September 3, 2014

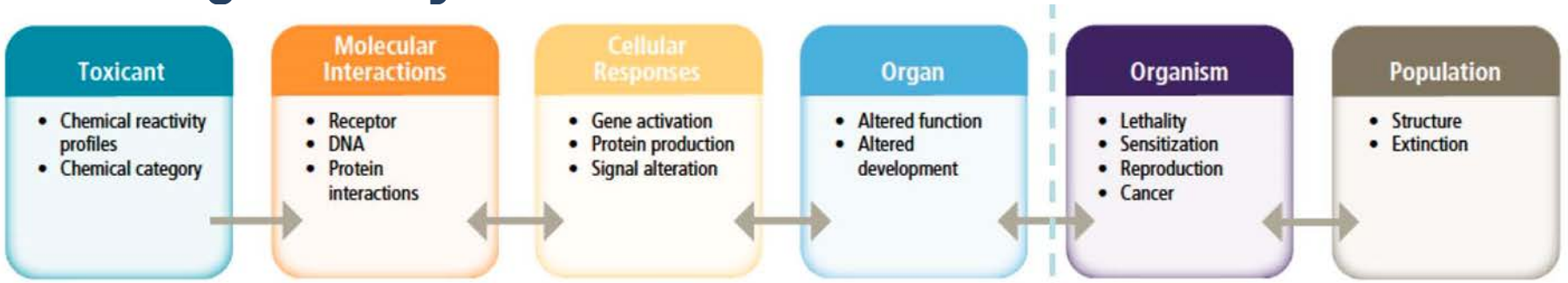
HOW DO WE ESTABLISH SCIENTIFIC CONFIDENCE IN AOPS?

Richard A Becker Ph.D. DABT

American® Chemistry Council

Scientific Confidence is a Necessity

- **Must have confidence to rely on AOPs for decision making in product stewardship & regulatory actions**



- **Problem Formulation --- degree of confidence depends on intended use:**
  - E.G., priority setting would require less confidence than risk assessment

Regulatory Toxicology and Pharmacology 65 (2013) 259–268

Contents lists available at SciVerse ScienceDirect

Regulatory Toxicology and Pharmacology

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

Use and validation of HT/HC assays to support 21st century toxicity evaluations

Grace Patlewicz [a,*], Ted Simon [b], Katy Goyak [c], Richard D. Phillips [c], J. Craig Rowlands [d], Shawn D. Seidel [e], Richard A. Becker [f]

- Considered how existing frameworks could be adapted:
  - OECD's "Validation Principles for (Q)SAR"
  - Inst. of Medicine's "Evaluation of Biomarkers and Surrogate Endpoints"

Regulatory Toxicology and Pharmacology 69 (2014) 443–450

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

Developing scientific confidence in HTS-derived prediction models: Lessons learned from an endocrine case study

Louis Anthony (Tony) Cox [a], Douglas Popken [a], M. Sue Marty [b], J. Craig Rowlands [b], Grace Patlewicz [c], Katy O. Goyak [d], Richard A. Becker [e,*]

open access:

http://www.ncbi.nlm.nih.gov/pubmed/24845243

A focus on this case example helped refine our thinking on a framework and extend it to AOPs

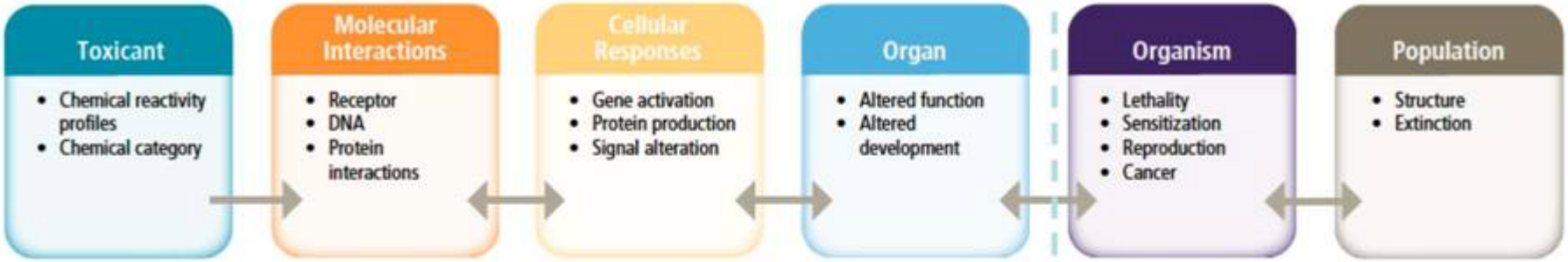
Scientific Confidence Framework for AOPs

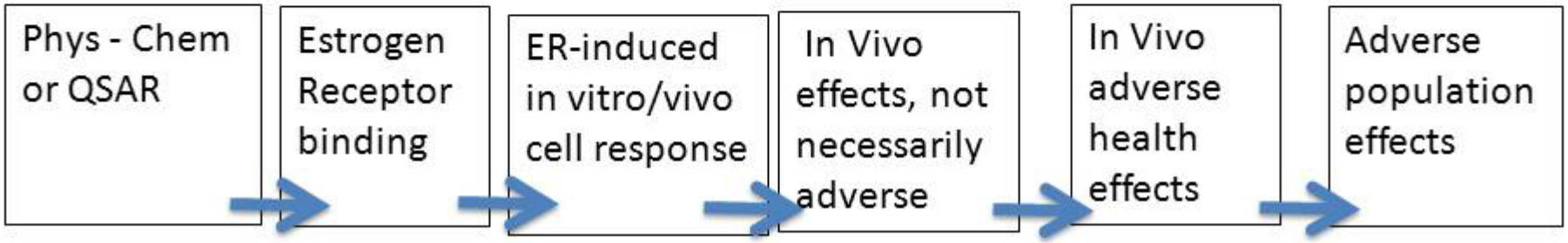
(adapted from Cox et al. 2014 Reg Tox Pharm)

1	**Develop the AOP**
2	**Develop new (or map existing) specific assays to key events within the AOP**
3	Conduct (or document) **Analytical Validation** of each assay
4	Develop new (or map existing) models that predict a specific key event from one or more pre-cursor key events. (The input data for the prediction models comes from the assays described in Steps 2 and 3 above.)
5	Conduct (or document) **Qualification** of the prediction models
6	**Utilization**: defining and documenting where there is sufficient scientific confidence to use one or more AOP-based prediction models **for a specific purpose** (e.g., priority setting, chemical category formation, integrated testing, predicting *in vivo* responses, etc.)
7	For regulatory acceptance and use, processes need to be agreed upon and utilized to ensure robust and transparent review and determination of fit for purpose uses of AOPs. This should include dissemination of all necessary datasets, model parameters, algorithms, etc., to enable stakeholder review and comment, fully independent verification and independent scientific peer review. While these processes have yet to be defined globally, in time, these should evolve to enable scientific confidence and credible and transparent use of AOPs.

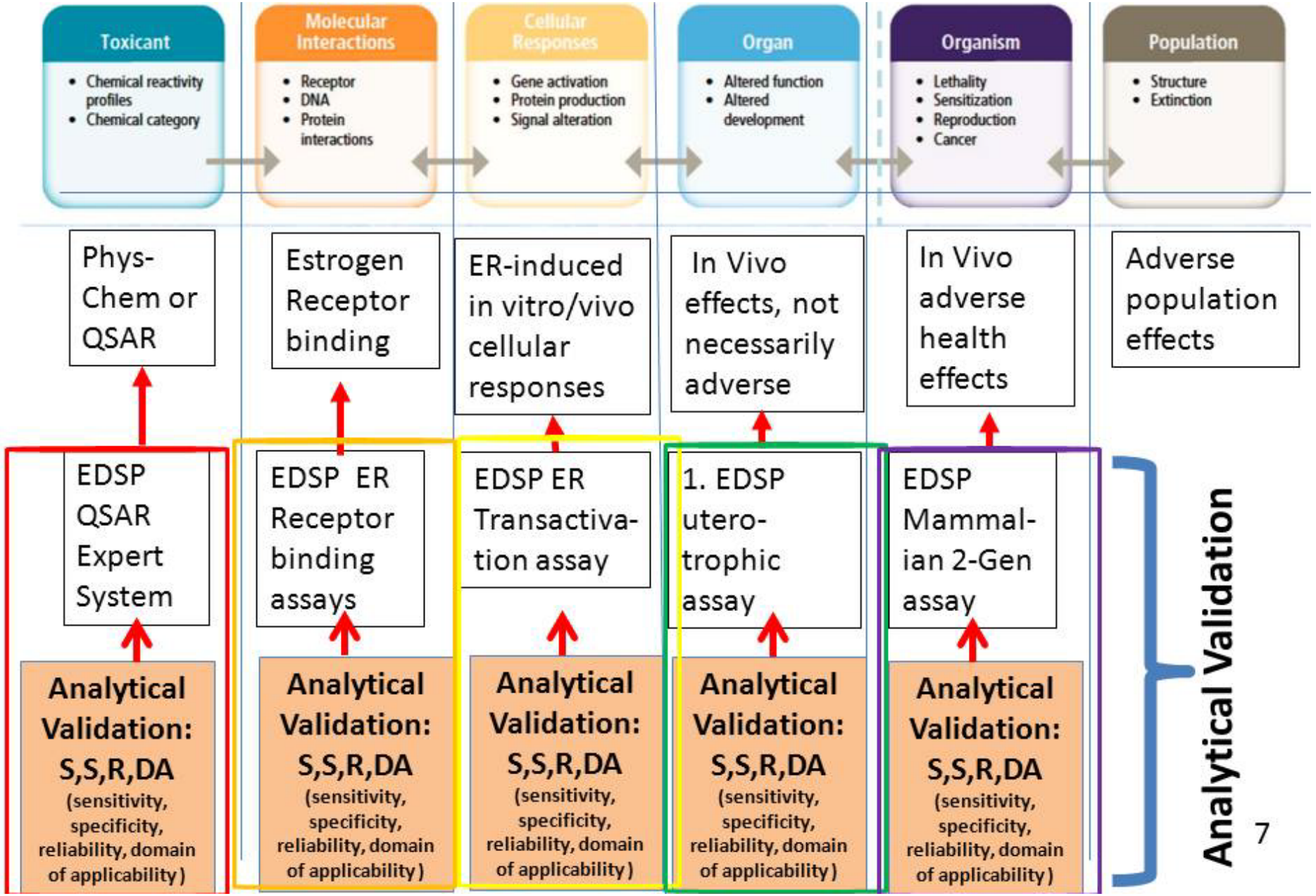
This framework was presented at 2014 SOT : The Toxicologist, Abstract 2253, page 602.

Prototype -- Estrogen AOP





Mapping assays to the AOP



Scientific Confidence Framework for AOPs

1. Analytical Validation

Assessment of the biological basis and analytical performance of assays.

- Each assay should map to a defined mechanistic endpoint (e.g., a key event in the mode of action or AOP).
- Documentation of assay performance characteristics (reliability, sensitivity, and specificity)
- A defined chemical domain of applicability
- Transparent data sets (to enable independent verification) should be readily available.

Note This framework was used to form the basis of the draft OECD guidance “Characterizing non-guideline in vitro test methods to facilitate their consideration in regulatory applications”

Scientific Confidence Framework for AOPs

2. Qualification

Assessment of the prediction model derived from the assays.

- A defined algorithm for each prediction model.
- Appropriate measures of goodness-of-fit, robustness and predictivity of the prediction models (models may be quantitative or qualitative).
- Known limitations of each prediction model should be summarized.
- Prediction models should be characterized in sufficient detail to facilitate review, reconstruction and independent verification of results.

Predicting E, A, T & S *in vivo* from ToxCast Results

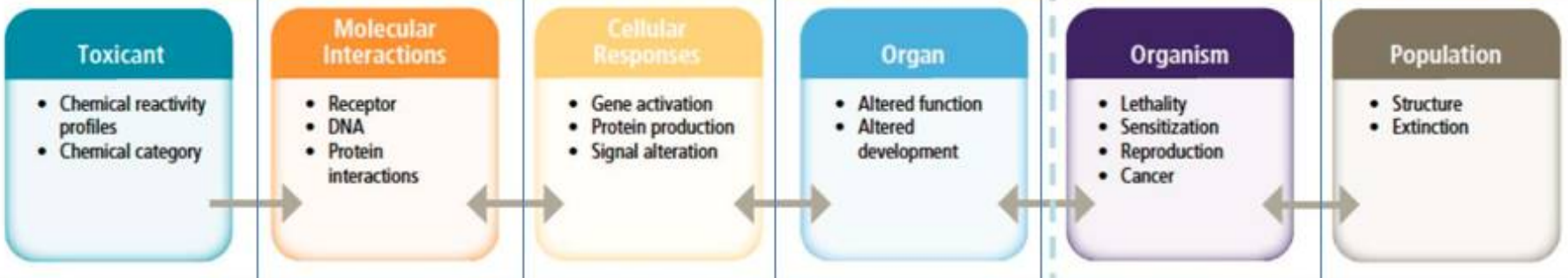
Using *in Vitro* High Throughput Screening Assays to Identify Potential Endocrine-Disrupting Chemicals Environ Health Perspect. 2013 Jan;121(1):7-14.

Daniel M. Rotroff,[1,2] David J. Dix,[2] Keith A. Houck,[2] Thomas B. Knudsen,[2] Matthew T. Martin,[2] Keith W. McLaurin,[2] David M. Reif,[2] Kevin M. Crofton,[3] Amar V. Singh,[4] Menghang Xia,[5] Ruili Huang,[5] and Richard S. Judson[2]

The authors concluded:

ToxCast estrogen receptor-mediated and androgen receptor-mediated assays predicted the results of relevant EDSP T1S assays with balanced accuracies of 0.91 ($p < 0.001$) and 0.92 ($p < 0.001$), respectively.

Models for steroidogenic and thyroid-related effects could not be developed with the currently published ToxCast data.



Phys - Chem or QSAR	Estrogen Receptor binding	ER-induced in vitro/vivo cell response	In Vivo effects, not necessarily adverse	In Vivo adverse health effects	Adverse population effects

Qualification

measurements

predict

Qualification:
Model uses measurements of molecular interactions and cellular responses to predict in vivo responses

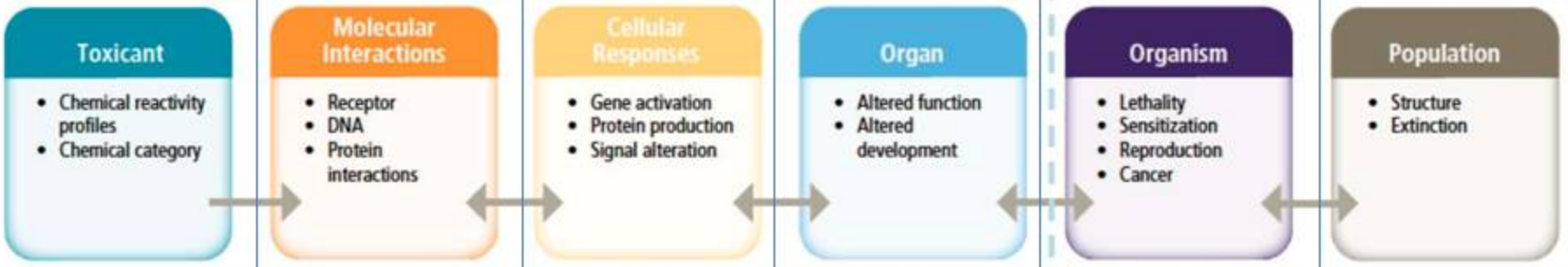
Class	Model-prediction (CV, Cox)
E	0.85
A	0.79
S	0.50
T	0.23

Scientific Confidence Framework for AOPs

3. Utilization

Contextual and weight-of-evidence analysis of the use (qualitative or quantitative) of the prediction model for a specific purpose.

- Defining the intended purpose of the prediction model
- Documenting/justifying applications, based on weight of evidence, of the scientific confidence to support the use of the AOP
  - (1) priority setting, where the model is used to identify priority substances for more detailed evaluation;
  - (2) chemical categorization for subsequent read-across
  - (3) screening level assessment of a biomarker, where model is used as a surrogate data point for a biochemical endpoint or a biomarker;
  - (4) integrated testing strategy, or where the model is used to describe/predict a hazard property in lieu of a traditional tox study
  - (5) to predict an adverse outcome.



E.g., Pchem or QSAR results

E.g., Receptor binding results

E.g., In vitro cell culture results

E.g., In vivo effects, not necessarily adverse

E.g., adverse health effects

E.g., adverse population effects

Qualification

measurements

predict

Qualification:
Model uses measurements of molecular interactions and cellular responses to predict in vivo responses

Class	Model-prediction (CV, Cox)
E	0.85
A	0.79
S	0.50
T	0.23

Prediction Model Evaluation: "I want to predict Y from assays that measure X." Starting from one step in the AOP and going to the "predicted response" step in AOP

For Utilization: Discussion is Needed by the Regulatory Science Community

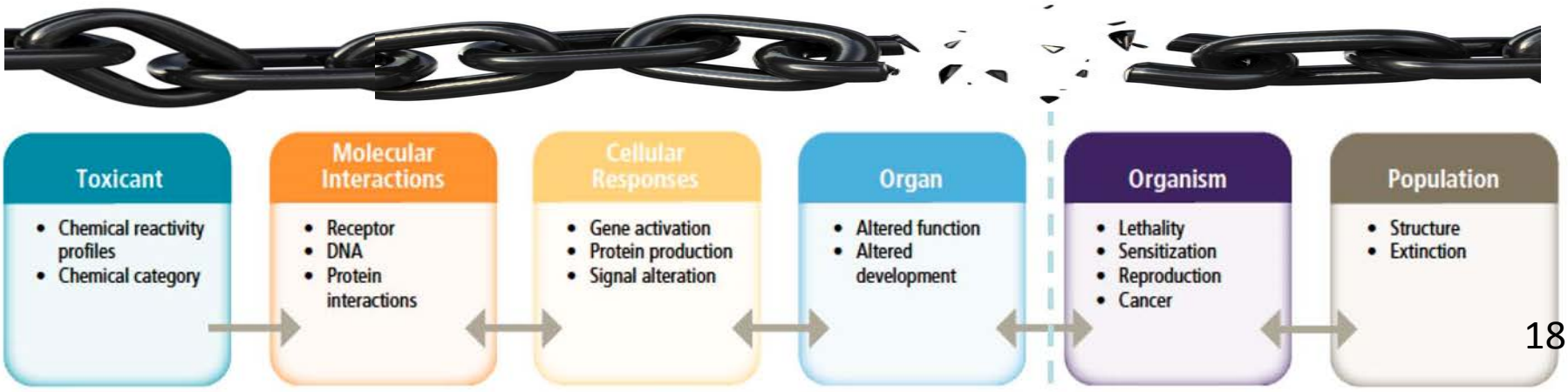
The balanced accuracies for prediction signal significant advancement in developing biologically-based HTS-derived models for E and A endocrine activities.

- *How accurate must predictivity be for prioritizing substances for E and A for screening? For bypassing certain E or A EDSP Tier 1 assays?*
- *Are different levels of uncertainty / confidence OK for different uses (e.g. priority setting vs. waiver), and, if so, what are these, in quantitative scientific terms?*
- *How can integration of exposure with HTS activity-based measures/predictions (Wetmore et al., 2012; Becker et al., 2014) provide an improved context for decision making?*

What Happens When….

- a specific KE or KER is judged, using the OECD Handbook WOE determination, to be weak?
  - one is likely to have low confidence proceeding along the AOP pathway beyond that KE or KER to predict subsequent KEs or the AO.

this weak level of confidence represents a weak
r a break in the causal chain.



Scientific Confidence Framework for AOPs

(adapted from Cox et al. 2014 Reg Tox Pharm)

1	**Develop the AOP**
2	**Develop new (or map existing) specific assays to key events within the AOP**
3	Conduct (or document) **Analytical Validation** of each assay
4	Develop new (or map existing) models that predict a specific key event from one or more pre-cursor key events. (The input data for the prediction models comes from the assays described in Steps 2 and 3 above.)
5	Conduct (or document) **Qualification** of the prediction models
6	**Utilization**: defining and documenting where there is sufficient scientific confidence to use one or more AOP-based prediction models **for a specific purpose** (e.g., priority setting, chemical category formation, integrated testing, predicting *in vivo* responses, etc.)
7	For regulatory acceptance and use, processes need to be agreed upon and utilized to ensure robust and transparent review and determination of fit for purpose uses of AOPs. This should include dissemination of all necessary datasets, model parameters, algorithms, etc., to enable stakeholder review and comment, fully independent verification and independent scientific peer review. While these processes have yet to be defined globally, in time, these should evolve to enable scientific confidence and credible and transparent use of AOPs.

This framework was presented at 2014 SOT : The Toxicologist, Abstract 2253, page 602.

How Will Confidence & Regulatory Acceptance Be "Officially" Established?

How should utilization and regulatory acceptance of an AOP be determined and communicated?

For example:

- a public, transparent vetting process by knowledge-able scientists representing all stakeholders?
- by an authoritative body (EPA?, OECD?, ECHA?, ICCVAM/ECVAM?) in a manner that includes rep of stakeholders?
- Wiki-type process?
- Other ????? Will a peer reviewed published paper be enough?