene	x	x	x	x
Cobalt	x	x	x	x
Cobalt sulfate heptahydrate	x	x	x	x
Coumarin			x	x
Cumene			x	x
Dimethyl hydrogen phosphite	x			
Estradiol mustard			x	x
Ethylbenzene			x	
Ethylene oxide			x	x
Gallium arsenide		x		
Glycidamide			x	x
Glycidol			x	
HC Blue 1		x		
Indium phosphide	x	x	x	x
Isobutyl nitrite	x	x	x	x
Metal Working Fluids: CIMSTAR 3800				x
Metal Working Fluids: TRIM® VX			x	x
Methylene chloride			x	x
Molybdenum trioxide			x	x
N,N-Dimethyl-p-toluidine				x
Naphthalene				x
Nickel (II) oxide	x	x		
Nickel subsulfide	x	x		
Nitromethane			x	x
N-Methylolacrylamide			x	x
o-Nitrotoluene	x			
Oxymetholone		x		
Ozone				x
p-Chloro-a,a,a-trifluorotoluene	x			
Phenesterin			x	x
Procarbazine hydrochloride			x	x
Riddelliine				x
Selenium sulfide				x
Sulfallate			x	
Talc		x		
Tetranitromethane	x	x	x	x
Trifluralin				x
tris(2,3-Dibromopropyl) phosphate			X	x
Urethane			X	x
Vanadium pentoxide	x		X	x
Vinylidene Chloride				x
Water disinfection byproducts (Dibromoacetic acid)			X	x

- Typical neoplasms: alveolar/bronchiolar adenoma or carcinoma
- 71 chemicals with positive calls in the lung; 61 tested in rats and mice
- Sex concordance 41/71 = 58%
- Species concordance 12/61 = 20%

Summary

- Sex Concordance > Species Concordance (expected)
  - Exemptions possible with sex specific tissues (e.g. mammary gland)
- Species concordance varies across tissues
  - Wide range of explanations with genetic differences related to ADME, sensitivity, etc.
- Is concordance necessarily good or bad?
  - Concordance across species will strengthen interpretation
  - Covering wider genomic background (good) can result in discordant findings between species

Questions?

ICCVAM Strategic Roadmap for Validating New Methods

Warren Casey, PHD, DABT

Division of Translational Toxicology, NIEHS

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 • Department of Veterans Affairs Office of Research and Development
Environmental Protection Agency • Food and Drug Administration • National Cancer Institute • National Institute for Occupational Safety and Health
National Institute of Environmental Health Sciences • National Institute of Standards and Technology • National Institutes of Health
National Library of Medicine • Occupational Safety and Health Administration

ICCVAM Strategic Roadmap for ~~Validating~~ *Establishing Confidence* in New Methods

Warren Casey, PHD, DABT

Division of Translational Toxicology, NIEHS

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 • Department of Veterans Affairs Office of Research and Development
Environmental Protection Agency • Food and Drug Administration • National Cancer Institute • National Institute for Occupational Safety and Health
National Institute of Environmental Health Sciences • National Institute of Standards and Technology • National Institutes of Health
National Library of Medicine • Occupational Safety and Health Administration

Try to avoid using the "V" word

Big "V" Validation

"Formal" Validation

Little "v" validation

ICH Validation

Technical Validation

EURL-ECVAM Validation

Air-quote "validation"

Process Validation

Qualification

OECD Validation

ICCVAM Validation

ISO Validation

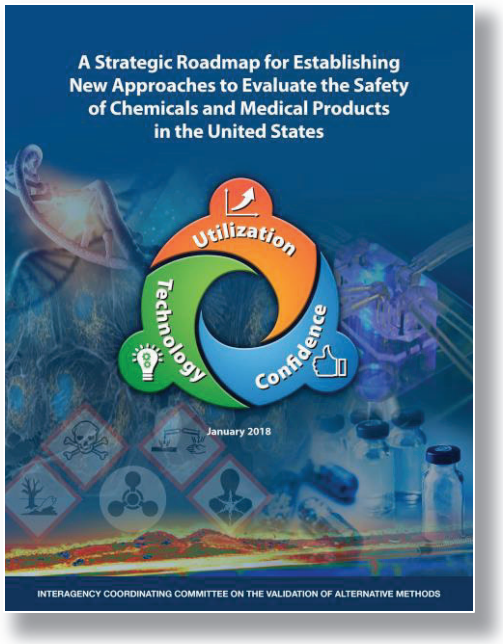
Establish confidence that new approaches are fit for their intended purpose

FIRST
THINGS
FIRST

Establish confidence that new approaches are fit for their intended purpose

FIRST
THINGS
FIRST

"Advances in science and technology have not been effectively leveraged to predict adverse human health effects"

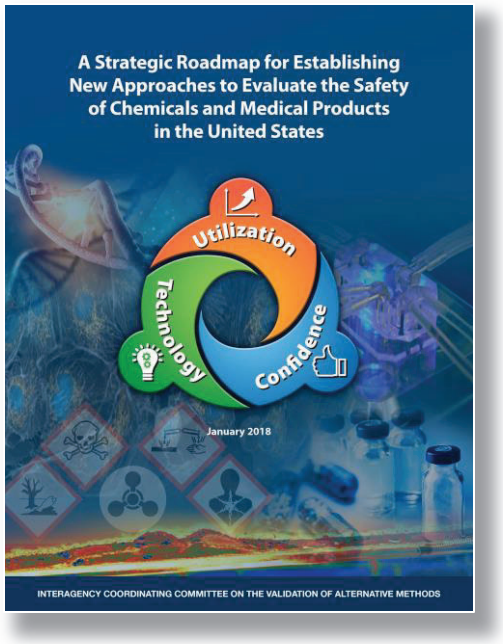


Help end-users guide the development of the new methods

Use efficient and flexible approaches to establish confidence in new methods

Encourage the adoption of new methods by federal Agencies and regulated industries

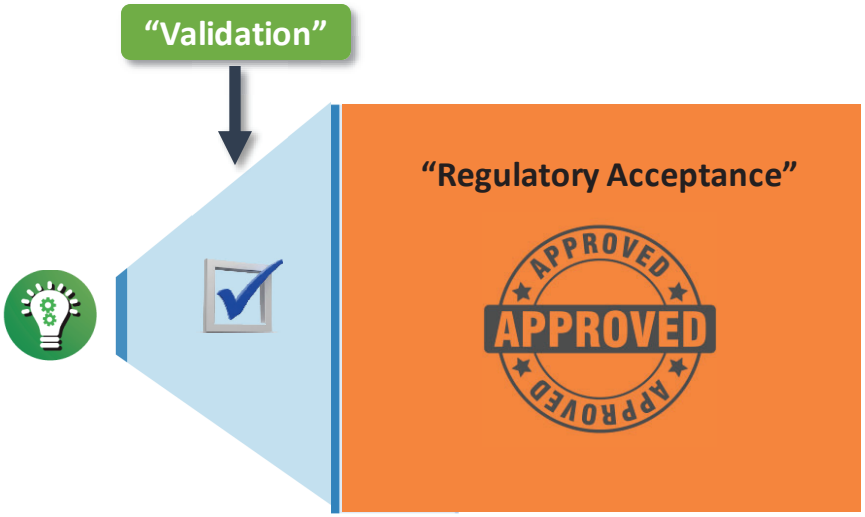
"Advances in science and technology have not been effectively leveraged to predict adverse human health effects"



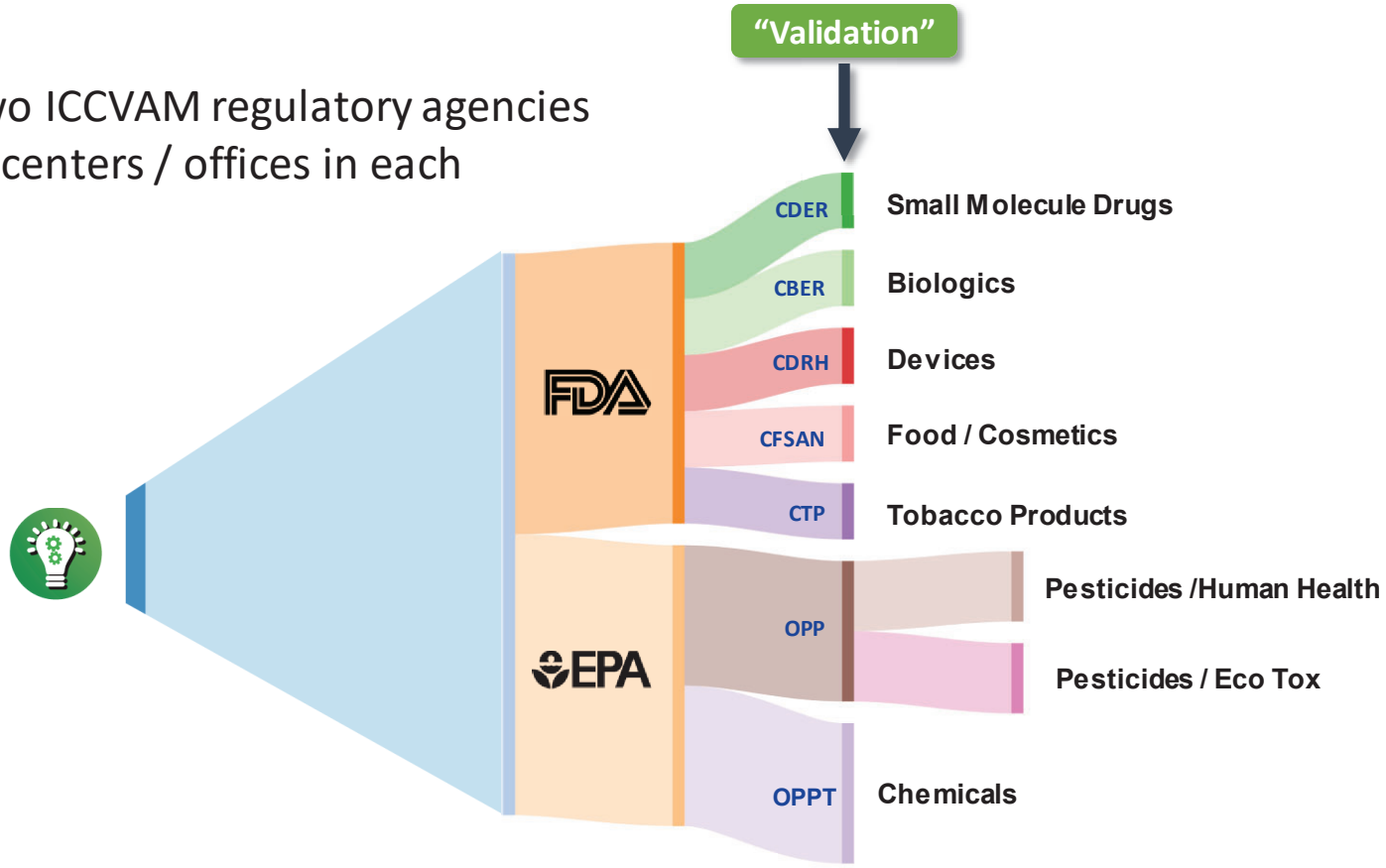
Help end users in the development of the new methods

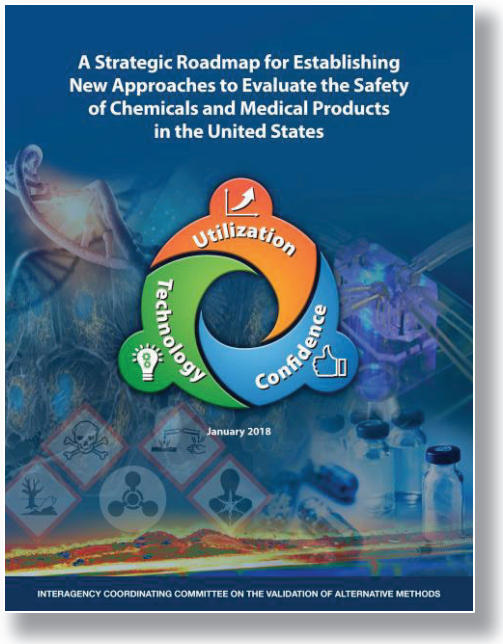
Use efficient and flexible approaches to establish confidence in new methods

Encourage the acceptance of new methods by federal Agencies and regulated industries


"Validation"
"Regulatory Acceptance"
APPROVED
APPROVED
APPROVED

Example of two ICCVAM regulatory agencies with multiple centers / offices in each





The "3Cs"

Communication

Collaboration

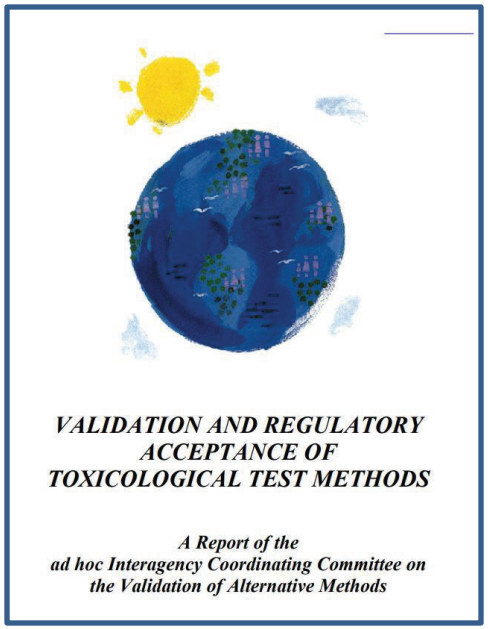
Commitment

ICCVAM: Validation Workgroup

Updating ICCVAM Guidance on Validation

ICCVAM Sponsor Agencies:
CPSC, FDA/CFSAN

Participating Agencies:
EPA/OPP, EPA/ORD, ATSDR, VA ORD, DOD, NIST, OSHA, NIEHS, NIH, FDA/CDER,/CTP,/OCS,/CDRH



NIH PUBLICATION NO: 97-3981

National Institute of Environmental Health Sciences
Research Triangle Park, North Carolina 27709

National Institutes of Health
U.S. Public Health Service
Department of Health and Human Services

March 1997

From

- Centralized ("VAMs")
- One Size Fits All
- Binary Status (Validated / Not)
- Stand Alone

TRANSITION

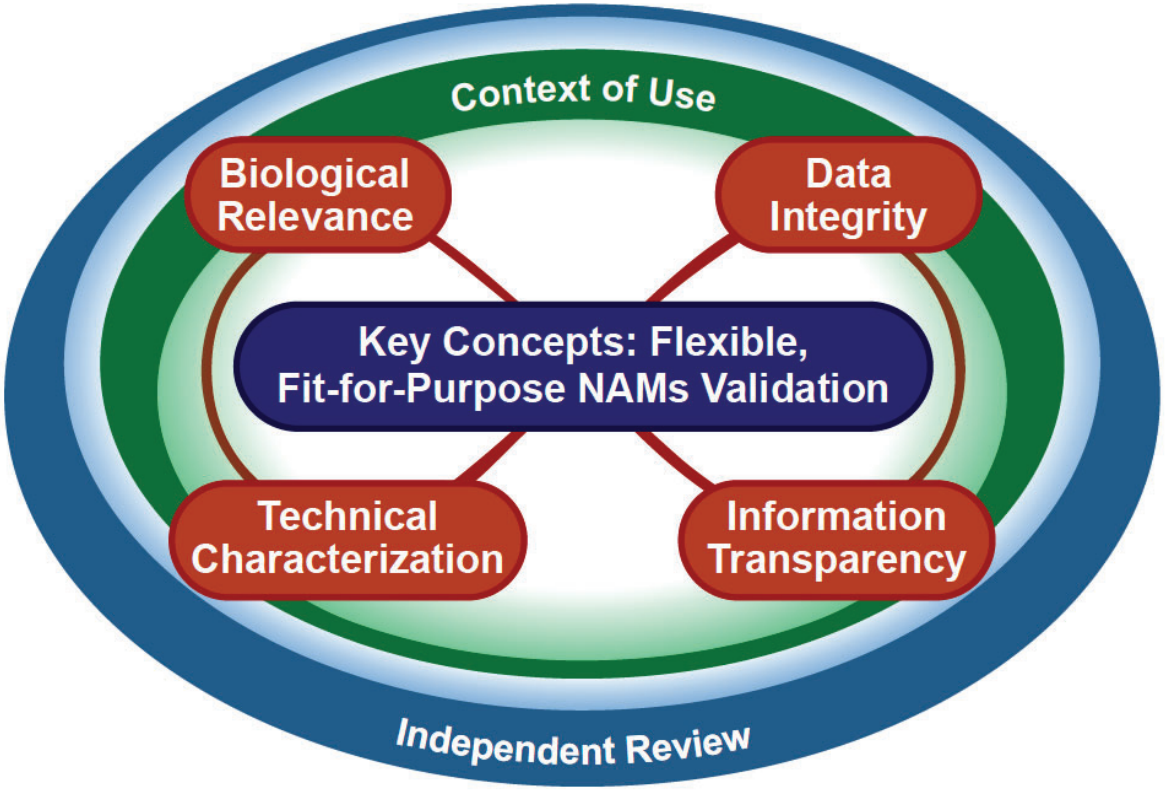
Towards

- Decentralized (End Users)
- Fit for Purpose
- Evolving Confidence
- Integrative

New Guidance from ICCVAM

- Underlying principles from OECD 34 remain the same in this new Guidance.
- Introduce the “context of use” terminology
- New guidance will emphasize that processes used to establish confidence should be flexible and adaptable.
- Emphasize the need for communication because regulatory needs may vary across the federal agencies

Guiding Pri


Context of Use
Biological Relevance
Data Integrity
Key Concepts: Flexible, Fit-for-Purpose NAMs Validation
Technical Characterization
Information Transparency
Independent Review

Topics Covered in the New Guidance

Foster the use of efficient, flexible, and robust practices to establish confidence in new methods

- Clearly delineate testing requirements and context of use
- Promote the use of new approaches for establishing confidence
- Utilize public workshops and/or public-private partnerships to promote cross-sector communication and cooperation

Topics Covered in the New Guidance

- Relevance of New Approach Methods
  - Biological Relevance
  - Biological Plausibility
  - Mechanistic Relevance
- Importance of Quality Reference Data
- Role of Legacy Animal Data

Topics Covered in the New Guidance

- Examination of best practices for quality and quality systems development
- Assessment of key sources of variability in the NAM
- Discussion of "Good or Better Standard" for qualification/validation.
- Incorporation of selected data quality tools such as:
  - Building a statistical model
  - Setting specifications

Topics Covered in the New Guidance

- How new principles for establishing confidence can fit into a globally harmonized approach to allow for continued mutual acceptance of data
- Reference to existing and well-vetted documents (e.g., GIVIMP, OECD GD34, GD69 on QSAR Validation, FDA Guidance for Industry, etc.)

Role of ICCVAM

- Assure an independent process for establishing confidence
- Advise federal agencies on different strategies for establishing confidence
- Facilitate cross-agency collaborations through work group/conferences
- Encourage global communication/harmonization on criteria used to establish confidence through conferences, seminars and meetings

Next Steps Prior to Finalization

- Format and organization of the document still under consideration.
- Input from the ICCVAM Federal Agencies still being incorporated through the VWG
- Draft document will be sent to ICCVAM agencies for review and sign off.
- Stakeholders will have opportunity to comment on the document.

Regulatory Question-Context of use:

What question needs to be answered and for what purpose?

Regulatory Question-Context of use:

What question needs to be answered and for what purpose?

"Predict" specific potential adverse health effects in humans

vs.

Identify "no biological effect" levels for human exposures

Let’s not allow idealized perfection to impede progress of approaches that are “good enough” for their intended purpose

United States
Consumer Product Safety Commission

Guidance for Industry and Test Method Developers:

Factors for CPSC Staff Evaluation of Alternative Test Methods and Integrated Testing Approaches to Support FHSA Labeling Requirements

EPA NAMs
October 12 and 13, 2022

Disclaimer: *This presentation was prepared by CPSC Staff and may not necessarily reflect the views of the Commission.*

Background

- The Federal Hazardous Substances Act (FHSA), 15 U.S.C. §1261-1275, requires appropriate cautionary labeling on certain hazardous household products to alert consumers to the potential hazard(s) that the products may present.
  - However, the FHSA does not require manufacturers to perform any specific toxicological tests to assess potential hazards (e.g., toxicity, corrosivity, sensitization, and irritation).

Background

- CPSC's 2012 Animal Testing Policy – Strongly encourages manufacturers to find alternatives to traditional animal testing that replace animals, reduce the number of animals tested, and decrease the pain and suffering in animals associated with testing household products.
- However, in the past CPSC had not issued any guidance describing what factors CPSC will consider in evaluating manufacturer's alternative test methods and resulting data submitted in support of a product's FHSA labeling.

Who Will Use this Guidance Document

- CPSC staff
- Manufacturers
- Test method developers
- Contract laboratories
- ICCVAM
- Other stakeholders, including the public

Purpose of Guidance Document

- Standardize the staff evaluation of alternative toxicological methods, and data generated by such methods, by providing factors staff should consider during technical review.

- Provide greater clarity to manufacturers, in particular, small businesses who lack toxicology expertise and have limited resources for their regulatory testing needs and strategies.

Guiding principles for evaluating methods and data

1. CPSC Staff Considers Scientific Validity and Defensibility of the Submitted Method and Data
   - Ensure that the method has been properly reviewed for accuracy and robustness.
   - Ensure that the data produced and submitted, pertains to CPSC regulatory needs to evaluate FHSA labeling.

2. Data on individual chemicals may not be sufficient for staff to determine FHSA labeling requirements for consumer products containing complex mixtures of chemicals.

Technical Factors:

1. The test method should have undergone independent scientific peer review by persons with no conflicts of interest.
2. There should be a detailed set of standard operating procedures (SOPs).
3. Data generated by the test method should adequately measure the endpoint of interest.
4. Applicability domain: There should be adequate test method data for chemicals and/or products representative of those administered by CPSC.
5. Limits of use should be specifically identified.
6. The test method should be robust (e.g., false positive and false negative rates).
7. Ideally, all data should be reported in accordance with Good Manufacturing Practices (GMP), Good Laboratory Practices (GLPs) or in the Spirit-of-GLP.

Guidance Overview

- Is not mandatory for the public and will not obligate CPSC to accept any particular alternative method.
- Explains that the evaluation of proposed test methods and data will be done on a case-by-case basis, and will require use of expert professional judgment.
- CPSC intends that the guidance will encourage a variety of viable test methods; it is not a blueprint or checklist for obtaining CPSC approval.
- If accepted, submitted method will be valid and acceptable for a specified purpose.

CPSC GUIDANCE DOCUMENT

- FR notice on proposed guidance published March 31, 2021
- Public Comment period ended June 14, 2021
  - Received 5 comments which were reviewed and addressed
  - Commission voted 4-0 to approve the final guidance document – April 2022
- Final version of the guidance document published April 11, 2022
  - https://www.regulations.gov/document/CPSC-2021-0006-0010
- Future Plans
  - Update web page with guidance document and any new methods reviewed and approved by the Commission.

Thank you

Final version of the guidance document:
https://www.regulations.gov/document/CPSC-2021-0006-0010

or e-mail me for the link
jgordon@cpsc.gov

CPSC.gov USCPSC

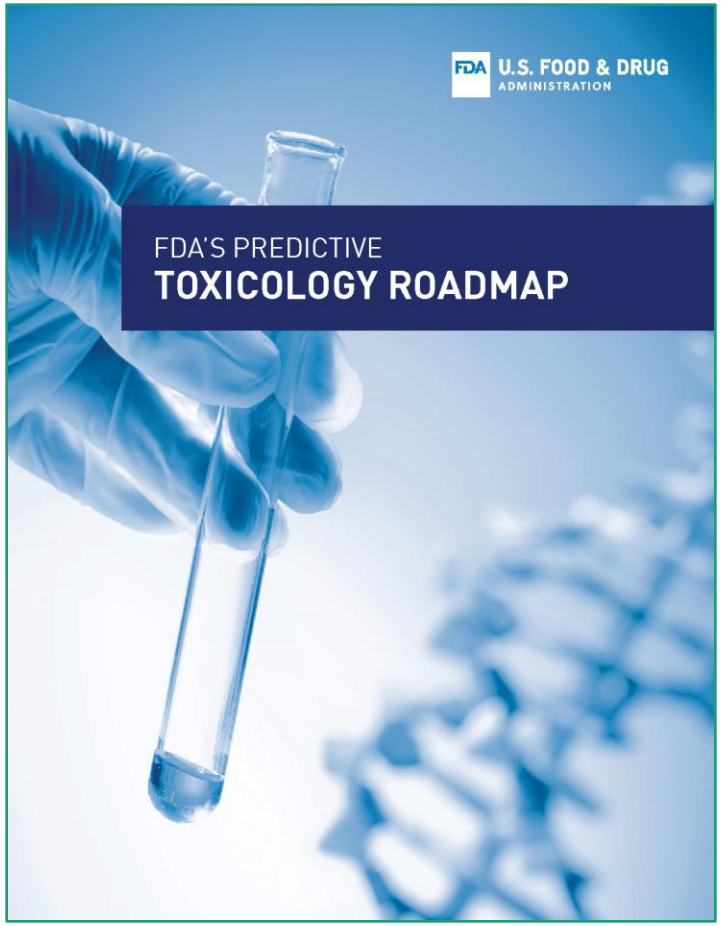
FDA Predictive Toxicology Road Map

Suzanne Fitzpatrick, PhD, DABT
CFSAN/FDA

National Academy of Science Meeting
October 13, 2022

FDA Predictive Toxicology Roadmap Announced December 6, 2017

- https://blogs.fda.gov/fdavoice/index.php/2017/12/fda-launches-predictive-toxicology-roadmap-to-enable-advances-in-toxicity-testing/



FDA Senior Level Toxicology Working Group

- Foster enhanced communication among FDA product centers and researchers
- Leverage FDA resources to advance the integration of emerging predictive toxicology methods and new technologies into regulatory safety and risk assessments.

Training of FDA regulators and researchers

- Continuing ongoing education in new predictive toxicology methods is essential for FDA regulators.
- Established an Agency-wide education calendar of events and a Toxicology Seminar Series to introduce concepts of new toxicology methodologies and updates in toxicology-related topics.

Collaborations with Stakeholders

- Foster collaborations across sectors and disciplines nationally and internationally.
- Pivotal to identifying the needs, maintaining momentum, and establishing a community to support delivery of new predictive toxicology methods.

Continued Communication

- Reaffirm FDA's commitment to incorporate data from newly qualified toxicology methods into regulatory missions
- Encourages discussions with stakeholders as part of the regulatory submission process.
- Encourage sponsors to submit a scientifically valid approach for using a new method early in the regulatory process

Leveraging Research

FDA’s research programs will identify data gaps and support intramural and extramural research to ensure that the most promising technologies are identified, developed, validated, and integrated into the product pipeline.

Oversight by Office of the Commissioner

- Track the progress of these recommendations and report to the Chief Scientist annually.
- Ensure transparency, fostering opportunities to share ideas and knowledge, showcase technologies, and highlight collaborations on developing and testing new methods

Start with a Regulatory Question- Context of Use

- What question needs to be answered and for what purpose?
- How much “validation/qualification” is needed for a particular assay will depend on the particular context of use

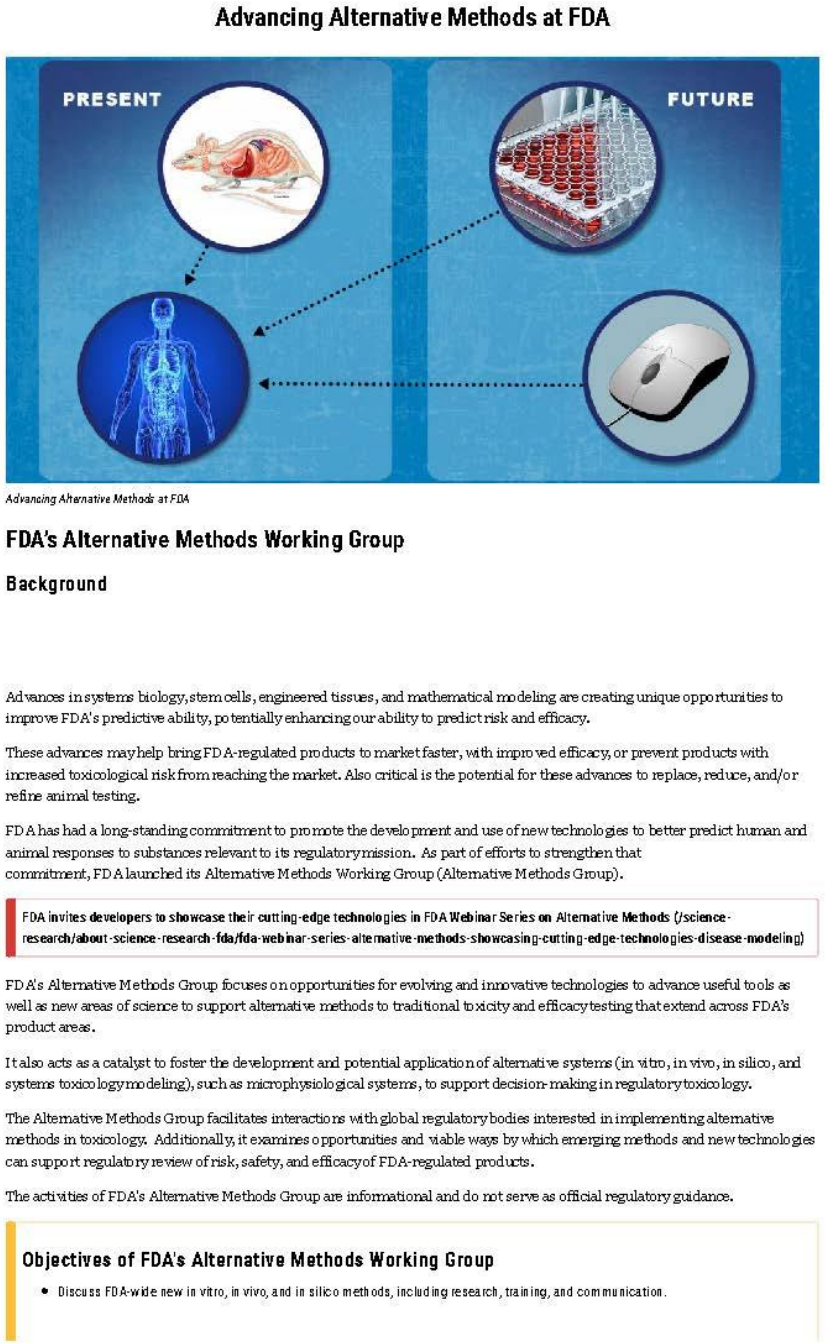
- Helps define acceptable applicability domain and limitations
- Additional context of use could be added at a later date

Alternative Methods Working Group (AMWG)

- Under Office of Chief Scientist, Office of Commissioner
  - Chaired by Drs. Fitzpatrick (CFSAN) and Mendrick (NCTR), regulatory members from each Center and OCS
- Strengthen FDA's long commitment to promoting the development and use of new technologies and to reduce animal testing
- Discuss new alternative *in vitro/in silico/in vivo methods* across FDA
- Interact with U.S. Federal partners and other global stakeholders to facilitate discussion and development of draft performance criteria for such assays.
- https://www.fda.gov/science-research/about-science-research-fda/advancing-alternative-methods-fda

Here now

- FDA now has an external webpage entitled **Advancing Alternative Methods at FDA**
- Essentially a webpage for the Alternatives Methods Working Group
  - Objectives
- Information on the FDA Webinar Series on Alternative Methods
- Page will be updated periodically
- Contact information: alternatives@fda.hhs.gov



Advancing Alternative Methods at FDA

Advancing Alternative Methods at FDA

FDA's Alternative Methods Working Group

Background

Advances in systems biology, stem cells, engineered tissues, and mathematical modeling are creating unique opportunities to improve FDA's predictive ability, potentially enhancing our ability to predict risk and efficacy.

These advances may help bring FDA-regulated products to market faster, with improved efficacy, or prevent products with increased toxicological risk from reaching the market. Also critical is the potential for these advances to replace, reduce, and/or refine animal testing.

FDA has had a long-standing commitment to promote the development and use of new technologies to better predict human and animal responses to substances relevant to its regulatory mission. As part of efforts to strengthen that commitment, FDA launched its Alternative Methods Working Group (Alternative Methods Group).

FDA invites developers to showcase their cutting-edge technologies in FDA Webinar Series on Alternative Methods (/science-research/about-science-research-fda/fda-webinar-series-alternative-methods-showcasing-cutting-edge-technologies-disease-modeling)

FDA's Alternative Methods Group focuses on opportunities for evolving and innovative technologies to advance useful tools as well as new areas of science to support alternative methods to traditional toxicity and efficacy testing that extend across FDA's product areas.

It also acts as a catalyst to foster the development and potential application of alternative systems (in vitro, in vivo, in silico, and systems toxicology modeling), such as microphysiological systems, to support decision-making in regulatory toxicology.

The Alternative Methods Group facilitates interactions with global regulatory bodies interested in implementing alternative methods in toxicology. Additionally, it examines opportunities and viable ways by which emerging methods and new technologies can support regulatory review of risk, safety, and efficacy of FDA-regulated products.

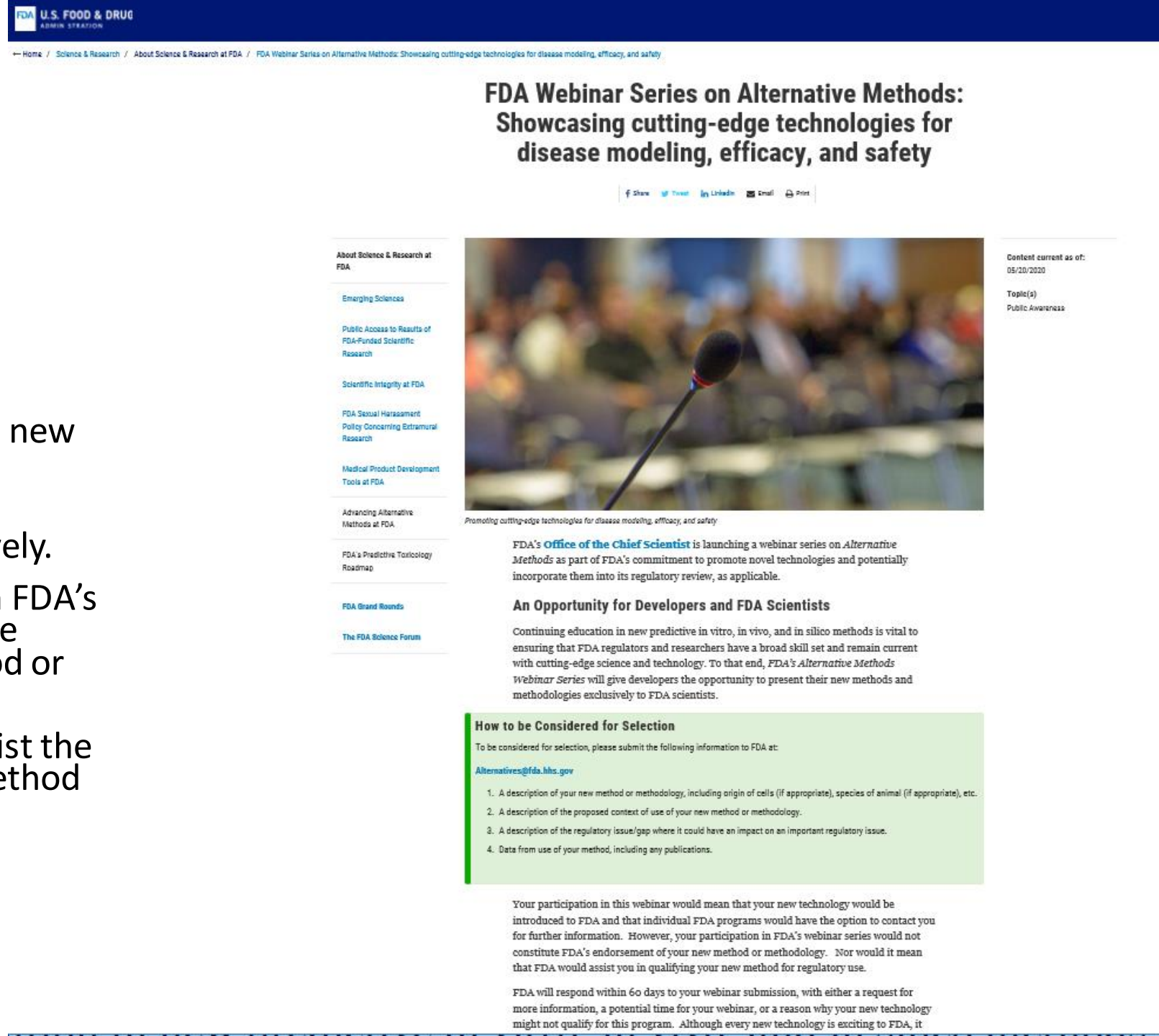
The activities of FDA's Alternative Methods Group are informational and do not serve as official regulatory guidance.

Objectives of FDA's Alternative Methods Working Group

- Discuss FDA-wide new in vitro, in vivo, and in silico methods, including research, training, and communication.

FDA Office of the Chief Scientist Webinar Series on Alternative Methods

- Opportunity for developers to present new methods and methodologies to FDA.
- Webinars will be held monthly and advertised to all FDA scientists exclusively.
- If selected, developers' participation in FDA's webinar series would not constitute the agency's endorsement of a new method or methodology.
- Nor would it mean that FDA would assist the developer in qualifying his/her new method for regulatory use.



U.S. FOOD & DRUG ADMINISTRATION

Home / Science & Research / About Science & Research at FDA / FDA Webinar Series on Alternative Methods: Showcasing cutting-edge technologies for disease modeling, efficacy, and safety

FDA Webinar Series on Alternative Methods: Showcasing cutting-edge technologies for disease modeling, efficacy, and safety

Share Tweet LinkedIn Email Print

About Science & Research at FDA

- Emerging Sciences
- Public Access to Results of FDA-Funded Scientific Research
- Scientific Integrity at FDA
- FDA Sexual Harassment Policy Concerning Extramural Research
- Medical Product Development Tools at FDA
- Advancing Alternative Methods at FDA
- FDA's Predictive Toxicology Roadmap
- FDA Grand Rounds
- The FDA Science Forum

Promoting cutting-edge technologies for disease modeling, efficacy, and safety

Content current as of: 05/20/2020

Topic(s) Public Awareness

FDA's **Office of the Chief Scientist** is launching a webinar series on *Alternative Methods* as part of FDA's commitment to promote novel technologies and potentially incorporate them into its regulatory review, as applicable.

An Opportunity for Developers and FDA Scientists

Continuing education in new predictive in vitro, in vivo, and in silico methods is vital to ensuring that FDA regulators and researchers have a broad skill set and remain current with cutting-edge science and technology. To that end, *FDA's Alternative Methods Webinar Series* will give developers the opportunity to present their new methods and methodologies exclusively to FDA scientists.

How to be Considered for Selection

To be considered for selection, please submit the following information to FDA at:

Alternatives@fda.hhs.gov

1. A description of your new method or methodology, including origin of cells (if appropriate), species of animal (if appropriate), etc.
2. A description of the proposed context of use of your new method or methodology.
3. A description of the regulatory issue/gap where it could have an impact on an important regulatory issue.
4. Data from use of your method, including any publications.

Your participation in this webinar would mean that your new technology would be introduced to FDA and that individual FDA programs would have the option to contact you for further information. However, your participation in FDA's webinar series would not constitute FDA's endorsement of your new method or methodology. Nor would it mean that FDA would assist you in qualifying your new method for regulatory use.

FDA will respond within 60 days to your webinar submission, with either a request for more information, a potential time for your webinar, or a reason why your new technology might not qualify for this program. Although every new technology is exciting to FDA, it

FDA's Alternative Report

Released January 5, 2021

FDA Tool Development Programs

Medical Device Development Tools (MDDT)

Share | Tweet | LinkedIn | Email | Print

Medical Device Development Tools (MDDT)

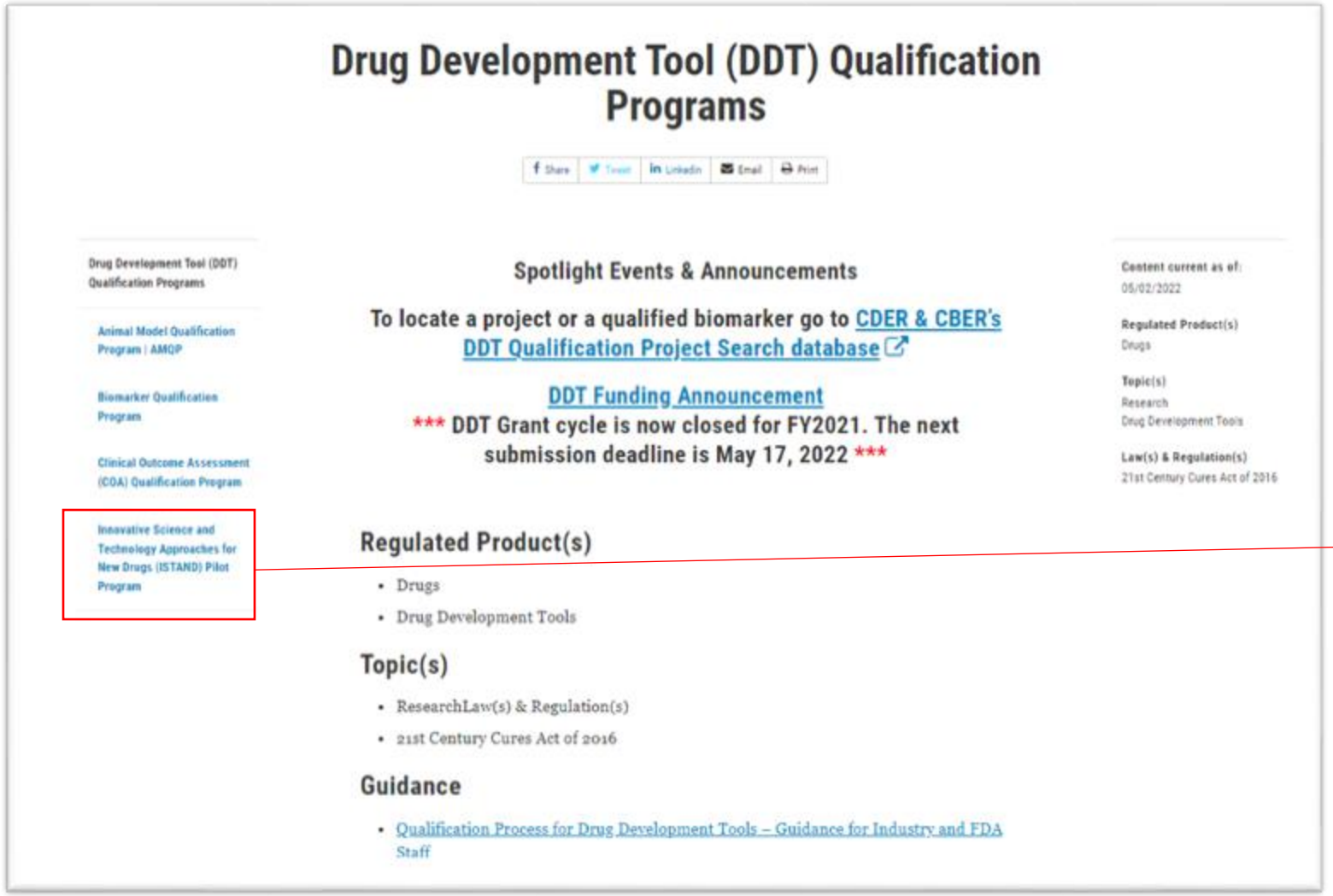
Content current as of:
05/05/2022

Regulated Product(s)
Medical Devices

On this page:

List of qualified tools includes “Nonclinical Assessment Models”

FDA Tool Development Programs



Drug Development Tool (DDT) Qualification Programs

Drug Development Tool (DDT) Qualification Programs

Animal Model Qualification Program | AMQP

Biomarker Qualification Program

Clinical Outcome Assessment (COA) Qualification Program

Innovative Science and Technology Approaches for New Drugs (ISTAND) Pilot Program

Spotlight Events & Announcements

To locate a project or a qualified biomarker go to CDER & CBER's DDT Qualification Project Search database

DDT Funding Announcement

*** DDT Grant cycle is now closed for FY2021. The next submission deadline is May 17, 2022 ***

Content current as of: 05/02/2022

Regulated Product(s): Drugs

Topic(s): Research, Drug Development Tools

Law(s) & Regulation(s): 21st Century Cures Act of 2016

Regulated Product(s)

- Drugs
- Drug Development Tools

Topic(s)

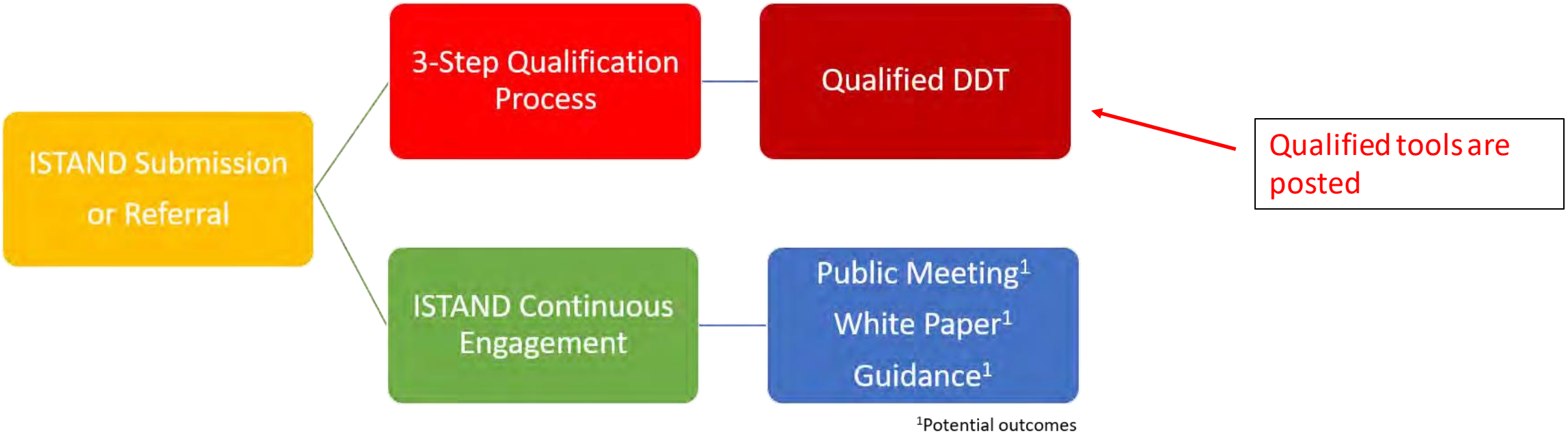
- ResearchLaw(s) & Regulation(s)
- 21st Century Cures Act of 2016

Guidance

- Qualification Process for Drug Development Tools – Guidance for Industry and FDA Staff

Innovative Science and Technology Approaches for New Drugs (ISTAND) Pilot Program

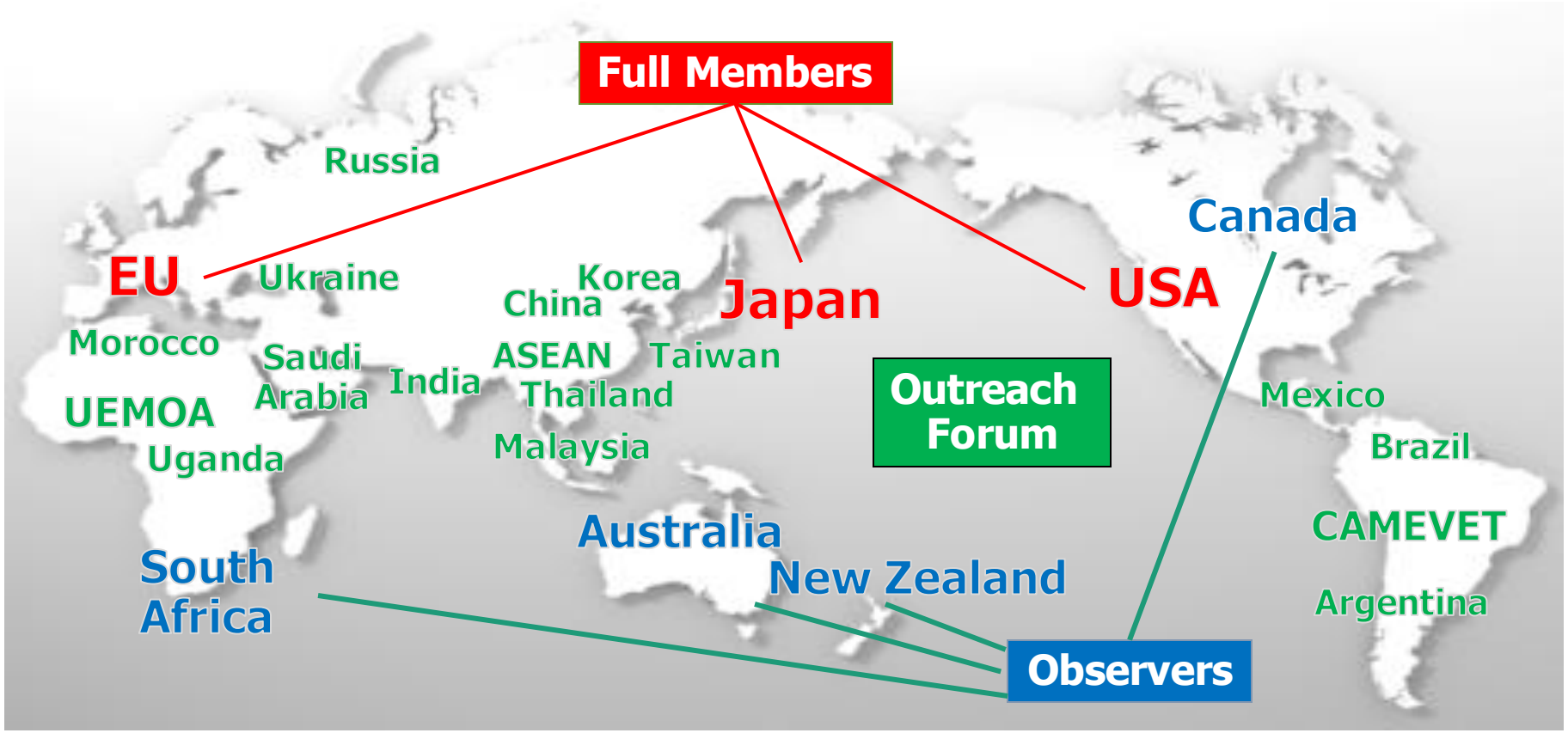
ISTAND Pilot Process



A transparent process – so all stakeholders are aware of tools in development, stage, and FDA determinations/recommendations

VICH =

International Cooperation
on Harmonisation of Technical Requirements
for Registration of Veterinary Medicinal Products (VMPs)



OIE : Associate Member, HealthforAnimals : Secretariat

International Liaison Group for Methods on Risk Assessment of Chemicals in Food (**ILMERAC**),

Organisation	Contact person
US FDA – Food and Drug Administration	Suzanne Fitzpatrick (co-chair) Goncalo Gamboa Steven Hermansky Jason Aungst Paul South
EFSA – European Food Safety Authority	Jose Tarazona (co-chair) Maria Chiara Astuto Irene Cataneo Jean-Lou Dorne Yann Devos Georges Kass Maria Bastaki
HC - Health Canada	Tara Barton-Maclaren Sonya Billiard John Field David Lefebvre Zoe Gillespie Marc Beal

Organisation	Contact person
RIVM	Esther de Jong Astrid Bulder Anne Kienhuis Ellen Hessel
JRC - Joint Research Centre	Sandra Coecke
BfR - German Federal Institute for Risk Assessment	Philip Marx-Stoelting Majlinda Lahaniatis
NVWA - the Netherlands Food and Consumer Product Safety Authority	Michiel den Braver
CFSA -China National Center for Food Safety Risk Assessment	Haixia Sui
OECD - Organisation for Economic Co-operation and Development	Patience Brown
NZFS - New Zealand Food Safety	Jeane Nicolas
KIT - Korean Institute of Toxicology	Yu WookJoon Lee Seung-Jin

Experts from non-ILMERAC organizations are invited for specific topics.

FDA's Proposed New Alternative Methods Program

- **Centrally coordinated through FDA's Office of the Chief Scientist with FDA Centers implementing Agency-wide programmatic objectives**

- **If this initiative is funded, FDA hopes to**
  - Expand processes to qualify alternative methods for regulatory use
  - Provide guidance to external stakeholders developing alternative methods
  - Fill information gaps with applied research to advance new policy and guidance development

- **Collaborations with external stakeholders are vital**
  - Federal partners, public-private partnerships, international regulators

Input From the FDA Science Board

FDA asked for input from the Science Board on how the agency can enhance its existing approaches to support the development, qualification, and implementation of alternative methods for regulatory use that can:

- Replace, reduce, and refine animal testing (the 3Rs)
- Improve predictivity of nonclinical testing

NAMs: Evolution of validation and scientific confidence building in Europe

Maurice Whelan

European Commission, Joint Research Centre (JRC)

3rd EPA NAM Workshop,

Oct 2022

The European Union Reference Laboratory for alternatives to animal testing

EURL ECVAM

Download it now !

The EU's Chemicals Strategy for Sustainability

Ensure most harmful chemicals are not contained in consumer products

Extend Generic Risk Assessment approach

"One substance one assessment" ambition

Common open data platform on chemicals

Address chemical mixtures

Promote safe and sustainable by design

Promote innovative testing and assessment methods

Better assessment of critical effects for more chemicals

Internationally recognised standards and tools

Make better use of 'academic' data in regulatory processes

https://ec.europa.eu/environment/strategy/chemicals-strategy_en

JRC Survey on NAMs

- Aimed primarily at method users (June ‘21 to March ’22).
- Supporting action to extend REACH info requirements
- Emphasis on regulatory applicability and deployability:
  - Derived No Effect Level (DNEL) for human health assessment
  - Predicted No Effect Concentration (PNEC) for env. assessment
  - Classification and Labelling
  - PBT or vPvB assessment
  - Assessment of (other) critical hazards

European Commission

General findings

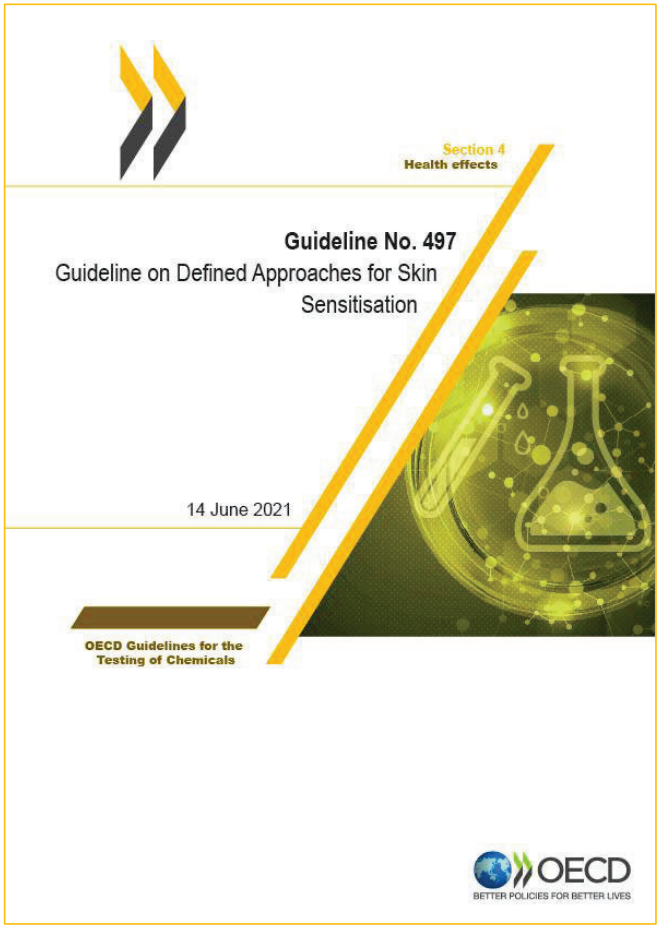
- **Many initiatives with different perspectives**
- **Many methods but fewer solutions** - impressive range of technologies and tools but little integration
- **Demonstration rather than validation** - case studies popular to show credibility and build confidence
- **A lot of variety but little standardisation** - multiple ways of generating similar information

Focus areas for the EU

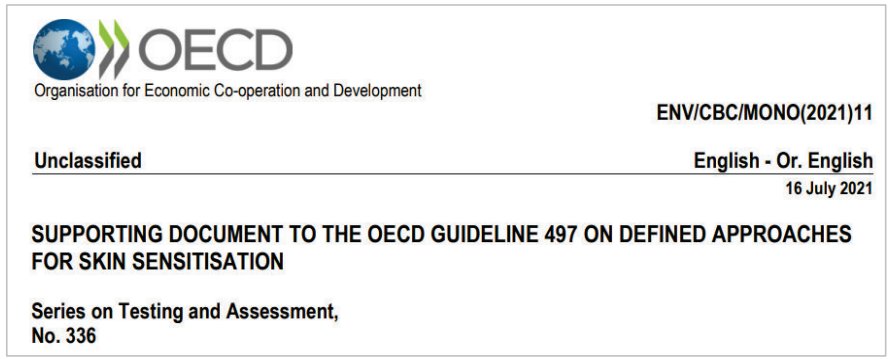
International Guidelines	• Mutual Acceptance of Data • Legal certainty & quality assurance • Efficiency and harmonisation
Technical standards	• Multiple uses including validation • Keep pace with NAM development • Important role in innovation
Academic studies	• Bespoke tools and design • Tackle complex problems • Best practices influence quality

European Commission

Defined Approaches for Skin Sensitisation

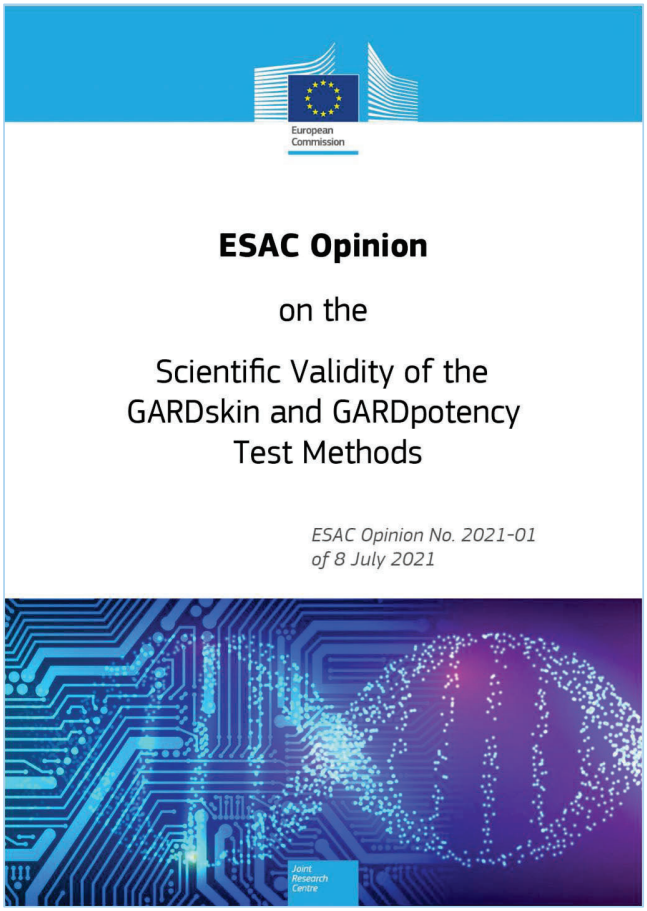


- First OECD Guideline to combine multiple alternative methods in a testing strategy
- First time to include computational methods (structural similarity algorithms) in a Guideline
- DAs for both hazard identification and potency based classification (GHS). The latter also provides a measure of confidence.



European Commission

Validation of ‘omics and machine learning



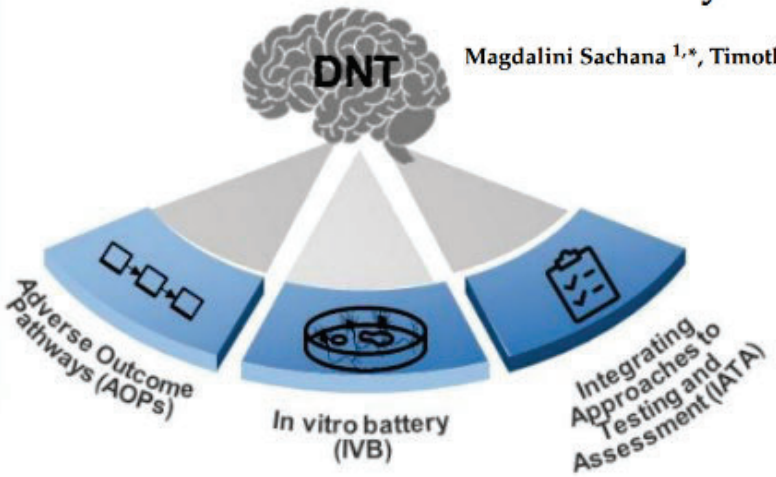
- Independent scientific **peer review by ESAC** of two Genomic Allergen Rapid Detection (SenzaGen GARD®) methods for skin sensitisation testing
- Methods combine cell-based test system with **transcriptomics** (~200 genes) and **SVM based algorithm** for hazard ID and potency classification
- ESAC **rebuilt and verified prediction models** (found that simpler model gave same results)
- TG development project at OECD triggered activities to deal with **IPR and GLP issues**
- **Sets a precedent**. Well worth a read!

European Commission

IATA for Developmental Neurotoxicity (DNT)

Highlights of work

- EFSA/OECD Workshop (Nov 2016)
- Formation of OECD DNT Expert Group (2017)
- Protocol for the implementation and interpretation of DNT in-vitro testing battery (November 2020)
- OECD DNT Guidance (first draft expected mid-2021)



biology

MDPI

Review

Toward a Better Testing Paradigm for Developmental Neurotoxicity: OECD Efforts and Regulatory Considerations

Magdalini Sachana [1,*], Timothy J. Shafer [2] and Andrea Terron [3]

EFSA JOURNAL

Scientific Opinion | Open Access

Development of Integrated Approaches to Testing and Assessment (IATA) case studies on developmental neurotoxicity (DNT) risk assessment

EFSA Panel on Plant Protection Products and their Residues (EFSA PPR Panel), Antonio Hernández-Jerez, Paulier Adriaanse, Annette Aldrich, Philippe Berny, Tamara Coja ... See all authors

First published: 18 June 2021 | https://doi.org/10.2903/j.efsa.2021.6599

Main goals of the OECD DNT project

- Improve DNT testing
- Incorporate mechanistic knowledge
- Provide regulatory relevant examples through case studies
- Accelerate regulatory uptake of the DNT IVB

European Commission

Table 1
Principles/criteria of different validation frameworks employed within the toxicology community.

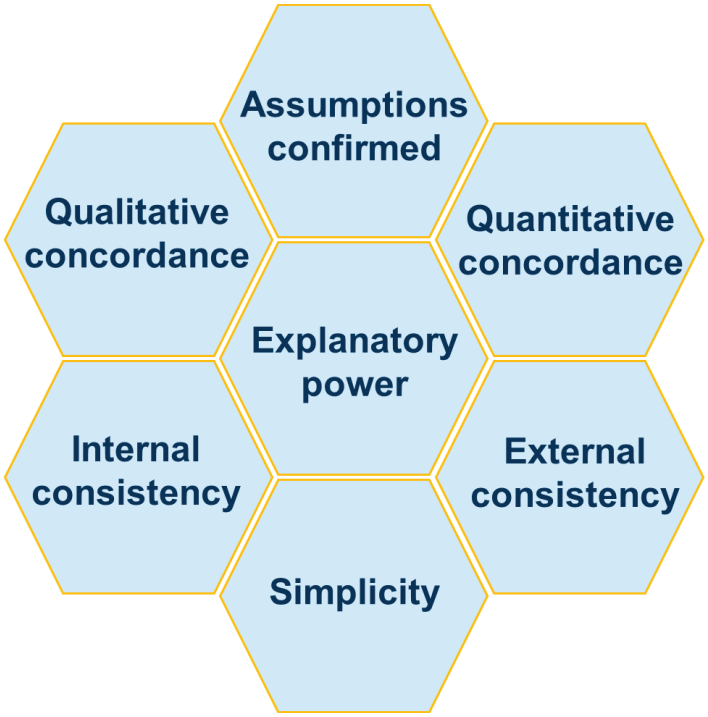
Minimum criteria for a valid test	ECVAM principles on test validity	QSAR validation principles	Defined Approaches	In vitro Developmental Neurotoxicity methods	Physiologically based kinetic models
OECD, 2005 [4]	Hartung et al, 2004 [5]	OECD, 2007 [6]	OECD 2016, 2017 [8,14,20]	Bal-Price et al, 2018 [22]	OECD Guidance Document
Rationale available for scientific need and regulatory purpose **Relevance**: relationship of test endpoint to in vivo biological effect **Protocol available**: subjected to independent peer-review **Repeatability and reproducibility shown**: intra-test, intra and inter-lab variability defined **Reference performance demonstrated** using reference chemicals **Toxicity performance** evaluated against existing relevant toxicity data **Validation available**: all data supporting assessment of validity available for review **Good Laboratory Practice** used to obtain data	**Test method definition**: endpoint, training set, prediction model (PM), applicability and mechanism **Within-laboratory variability:** assessment of reproducibility of data **Transferability**: confirmation by second operator (facility) **Between-laboratory variability**: assessment of reproducibility in 2 to 4 laboratories **Predictive capacity**: ability to predict beyond training set based on comparisons **Applicability domain**: definition of chemical classes and/or ranges for which predictions are reliable **Performance standards**: reference chemicals defined for equivalence between original and new (similar) tests	**A defined endpoint:** transparency of effect being predicted **An unambiguous algorithm:** transparency of description of an unambiguous model **A defined applicability domain:** recognising QSARs are reductionist and inevitably limited to subsets of chemical space **Appropriate measures of goodness-of-fit, robustness & predictivity:** performance when using training set or test set **A mechanistic interpretation:** an assessment of mechanistic associations between descriptors and end-points	**Structure:** elements of defined approach, information provided: **Relevance:** mechanistic basis **Predictive Capacity:** performance compared to reference data **Reliability:** reproducibility **Applicability domain:** technical limitations and chemical space **Complexity** of the Data Interpretation Procedure **Transparency:** availability of elements	**Test system:** definition, stability and biological relevance of cell-based system **Exposure scheme:** details of chemical treatment and incubation conditions **Documentation / SOP:** transparency in method protocol **Endpoint(s):** transparency of effect(s) being measured **Test method controls:** chemicals used to determine whether effects are positive or negative, and endpoint-specific **Data evaluation:** statistical analysis of concentration–response data **Testing strategy:** role in test battery **Robustness:** reproducibility within and between labs and over time **Test benchmarks:** sensitivity and specificity, data acceptance criteria **Prediction model:** how to extrapolate the in vitro data **Applicability domain:** chemistry and biological pathways **Screening hits:** definition of positive vs negative response	**Biological basis:** physiologically relevant model structure and parameters **Theoretical basis of model equations:** established mathematical basis such as Michaelis-Menten kinetics **Reliability of input parameters:** reproducibility **Sensitivity of output to input parameters:** relative importance of input parameters in determining simulation outcome **Goodness-of-fit and predictivity:** performance when using training set or test set

E.A. Patterson, M.P. Whelan, A.P. Worth (2021) The role of validation in establishing the scientific credibility of predictive toxicology approaches intended for regulatory application, *Comp. Tox*, 17, 100144.

Validation and scientific credibility

Scientific Credibility* is the **willingness of others to use predictions to inform their decisions**.

Requires a process of **social epistemology** to develop a ***shared knowledge and understanding*** between developers, users, and decision-makers.



7 Credibility Factors

ELSEVIER

Computational Toxicology
Volume 17, February 2021, 100144

The role of validation in establishing the scientific credibility of predictive toxicology approaches intended for regulatory application

Eann A. Patterson [a], Maurice P. Whelan [b], Andrew P. Worth [b]

*LW Schruben, *Simulation*, 34:101-105, 1980

European Commission

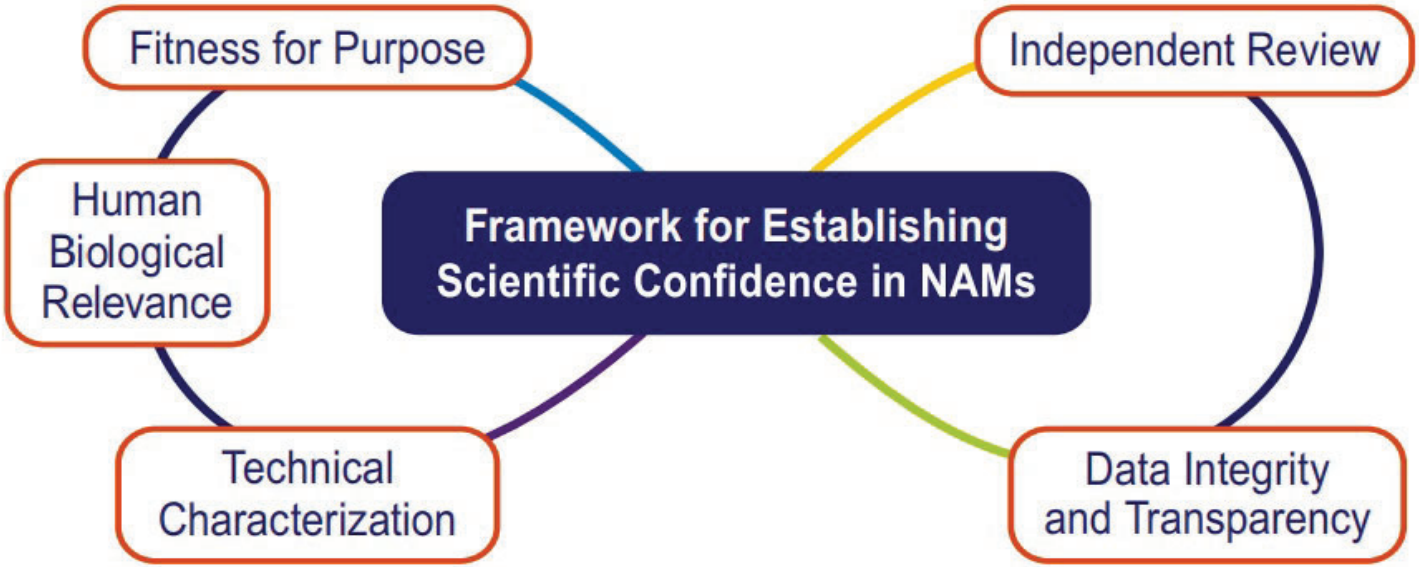
Archives of Toxicology
https://doi.org/10.1007/s00204-022-03365-4

August 2022

REVIEW ARTICLE

A framework for establishing scientific confidence in new approach methodologies

Anna J. van der Zalm[1] · João Barroso[2] · Patience Browne[3] · Warren Casey[4] · John Gordon[5] · Tala R. Henry[6] · Nicole C. Kleinstreuer[7] · Anna B. Lowit[6] · Monique Perron[8] · Amy J. Clippinger[1]



ADDRESSING EVIDENCE NEEDS IN CHEMICALS POLICY AND REGULATION

Trust and transparency+

JRC Science for Policy Report (Feb 2022)

https://publications.jrc.ec.europa.eu/repository/handle/JRC126724

European Commission

standard

noun
plural noun: **standards**

1. a level of quality or attainment.
2. something used as a measure, norm, or model in comparative evaluations.

OECD **performance standards** for test methods

41. The purpose of performance standards is to communicate the basis by which new test methods, both proprietary (*i.e.*, copyrighted, trademarked, registered) and non-proprietary can be determined to have sufficient accuracy and reliability for specific testing purposes.

These performance standards, based on validated and accepted test methods, can be used to evaluate the accuracy and reliability of other analogous test methods (colloquially referred to as "me-too" tests) that are based on similar scientific principles and measure or predict the same biological or toxic effect

Already exist e.g. skin corrosion, skin irritation, eye damage, ERTA

European Commission

28-29 April 2021

Organ-on-chip

Putting Science into Standards

EUROOCS
EUROPEAN ORGAN-ON-CHIP SOCIETY

CEN-CENELEC **Focus Group** on Organ on chip

Stem Cell Reports
Meeting Report
ISSCR
OPEN ACCESS

Putting Science into Standards workshop on standards for organ-on-chip

Monica Piergiovanni,[1,*] Ozlem Cangar,[2] Sofia B. Leite,[1] Livia Mian,[3] Andreas Jenet,[4] Raffaella Corvi,[1] Maurice Whelan,[1] Fabio Taucer,[4] and Ashok Ganesh[3]
[1]European Commission, Joint Research Centre (JRC), Ispra, Italy
[2]European Health and Digital Executive Agency (HaDEA), Brussels, Belgium
[3]CEN-CENELEC, Market Perspective and Innovation, Brussels, Belgium
[4]European Commission, Joint Research Centre (JRC), Brussels, Belgium
*Correspondence: monica.piergiovanni@ec.europa.eu
https://doi.org/10.1016/j.stemcr.2021.07.010

The European Commission Joint Research Centre and the European Standardization Organizations CEN and CENELEC organized the "Putting Science into Standards" workshop, focusing on organ-on-chip technologies. The workshop, held online on 28–29 April, 2021, aimed at identifying needs and priorities for standards development and suggesting possible ways forward.

CEN CENELEC European Commission

Better use of ‘academic’ data

- *International **Workshop** at JRC on 25-26 Oct 2022*
- *Proposal to develop **Guidance** submitted to OECD*

European Commission

Thank you

Maurice Whelan

Head of Unit, Chemical Safety and Alternative Methods,

Directorate for Health, Consumers and Reference Materials,

European Commission, Joint Research Centre (JRC).

maurice.whelan@ec.europa.eu

@MauriceAtEcvam

www.linkedin.com/in/maurice-whelan-ec-jrc

European Commission

OECD PERSPECTIVE* ON THE FUTURE OF NAMS, MAD, AND TGS

Patience Browne, OECD
EPA NAM Workshop
12-13 October 2022

BETTER POLICIES FOR BETTER LIVES

OECD and Mutual Acceptance of Data

OECD: 38 member countries
MAD-adhering countries: Argentina, Brazil, India, Malaysia, Singapore, South Africa, Thailand
(partner countries)
2022 Accession Countries: Argentina, Brazil, Bulgaria, Croatia, Peru, Romania

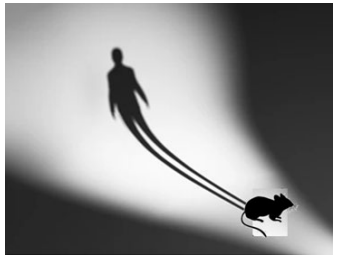
Test Guideline	+	GLP	=	MAD
Internationally harmonised **methods** for evaluating chemical safety		**Principles** and **conditions** under which laboratory studies are conducted, reported and recorded		Studies conducted using OECD TG and according to GLP fall under the **Mutual Acceptance of Data**

MAD is a legal agreement among all member and partner countries that share a common data requirement to accept the data generated by other member countries

MAD saves
- **€ 309 million/yr**
- **10,000s of animals**

Drivers for increase uptake of NAMs

- **Throughput**
  - Testing requirements vary may include a number of (sequential) experiments = months to years to produce and analyse data
  - Using traditional (mostly animal-based) methods for assessing safety, only 10s/100s/1000s of chemicals can be evaluated each year
- **Costs**
  - Bringing new products to market estimated:
    - Average for new drugs 1.3B USD
    - New pesticide active ingredients 250M USD
    - Cosmetics R&D in Europe 2.35B Euro/yr



- **Relevance**
  - There is increasing recognition that the animal tests may not be good predictors of effects in humans
- **Changing regulations** which **reduce or prohibit animal testing** to evaluate chemical safety, e.g.:

Australia	Israel
Columbia	Mexico
Guatemala	New Zealand
European Union	South Korea
Iceland	Switzerland
India	Türkiye

OECD support of New Approach Methods (NAMs)

*… supports use of **New Approach Methods** when suitability can be demonstrated (to be as good or better than existing approaches)*

- **Q1: What counts as "new"?**
  - "New Approach Methods" include **everything that is not an "old approach"**
    - *in chemico*, *in vitro*, computational, *in vivo* methods
    - stand-alone or (more often) integrated approaches to testing and assessment (IATAs)
    - use data science/machine learning/AI
  - Not "non-animal methods", but aligned with the 3Rs
    - Faster time to safety decisions
    - Less resources intensive
      - e.g. cheaper, less time for testing/analyses, fewer/no animals used

OECD support of New Approach Methods (NAMs)

*… supports use of New Approach Methods when suitability can be demonstrated (to be as **good or better** than existing approaches)*

- **Q2: What counts as "as good or better"?**
  - Results must be **reproducible**
  - The test system must be **relevant**
    - "**Relevance**" may vary with a specific **regulatory application**; e.g.
      - Sensitive to chemical-changes
      - Has a demonstrated relationship to the toxicological endpoint
      - Is biologically relevant to the target species
      - Should include a consideration of **approaches** that are **currently in use**
        - » e.g. >80% do not have full suite of chemical safety data
- **FIT FOR PURPOSE**

How the use of NAMs changes testing paradigms

- OECD Test Guidelines include that NAMs (not exhaustive)

Acute Toxicity	OECD publications
Oral	GD 237 ; TG 420, 423, 425
Dermal	GD 237; TG 402
Inhalation	GD 237, GD 39; TG 403, 433, 436
Eye Irritation and damage	GD 263; TG 437, 438, 460, 491, 492
Skin Irritation and corrosion	GD 203; TG 430, 431, 435, 439, 460
Skin sensitisation	GD 256; TG 442C, 442D, 442E, GL 497

General Guidance	OECD publications
Grouping chemicals /read across	GD 194
Waving or bridging (read-across) acute toxicity tests	GD 237
Use of AOPs for Developing IATA	GD 260
Reporting DA to be used within IATA	GD 255
Describing non-guideline in vitro test methods	GD 211
Workshop report on framework for development and use of IATA	GD 215

- MAD regards information sharing among Member Countries that have **the same data requirement**

Global initiatives for NAMs

EXTERNAL SCIENTIFIC REPORT

APPROVED: 2 May 2022

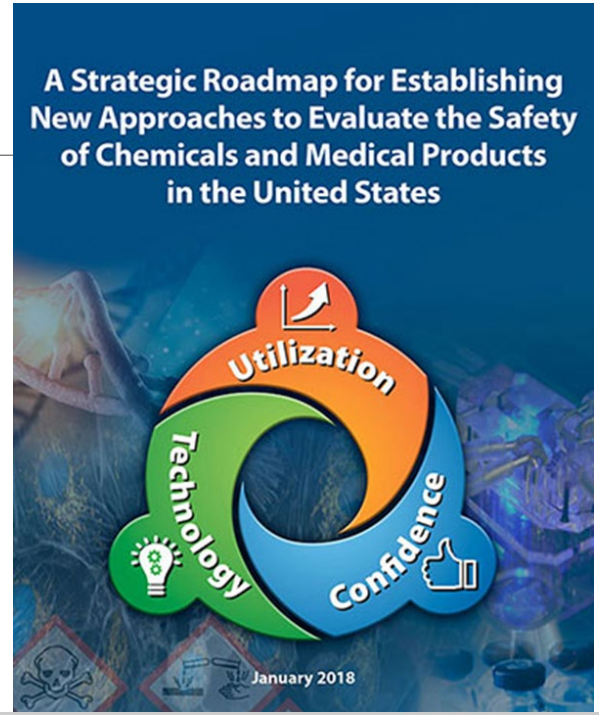
doi:10.2903/sp.efsa.2022.EN-7341

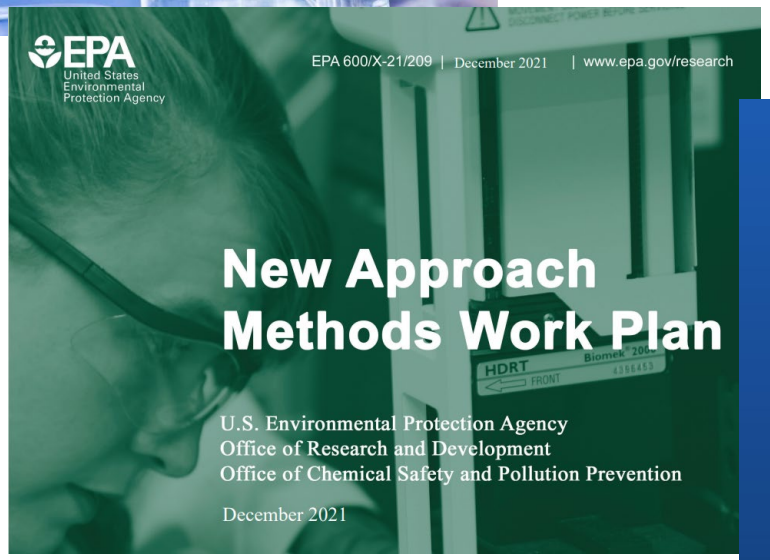
Development of a Roadmap for Action on

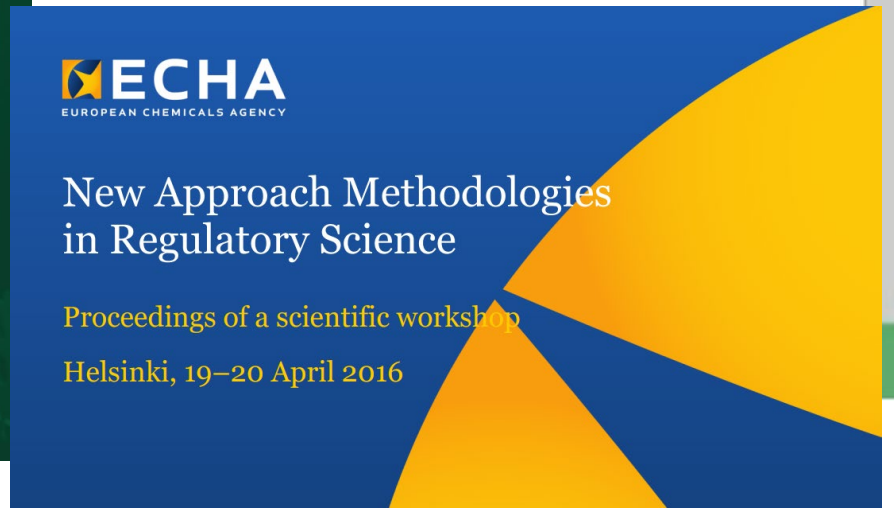
New Approach Methodologies in Risk Assessment

Sylvia E. Escher[1], Falko Partosch[1], Sebastian Konzok[1], Paul Jennings[2], Mirjam Luijten[3], Anne Kienhuis[3], Victoria de Leeuw[3], Rosmarie Reuss[4], Katrina-Magdalena Lindemann[4], Susanne Hougaard Bennekou[5]

[1] Fraunhofer ITEM, [2] Vrije Universiteit Amsterdam, [3] National Institute for Public Health and the Environment, [4] Eura AG, [5] The National Food Institute Denmark

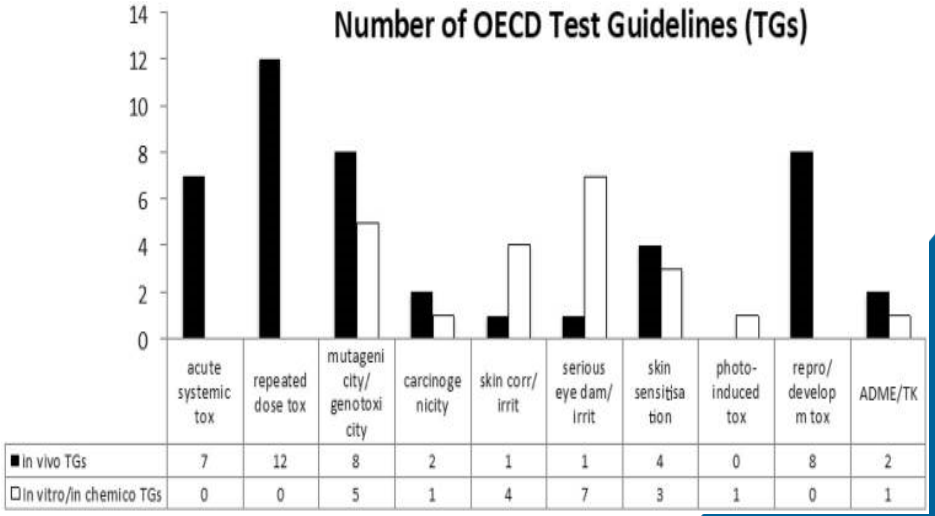






How the use of NAMs changes testing paradigms

- Regulations **vary** in:
  - Specific data requirements defined in regulations
  - Flexibility to fulfil requirements
  - Explicit national/organisational mandates to use NAMs
- **Creates potential divergence among countries & regulatory authorities**
  - A variety of NAM roadmaps
  - Acceptance of NAMs is not harmonised
  - Potential threat to MAD



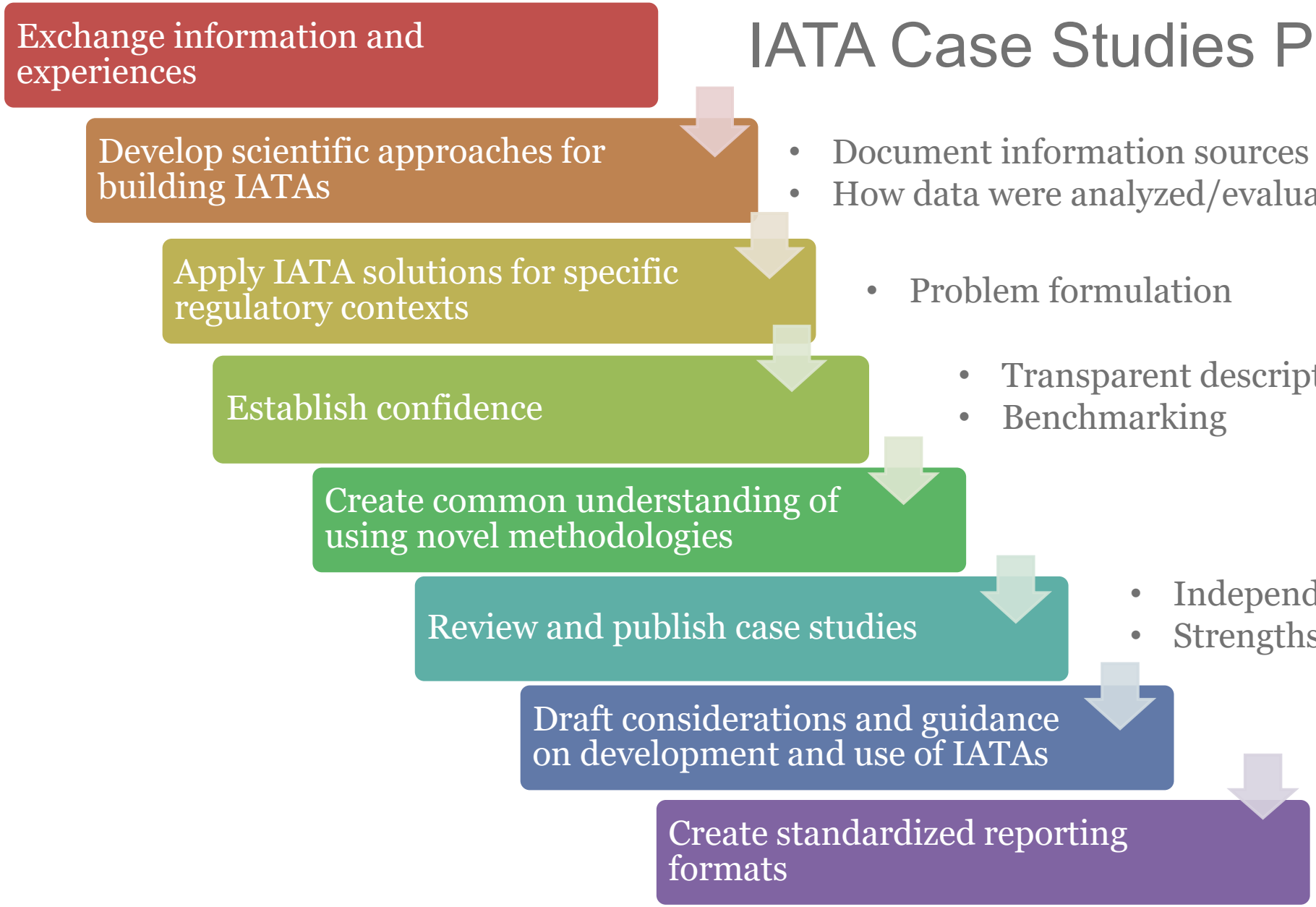
From Pistollato et al., 2021 – focus on human health TGs

OECD Hazard Assessment Programme: Innovative approaches to evaluate chemical hazards

Best approaches and practices for **integrating information** to come to a regulatory decision

- Discussion of use of NAMs in a regulatory context +**identification of aspects that can be harmonised**
- Projects on
  - **IATA Case Studies**
  - Chemical grouping
  - QSAR Toolbox + other electronic tools
  - Omics approaches
  - Various topic-specific guidance documents
- Forum to discuss how to **build confidence in NAMs**
- **Not bound by MAD**
  - thus flexible, innovate approaches, some of which **may become TGs**

IATA Case Studies Project



- Document information sources used,
- How data were analyzed/evaluated/decision criteria

- Problem formulation

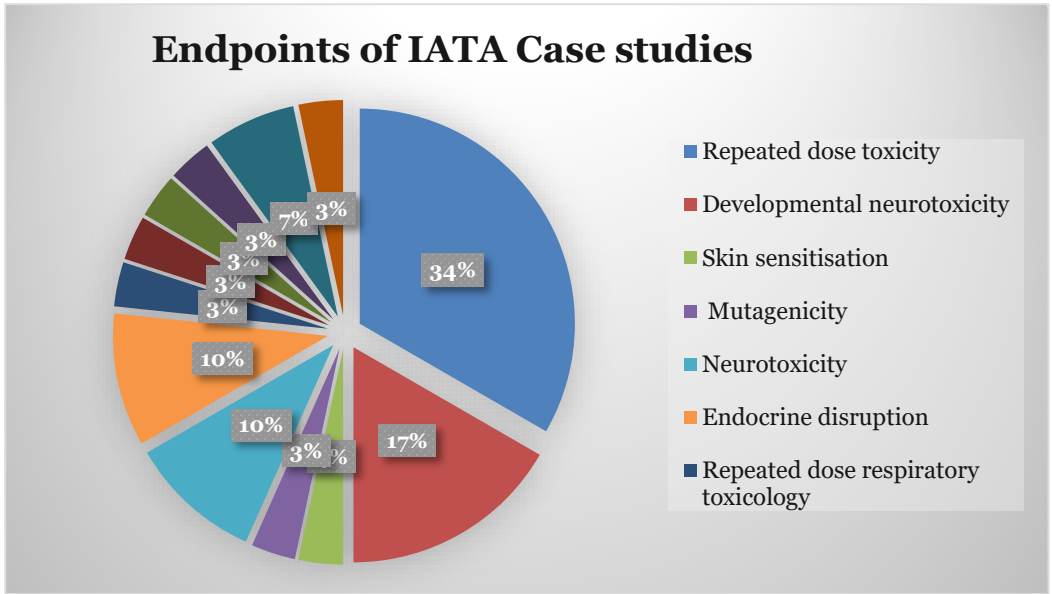
- Transparent description of strengths/limitations
- Benchmarking

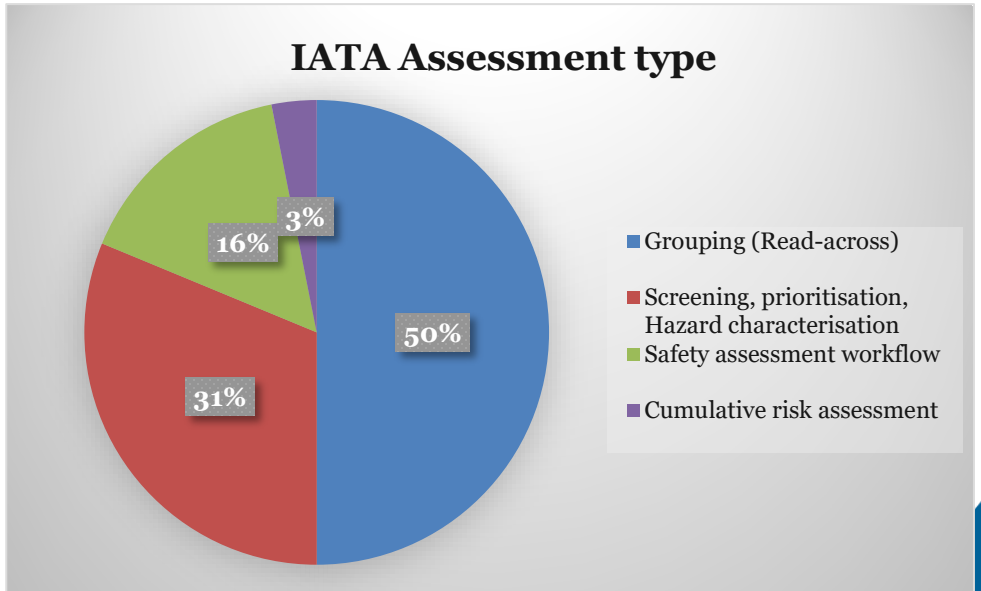
- Independent peer review
- Strengths/limitations/uncertainties

DETERMINE SUITABILITY / FIT FOR PURPOSE

OECD IATA Case Studies Project

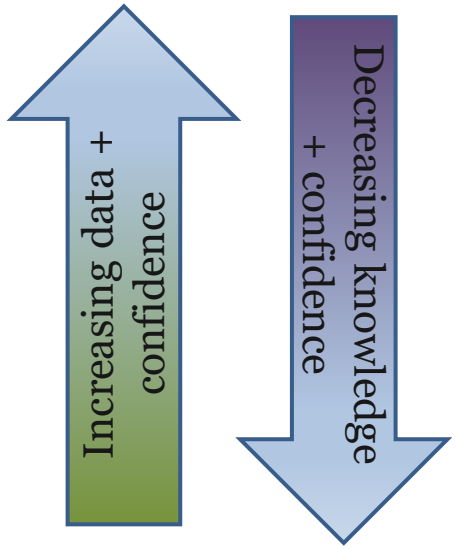
- 8 cycles = 35 cases studies (as of September 2022)
  - use a variety of approaches
  - address various endpoints
- Finalised case studies are published on [OECD website]
- Experiences have led to:
  - New and revised Guidance Documents
  - Data templates and reporting formats to standardise and facilitate exchange of information
  - TG 497 on Defined Approaches for Skin Sensitisation

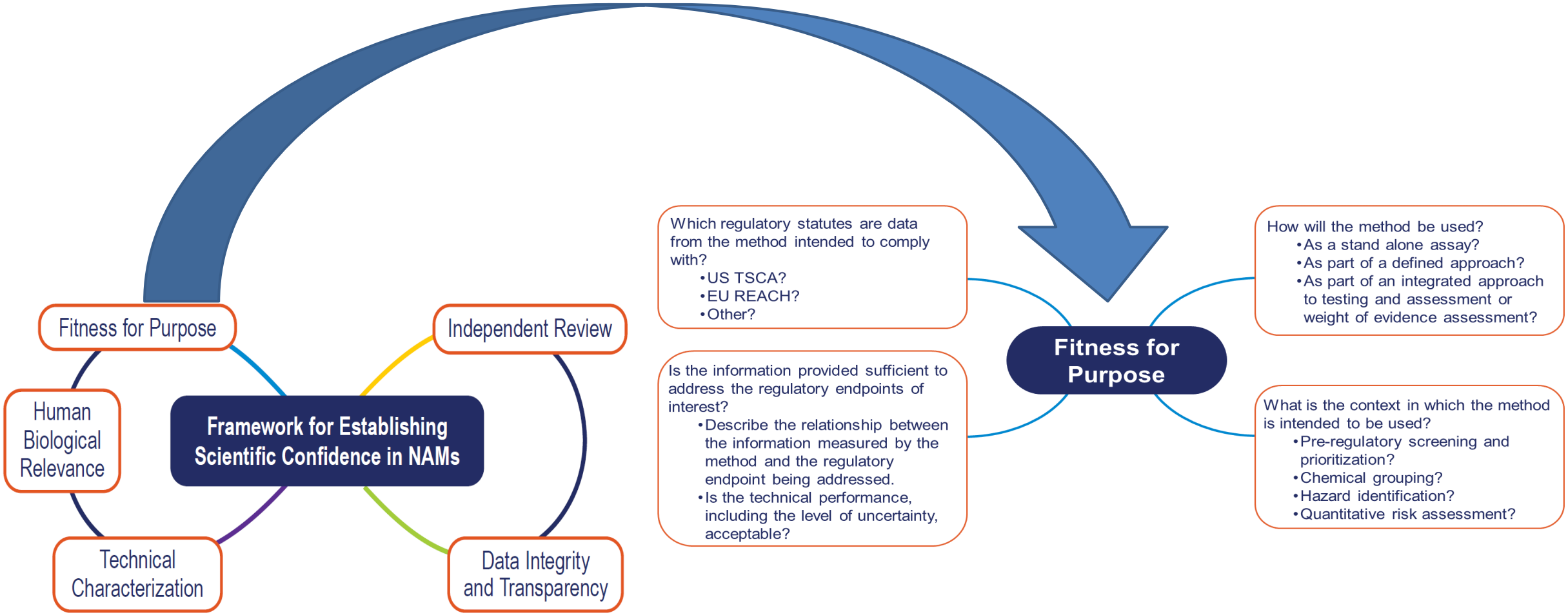




Internationally applicable solutions

- Solutions for a variety of regulatory contexts
  - data rich/data poor chemicals
  - across chemical sectors/regulations
  - various regulatory problem formulations
    - Risk assessment
    - POD
    - Hazard characterisation
    - Hazard identification
    - Prioritisation
- Likely to be a continuum
  - progress towards regulatory application that require more data/less uncertainty as more experience/knowledge is acquired

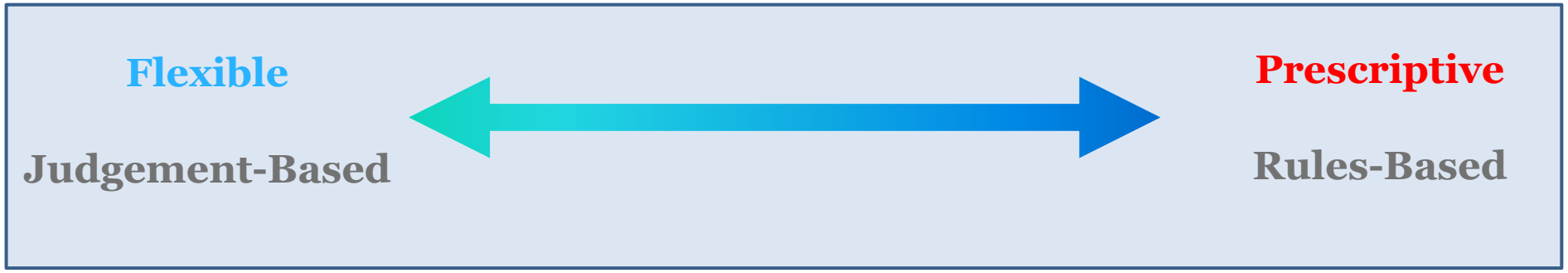




Van der Zalm, AJ; Barroso, J; Browne, P; Casey, W; Gordon, J; Henry, TR; Kleinstreuer, NC; Lowit, AB; Perron, M; Clippinger, AJ. 2022. A framework for establishing scientific confidence in new approach methodologies. A*rchives of Toxicology 2022*, Vol. 1, pp. 1-15

Identification on aspects of IATA that can be standardised: NAMs/IATAs and TG DAs

IATA	Defined Approaches
Designed in response to problem formulation	Designed to address pre-defined endpoint/prediction
Inputs are defined by user	Defined information sources
Sequence of input, next steps, decision context defined by user	Sequence defined and next steps are rule-based
Expert judgement for weighting data, interpreting data	Fixed data interpretation procedure
Conclusion may be open to interpretation	Regulatory conclusion is clear

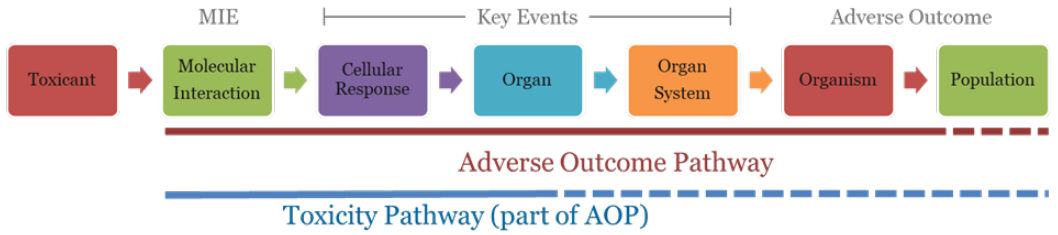


Others are NAMs not under MAD, but a high level of confidence

- Setting up circumstances for **opt-in** use
  - IATA/NAM examples with:
    - Defined context of use
    - Transparent documentation
    - Descriptions of strengths and limitations
    - Peer Review reviews
    - [Maybe met/not met criteria]
      - E.g. criteria for determining state of "readiness" for use in regulatory decisions
        - » WPHA project to **develop assessment framework for QSAR models and predictions**
        - » Establish checklist and criteria for evaluation
        - » Determine aspects that are relevant to other NAMs
  - What else may be needed?

The first wave of NAMs: Mechanistic understanding and AOPs

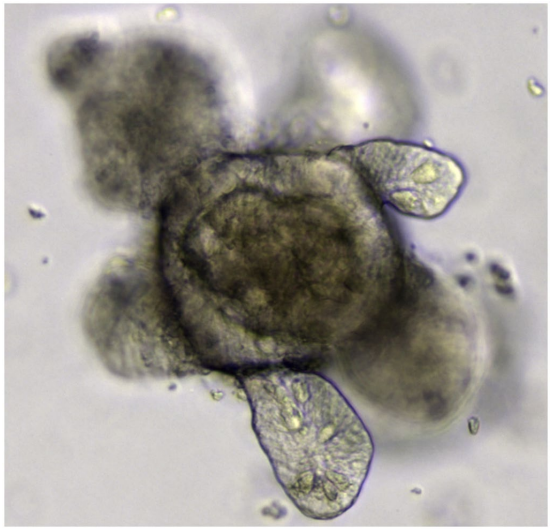
- Pathway defined NAMs (i.e. AOP-amenable):
  - good understanding of mechanisms and key events
  - Establish plausible links between mechanistic and apical responses using existing test data and biological knowledge
  - approaches **predict** an **apical outcome**(s)



- **Not the only option, but proof of concept**

The next wave of NAMs: Physiological understanding

- Pathway undefined NAMs:
  - test systems that mimic biology;
  - perturbation of signalling *could* lead to a variety of outcomes
  - changes are assumed to be undesirable
  - approaches **protective** against **potentially adverse effects**



*approaches not mutually exclusive

Other future evolutions in hazard assessment: 2nd "A" in IATA

- **Mutual Acceptance of Assessments**
  - Complex NAMs/IATAs are beyond just data
  - OECD consideration of opportunities for MAD-like approach for assessment
    - Already experiences and additional opportunities
      - » Biocides
      - » Interest in Joint Reviews of Minor Use Pesticides
      - » Some authorities accept human health risk assessments from trusted authorities for biopesticides

What we need to get there

- Available **data** for review
  - Examples of hazard assessments comparing IATAs to traditional animal test data
  - First Defined Approach Test Guideline was made possible by Cosmetics Europe Database for Skin Sensitisation
    - Hoffman et al. 2017, Kleinstreuer et al. 2017
- Continued engagement
  - IATA Case Study authors and reviewers
  - Communities of practice
- Clusters of Case Studies
  - Using the same approach
  - Evaluating the same endpoint(s)
  - Case Study authors and expert reviewers willing to contribute to guidance for use
- Engagement of regulators and data submitters to provide feedback
  - Retrospective engagement
    - NAMs that are submitted/reviewed
    - challenges/road blocks
    - possible solutions

Evolution of the Test Guidelines Programme

- Workshop in **Dec 2022** on evolving validation practices
  - Opportunity to advance the concept of (performance) standards
  - Discussion of how to validate test systems that are “difficult” to transfer as a block
  - Discussions around steps needed for regulatory application of non-stand alone method(s)
- Goal is to facilitate TGP uptake of emerging technologies

Find out more

Thank You For Listening

Patience.BROWNE@oecd.org

https://www.oecd.org/env/ehs/

Twitter: https://twitter.com/OECD_ENV
YouTube: http://bit.ly/youtube-chemical-safety
Subscribe to our newsletter: http://bit.ly/newsletter-chemical-safety

Draft Outline for the EPA Scientific Confidence Framework

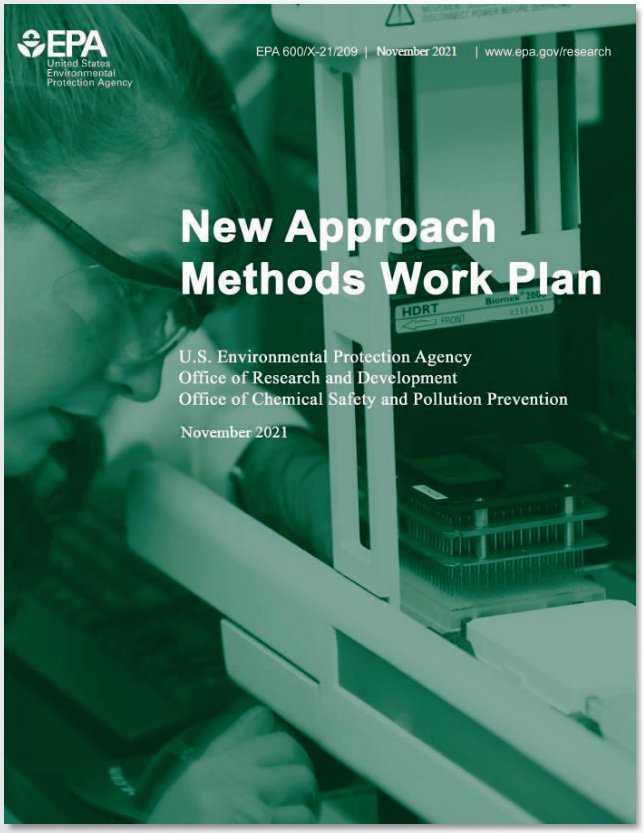
Alison Harrill, PhD

Center for Computational Toxicology and Exposure, Office of Research and Development, US Environmental Protection Agency

Research Triangle Park, NC

The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA

The release of the EPA NAM Work Plan provided clear objectives, strategies and deliverables



- Five objectives for achieving the reduction goals while ensuring that Agency decisions remain fully protective of human health and the environment
  - Evaluate regulatory flexibility
  - Develop baselines and metrics
  - **Establish scientific confidence** and demonstrate application
  - Develop NAMs to address information gaps
  - Engage and communicate with stakeholders
- Changes in 2021 updated work plan:
  - Modified timelines & deliverables through 2024; two case studies
  - Covered species now includes all vertebrate animals, consistent with TSCA
  - Pilot study to develop NAMs training courses for a broad range of stakeholders

Goal of Scientific Confidence Framework

To develop a more generalizable scientific confidence framework that is applicable across a broad range of NAMs and Agency decision contexts.

What is a NAM?

- **NAMs include any technology, methodology, approach, or combination that provides information on chemical hazard and risk assessment while avoiding the use of animal testing.** Examples include *in silico, in vitro,* and *in chemico* approaches.
  - The definition of a NAM has expanded to include new approaches for assessing: hazard, dose response, toxicokinetics, and exposure.

- Use of NAMs allows the **Agency to meet its objective to reduce the reliance on vertebrate animals to test chemicals in evaluating the risks of chemicals, where scientifically justifiable**. The EPA has multiple statutory requirements and policy initiatives that prioritize reduction of animal testing (*e.g.*, the 2018 Toxic Substances Control Act (TSCA) Alternatives Strategic Plan, the Endocrine Disruptor Screening Program for the 21st Century, and the Office of Pesticides Program guidance on waiving acute toxicity studies).

Process to a 2024 Deliverable

Including (but not limited to) NAS report on variability and relevance of current laboratory mammalian toxicity tests and expectations for NAMs for use in human health risk assessment

Initial Framing of Confidence Framework

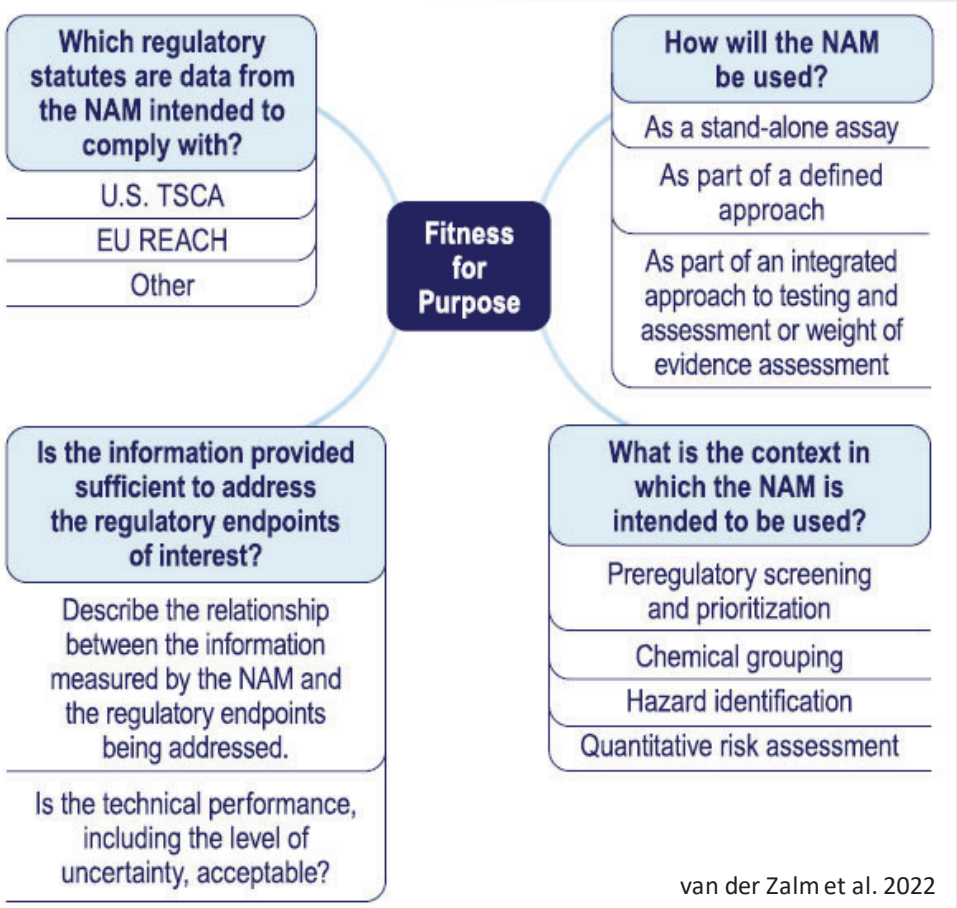
- Many scientific resources emerging, tend to focus on a specific NAM type or applicability domain:
  - OECD guidance document on the validation of (Quantitative)Structure-Activity Relationships [(Q)SAR] models
  - OECD guidance document on good *in vitro* method practices (GIVIMP)
  - Casati, S., et al., *Standardisation of defined approaches for skin sensitisation testing to support regulatory use and international adoption: position of the International Cooperation on Alternative Test Methods.* Arch Toxicol, 2018. **92**(2): p. 611-617.
  - Patlewicz, G., et al., *Proposing a scientific confidence framework to help support the application of adverse outcome pathways for regulatory purposes.* Regul Toxicol Pharmacol, 2015. **71**(3): p. 463-77.
  - van der Zalm, A.J., et al., *A framework for establishing scientific confidence in new approach methodologies.* Arch Toxicol, 2022.
  - Etc!

Essential Elements of Framework

Graphic inspired by figure presented in van der Zalm *et al.* 2022.

Fit for Purpose

The NAM should be **fit-for-purpose** for a specific decision context and the context of use for the NAM should be clearly defined.



van der Zalm et al. 2022

Transparent

The technology, method, and/or analysis procedure associated with the NAM should be **transparently described and sufficiently detailed to enable independent review and evaluation.**

- Depending on the type of NAM, the description of the technology, methods, and analysis procedures should follow scientific best practices and applicable guidance, where available. The underlying principle, technology, and methods for the NAMs should be clearly documented and published in open-access journals or released to public access, made public via government repositories or accessible online servers, and/or summarized in public-facing regulatory or policy documents.
- For commercial NAMs, the computer code, models, or assay system should be available as a commercial service, product, or license.

Transparent

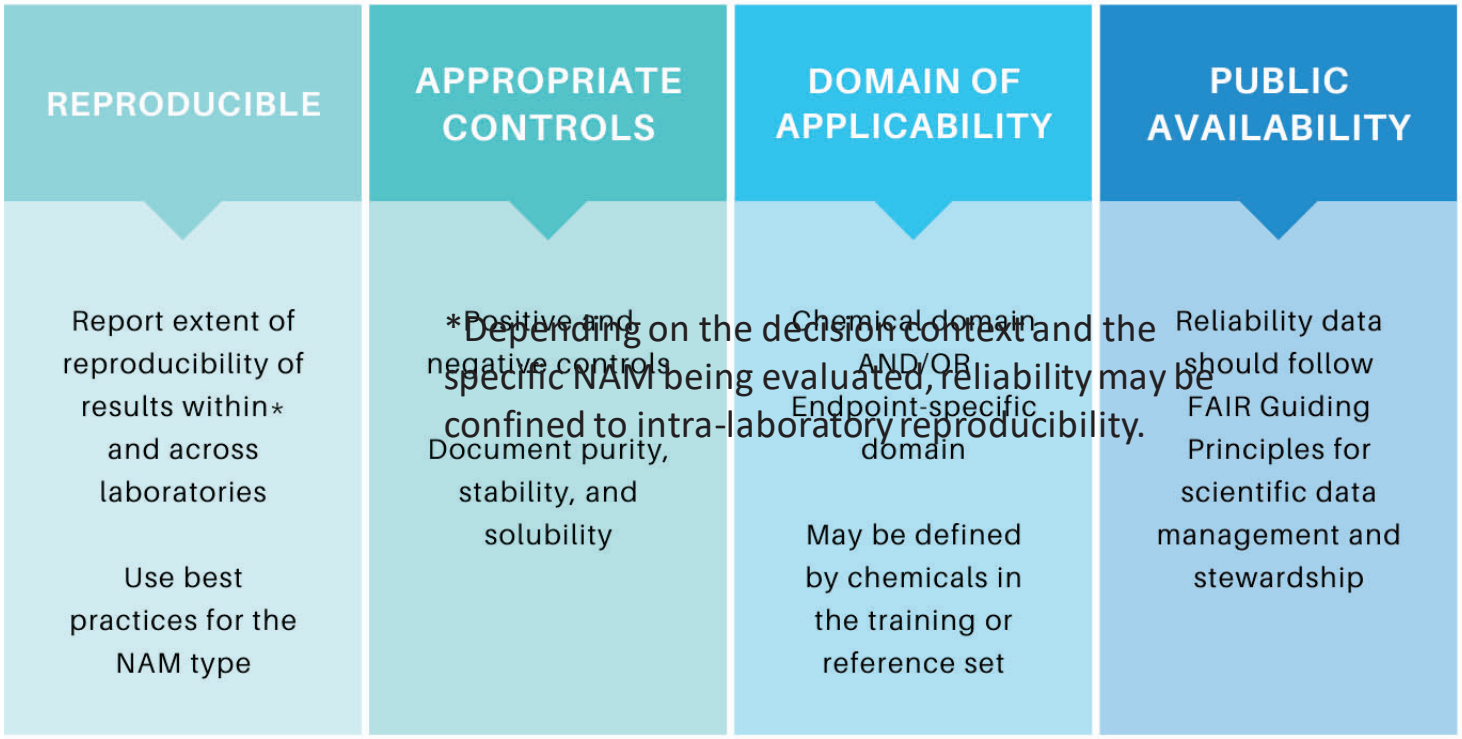
The technology, method, and/or analysis procedure associated with the NAM should be **transparently described and sufficiently detailed to enable independent review and evaluation.**

The NAM(s) should undergo an appropriate level of independent, external review necessary to raise confidence in the approach. Peer review and publication of a NAM's context-informed relevance, fitness-for-purpose, and/or technical characterization is encouraged.

If NAMs are subjected to an independent review, the results of the review should be made publicly available.

Reliable

The **reliability of the NAM** should be characterized, clearly described, and considered within the context of intended use.



Reliable

The **reliability of the NAM** should be characterized, clearly described, and considered within the context of intended use.

REPRODUCIBLE	APPROPRIATE CONTROLS
Report extent of reproducibility of results within* and across laboratories Use best practices for the NAM type	Positive and negative controls Document purity, stability, and solubility

Reliable

The **reliability of the NAM** should be characterized, clearly described, and considered within the context of intended use.

REPRODUCIBLE	APPROPRIATE CONTROLS	DOMAIN OF APPLICABILITY
Report extent of reproducibility of results within* and across laboratories Use best practices for the NAM type	Positive and negative controls Document purity, stability, and solubility	Chemical domain AND/OR Endpoint-specific domain May be defined by chemicals in the training or reference set

Chemical domain of applicability includes chemical structural features, chemical classes, and/or physical-chemical properties that can be confidently evaluated by the NAM as well as those structural features, classes, or physical-chemical properties that may not be confidently evaluated.

Endpoint-specific domain of applicability may include biological-, mechanistic-, temporal-, or process-specific constraints on the use of the NAM. For example, a NAM may be applicable to only certain species, potency classes, or exposure scenarios.

Reliable

The **reliability of the NAM** should be characterized, clearly described, and considered within the context of intended use.

REPRODUCIBLE	APPROPRIATE CONTROLS	DOMAIN OF APPLICABILITY	PUBLIC AVAILABILITY
Report extent of reproducibility of results within* and across laboratories Use best practices for the NAM type	Positive and negative controls Document purity, stability, and solubility	Chemical domain AND/OR Endpoint-specific domain May be defined by chemicals in the training or reference set	Reliability data should follow FAIR Guiding Principles for scientific data management and stewardship

Relevance

The **relevance of the NAM for the intended use should be described** to the extent possible**.**

Relevance to the endpoint being evaluated should be clearly described.

The mechanistic interpretability of the NAM and direct scientific linkage to the regulatory endpoint being assessed is desirable and reduces uncertainty in the applicability of NAM.

Uncertainty

Uncertainties relating to the NAM should be well-described.

a. Uncertainty refers to a lack of data or an incomplete understanding of NAM components, inputs, or outputs and their relationship to the regulatory decision. Uncertainty can be qualitative or quantitative. During evaluation, the uncertainties of the NAM should be described and reported relative to the chemical- and endpoint-specific domains of applicability.
b. Where appropriate, applicable uncertainties for the NAM should be presented relative to uncertainties associated with standard or traditional approaches that the NAM seeks to replace.
c. Depending on the NAM and its context of use, the acceptable level of uncertainty associated with the NAM may vary.

Acknowledgements

Implementation Team

EPA/ORD
Annette Giuseppe-Elie
Monica Linnenbrink
Rusty Thomas

EPA/OCSPP
Tala Henry
Anna Lowit
Monique Perron
Krystle Yozzo

With expert feedback from (ORD):

Gary Ankley
Tim Buckley
Kristin Isaacs
Jason Lambert
Grace Patlewicz
John Wambaugh
Tony Williams
Dan Villeneuve