

Unclassified

ENV/JM/MONO(2015)22

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

22-Jul-2015

English - Or. English

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

Cancels & replaces the same document of 29 June 2015

REPORT OF THE WORKSHOP ON A FRAMEWORK FOR THE DEVELOPMENT AND USE OF
INTEGRATED APPROACHES TO TESTING AND ASSESSMENT

Series on Testing and Assessment
No. 215

17-19 November 2014, Crystal City VA, USA

Mr. Joop DE KNECHT
Tel: +33 (0) 1 45 24 82 57; Email: joop.deknecht@oecd.org

JT03380373

Complete document available on OLIS in its original format

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

ENV/JM/MONO(2015)22
Unclassified

English - Or. English

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 215

**REPORT OF THE WORKSHOP ON A FRAMEWORK FOR THE DEVELOPMENT AND USE
OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT**

IOMC

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

A cooperative agreement among **FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD**

Environment Directorate
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT
Paris 2015

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 34 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in eleven different series: **Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides; Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; and Safety of Manufactured Nanomaterials.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (<http://www.oecd.org/chemicalsafety/>).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organisations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

This publication is available electronically, at no charge.

Also published in the Series on Testing and Assessment link

**For this and many other Environment,
Health and Safety publications, consult the OECD's
World Wide Web site (www.oecd.org/chemicalsafety/)**

or contact:

**OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France**

Fax: (33-1) 44 30 61 80

E-mail: ehscont@oecd.org

© OECD 2015

Applications for permission to reproduce or translate all or part of this material should be made to: Head of Publications Service, RIGHTS@oecd.org, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France

FOREWORD

This document is a report of the Workshop on a framework for the development and use of integrated approaches to testing and assessment which was held on 17-19 November 2014 in Crystal City VA, USA. The workshop was organised in close cooperation with the World Health Organisation following the proposal from the 50th OECD Joint Meeting of the Chemicals Committee and Working Party on Chemicals, Pesticides and Biotechnology in June 2013.

This document is published under the responsibility of the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology of the OECD.

TABLE OF CONTENTS

background.....	8
Workshop.....	8
Opening.....	8
Purpose, Objectives And Specific Aims	8
Format Of The Workshop	9
Case Study 1: Use Of The Aop For Ar / Er Binding / Thyroid Effect To Prioritise And Screen Chemicals.....	9
Case Study 2: Adverse Outcome Pathway For Skin Sensitization.....	9
Case Study 3: Development Of Iatas Based On The Aop Of Sustained Aryl Hydrocarbon Receptor (Ahr) Activation Leading To Rodent Liver Tumour Promotion.....	10
Case Study 4: Aromatase Inhibition Leading To Reproductive Dysfunction (In Fish). A Quantitative Aop Case Study.....	10
Outcome Of Aop Related Workshops.....	10
Working Definitions Of The Workshop.....	10
General Questions	11
Breakout Group Conclusions And Recommendations	12
Conclusions And Recommendation Derived From Additional Questions.....	13
Annex 1 Workshop Agenda	17
Annex 2 The Adverse Outcome Pathway (AOP)/Mode Of Action (MOA) Concept: A Framework For The Development And Use Of Integrated Approaches To Testing And Assessment (IATA).....	18
Annex 3 General Principles For Using IATA Based On The AOP/MOA Concept.....	26
Annex 4 Presentation: Oecd Framework For Integrated Approaches For Testing And Assessment.....	30
Annex 5 Presentation: The Who Framework On Mode Of Action/Species Concordance Analysis Implications For AOPs/IATA	50
Annex 6 Presentation: TK And TD Modelling Approaches Relevance For Using MOA/AOP In IATA.....	65
Annex 7 Presentation Case Study 1: The Use Of The AOP For AR / ER Binding Thyroid Effect To Prioritise And Screen Chemicals.....	75
Annex 8 Presentation Case Study 2: Adverse Outcome Pathway For Skin Sensitization	88
Annex 9 Presentation Case Study 3: Development Of IATAs Based On The Adverse Outcome Pathway (AOP) Of Sustained Ahr Activation Leading To Rodent Liver Tumor Promotion	101
Annex 10 Presentation Case Study 4: Aromatase Inhibition Leading To Reproductive Dysfunction (In Fish). A Quantitative Aop Case Study	114
Annex 11 Presentation Outcomes From The Somma Lombardo Workshop: Advancing Adverse Outcome Pathways For Integrated Toxicology And Regulatory Applications.....	129
Annex 12 Presentation Outcomes From The Workshop On The Adverse Outcome Pathways: From Research To Regulation	144
Annex 13 Participants List For Workshop On A Framework For The Development And Use Of Integrated Approaches To Testing And Assessment.....	150

BACKGROUND

1. In order to improve the harmonisation of integrated approaches to testing and assessment (IATA), the 47th Joint Meeting recommended to elaborate an OECD agreed framework for developing and using IATAs, building on current activities on Mode of Action (MoA) and Adverse Outcome Pathways (AOP). This framework should provide guiding principles, and technical guidance on how results from alternative approaches (*in silico*, *in chemico*, *in vitro* including high throughput and high content test methods) should be interpreted for characterising (both qualitatively and quantitatively) the adverse effects in animals and humans and/ or the environment, so that they can be used for hazard identification, hazard characterisation and risk assessment.

2. The 50th Joint Meeting agreed to organise a workshop in close cooperation with the World Health Organisation, bringing regulators, scientists, industry and NGOs together to define in a practical manner the applicability of the concept of AOP/MoA in a framework for the development and use of IATAs.

WORKSHOP

3. The workshop was held on 17-19 November 2014 in Crystal City, VA, hosted by the United States Environmental Protection Agency (USEPA). The agenda is outlined in Annex 1.

4. The workshop was attended by experts nominated by Canada, Denmark, France, Japan, Germany, Switzerland, the United Kingdom, the Netherlands, the United States, the European Commission, BIAC, ICAPO, and the OECD Secretariat. The list of the participants is attached to this document as Annex 13.

5. The workshop was chaired by Terry Schultz (University of Tennessee, Knoxville) and Mark Cronin (Liverpool John Moores University).

Opening

6. Tala Henry (USEPA) welcomed the participants on behalf of the United States.

Purpose, Objectives and Specific Aims

7. The objective of this workshop was to discuss the applicability of the AOP / MoA concept as a framework for developing and using IATAs and to refine the framework as far as possible and define the degree of confidence in an AOP / MoA needed to inform an IATA in a specific regulatory context that can then be communicated throughout the decision making process.

8. To meet these objectives a background document on the AOP / MoA concept as the basis for developing and using IATAs was prepared together with a document outlining the general principles for using IATA based on the AOP / MoA concept including an initial definition and tentative set of considerations for different stages of development of AOPs / MoAs. These documents are presented in Annex 2 and 3.

9. The conclusions and recommendations from the workshop will subsequently be used to revise the framework as a basis for further testing in case studies. Furthermore the outcome of the workshop will be used by the OECD Task Force on Hazard Assessment within the cooperative work on the hazard

assessment of chemicals to develop further case studies and provide guidance on how IATAs based on the framework can be used for regulatory purposes.

Format of the Workshop

10. The workshop started with a general introduction into the framework for the development and use of integrated approaches to testing and assessment by the OECD Secretariat and the WHO Framework on Mode of Action/Species Concordance Analysis Implications for AOPs/IATA by Bette Meek (University of Ottawa). These presentations are found in Annex 4 and 5. Jos Bessems of the EC Joint Research Centre gave an overview on how toxicokinetic and toxicodynamic data could be integrated in an AOP-informed IATA (see Annex 6). These presentations were followed by a series of case study presentations. Subsequently, participants, in three breakout groups, were asked to consider a set of questions designed to:

- a. build consensus that the AOP / MoA concept is a good basis for developing and using IATAs,
- b. determine which type of IATA, or their respective elements, can be informed by an AOP / MoA,
- c. to discuss and refine the proposed definitions and characteristics of the different stages of development of AOPs / MoAs, and to
- d. outline / propose which stages of development of AOPs/MOAs are most suitable to inform the development and use of different types of IATA or their respective elements and their regulatory purposes.

Case Study 1: Use of the AOP for AR / ER binding / thyroid effect to prioritise and screen chemicals

11. Kevin Crofton of the USEPA presented the first case study, on the use of the AOP for estrogen receptor (ER) binding / thyroid effect as a potential basis to prioritise and screen chemicals. With respect to the use of an AOP for ER binding in an IATA, multiple assays integrated into a consensus model in combination with exposure estimates provides a prioritization tool for follow-up testing. Due to the complexity and multiple AOPs and the lack of assays to measure the different MIEs involved, there is at present no confidence in making regulatory decisions related to thyroid effects. IATA process provides unique opportunity to fine-tune data needs to predict the adverse outcome. This presentation is found in Annex 7.

Case Study 2: Adverse Outcome Pathway for Skin Sensitization

12. Frank Faulhammer of BASF presented case study two on the proposed use of the Adverse Outcome Pathway for Skin Sensitization for grouping of substances, read-across and regulatory application. In this case study it is shown that the AOP for skin sensitization fulfils aspects of semi-quantitative and quantitative AOPs and that single assays that cover key events of the AOP may be used to prioritize chemicals for testing, add confidence to a read-across approach and help to group chemicals. It was also illustrated how the AOP for skin sensitization could be used to develop an *in vitro* testing strategy, which could be used to determine whether a chemical could be a sensitizer or not but is not yet applicable to determine potency (e.g. strong, moderate and weak), to sub-categorize according to GHS (e.g. Cat. 1A or 1B) or to assess complex mixtures/substances such as polymers and formulations. This presentation is found in Annex 8.

Case Study 3: Development of IATAs Based on the AOP of Sustained Aryl Hydrocarbon Receptor (AHR) Activation Leading to Rodent Liver Tumour Promotion

13. Katy Goyak of ExxonMobil Biomedical Sciences presented case study three, development of IATAs based on the AOP of Sustained AHR Activation Leading to Rodent Liver Tumour Promotion. In this case study it is shown how a semi-quantitative AOP could potentially inform the development of an integrated testing strategy using exposure considerations. This presentation is found in Annex 9.

Case Study 4: Aromatase inhibition leading to reproductive dysfunction (in fish). A quantitative AOP case study

14. Daniel Villeneuve and Rory Conolly of the USEPA presented case study four, on a quantitative AOP for Aromatase inhibition leading to reproductive dysfunction (in fish). The case study showed that there is strong scientific confidence that *in vivo* observation of sustained reductions in ovarian aromatase activity, circulating estradiol, and circulating vitellogenin are indicative of reproductive effects (in fish). Structural alerts and *in vitro* measures indicative of aromatase inhibition provide strong evidence of potential reproductive effects (in fish) and could be viable alternatives to fish short-term reproduction assay as a Tier 1 screen. The AOP could also support the development of a tiered or sequential testing strategy for aromatase inhibitors. Overall it is proposed that mature quantitative AOPs could serve as an “*in silico*” description of *in vivo* biology to aid in the design of *in vitro* tests and interpretation of *in vitro* data. This presentation is found in Annex 10.

Outcome of AOP related workshops

15. Ed Perkins from the US Army Engineer Research & Development Centre presented the outcomes from the workshop "Advancing AOPs for Integrated Toxicology and Regulatory Applications" held 2nd-7th March 2014 at Somma Lombardo 2014 (see annex 11). Michelle Embry from the Health & Environmental Sciences Institute presented the Outcomes from the workshop "AOPs: From Research to Regulation" held 3rd-5th September 2014 at Bethesda (see annex 12).

Working definitions of the workshop

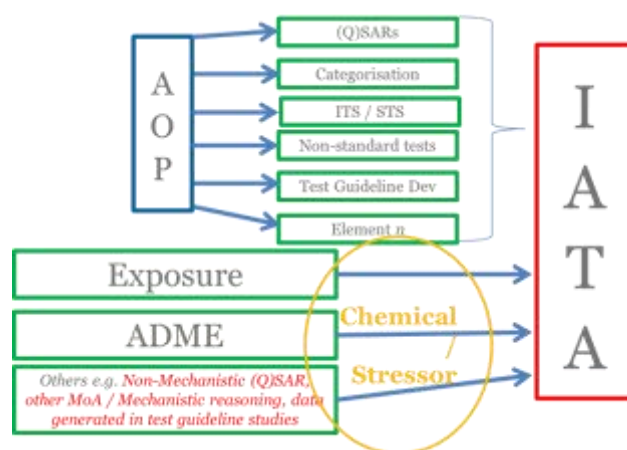
16. After the presentations of the case studies the participants divided into breakout groups initially for a general discussion of the development of IATAs based on the AOPs with later discussions leading to the answers to the questions posed to the participants.

17. At the start of the breakout group discussion, participants were presented with the following working definition of IATA: *a structured approach that strategically integrates and weights all relevant data to inform regulatory decisions regarding potential hazard and/or risk and/or the need for further targeted testing and therefore optimising and potentially reducing the number of tests that need to be conducted.*

18. The participants agreed that an IATA logically starts with problem formulation, including the decision context, consideration of plausible and testable hypotheses about the hazard profile of a substance or group of substances, and, in some cases exposure information. The hazard information together with the exposure information would then be used to determine what data gaps exist and what testing if any would be most appropriate to undertake in order to elucidate the hazard profile of that substance for a given use context. Thus the extent to which testing approaches are needed therefore depends on the problem formulation which in turn is defined by the end purpose under consideration and the scientific confidence needed.

19. While the strategy for gathering existing information and use of information from non-testing methods in an IATA may be similar regardless of the decision context, generation of new test data (e.g. data from *in chemico* or *in vitro* methods) may differ considering the scope of the IATA and the collected evidence and should be tailored to reduce uncertainty in the initial conclusion. Evaluation of existing information or generation of additional data within an IATA can be performed on the basis of a non-formalised Weight of Evidence (WoE) approach or by using predefined, structured approaches such as Sequential Testing Strategies (STS), Integrated Testing Strategies (ITS) or their combination.

20. An IATA may be comprised of one or more elements. These elements can be informed by an AOP, e.g. SAR / QSAR, testing assays etc., or could also contain elements that are not informed by an AOP, such as exposure, ADME, use profiling, etc. It was noted that AOPs are one means to structure an IATA but not necessarily required. It was recommended to develop more case studies for different decision contexts (or problem formulations) to better understand how IATA can be constructed based on these different elements (see figure 1).



21. It was agreed that Modes of Action and Adverse Outcome Pathways are conceptually similar, dividing the path between exposure and effect into key events. MOAs include some chemical specific elements such as metabolism, whereas AOPs are restricted to the non-chemical specific biological pathway, and the final outcome of a MOA is not necessarily adverse. MOA analysis has been chemical specific, including both kinetics and metabolism.

General questions

22. All breakout groups were asked the following general questions:

- Five types of IATA or their respective elements have been identified (see Annex 1 and 2) for which AOPs / MoAs could be used for regulatory purposes.
 - a. Do you agree that the AOP / MOA concept is a good basis for the development of these types of IATA or the respective IATA elements?
 - b. Are there elements for which they would be unsuitable?
 - c. Are there other types of IATA or their respective elements for which the AOP / MoA concept could be applied?

- It is recognized that knowledge of an AOP / MoA evolves and, as such, its development represents a continuum from less to more complete.
 - a. Is the categorisation into four stages of completeness/maturity of an AOP / MoA, as indicated in table 1 in annex 2, sufficient as an initial pragmatic approach for covering the continuum of the development of AOPs?
 - b. Where would you envisage the need for additional stages of development to meet potential application needs, especially when addressing more complex toxicological endpoints?
 - c. Where might there be alternative approaches that could be identified, please give examples?
- Each of the four stages of maturity of AOPs / MoAs are described.
 - a. Are these descriptions sufficiently adequate to characterise each of these stages to help an end-user to appropriately use and apply an AOP for developing IATAs, or their respective IATA elements, for their intended purposes?
 - b. Where and how might they be modified and improved to facilitate understanding of the different stages of maturity?
- In which ways do you think that the AOP stages shown in Figure 1 in annex 2 correlate with the IATA elements and regulatory applications? Can you give any potential examples/illustrations from your experience? Are there cases you can identify where and why the correlation may not be in accord?

Breakout Group Conclusions and Recommendations

23. The Workshop participants agreed on the following responses to the general questions outlined above.

24. AOPs have the potential to provide mechanistic support, credibility and transparency to the IATA and its elements. Any IATA element that is informed by an AOP and as such will likely provide greater confidence in any regulatory decision. The following IATA elements that could be informed by an AOP have been identified:

- Use of non-standard (non-animal) test methods;
- SAR/QSAR modelling;
- Chemical categorization and read-across,
- Test guideline development (both *in vivo* and *in vitro*)
- Integrated testing strategies.

25. The level of AOP development represents a continuum with increasing confidence in support for and degree of quantitation of key events (KEs) and key events relationship (KERs). Within this continuum different types of AOP can be distinguished: qualitative, semi-quantitative and quantitative. A further type may be desired reflecting AOPs that are not yet described fully enough to be classified into one of these 3 types (sometimes termed correlative or putative). Suggestions of potential regulatory purposes were made

to provide some perspective of the level of maturity an AOP might be desired depending on the application that it would be used for. It was concluded that while descriptors of the continuum and a distinction between different types of AOPs may be useful for developers (e.g. those completing the AOP wiki), this distinction may not be as helpful for users. The use of AOPs to inform an IATA and its regulatory use is context dependent (i.e. taking account of problem formulation, data availability etc.).

26. Characterization of confidence and degree of quantification in AOP elements (i.e the KEs and KERs) contributes to the required flexibility for their application in the development and use of IATAs. As such the different types of AOP identified are not necessarily related directly to a specific application in an IATA or type of regulatory use. The regulatory decision will also determine which type of IATA whether or not informed by AOP is most appropriate and the level of uncertainty that can be tolerated. In order to determine confidence in an AOP, the OECD AOP Handbook guidance on conducting a Weight of Evidence (WoE) evaluation on an AOP can be used (published in September 2014 on the AOP Wiki).

27. Likewise the degree to which an IATA needs to be populated by a full complement of methods addressing each of the key events in the corresponding AOP will be dependent on its envisaged application. For chemical categorization purposes, e.g. to facilitate read-across, it is conceivable that using approaches to characterise the MIE, might suffice whereas if a risk assessment decision is being made where uncertainty needs to be minimized as far as possible, generating information to address a number of other key events and their quantitative relationship with the adverse outcome as well as information on the expected exposure may be necessary. Thus flexibility is needed in the choice of the various information sources depending on the purpose of the IATA and the chemical under investigation. On the other hand there is also a need to provide regulators with some degree of consistency and understanding of the assumptions on which the IATA is based.

28. The four case studies presented illustrate potential applications of various AOP-informed IATAs to make or support various decisions. These range from priority setting and screening to (quantitative) hazard assessment. It is noted further that in some cases, including quantitative risk assessment, AOP independent elements (e.g. exposure, ADME) may be required. Depending on the regulatory context, AOPs can already be used in an IATA at early stages of development. In that respect, it was recommended that the selection of AOPs for review by the sub-bodies of the Joint Meeting (i.e. WNT and TFHA) should not be restricted to certain types of well-developed AOPs. Furthermore it was recommended that these sub-bodies consult the AOP Wiki as a basis to stimulate the dialogue between users and developers on which AOPs require further evaluation.

29. It is possible that an AOP may be suitable to inform IATAs for every type of regulatory application. It was recommended to develop case studies of AOP-informed IATAs having multiple purposes. Further case studies of AOPs informing IATAs for different decision contexts (e.g., priority setting, hazard identification, characterization, quantitative risk assessment) could be useful. It was recommended to engage the AOP development community and those contributing to IATA development and use in this activity. This includes cases documented through the IPCS programme engaging both and the research and risk assessment communities. AOPs could be used in informing IATAs in a comparable way as done in the OncoLogic cancer expert system which uses a mechanism-based approach to organize and integrate all available non-cancer short/medium-term predictive test/data of a chemical as a tool/basis for predicting the carcinogenic potential of that chemical. Qualitative AOPs have also been used to inform both cumulative assessments and testing strategies for pyrethroids and organophosphates.

Conclusions and recommendation derived from additional questions

30. The Workshop participants agreed on the following response to the additional questions outlined above.

31. **Additional question 1:** As different types of standard or non-standard test methods (e.g. high-throughput screening (HTS) assays and toxicogenomics) could be available to measure the outcome of a key event, how can we decide which method(s) will be most suitable for the development of different types of IATA or the respective IATA elements? What might be key considerations or criteria?

32. **Answer:** The participants agreed that issues for establishing “suitability”, especially mechanistic plausibility, for the “traditional” *in vitro* assays are the same as for HTS/-omics assays. Several guidance/best practices documents already exist, e.g. the OECD guidance for characterising non-guideline *in vitro* test methods (OECD, 2014¹) to facilitate their consideration in regulatory application, which articulate that the scientific validity of the assays needs to be characterised by considering issues such as chemical applicability domain, technical limitation of the test system (water solubility, metabolic competence), performance of the test method (sensitivity, specificity etc.), and their relationship to key events in the AOP.

33. If data from multiple assays are available, then all these data could be used, provided that conflicting data are properly addressed. If multiple assays provide the same type of information then their usefulness should be weighted with respect to their feasibility, efficacy and cost. Assays that provide different types of information may be more useful. The qualitative vs. quantitative outputs from the tests will dictate potential regulatory use i.e. prioritisation vs. safety assessment. It is recommended to develop guidance for assay selection.

34. **Additional question 2:** How can we establish confidence in key events that are based on different lines of evidence (e.g. *in vivo*, *in vitro*, HTS, toxicogenomics) in the process of the development of AOPs?

35. **Answer:** The modified Bradford Hill considerations as outlined in the AOP Wiki User Handbook (https://aopkb.org/common/AOP_Handbook.pdf) provide a framework to establish confidence for key events and key event relationships with different lines of evidence and as such can be used to assess the robustness and reliability of AOPs, e.g. the ER data from ToxCast/Tox21 expand the applicability domain and give greater confidence in the ER binding and gene induction KEs of the AOP (Cox et al., 2014²; Judson et al, 2013³).

36. **Additional question 3:** For which IATA applications does the AOP concept currently have the most merit? Potential promise in the near term (i.e., next 3-5 years)? Can you provide examples?

37. **Answer:** There is a need for greater understanding of how AOP-informed IATAs can be used for different regulatory decisions. More experience in applying these IATA should be acquired to increase common understanding of appropriate supporting information and associated confidence. It is envisaged

¹ Guidance document for describing non-guideline *in vitro* test methods. Series on Testing and Assessment No. 211; [ENV/JM/MONO\(2014\)35](#)

² Cox L.A., Popken D., Marty M. S., Rowlands J.C., Patlewicz G., Goyak K.O., Becker R.A. (2014). Developing scientific confidence in HTS-derived prediction models: Lessons learned from an endocrine case study. *Reg. Toxicol. Pharm.* Vol 69 (3), 443–450

³ Judson, R.J., Kavlock, R., Martin, M., Reif, D., Houck, K., Knudsen, T., Richard, A., Tice, R.R., Whelan, M., Xia, M., Huang, R., Austin, C., Daston, G., Hartung, T., Fowle III, J.R., Wooge, W., Tong, W. Dix, D. (2013) Perspectives on validation of high-throughput assays supporting 21st Century Toxicity Testing. *ALTEX* 30(1): 51-56

that some AOPs or sections of AOPs could be used for some regulatory applications. It was recommended to develop case studies in which an IATA is used for different decision contexts, focusing on AOPs where there has already been considerable progress e.g. skin sensitization, Thyroid, Aromatase Inhibition or Estrogen Receptor. Collective consideration of these case studies could serve as the basis for development of guidance on IATA. At present OECD guidance is in development towards a harmonised approach for the reporting of IATA by delivering a set of principles for describing and evaluating IATA to facilitate the consideration of IATA's assessments in regulatory decision-making and templates for reporting structured approaches to data integration and individual information sources used within IATA so that the same documentation format for describing and evaluating IATA and its elements. The suitability and workability of the templates proposed is evaluated by documenting a number of structured approaches for skin sensitisation hazard and potency prediction and by describing the individual information sources used within such approaches.

38. **Additional question 4:** What kind of quantitative/kinetic information would be needed to develop an IATA for quantitative hazard assessment and what are the guiding principles on what level of uncertainty can be tolerated for a specific decision context?

39. **Answer:** Establishing scientific confidence is critical for both the elements making up the IATA as well as the IATA as a whole. The level of uncertainty that can be accepted will be context dependent. Transparency on the extent of support and uncertainty and consistency of reporting will increase trust and acceptance of the outcome of these approaches. The AOP Wiki User Handbook is a useful guide to describe the confidence and uncertainty in individual elements in order to increase confidence/trust. Confidence will increase as we acquire experience in application. In that respect it is considered vital to develop further case studies to solicit input from both the scientific and regulatory community on those aspects that are important as a basis to illustrate how the scientific confidence can be built. Lessons can be learned from the existing examples analysed by the WHO MOA group.

40. **Additional question 5:** What are the guiding principles for consideration for the development of test guidelines for specific key events within AOPs to be used in specific regulatory context?

41. **Answer:** Regulators may prefer results on KE “near” the apical endpoint. This will however depend on the application context. For prioritization it might be more useful to develop test guidelines for “upstream” key events, whereas for risk assessment it might be more desired to develop test guidelines for assays that are related to “downstream” key events for which there is more confidence in the causal relationship with the adverse outcome. In that respect, if a KE is clearly linked to an endpoint, it matters less where the KE is located along the AOP.

42. Obviously, there should be a clear regulatory need to develop a test guideline. The way in which an AOP can inform the development of an IATA, might also determine for which KE it is most meaningful to develop a test guideline. In that respect it might not be necessary to develop a test guideline for every KE. The use of non-guideline test methods will necessarily increase and therefore it is very useful that the OECD is developing guidance for evaluating non-guideline methods. For the mutual acceptance of the results generated by an IATA, it might be necessary that these are constructed based on OECD Test Guidelines.

43. **Additional question 6:** How can AOPs for systemic toxicity effects be used to refine the existing *in vivo* test methods and how can the existing test methods, and the test results we have obtained, be used to inform how we build AOPs?

44. **Answer:** Test Guidelines, in general, are not designed to inform or be informed by AOPs. In that respect, they do not necessarily easily accommodate the generation of additional information that could

inform the development of AOPs. Adjustment of Test Guidelines to accommodate this may affect the statistical power of the assay. On the other hand *in vitro* screening data are useful to focus *in vivo* study development and modifications. Critical information needs could be identified in the AOP Wiki which could inform development. Obviously, for any adjustment to an existing Test Guideline there must be a benefit for the regulatory community as well as for the AOP developers. Via Public Crowd sourcing new AOPs and TG related events within these pathways could be identified. It might be useful to look at existing AOPs and compare those to current TGs and identify which TG would help further develop the AOPs. To further support this activity, research questions could be formulated by developers within the AOP wiki. The development and integration of toxicokinetic information in existing TGs (e.g. on repeated dose toxicity) is important and should be prioritised. The integration of parameters from Test Guidelines that are currently performed separately into one Test Guideline is needed to assist with the generation of more information with one *in vivo* test.

45. For complex toxicological effects, the effects are not related to a single AOP but a network of AOPs in which a multitude of KEs are involved. The framework needs be to more flexible to deal with this complexity. It was recommended to develop case studies to illustrate which key events are most appropriate to investigate (e.g. those that are rate limiting with respect to the adverse outcome).

ANNEX 1 WORKSHOP AGENDA

Day 1	
09h00	Opening and Welcome Representative of US Environmental Protection Agency will provide the welcome The meeting will be opened by the Chairs Terry Schultz (University of Kentucky) and Mark Cronin (Liverpool John Moores University). The meeting participants will briefly introduce themselves to the meeting (Tour de Table). The Chair will explain housekeeping items.
09h20	Presentation of the draft framework for developing and use of IATAs and Objectives of the Workshop The OECD Secretariat will explain why a framework for developing and use of IATAs is needed, what it should do and how AOPs / MoAs fulfil the objectives of the framework The Chair will describe the objectives of the Workshop The participants are invited to comment on the objectives of the workshop
10h15	General introduction on format of the workshop The OECD Secretariat will describe the format of the workshop
10h30	Coffee break
11h00	Presentation of the WHO/IPCS framework on mode of action/species concordance analysis (Bette Meek, University of Ottawa)
11h45	Overview of toxicokinetics and modelling approaches and relevance for using MoA/AOP (Jos Bessems, EU Joint Research Center)
12h30	Lunch
13h30	Presentation of case studies for different types of IATA and Q & A <ul style="list-style-type: none"> • The use of the AOP for AR / ER binding / thyroid effect to prioritise chemicals (Kevin Crofton, US EPA) • The use of the AOP for skin sensitisation to group chemicals (Frank Faulhammer, BASF) • Development of a testing strategy based on the AOP for Sustained AhR Activation leading to Rodent Liver Tumours (Grace Patlewicz, Dupont) • The use of the AOPs for Aromatase inhibition leading to reproductive dysfunction (in fish) to make a quantitative prediction for hazard characterisation (Dan Villeneuve, US EPA)
16h00	Break
16h20	Breakout sessions The questions for the breakout session are presented in the annex
18h00	Adjourn for the day
Day 2	
9h00	Chair's Summary of first day activity
9h30	Outcomes from the AOP workshop at Somma Lombardo, March 2014 (Ed Perkins, US Army Engineer Research & Development Center) Outcomes from the AOP workshop at Bethesda, September 2014 (Michelle Embry, Health & Environmental Sciences Institute)
10h00	Continuation of the Breakout sessions
12h30	Lunch
16h15	Rapporteurs' reports (60min; 20 minutes per group)
17h30	Adjourn for the day
Day 3	Plenary session
9h00	Conclusions and recommendations
13h00	Closure of the Workshop

**ANNEX 2 THE ADVERSE OUTCOME PATHWAY (AOP)/MODE OF ACTION (MOA)
CONCEPT: A FRAMEWORK FOR THE DEVELOPMENT AND USE OF INTEGRATED
APPROACHES TO TESTING AND ASSESSMENT (IATA)**

Introduction

Current regulatory toxicity testing and assessment approaches remain to a large extent chemical specific evaluations based on a checklist of typically *in vivo* tests conducted in accordance with standardised test guidelines or protocols. While this approach has evolved over the past half century, it is unlikely to efficiently meet future legislative mandates that will require increased numbers of chemical assessments to be undertaken without a concomitant increase in the use of animals. Significant advances in high throughput (HT) and high content (HC) methods offer new opportunities for gathering relevant information which quantify and characterise molecular and cellular responses to substances. For some endpoints, progress has been made in developing *in vitro* test methods; OECD Test Guidelines are available for skin / eye corrosion and irritation, genotoxicity and endocrine disruption. In recent years, these alternative test methods have influenced regulatory decisions especially when coupled with *in silico* approaches such as grouping of substances into chemical categories. Thus a shift from a scheme basing toxicity assessments (and other related chemical management decisions) largely on *in vivo* test results to one incorporating results from alternative approaches (e.g. *in silico*, *in chemico*, *in vitro* including HT/HC test methods) is already occurring.

At present, many non-animal approaches, irrespective of the particular methodology still suffer from a lack of clarity regarding the relationship (relevance and reliability) between the tested property and the apical toxicity endpoint being assessed and/or its adequacy for an intended purpose. This is perhaps one of the reasons why results from alternative approaches are not yet widely and consistently used for regulatory decision-making.

An objective and systematic framework that provides the biological anchor to help interpret the results from novel test and non-test approaches and facilitate their application in regulatory decision-making is needed. Such a framework should comprise guiding principles as well as technical guidance on how the methods and their outcomes can be interpreted to characterise (both qualitatively and quantitatively) the adverse effects for particular regulatory contexts.

What are integrated approaches to testing and assessment?

Integrated approaches to testing and assessment (IATA) are pragmatic, science-based approaches for chemical hazard characterization that rely on an integrated analysis of existing information coupled with the generation of new information using testing strategies. The vision is that the outcomes from appropriate combinations of *in silico*, *in chemico*, and *in vitro* approaches that target key events (KEs) along well defined toxicity pathways should ideally provide sufficient information for hazard and risk assessments with no or minimal *in vivo* testing. An IATA is envisioned as an iterative hypothesis generating and testing process that defines how to assess or test strategically based on regulatory needs (Meek et al., 2014).

An IATA logically starts with the formulation of plausible and testable hypotheses about the hazard profile of a substance or group of substances based on existing information and/or information derived from lower tier testing (such as *in chemico* and *in vitro* approaches). The hazard information together with the exposure information would then be used to determine what data gaps exist and what

testing if any would be most appropriate to undertake in order to elucidate the hazard profile of that substance for a given use context. Thus the extent to which testing approaches are needed therefore depend on the problem formulation which in turn is defined by the end purpose under consideration and the scientific confidence needed.

Understanding the likelihood of effects (i.e., initiation of a toxicity pathway) at lower levels of biological organisation (e.g., from structure-activity relationships (SAR) and *in vitro* models), can help inform whether more resource intensive testing (i.e., *in vivo*) is warranted. This then contributes to an increased efficiency in amount and type of hazard testing for substances undertaken. For some endpoints, some *in vivo* testing may be waived depending on the regulatory context for substances that show no potential to initiate the chain of events leading to an (adverse) outcome.

Thus, IATA is a means of organising and analysing all the available relevant data on a given substance or group of substances coupled with mechanistic, exposure, and dosimetry information where possible, to focus testing when needed and facilitate an assessment conclusion.

Its strength lies in the breadth of information that can be collated to understand the hazard and exposure profile of a substance that forms the foundation of the ultimate regulatory decision.

Framework for developing and using integrated approaches to testing and assessment

One of the reasons for the lack of uptake of some of the novel approaches in toxicity testing is the absence of a systematic framework to characterise their biological relevance in predicting an adverse effect. Therefore a more comprehensive understanding of pathways leading to toxicity, from molecular initiating events to adverse outcomes at the whole organism and/or population level is needed. The knowledge of an adverse outcome pathway (AOP, as defined by OECD, 2013) or mode of action (MoA, as defined by WHO, Meek et al., 2014) can provide the scientific, mechanistically based, framework for developing and using IATA.

The AOP and MoA concepts as defined by OECD and WHO/IPCS respectively are based on the principle that chemical interactions are at the molecular level (the so-called molecular initiating event or early key event) and not at the whole animal level. Both concepts describe the linkages between this chemical interaction with a biological system at the molecular level and the subsequent biological effects at the subcellular, cellular, tissue, organ, and whole animal and population levels. In fish for example, estrogen agonists bind to the estrogen receptor, measurable *in chemico*, which may set off a cascade of responses (depending on dose/time considerations) including the up-regulation of vitellogenin production in the liver (which can be measured *in vitro*), the conversion of testes to ova and the feminisation of males (which can be observed *in vivo*), leading to reproductive impairment in the individual animal and potentially, subsequently, a decrease in the population.

An AOP / MoA should be based on a single, defined ‘molecular initiating event’ (hereafter designated MIE) or if not possible, an ‘initial key event’ and linked to a stated *in vivo* hazard outcome (Figure 1). To establish an AOP / MoA, three blocks of information are used. The first block is the chemical-induced perturbations of biological systems at the molecular level (anchor 1). While a number of steps are required for an adverse outcome to be realised, the MIE is a prerequisite for all subsequent steps. It should be noted however that depending on the stage of development of the AOP/ MoA, the MIE (or other critical key events) may not always be defined, which inevitably may have implications for how the AOP / MoA can be applied. Indeed this is also true for other key events in the AOP. The last block is typically the *in vivo* adverse outcome (AO) of regulatory interest (anchor 2). These are often the reported endpoints from standard OECD Test Guidelines or may be observations in other toxicological or epidemiological investigations. Key events (KEs) which are essential intermediate steps along the pathway that represent pivotal events form the second block. These are usually at the different levels of biological organisation and that are relevant to the AO under consideration. To be a KE, the intermediate step must be

able to be evaluated experimentally. That is to say, the event must be able to be used in a hypothesis which can then be tested. There are no rules as to which types of data have to, or can be used to support a KE. However, such data should be reliable and relevant to the AO under consideration. Key events may of course be of relevance to other AOs as part of other AOPs / MoAs. There is no specification as to how many KEs have to be defined. The number of KEs clearly depends on where in the biological organisation the apical outcome is located (e.g., cell, organ or population level) and on the number of biological processes involved and the potential for interaction between these. The extent to which KEs need to be characterised experimentally will also depend on the end IATA application under consideration.

In order to use the AOP / MoA concept to inform an IATA it is also of importance to consider the toxicokinetics (i.e., absorption, distribution, metabolism and elimination) of a chemical to determine the likelihood that it and/or its ultimate toxicant(s) can reach the target organ(s) in the species of interest. The toxicokinetics determine the relevant structural moiety (i.e., parent compound and/or metabolite(s) and site of the molecular initiating event(s) of the toxic action). Examining the physicochemical properties and structural features, suggestive of labile structural moiety and potential activation / detoxification, as well considering available toxicokinetic data or generating such data with integrated approaches may be helpful in such an assessment.

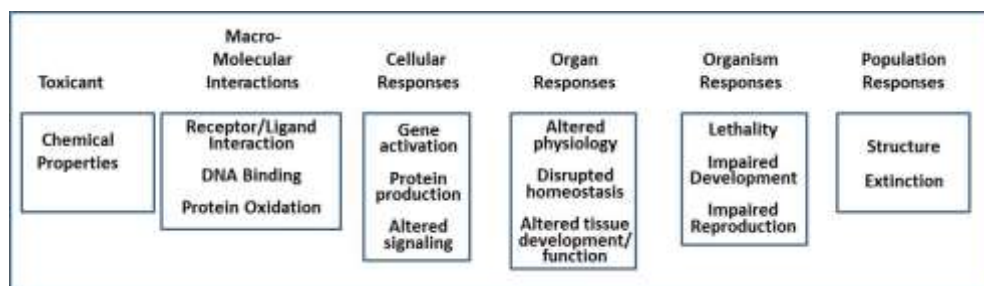


Figure 1. A schematic representation of the AOP illustrated with reference to a number of pathways.

Development of an IATA based on the AOP / MoA concept

An IATA can be constructed by using one or many methodological approaches ((Q)SAR, read-across, *in chemico*, *in vitro*, *ex vivo*, *in vivo*) or technologies (e.g. HTS). When underpinned by an AOP / MoA, these building blocks are related to the KEs they measure or compute, and the AOs they intend to predict. The following sections aim to illustrate how AOPs / MoAs may inform the development of these building blocks and how these are then used to develop IATA for different purposes e.g. priority setting, developing categories, testing strategies or hazard assessment. It is important to note that the IATA will depend on the level of completeness or maturity of their associated AOPs / MoAs upon which they are based.

Use of AOPs to develop (Q)SARs

As the MIE in each AOP/MoA involves a rather specific interaction of chemicals with biological systems, it may be used as the basis for generating structure–activity relationships (SARs). In turn, these SARs can be used for chemical grouping to facilitate associated read-across or testing strategies (OECD, 2012).

If *in vitro* assays have been developed for one or more KEs along the AOP and have been tested for a certain number of chemicals, then these results can be used to develop SARs, or when quantifiable, to develop QSARs. These could be used to enable the development of chemical categories containing chemicals that are as similar as possible in terms of presumed AOP / MoA, based on these KEs, in addition to the MIE. This has been actualised within the OECD Toolbox as part of the implementation of the AOP for skin sensitisation. Profilers exist that characterise SARs derived based on the experimental *in vivo* sensitisation test methods. Profilers that have specific numeric thresholds have also been developed based on substances that have been experimentally tested in assays characterising the different KEs. Thus a substance can trigger a structural alert and be categorised as moderately reactive based on a profiler derived on the basis of an assay measuring the MIE.

In case the sequence of KEs leading to a specific (adverse) effect is known at a sufficient level of detail, and the response–response relationships between the MIE, the KEs and the AO are well characterised by *in chemico*, *in vitro*, *ex vivo* and *in vivo* assays, the toxicity of many other chemicals acting through the same AOP / MoA may be practically determined by predicting the MIE or any of the KEs, as illustrated in Figure 2. As mentioned earlier, it will be important to factor the comparative kinetics and metabolism of the chemicals in question. Of course, some MIEs might not readily lend themselves to measurement *in vitro* assays and for some AOPs, the MIE might not be identifiable which would hinder the potential development of QSARs.

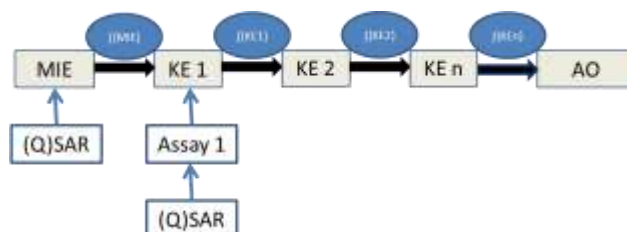


Fig 2. Use of an AOP / MoA to develop QSARs

Use of AOPs / MoAs to support grouping of chemicals into chemical categories for subsequent data gap filling for a specific endpoint using read-across or trend analysis

AOPs / MoAs can inform chemical grouping and subsequent data gap filling by read-across or trend analysis. Chemicals that are presumed to act by the same AOP / MoA on the basis of the MIE or other KEs can be grouped together, thereby improving the robustness of the data gap filling approach for the AO, compared to grouping chemicals solely based on their chemical structural similarity. AOPs / MoAs thus provide an opportunity to group chemicals based on their intrinsic chemical properties as well as their biological activity at different levels of biological organisation. Such categorisations of chemicals based on MIEs and/or subsequent KEs offer greater confidence that all chemicals in the category induce a specific AO.

Whilst a complete AOP / MoA from the molecular initiating event to the final adverse outcome is not considered critical for the purposes of grouping substances around a common MIE or KE, establishing the linkages between the MIE or KEs and the AO will be needed to justify the data gap filling (such as read-across) performed. Figure 3 illustrates how a category of chemicals presumed to trigger the same AOP / MoA can be used for a read-across.

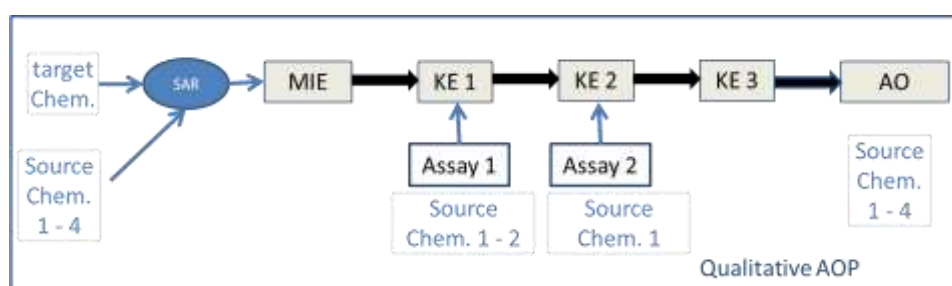


Figure 3. Use of the AOP concept to categorise chemicals for a specific endpoint

In the example outlined in Figure 3, it is assumed that the four source substances (1-4) exert an AO. The same information is lacking for a target substance which is structurally similar. An AOP / MoA has been developed where three KEs have been identified. In addition, a SAR has been developed that predicts the MIE (e.g protein binding). For two KEs, identified assays are also available. Based on the SAR, it can be shown that both the source and the target substance will trigger the MIE. Based on the commonality in the MIE, it can be hypothesised that the target substance will exert a similar AO. For two of the source substances, *in vitro* test results show that they elicit KE 1, while one of these two substances also triggers KE 2. Based on these observations, it is likely that all four source substance exert their effects through this common AOP / MoA. This suggests that the target substance will also follow the same pathway resulting in the same AO thereby strengthening the read-across between the source substances and the target substance. Depending on the potential use of this read-across prediction, the confidence could be strengthened by testing the target substance in assays which measure KE 1 or KE 2. Additional consideration of toxicokinetic aspects may permit a relative ranking of potency amongst the substances to be made as well as establishing KE temporal concordance.

Use of AOPs/MoAs to develop testing strategies

The AOP / MoA concept can be used to develop more efficient testing strategies for endpoints of interest by combining results from assays that evaluate specific KEs along a particular AOP / MoA. Depending on the regulatory context, for an endpoint for which no quantitative estimation is needed a qualitative understanding of the AOP / MoA might be sufficient. However the assays and their combinations or prediction models should be well characterised in terms of their performance characteristics and combined in a transparent manner so that conclusions can be independently verified.

Figure 4 outlines an example of how an AOP / MoA can be potentially used to inform a sequential testing strategy for the identification of a discriminant (positive/negative) endpoint. In this example the MIE and two KEs are well characterised and *in silico*, *in chemico* and *in vitro* approaches are available. In addition, the individual performance of the non-animal tests have been compared to a standard *in vivo* test. This is an illustrative example since not all testing strategies will necessarily include a QSAR component upfront.

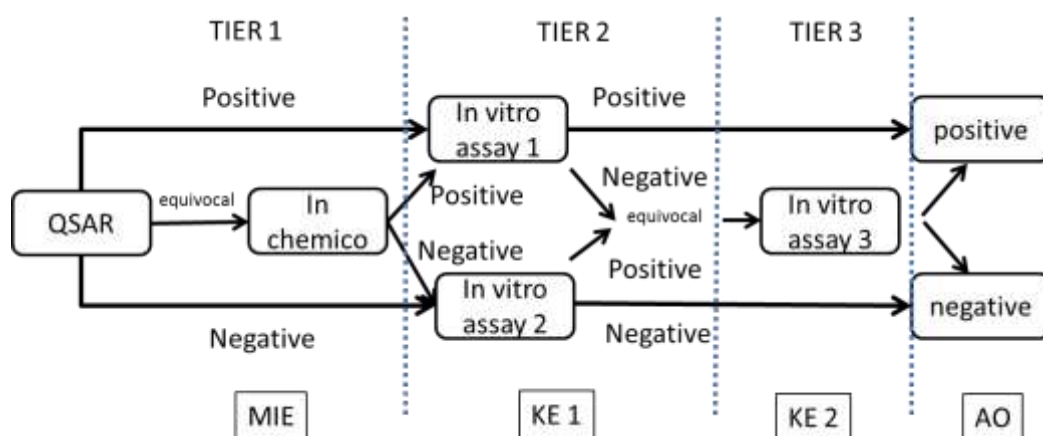


Fig 4. Use of an AOP/MoA in a testing strategy

The desire of future testing strategies will be to gather information from a combination of non-animal tests that address different KEs along the AOP / MoA in a tiered-approach. Information from each tier is used to decide what test systems will generate the most relevant information in the next tier and overall for the decision context in mind.

In Figure 4, the MIE is known and can be characterised using a QSAR approach. The prediction made determines what subsequent testing is warranted. A positive prediction from the QSAR (Tier 1) triggers testing (Tier 2) with an assay that addresses KE 1 and has high a positive predictivity (low false positives) whereas a negative prediction from the QSAR triggers testing (Tier 2) with an assay that addresses KE 1 and has high a negative predictivity (low false negatives). The final decision for the substances with a definitive positive or negative prediction in the Tier 1 analysis can be made in Tier 2 if the results in Tier 1 and 2 are concordant.

Substances for which the QSAR cannot generate an unambiguous prediction can be resolved in Tier 1 by testing in an assay that addresses the MIE. A positive or negative result from this assay determines which type of KE 1 assay should be used in Tier 2, namely one with a high positive or high negative prediction rate. Note the validity of the QSAR and its prediction are evaluated with reference to the QSAR Validation Principles.

Substances with conflicting results from Tier 1 and 2 are tested in Tier 3 by an assay addressing KE 2 and a weight-of-evidence approach is used to arrive at a final decision depending on the purpose and the stage of the associated AOP / MoA.

Use of AOPs / MoAs to help interpret results from non-standard test methods

Omic data (including toxicogenomics, transcriptomics, proteomics, and metabolomics) allow for more detailed insights into mechanisms of action, and can be applied to more efficiently survey the breadth of molecular/cellular effects elicited (*in vivo* or *in vitro*) by specific substances. Omic data could serve as either direct markers or indirect evidence hinting at a particular KE along an AOP / MoA leading to an adverse effect in the whole organism. Any omic dataset could potentially be associated with any KE, depending on the actual design of the experiment that was used to generate such data.

HTS data generated through *in chemico* methods, receptor binding or receptor transactivation assays, cellular reporter assay, may also serve to enhance identification of the chemical space associated with a particular KE. High throughput screening (HTS) approaches have the potential to provide data on large numbers of chemicals in a cost efficient manner (Judson et al., 2013). Scientific confidence in this assays needs to be established in terms of the analytical validation of the assays and their prediction models in the appropriate biological context, the latter being the associated AOP / MoA (Cox et al., 2014).

In a prioritisation approach aiming at screening thousands of chemicals, HTS could be well positioned to identify new/novel chemicals that would be expected to initiate specific molecular targets or perturbation of cellular response pathways within AOPs / MoAs. HTS or *in vitro* methods closely linked to a KE within a well characterised AOP / MoA would provide high confidence for high predictive capacity (i.e., with a low need for subsequent higher tier testing demands – regardless whether they would be used in prioritisation or for subsequent testing). However, there is also scope to use AOPs / MoAs that are not completely described and that provide limited confidence with respect to their relevance for adverse effects. It is noted however that such screening can only be used to identify substances with a likely (adverse) effect and generally not to identify substances with no effects because it cannot be excluded that other KEs than those known and described in the AOP / MoA are in fact also leading to the final AO. This might then motivate the development of detailed and predictive AOPs / MoAs as they would reduce subsequent higher tier testing. If the AOP / MoA consists of a clearly, quantitatively linked sequence of events (i.e., a chain of causative KEs), HTS assays only needs to target one of these events to be predictive.

Use of AOPs/MoAs to select methods for Test Guideline (method) development/refinement

By linking KEs in an AOP / MoA to *in vitro* test methods (or refined *in vivo* methods with integrated kinetic information), the relationship between the results of the methods to hazard endpoints can be established. In practice, it makes most sense to develop test methods for a KE, or a set of KEs, that are sufficient to infer that an AO will occur following chemical exposure. In principle, triggering all KEs along the AOP / MoA is necessary for triggering an AO, but none them individually are necessarily sufficient to infer an AO. Identifying KEs that are essential to induce the AO will allow those who develop alternative methods to direct resources to the development of testing methods targeted to these specific informative KEs. This will also decrease the overall number of assays required for hazard identification. By reference to a (semi)quantitative AOP, Figure 5 aims to illustrate how the most appropriate assays can be selected for test guideline development.

On the other hand, we may also see more rapid development of assays that are measuring the same KE, albeit in a different way or which are applicable for different chemical classes. The benefit of having more assays measuring the same KE within the AOP / MoA, will be to minimise the false positives and false negatives that are potentially generated by individual assays and to increase the overall weight of evidence. E.g. the DRPA (Gerberick et al., 2004; Gerberick et al., 2007) and the GSH assay (Schultz et al., 2005) both measure reactivity, the presumed MIE within the AOP for skin sensitisation. A (semi)quantitative AOP / MoA for which all KEs have been well characterised and for which response-to-response relationships between the KEs have been determined are most likely to help identify test guideline development/refinement needs.

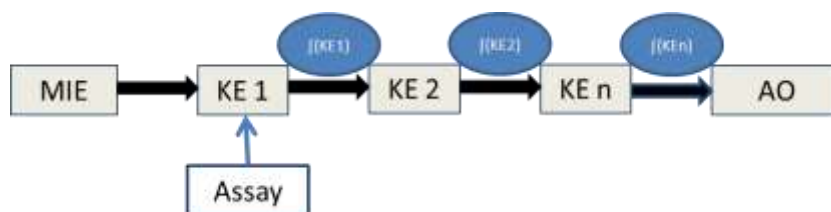


Fig 5. Illustrative example to show how a (semi)quantitative AOP can be used to help select the most suitable KE for which a Test Guideline could be developed or refined.

References

Cox, L. A., Douglas, D., Marty, S., Rowlands, J. C., Patlewicz, G., Goyak, K. O., Becker, R. A. (2014). Applying a Scientific Confidence Framework to a HTS-Derived Prediction Model for Endocrine Endpoints: Lessons Learned from a Case Study. *Reg. Toxicol. Pharmacol.* 69(3): 443–450.

Gerberick, G.F., Vassallo, J. D., Bailey, R. E., Chaney, J. G., Morrall, S. W., Lepoittevin, J.-P. (2004). Development of a Peptide Reactivity Assay for Screening Contact Allergens. *Toxicol. Sci.* 81: 332-343.

Gerberick, G. F., Vassallo, J. D., Foertsch, L. M., Price, B. B., Chaney, J. G., Lepoittevin, J. P. (2007). Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. *Toxicol. Sci.* 97: 417–427.

Judson, R.J., Kavlock, R., Martin, M., Reif, D., Houck, K., Knudsen, T., Richard, A., Tice, R.R., Whelan, M., Xia, M., Huang, R., Austin, C., Daston, G., Hartung, T., Fowle III, J.R., Wooge, W., Tong, W. Dix, D. (2013) Perspectives on validation of high-throughput assays supporting 21st Century Toxicity Testing. *ALTEX* 30(1): 51-56

Meek, M. E., Boobis, A., Cote, I., Dellarco, V., Fotakis, G., Munn, S., Seed, J. and Vickers, C. (2014). New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis. *J. Appl. Toxicol.* 34: 1–18. Doi: 10.1002/jat.2949

Organisation for Economic Co-operation and Development. (2013). Guidance document on Developing and Assessing Adverse Outcome Pathways Series on Testing and Assessment No. 184 ENV/JM/MONO(2013)6

Schultz, T. W., Yarbrough, J. W., Johnson, E. L. (2005). Structure-activity relationships for reactivity of carbonyl compounds with glutathione. *SAR and QSAR Environ. Res.* 16: 313-322.

ANNEX 3 GENERAL PRINCIPLES FOR USING IATA BASED ON THE AOP/MOA CONCEPT

Introduction

IATA is a means of organising and analysing all the available relevant data on a given substance or group of substances coupled with mechanistic, exposure, and dosimetry information where possible, to focus testing when needed and facilitate an assessment conclusion. An IATA may be utilized to address a wide variety of regulatory needs that range from simple hazard identification for priority setting to complex quantitative-based risk/safety assessments. Development of IATAs can be aided and informed by inclusion of Adverse Outcome Pathways (AOPs)/Mode of Action (MoAs). The objective of this workshop is to test the applicability of the AOP / MoA concept as a framework for developing and using IATAs and to refine the framework as far as possible. It will also define the degree of confidence in an AOP/MoA needed to inform an IATA in a specific regulatory context that can then be communicated throughout the decision making process.

The utility of AOPs/.MoAs for regulatory application is defined to a large extent by the confidence and precision with which they facilitate extrapolation of data measured at low levels of biological organisation (often in vitro) to predicted outcomes at higher levels of organisation and the specificity with which they can link biological effect measurements to their specific causes. The confidence in the AOP / MoA is based on the following considerations which will determine their applicability for a variety of regulatory purposes: (1) the extent of support for the biological plausibility of Key Event Relationships⁴ (KERs) and KE, and the apical outcome (AO); (2) the extent of support for the Essentiality of the Molecular Initiating Event (MIE) and KEs and (3) the extent of Empirical support for the KERs. The degree of support for each of these factors is determined based on consideration of the comparative extent of the evidence including identification of inconsistencies in the empirical evidence or significant knowledge gaps or uncertainties with regard to the essentiality of or relationship between the KEs. It is important to note that the KER descriptions and evaluation of the level of confidence in the relationship are designed to stand alone for a given pair of KEs without reference to or consideration of all the other KEs in the pathway, whereas the essentiality of upstream KEs is relevant to all downstream KEs in the AOP.

For some AOPs / MoAs, the relationship between specific KEs may be described quantitatively, while for others, the level of understanding might be such that only qualitative or semi-quantitative descriptions may be possible. AOPs / MoAs can be arbitrarily divided into the following four stages of maturity, which is described in more detailed in Table 1:

- **Correlative AOPs / MoAs** have only qualitative or limited quantitative understanding of one or two cause and effect linkages between KEs or a KE and the AO. These pathways are often based on a few stressors tested in a limited number of assays with a low level of confidence in the AOP.
- **Qualitative AOPs / MoAs** have qualitative understanding of critical components of the AOP / MoA. These pathways are often based on one or a few well-studied stressors where there is experimental evidence for the most critical KEs and the AO. The level of confidence in the AOP is moderate.
- **Semi-Quantitative AOPs / MoAs** have, in addition to qualitative understanding of the entire AOP / MoA, semi-quantitative understanding of some of the KEs. These pathways are based on multiple

⁴ A Key Event Relationship is a scientifically-based relationship that connects one key event to another, defines a directed relationship between the two (i.e., identifies one as upstream and the other as downstream), and facilitates inference or extrapolation of the state of the downstream key event from the known, measured, or predicted state of the upstream key event.

compounds and/or stressors evaluated at several KEs and the AO. The level of confidence in the AOP is moderate to high.

- **Quantitative AOPs / MoAs** have in addition to quantitative understanding of critical components of the AOP, empirical data across the spectrum of KEs and AO. These pathways are based on many compounds evaluated for all KEs and the AO so *in vitro* effects can be scaled to *in vivo* effects for risk assessment. The level of confidence in the AOP is high.

It should be noted that these proposed stages are arbitrary categorisations and it is well recognised that these form part of a continuum of the level of knowledge of AOPs / MoAs. The division into “stages” is proposed as a practical way forward for the development of AOPs and regulatory applications within the OECD work programme. The description of the level of confidence in the AOP is consistent with User Handbook that has been developed as a supplement to the Guidance Document for developing and assessing Adverse Outcome Pathways (AOPs) [ENV/JM/MONO(2013)6]. This handbook contains an updated template for AOP development and provides focused and practical instructions for both AOP developers and reviewers and is intended to assist in identifying, organising and evaluating critical information on KEs as well as linkages between KEs within the AOP (i.e., AOP development). The use of the proposed “stages” of AOP development should be in the context of problem formulation (what is the proposed regulatory application for the context where the AOP will be used).

A higher level of maturity is not automatically preferred, as the requirements for an AOP / MoA will depend on its regulatory application. The objective of development of AOPs / MoAs is to support inference or extrapolation from one KE to another, most notably from KE measurements that may be made efficiently and cost-effectively to adverse effects that are relevant to regulatory protection goals and decision-making. The overall weight of evidence (WoE) and level of scientific confidence underlying the inference and extrapolation in turn dictates the suitable applications of the AOP / MoA knowledge.

Table 1. Stage of development of an AOP

<p>Correlative AOP</p> <p>Has only <u>qualitative or limited quantitative understanding</u> of one or two cause and effect linkages between key events; <u>often based on a few stressors tested in a limited number of assays</u>. Information on early or late key events could be insufficient. Low level of confidence in the AOP:</p> <ul style="list-style-type: none"> • Limited support for the biological plausibility of KERs (Structural or functional relationship between KEs between them is not understood) • Limited support for the Essentiality of KE (No or contradictory experimental evidence of the essentiality of any of the KEs). Limited Empirical support for the KERs (Limited or no studies reporting dependent change in both events following exposure to a specific stressor (i.e., endpoints never measured in the same study or not at all) and/or significant inconsistencies in empirical support across taxa and species which don't align with expected pattern for hypothesized AOP) <p><u>Not able to determine the response-to-response relationships required to scale <i>in vitro</i> effect to <i>in vivo</i> outcome for relevant KEs.</u></p>
<p>Qualitative AOP</p> <p>Has <u>qualitative understanding</u> of the AOP with assessment of the experimental evidence and empirical data across the key events; <u>often based on one or a few well-studied stressors</u>. The critical (early and late) KEs are identified Moderate level of confidence in the AOP:</p> <ul style="list-style-type: none"> • Moderate support for the biological plausibility of some of the KERs (The KER is plausible based on analogy to accepted biological relationships but scientific understanding is not completely established.) • Moderate support for the Essentiality of KE (Indirect evidence that sufficient modification of an expected modulating factor attenuates or augments a KE (e.g., augmentation of

<p>proliferative response (Keup) leading to increase in Kedown or AO.)</p> <ul style="list-style-type: none"> • Limited to Moderate Empirical support for the KERs (Demonstrated dependent change in both events following exposure to a number of specific stressors but some evidence inconsistent with expected pattern which can be explained by factors such as experimental design, technical considerations, differences among laboratories, etc.) • Not able to establish quantitative understanding of any of the KEs. <p><u>Not able</u> to determine the response-to-response relationships required to scale <i>in vitro</i> effect to <i>in vivo</i> outcome for relevant KEs.</p>
<p>Semi-Quantitative AOP</p> <p>Has <u>semi-quantitative understanding</u> of the AOP- assessment of the experimental evidence and empirical data across the key events <u>based on multiple compounds and/or stressors studied at the KEs</u>. The critical (early and late) KEs are identified.</p> <p>Moderate to strong level of confidence in the AOP:</p> <ul style="list-style-type: none"> • Moderate to Strong support for the biological plausibility of some of the KERs (Good understanding of the KER based on previous documentation and broad acceptance; established mechanistic basis (e.g., mutation leading to tumours.) • Moderate to strong support for the Essentiality of KE (Direct evidence from specifically designed experimental studies illustrating essentiality for at least one of the important KEs.) • Moderate Empirical support for the KERs (Demonstrated dependent change in both events following exposure to a multiple number of specific stressors and some evidence inconsistent with expected pattern which can be explained by factors such as experimental design, technical considerations, differences among laboratories, etc.) • Establishes quantitative understanding of some of the KEs. <p><u>Not able</u> to determine the response-to-response relationships required to scale <i>in vitro</i> effect to <i>in vivo</i> outcome.</p>
<p>Quantitative AOP</p> <p>Has <u>quantitative understanding</u> of the AOP- assessment of the experimental evidence and empirical data across the KEs; <u>based on many compounds studied at the KEs</u>. The critical (early and late) KEs are identified.</p> <p>High level of confidence in the AOP:</p> <ul style="list-style-type: none"> • Strong support for the biological plausibility of some of the KERs (Understanding of the KER based on extensive previous documentation and broad acceptance; established mechanistic basis (e.g., mutation leading to tumours.) • Strong support for the Essentiality of KE (Direct evidence from specifically designed experimental studies illustrating essentiality for most of the important KEs [e.g., stop/reversibility studies, antagonism, knock out models, etc.) • Strong empirical support for the KERs (Multiple studies showing dependent change in both events following exposure to a wide range of specific stressors. Extensive evidence for temporal, dose-response and incidence concordance and no or few critical data gaps or conflicting data.) • Establishes quantitative understanding of the relationship between critical KEs. <p><u>Able</u> to determine the response-to-response relationships required to scale <i>in vitro</i> effect to <i>in vivo</i> outcome.</p>

Potential Regulatory application

Figure 1 illustrates at which stage of development an AOP / MoA can inform different types of IATA or their respective elements and which regulatory purposes these approaches can inform. Here the

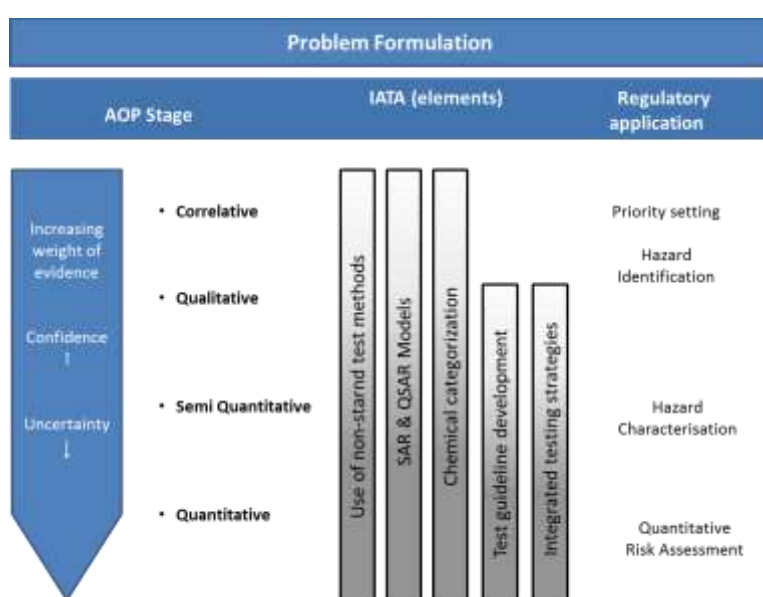
following elements are considered: (1) Interpretation of non-standard test result as part of a WoE assessment or for screening purposes, (2) Development of (Q)SAR models, (3) Chemical category formation for data gap filling, (4) Test guideline development/refinement, and (5) integrated testing strategies.

As indicated in the WHO/IPCS mode of action roadmap and revised Framework, areas of potential application of MOA / AOP analysis are interconnected with feedback loops that allow continuous refinement of fit for purpose risk assessment and testing strategies (Meek et al., 2014)

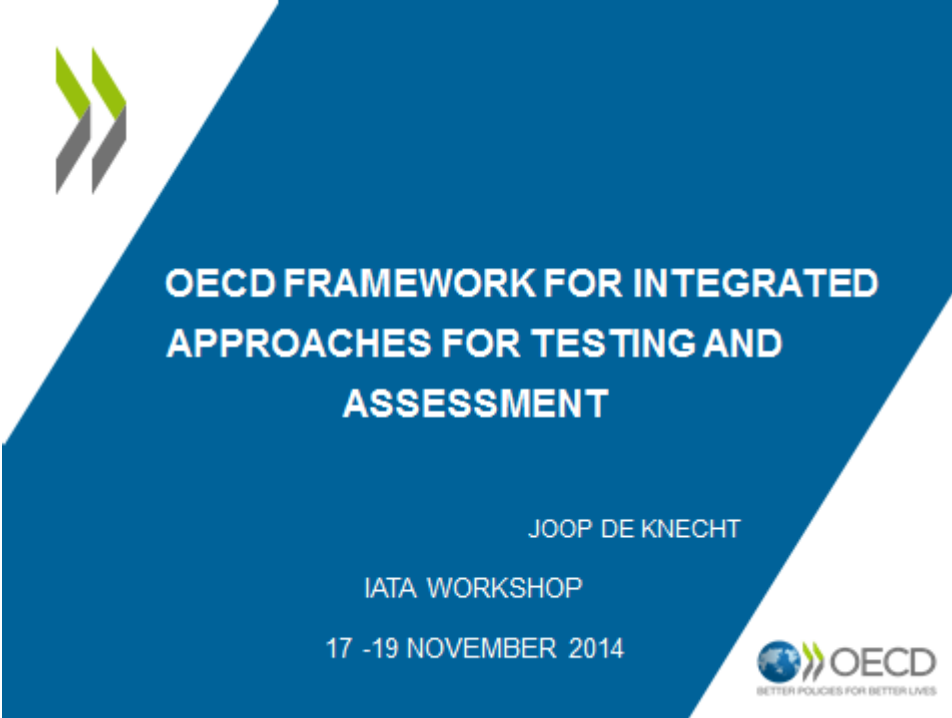

Similarly in Figure 1 below, depending on the problem formulation and the potential application of AOP / MoA within a specific regulatory context, updated/advanced AOPs should be reconsidered as more information is acquired (i.e., the basis for IATA).

For screening purposes, correlative AOPs with a minimum level of confidence in the relationship between specific KEs and the AO might be sufficient to interpret non-standard test results and to prioritise substances for further assessment or testing. At this stage of development, the MIE could also be used to develop *in silico* methods (e.g., structural alerts, SARs) to group chemicals into chemical categories, which in turn can be used to fill data gaps for hazard identification / classification and labelling. The reliability of these *in silico* methods will depend on the underlying data and the breadth of the applicability domain. Qualitative or (semi)quantitative AOPs for which the early key events (including the MIE) are identified can lead to the development of (Q)SARs that can be used for different regulatory applications. To inform the development of testing strategies, at least the early and late KEs within an AOP / MoA need to be identified/understood and a qualitative understanding of the AOP / MoA is required in order to provide a level of certainty for the KER and their relation to the AO. To use a result from a KE for estimating a quantitative effect level for the AO to be applied in a risk assessment, an understanding of the quantitative nature of the KERs and their relation to the AO is necessary. In the future when many more AOPs / MoAs are known, and sufficiently linked, it might be possible to develop a predictive toxicological system, taking into account quantitative relationships, where one AO vs. another can be predicted from KEs proximal to the MIE.

Fig 1. Relationship between different stages of AOP/MoA development and the informed types of IATA and its elements. Although this figure portrays a linear relation between the stage of development of an AOP and the regulatory application, it might be possible to use less developed AOPs in conjunction with other information in an IATA for more quantitative assessments (e.g., in grouping of chemicals).




ANNEX 4 PRESENTATION: OECD FRAMEWORK FOR INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT



**OECD FRAMEWORK FOR INTEGRATED
APPROACHES FOR TESTING AND
ASSESSMENT**

JOOP DE KNECHT
IATA WORKSHOP
17 -19 NOVEMBER 2014



**Integrated Approach to
Testing and Assessment (IATA)**

Are structured approaches that integrate and weigh different types of data for the purposes of performing hazard identification, hazard characterization and/or safety assessment of a chemical or group of chemicals





Integrated Approach to Testing and Assessment (IATA)

1. Formulation of plausible and testable hypotheses about the hazard profile based on existing information and/or information derived from lower tier testing
2. Evaluate data gaps
3. Identify non-testing and experimental approaches that would be most appropriate to undertake in order to elucidate the toxicological profile



Integrated Approach to Testing and Assessment (IATA)

The IATA may encompass

- category approaches
- testing strategies such as ITS and STS
- WoE considerations





Promoting the regulatory use of alternative test methods

- The OECD is actively working towards the development of methods to replace animal tests, as chemical management based on a battery of *in vivo* tests is unlikely to efficiently meet future legislative mandates
- A common approach to IATA ensure consistency in how information from alternative methods is used in regulatory decision-making process.
- Allow sharing of assessments between countries and avoid duplicative efforts

7



Need for scientific framework

Many non-animal approaches still suffer from a lack of clarity regarding the relationship between the tested property and the apical toxicity endpoint being assessed and/or its adequacy for an intended purpose.

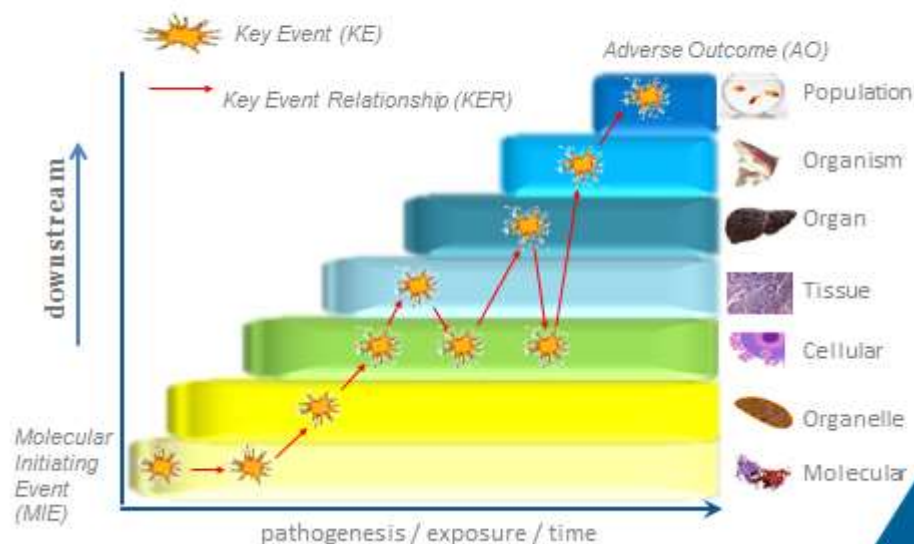


Need for scientific framework

- The **Adverse Outcome Pathway** is an **objective and systematic mechanistic based** framework that provides the **biological context** to facilitate the interpretation of results from alternative test and non-test approaches in predicting an adverse effect and facilitates their application in regulatory decision-making.



Anatomy of an AOP



Molecular initiating event (MIE)	A specialised type of key event that represents the initial point of chemical interaction on molecular level within the organism that results in a perturbation that starts the AOP.
Key event (KE)	A change in biological state that is both measurable and essential to the progression of a defined biological perturbation leading to a specific adverse outcome.
Key event relationship (KER)	A scientifically-based relationship that connects one key event to another, defines a directed relationship between the two, and facilitates inference or extrapolation of the state of the downstream key event from the known, measured, or predicted state of the upstream key event.
Adverse Outcome (AO)	A specialised type of key event that is generally accepted as being of regulatory significance on the basis of correspondence to an established protection goal or equivalence to an apical endpoint in an accepted regulatory guideline toxicity test.

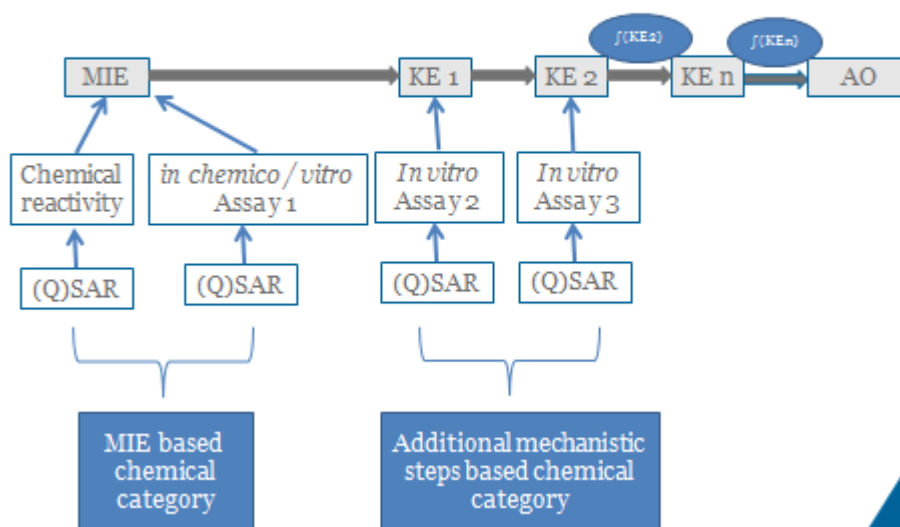


Mode of Action versus Adverse Outcome Pathway

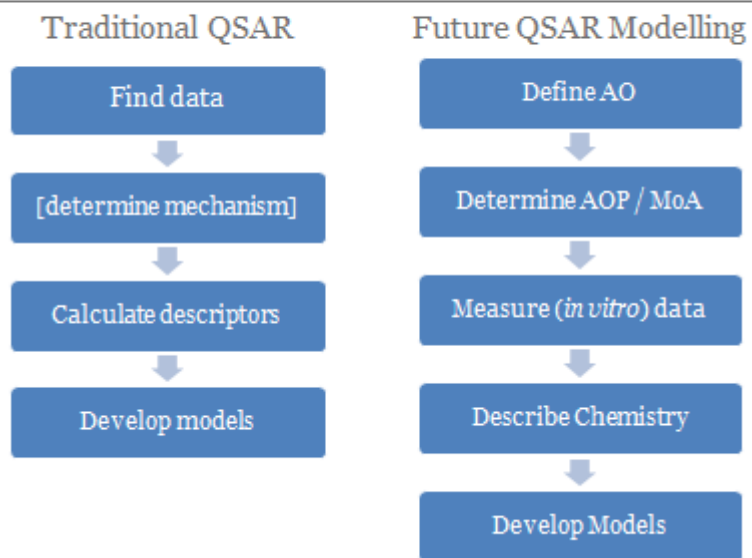
- The updated Mode of Action (MoA) framework (Meek et al, 2013) more explicitly describes the contribution of information at different levels of biological complexity
- Early key events as being more related to chemical characteristics and later key events as being less chemical-specific and more of a common expected consequence of earlier key events.
- In this framework, MoA and AOP are conceptually similar



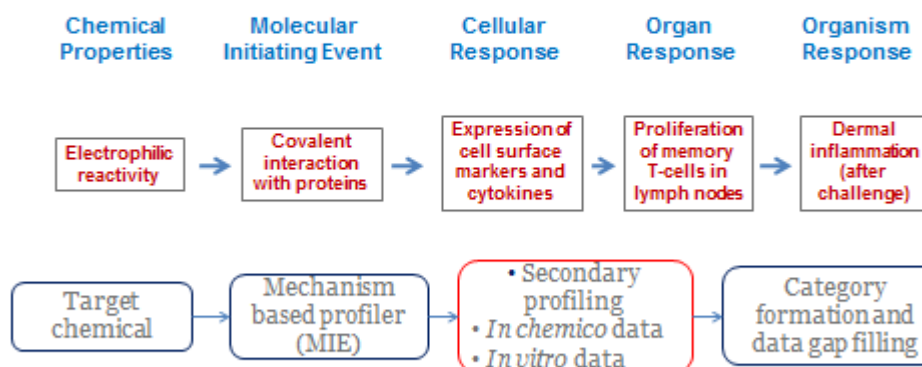
Use of an AOP to develop QSARs



QSAR development



AOP: An example for skin sensitisation

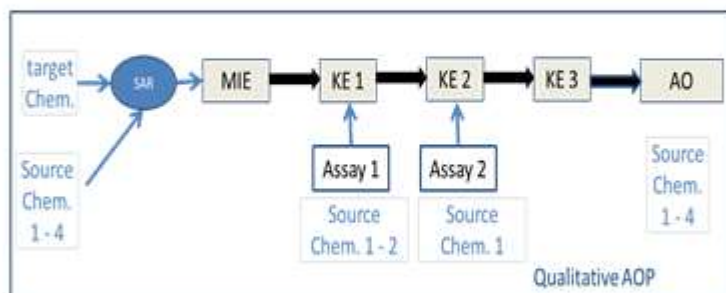


Use of the AOP concept to categorise chemicals for a specific endpoint

- AOPs / MoAs provide an opportunity to group chemicals based on their intrinsic chemical properties as well as their biological activity at different levels of biological organisation.
- Such categorisations of chemicals based on MIEs and/or subsequent KEs improve the robustness of the data gap filling approach and offer greater confidence that all chemicals in the category induce a specific AO.



Use of the AOP concept to categorise chemicals for a specific endpoint

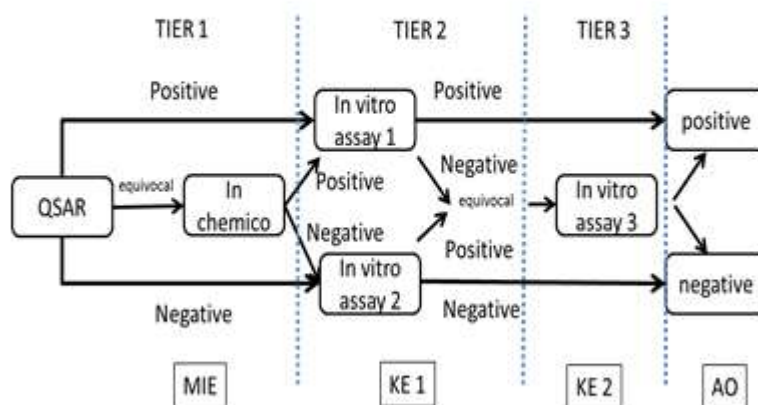


Use of an AOP/MoA in a testing strategy

- The AOP / MoA concept can be used to develop more efficient testing strategies for endpoints of interest by combining results from assays that evaluate specific KEs along a particular AOP / MoA.
- Depending on the regulatory context, for an endpoint for which no quantitative estimation is needed a qualitative understanding of the AOP / MoA might be sufficient.
- The assays should be well characterised in terms of their performance characteristics and combined in a transparent manner so that conclusions can be independently verified.



Use of an AOP/MoA in a testing strategy



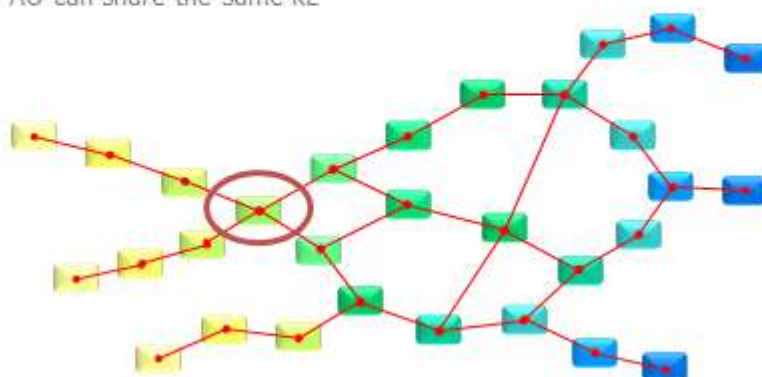
Use of AOPs/MoAs to select methods for Test Guideline development

- Identifying KE(s) that are essential to induce the AO will allow those who develop alternative methods to direct resources to the development of testing methods targeted to these specific informative KEs.
- The benefit of having more assays measuring the same KE within the AOP / MoA, will be to minimise the false positives and false negatives that are potentially generated by individual assays and to increase the overall weight-of-evidence
- Assays that linked to KE that appear to be present in more AOPs are especially good candidates to be standardised in the form of an OECD TG.



AOP Network

- AOPs are elements of an AOP causality network
- Multiple AO can be caused by a single MIE
- Multiple AO can share the same KE



Use of AOPs / MoAs to help interpret results from non-standard test methods

- Omic data allow for more detailed insights into mechanisms of action
- Omic data could serve as either direct markers or indirect evidence hinting at a particular KE along an AOP / MoA
- HTS assays which can be linked to KEs have the potential to provide data on large numbers of chemicals in a cost efficient manner



General principles for using IATA based on the AOP/MoA concept

- The utility of AOPs/MoAs for regulatory application is defined to a large extent by the confidence and precision with which they facilitate extrapolation of data measured at low levels of biological organisation (often in vitro) to predicted outcomes
- A high level of maturity is not automatically needed, as the requirements for an AOP / MoA will depend on its regulatory application.



SEMINAL ASPECTS OF A KEY EVENT

- Wherever possible one KE at each major level of biological organisation.
- Selection of a limited number of KEs which are measurable and for which evidence supports plausibility and relevance in a regulatory context.
- Preference to KEs that could be measured in a relatively routine manner over those that would require highly specialised expertise, equipment, or supplies to measure.
- Taxonomic applicability/Species Concordance



Weight of Evidence for the KER

Biological Plausibility

Mechanistic (i.e., structural or functional) relationships between the KEs consistent with established biological knowledge

Empirical support for the KER

- Evidence showing that stressors that perturb KE-upstream also perturb KE-downstream
- Evidence showing the temporal concordance of the KEs
- Evidence of dose response and/or response-response relationships (later KEs) and dose-dependent- and time-dependent transitions from KE-upstream to KE-downstream

ASSESSMENT OF THE AOP

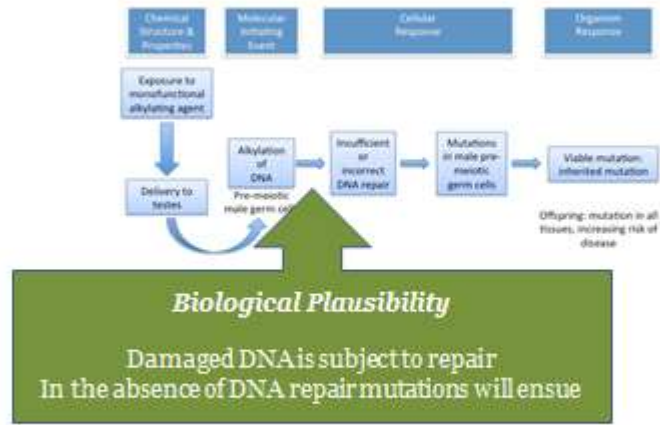
Domain of Applicability of the AOP

applicability in terms of sex, life-stage, taxa, and other aspects of biological context

Relative Level of Confidence in the AOP

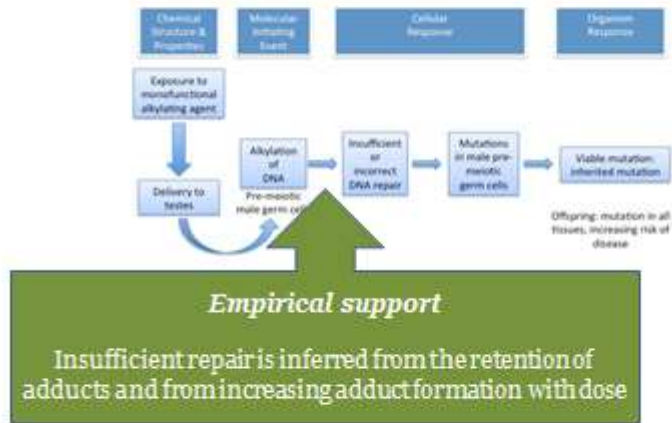
- Level of Confidence in the KERs
(Biological plausibility of each of the KERs)
- Essentiality of the KEs
Are downstream KEs (incl. AO) prevented if an upstream KE is blocked?
- Empirical Support for each of the KERs
- Degree of Quantitative Understanding

Alkylation of DNA in male pre-meiotic germ cells leading to heritable mutations



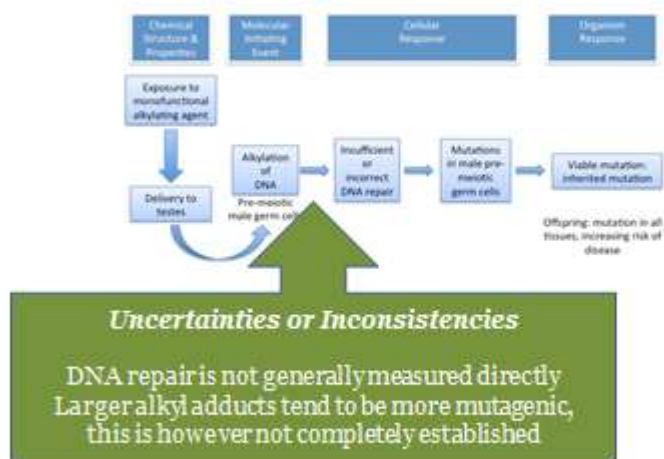
Yauket et al.

Alkylation of DNA in male pre-meiotic germ cells leading to heritable mutations



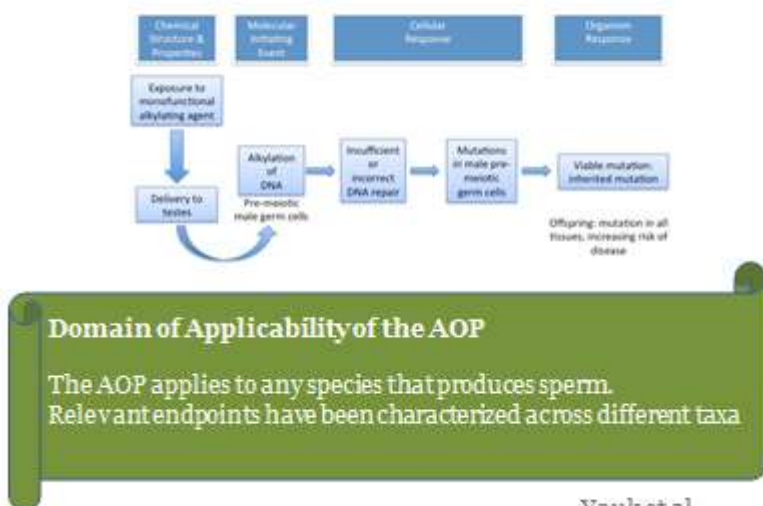
Yauket et al.

Alkylation of DNA in male pre-meiotic germ cells leading to heritable mutations



Yauk et al.

Alkylation of DNA in male pre-meiotic germ cells leading to heritable mutations



Yauk et al.

Alkylation of DNA in male pre-meiotic germ cells leading to heritable mutations

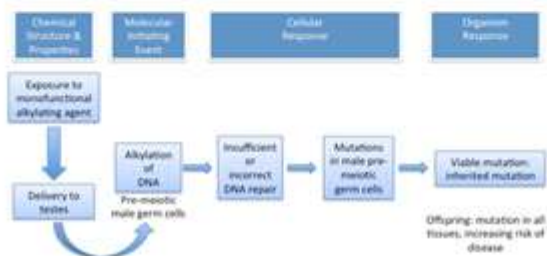


Essentiality of the KEs

Each of the key events is essential.
 Enhanced DNA repair will reduce mutation frequencies
 Correct repair of the alkylated DNA (a block of KE1) will not lead to mutation

Yauk et al.

Alkylation of DNA in male pre-meiotic germ cells leading to heritable mutations



Empirical Support for each of the KERs

Although the support for the direct linkages between KEs is considered moderate, the indirect associations between the MIE and KE2 and the AO are all strong

Yauk et al.

Alkylation of DNA in male pre-meiotic germ cells leading to heritable mutations



Degree of Quantitative Understanding

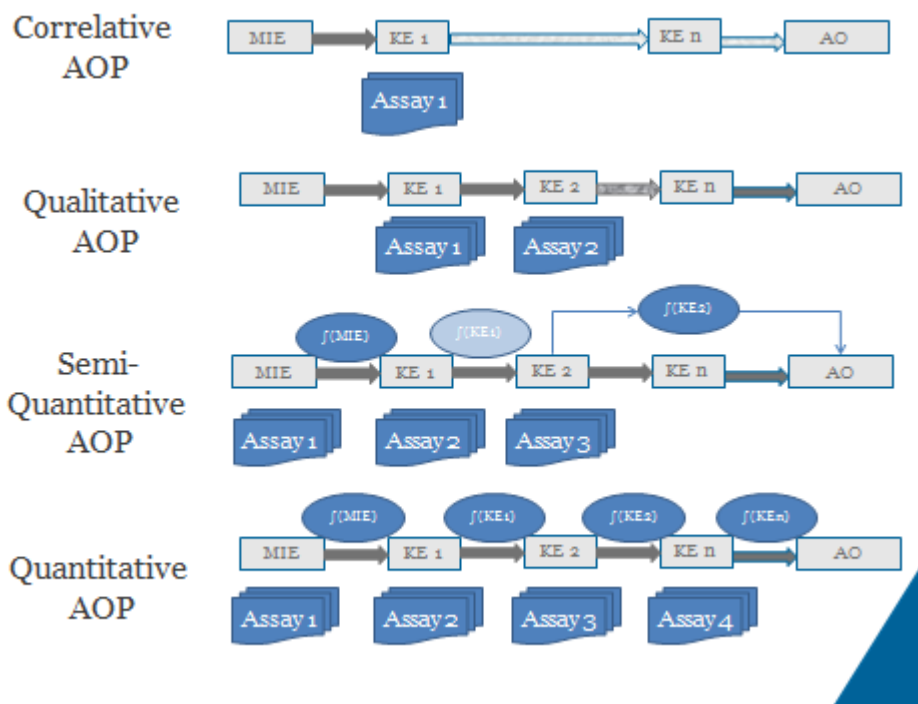
- Alkylation must exceed a threshold before alkylated DNA is retained in cells, and mutations subsequently begin to occur.
- However, the precise quantitative relationship has not been modelled.

Yauk et al.

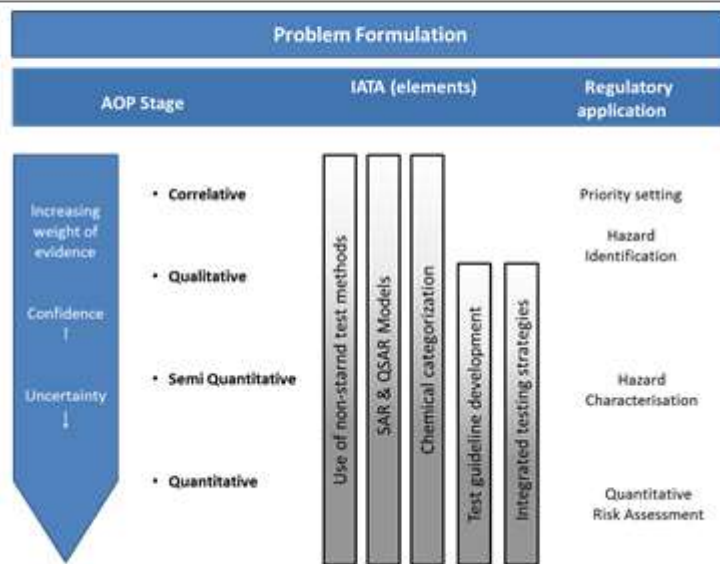


STAGES OF AN AOP

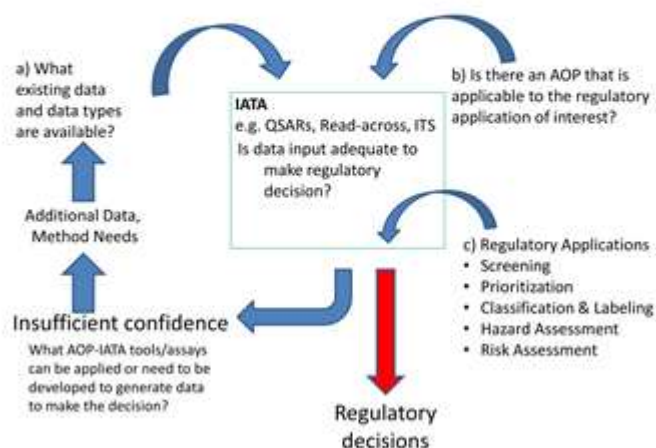
- Correlative AOP
- Qualitative AOP
- Semi-Quantitative AOP
- Quantitative AOP



Relationship between different stages of AOP/MoA development and the informed types of IATA and its elements.



Applying AOPs to support IATA

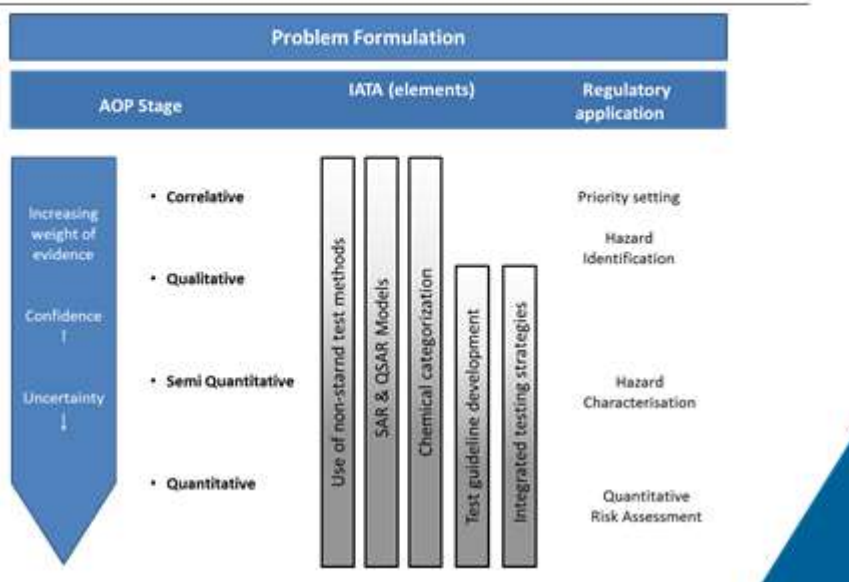


The objective of the workshop

- To investigate the applicability of the AOP / MoA concept as a framework for developing and using IATAs and to refine the framework as far as possible.
- To define the degree of confidence in an AOP / MoA needed to inform an IATA in a specific regulatory context that can then be communicated throughout the decision making process.



Relationship between different stages of AOP/MoA development and the informed types of IATA and its elements.



Discussion in Breakout groups

Participants will be asked to consider a set of questions designed to

- Build consensus that the AOP / MoA concept is a good basis for developing and using IATAs,
- Determine for which type of IATA or their respective elements, can be informed by an AOP / MoA,
- To discuss and refine the proposed definitions and characteristics of the different stages of development of AOPs / MoAs
- To outline/propose which stages are most suitable for what types of IATA or their respective elements and their regulatory purposes.



Breakout sessions

All three breakout groups will be asked four sets of general questions to further discuss and define the overall framework

Two additional questions per breakout group will be discussed the following themes:

- Confidence in the KE and KER of the AOP
- Current status and work needed to broaden the scope of the AOP / MoA application
- Identifying test methods and data gaps for further development of AOPs / MoAs



Follow-up

Recommendations and conclusions of the workshop will be used

- to finalise the framework.
- to develop further cases studies by the TFHA and provide guidance on how IATAs based on the framework can be used for regulatory purposes
- to identify new alternative test methods by the WNT that are candidates to become OECD Test Guidelines or to refine current test guidelines



ANNEX 5 PRESENTATION: THE WHO FRAMEWORK ON MODE OF ACTION/SPECIES CONCORDANCE ANALYSIS IMPLICATIONS FOR AOPS/IATA

The WHO Framework on Mode of Action/Species Concordance Analysis Implications for AOPs/IATA

uOttawa
L'Université canadienne
Canada's university

Workshop on a Framework for the Development and Use of Integrated Approaches to Testing and Assessment
November 17th - 19th, 2014, Arlington

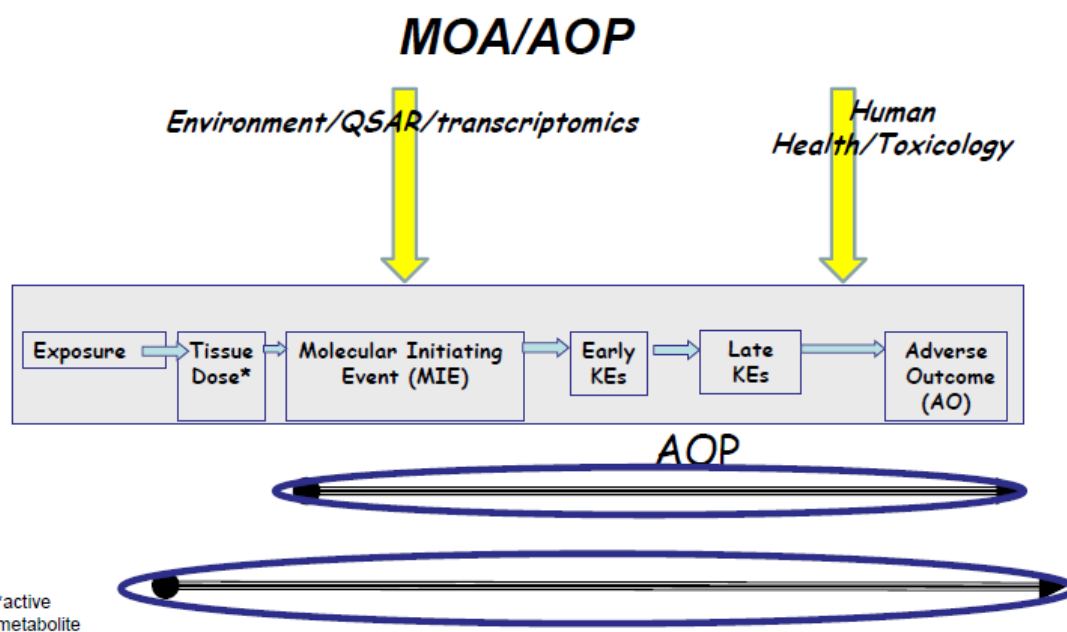
*Presented by: M.E. (Bette) Meek
bmeek@uottawa.ca*

Université d'Ottawa | University of Ottawa

uOttawa.ca

Outline

- Update on MOA Analysis
 - Problem Formulation/Roadmap
 - Weight of Evidence Analysis
 - Examples
- Implications for AOP Development/Analysis and IATA
 - Considering Confidence
 - Supplement to the AOP Guidance/wiki
 - Illustrative output to facilitate application
- Conclusions/Recommendations



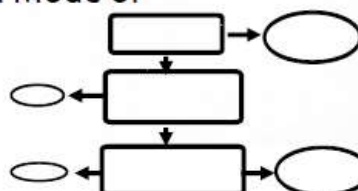
Mode of Action Analysis

Conceptually, Adverse Outcome Pathways (AOPs) and MOA are identical. Divide a path between exposure and effect into key events, but MOA has been considered on a chemical specific basis.

3

Mode of Action/Species Concordance Analysis

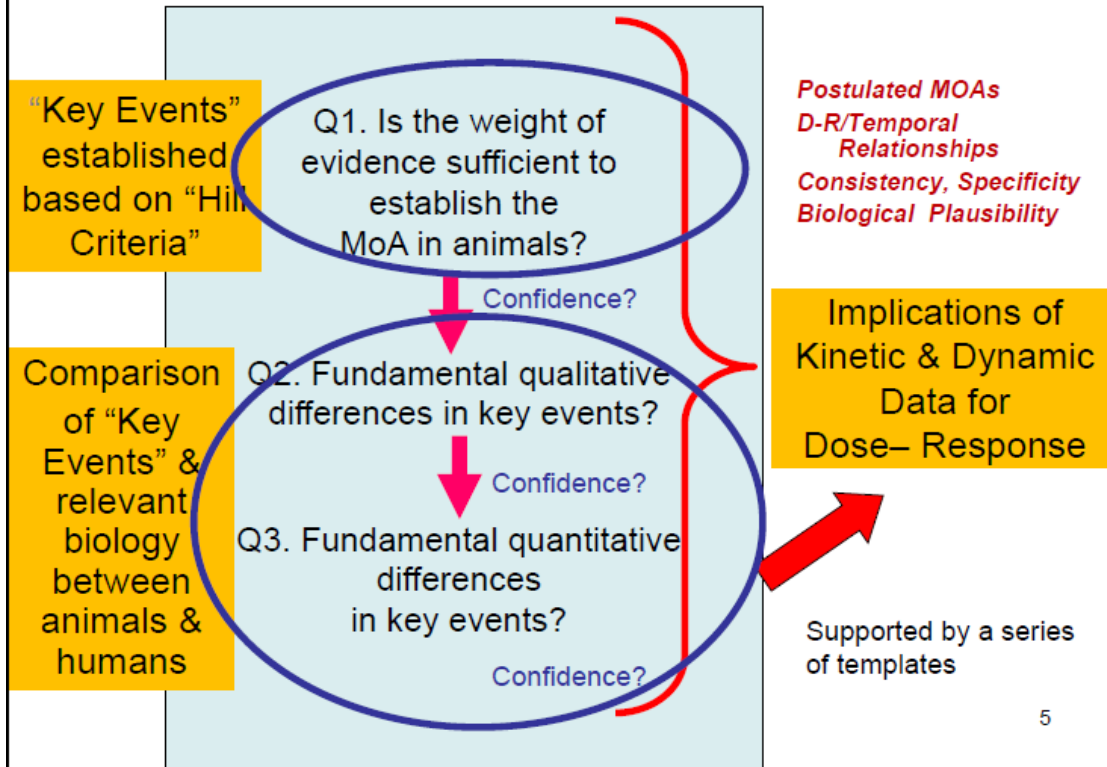
- World Health Organization (WHO)/International Programme on Chemical Safety (IPCS) Framework on Mode of Action/Human Relevance (MOA/HR)



- Derived from early US EPA/ILSI work
- since 1999, 100s of experts internationally involved in its development
- widely incorporated in program guidance internationally (US EPA, EFSA, EU TGD, JMPR, OECD)/adopted in risk assessments, training
- Recent update that extends and builds on international regulatory experience (Meek et al., 2014a)

4

WHO MOA/HR (WOE) Analysis Framework

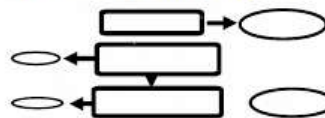


Focus on MOA/HR Analysis

Increasing predictive capacity and utility of risk assessment

Objectives

- Drawing maximally and early on mechanistic information
- Transparency
- Doing the right research/testing
 - Bridging regulatory/research



Issues:

- Perceived as a “labour intensive” add on
- Hypothesized MOAs often not well defined
 - Regulatory/research input
- Inconsistent use/interpretation of weight of evidence considerations
 - *Need for simplicity for broad applicability, including evolving technology*

New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis†

M. E. Meek^a, A. Boobis^b, I. Cote^c, V. Dellarco^d, G. Fotakis^e, S. Munn^f, J. Seed^g and C. Vickers^{h*}

ABSTRACT: The World Health Organization/International Programme on Chemical Safety mode of action/human relevance framework has been updated to reflect the experience acquired in its application and extend its utility to emerging areas in toxicity testing and non-testing methods. The underlying principles have not changed, but the framework's scope has been extended to enable integration of information at different levels of biological organization and reflect evolving experience in a much broader range of potential applications. Mode of action/species concordance analysis can also inform hypothesis-based data generation and research priorities in support of risk assessment. The modified framework is incorporated within a roadmap, with feedback loops encouraging continuous refinement of fit-for-purpose testing strategies and risk assessment. Important to this construct is consideration of dose-response relationships and species concordance analysis in weight of evidence. The modified Bradford Hill considerations have been updated and additionally articulated to reflect increasing experience in application for cases where the toxicological outcome of chemical exposure is known. The modified framework can be used as originally intended, where the toxicological effects of chemical exposure are known, or in hypothesizing effects resulting from chemical exposure, using information on putative key events in established modes of action from appropriate *in vitro* or *in silico* systems and other lines of evidence. This modified mode of action framework and accompanying roadmap and case examples are expected to contribute to improving transparency in explicitly addressing weight of evidence considerations in mode of action/species concordance analysis based on both conventional data sources and evolving methods. Copyright © 2013 John Wiley & Sons, Ltd. The World Health Organization retains copyright and all other rights in the manuscript of this article as submitted for publication.

Keywords: key events; mode of action; adverse outcome pathway; human relevance framework; modified Bradford Hill considerations; weight of evidence approach; species concordance analysis; cellular response; dose response; molecular target

Introduction

The mode of action/human relevance framework was developed in initiatives of the International Programme on Chemical Safety (IPCS) of the World Health Organization (WHO) (Boobis et al., 2006, 2008; Sarrich-Kullin et al., 2001) and the International Life Sciences Institute Risk Sciences Institute (ILSI-RSI) (Meek et al., 2002; Seed et al., 2005). It derives from earlier work on mode of action in animals by the US Environmental Protection Agency (US EPA, 1996, 2005) and has involved large numbers of scientists internationally.

Further development of the mode of action/human relevance framework is described in the publications mentioned above and summarized more recently in Meek and Klasing (2010). The framework has been illustrated by an increasing number of case studies (more than 30 currently) demonstrating the value of mode of action in evaluating human relevance and its stage susceptibility and guiding dose-response assessment. Documented examples are presented in Table 1. The contribution of the framework has been recognized by the Society of Toxicology and the framework has been adopted by several international and national organizations and agencies to increase transparency in the assessment of weight of evidence and identification of critical data needs (Meek, 2008, 2009; Meek et al., 2009).

The framework continues to evolve as experience identifies in its application to consider systematically the weight of evidence

*Correspondence to: C. Vickers, International Programme on Chemical Safety, World Health Organization, 20 Avenue Appia, CH-1211 Geneva 27, Switzerland. Email: vickers@ipcs.who.int

†This publication includes the collective views of an international group of experts and does not necessarily represent the decisions or the stated policy of the World Health Organization or its affiliated organizations.

^aChemical Risk Assessment, McCaig Centre for Sustainable Health Risk Assessment, 1 Stuart Street, Ottawa, Canada, K1N 6N7

^bDepartment of Medicine, Imperial College London, Hammersmith Campus, Duane Road, London, W12 0AZ, UK

^cValencia Center for Environmental Assessment, US Environmental Protection Agency, 1209 Pennsylvania Avenue NW, Washington DC, 20460 USA

^dPriority Office of Toxicologic Programs, US Environmental Protection Agency, 1209 Pennsylvania Avenue, NW, Washington DC, 20460 USA

^eEuropean Chemical Agency, Assessment, DE-81259 Heilbronn, Heilbronn

^fScience Technology Unit, Institute for Health and Consumer Protection, Joint Research Centre, Commission European, Via C. Cornei, 209A, I-20133, Milan, Italy

^gOffice of Pollution Prevention and Control, US Environmental Protection Agency, 100 Pennsylvania Avenue, NW, Washington DC, 20460 USA

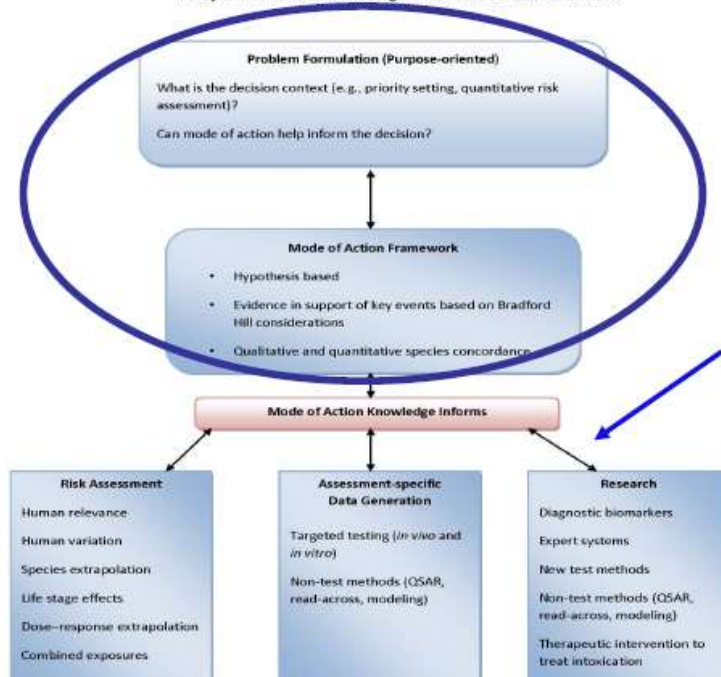
^hInternational Programme on Chemical Safety, World Health Organization, 20 Avenue Appia, CH-1211, Geneva 27, Switzerland

Objectives

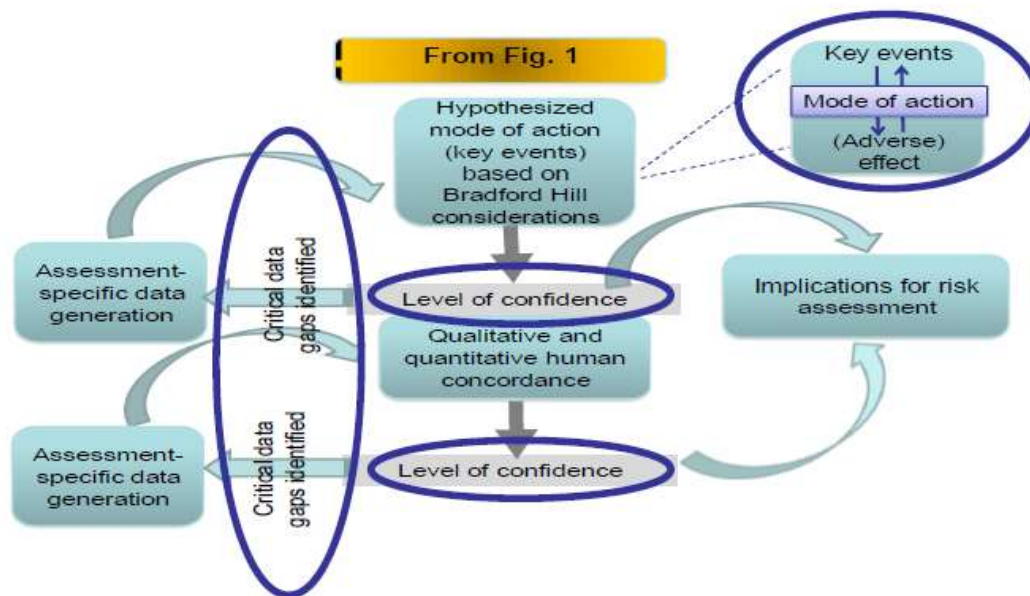
Meek et al. *Applied Toxicol.* 34: 1-18 (2014 a).

- Clarify terminology (**MOA conceptually = AOP**)
- To tailor analysis to issue at hand
 - Problem formulation
- To extend utility to new areas in toxicity and non-toxicity testing, providing practical examples
- Need for simplicity for broad applicability, including evolving technology**
- Simplifying /“codifying” experience in WOE analysis and quantitation
 - **modified Bradford Hill considerations for WOE for MOA wiki**
 - **Defining questions**
- incorporating dose-response analysis (quantitation)

Mode of Action Roadmap
Utility of Mode of Action Knowledge in Human Health Risk Assessment



Modified MOA Framework



(c) World Health Organization 2013

9

Contents of the WHO Update

Case Examples illustrating application of MOA analysis in:

- 1: Lack of human concordance
- 2: Contribution of well-designed genomic studies to species concordance and dose-response analysis
- 3: The evaluation of epidemiological data
- 4: Development of more efficient testing strategies
- 5: Prioritizing substances for further testing (including genomic data)
- 6: Creation of chemical categories (including genomic data)
- 7: Identifying critical data gaps and testing strategies in read-across

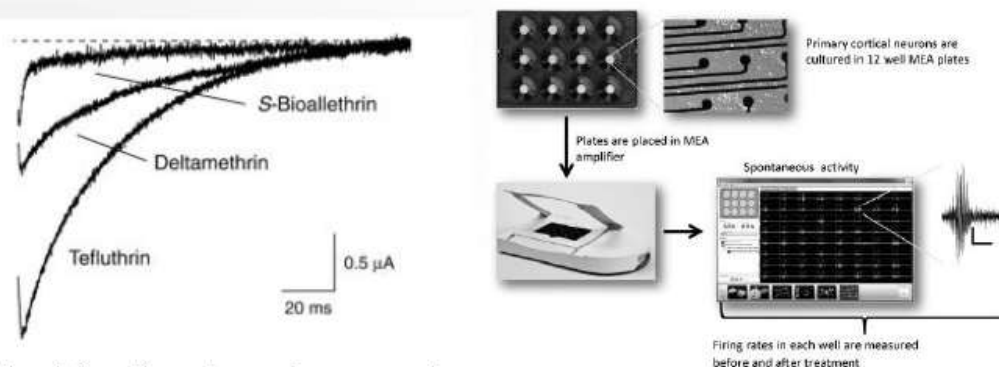
Case example 6: MOA in grouping and potency estimates for categories/combined exposures

Anchoring the results of *in vitro* approaches to relevant outcomes based on existing knowledge and concepts:

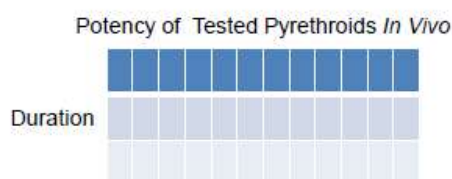
- Class of pesticides, same well established mode of action and insecticidal effects
 - reversible neurotoxicity through interaction with neuronal sodium channels
- Members of the class expected to share key events
- Consider grouping and rank for potency for untested compounds in suitable *in vitro* system for this key event
- Consider toxicokinetic aspects
- Choose reference point from amongst those class members tested in *in vivo* assays

11

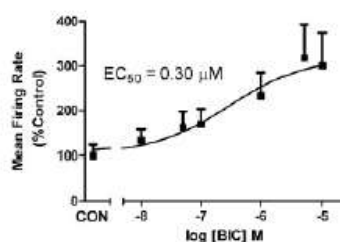
Evaluating key event *in vitro*



Rat Na_v 1.6 sodium channel expressed in oocytes



McConnell et al, 2012



Mode of action human relevance (species concordance) framework: Evolution of the Bradford Hill considerations and comparative analysis of weight of evidence

M. E. (Bette) Meek*, Christine M. Palermo, Ammie N. Bachman, Colin M. North and R. Jeffrey Lewis

ABSTRACT: The mode of action human relevance (MOAHR) framework increases transparency in systematically considering data on MOA for end (adverse) effects and their relevance to humans. This framework continues to evolve as experience increases in its application. Though the Bradford Hill framework is not designed to address the question of “how much information is enough” to support a hypothesized MOA in animals or its relevance to humans, its organizing construct has potential value in considering relative weight of evidence (WOE) among different cases and hypothesized MOAs. This content is explored based on MOA analysis in published assessments to illustrate the relative robustness of supporting data and their implications for dose-response analysis and involved comparisons for chemical assessment on trichloroethylene, and carbon tetrachloride with several hypothesized MOAs for cancer. The WOE for each hypothesized MOA was summarized in narrative tables based on comparison and contrast of the extent and nature of the supporting database versus potentially inconsistent or missing information. The comparison was based on evolved Bradford Hill considerations rank ordered to reflect their relative contribution to WOE determinations of MOA, taking into account increasing experience in their application internationally. This clarification of considerations for WOE determinations as a basis for comparative analysis is anticipated to contribute to increasing consistency in the application of MOAHR analysis and potentially, transparency in reporting science/policy from public policy considerations in regulatory risk assessment. Copyright © 2014, The Authors. Journal of Applied Toxicology. Published by John Wiley & Sons Ltd.

Keywords: human relevance framework, mode of action, weight of evidence, log severity, evolved Bradford Hill considerations

Introduction

The mode of action human relevance (MOAHR) framework is an analytical framework designed to increase transparency in the systematic consideration of the weight of evidence (WOE) of hypothesized MOAs for critical effects and their relevance to humans. It was developed in reliance on the International Life Sciences Institute Risk Services Institute (ILSI) and the International Programme on Chemical Safety (IPCS) and derives from earlier work on MOA by the US Environmental Protection Agency (USEPA) and IPCS (Danch-Mullin et al., 2001).

The development and evolution of the IPCS ILSI MOAHR framework which has involved large numbers of scientists internationally, is described in several publications (Bachman et al., 2001, 2005; Meek, 2010; Meek et al., 2011; Meek et al., 2011). Potential application in a broader range of relevant contexts has been considered more recently (Cassidy et al., 2011; Meek and Blasing, 2010). The framework has been illustrated by an increasing number of case studies (n = 30, currently), and is widely adopted in international and national guidance and assessments (Meek et al., 2008), including those of the USEPA (Dufresne and Bartsch, 2005; Markaverich et al., 2007; 548, 1999, 2007; 549, 2009; USEPA, 2003a), building on the reflective experience, the framework has been updated recently to address uncertainty additionally and to extend its utility to emerging

sciences in toxicity testing and monitoring methods. The updated MOAHR framework offers a road map, incorporating continuous refinement efforts for purpose testing strategies and risk assessment (Meek et al., 2014).

In addition to increasing transparency through structured evaluation of the evidence and uncertainties upon which conclusions are based, MOAHR analysis also contributes to the transparent assessment of a wide range of data to best inform and research context. This is important because it facilitates identification of critical data needs and contributes to transparency in the separation of science/policy, i.e., weighting of options based on systematic consideration of available scientific support from public health protection policy, the latter

*Correspondence to: M. E. Bette Meek, University of Ottawa, One Jean Jacques, Suite 80, Ottawa, Ontario, Canada, K1N 6N5. Email: bette@uottawa.ca

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited. This article is a U.S. Government work and, as such, is in the public domain in the United States of America.

University of Ottawa, One Jean Jacques, Suite 80, Ottawa, Ontario, K1N 6N5, Canada

**Objectives/Approach
 Meek et al. DOI 10.1002/jat.2984
 (2014b).**

•Application of B/H Considerations for WOE in MOA Analysis

•Evolved (simplified & rank ordered) B/H considerations based on acquired experience to increase:

- Transparency
- Consistency

•Illustration through application to existing regulatory risk assessments in comparative WOE analysis

Weight of Evidence for Stressor Specific Hypothesized MOAs/AOPs

Evolved BH Considerations	Defining Questions
Biological Concordance	Does the hypothesized AOP conflict with broader biological knowledge? How well established is the AOP?
Essentiality of Key events	Is the sequence of events reversible if dosing is stopped or a key event prevented?
Concordance of Empirical Observations	Dose response – Are the key events observed at doses below or similar to those associated with the apical effect? Temporality – Are the key events observed in hypothesized order? Incidence – Is the frequency of occurrence of the adverse effect less than that for the key events?
Consistency	Is the pattern of effects across species/strains/organs/test systems what would be expected based on the hypothesized AOP?
Analogy	Would the mode of action be anticipated based on broader chemical specific knowledge?

Evolving Guidance for WOE – Stressor Specific MOA/AOP

Evolved BH Considerations	Stronger	Weaker
Biological Concordance	Well established	Novel biological processes
Essentiality of Key events	<i>Direct</i> experimental evidence	Data on reversibility only. Indirect or no data.
Concordance of Empirical Observations	Dose Response & Temporality – expected pattern based on <i>robust</i> database Incidence – incidence of early KEs > than for later KEs	All key events at all dose levels and time points and/or <i>limited</i> data
Consistency/ Uncertainty	Pattern of effects are what you would expect across species, strains, organs, and/or test systems	Significantly inconsistent or limited data available to assess (e.g., observed in single test system)
Analogy	Observations are consistent with those for other (related) chemicals having well defined MOA	Pattern of effects for other (related) chemicals is distinctly different. Insufficient data to evaluate whether chemical behaves like related chemicals with similar proposed MOA

Meek et al., 2014b

Refined AOP Template

Background

AOP Identifier

Authors

Date of Updating

Abstract/Background (Optional)

Summary of AOP and Key Event Descriptions

KER Descriptions

- Weight of Evidence for KERs
- Quantitative Understanding

Assessment of the AOP

- Domain of Applicability
- Relative Level of Confidence
 1. Biological Plausibility – KERs
 2. Essentiality – KEs
 3. Empirical Support for the KERs/AOP
 4. Quantitative Understanding of Each of the KERs

Potential Application (Optional)

(OECD, 2014) Users' Handbook Supplement To The Guidance Document For Developing And Assessing AOP

16

Objectives of WOE Guidance in the Supplement/Wiki

- To *simplify*, clarify and “codify” to the extent possible, WOE application (B/H considerations) in an AOP context, addressing:
 - Focus (a limited no. of critical elements)
 - Consistency/simplicity of terminology
 - Clarification of the nature of the data that informs through inclusion of:
 - defining questions &
 - provision of criteria & examples to bound categories of confidence (low, medium, high)
- To increase understanding of more versus less influential components (ranking/weighting) for qualitative WOE
- To integrate quantitation to inform specific application

17

Why Bother?

- To provide a consistent representation of degree of confidence in various aspects of (incomplete) AOPs as a basis to facilitate their consideration for specific application (e.g., IATA)
 - Essential to coordinate the critical *research/regulatory* interface
- But, we’ve just *started* down the path
 - *Requires application and feedback*
 - *i.e., development/refinement of AOPs*
 - Need for a *rolling forward agenda* for additional development of the wiki/Supplement
 - Several aspects flagged, currently
 - Training

18

Focus/Consistent Terminology for Qualitative/Quantitative Elements

Qualitative WOE/Confidence

- Biological Plausibility – **KERs**
 - Biology of the pathway
- Essentiality – KEs
 - Experimental support from specialized studies to block or modify key events, stop/recovery studies
- Empirical Support – **KERs**
 - Dose-Response, Temporal and Incidence Concordance
 - Availability of Simple Template to Illustrate

Degree of Quantitative Understanding - KERs

19

Annex 1 – Assessing Confidence Definition, Basis for Calls, Examples

Consideration	Defining Questions	High (Strong)	Moderate	Low (Weak)
Biological Plausibility of KERs	Mechanistic basis dogma	Established mechanistic basis	Incomplete understanding	Empirical, association, only
Support for Essentiality of KEs	Impact of preventing a KE/ reversibility	Direct support	Indirect (impact of change in modifying factor)	No or contradictory evidence
Empirical Support for KERs	Use the template !!! Includes Consistency	Extensive consistent evidence with a wide range of stressors	More limited database with outliers that can be explained	No studies or evidence not supporting expected pattern

20

Template: Dose – Response and Temporality

Dose (mg/kg bw/day)	Key event 1	Key event 2	Key event 3
0.2 (2 ppm)	+ 4 weeks	+ 52 weeks	
1 (10 ppm)	++ 4 weeks	++ 52 weeks	+ 107 weeks
4 (40 ppm)	+++ 4 weeks	+++ 13 weeks	++ 52 weeks

Dose-Response

+ = severity

Contents of the WHO Update

Case Examples illustrating application of MOA analysis in:

- 1: Lack of human concordance
- 2: Contribution of well-designed omic studies to species concordance and dose-response analysis
- 3: The evaluation of epidemiological data
- 4: Development of more efficient testing strategies
- 5: Prioritizing substances for further testing (including genomic data)
- 6: Creation of chemical categories (including genomic data)
- 7: Identifying critical data gaps and testing strategies in read-across

**Case Example 2: Contribution of Well Designed “Omic” studies to Concordance and Dose-Response Analysis
Cancer**

Bladder tumours at high doses in rats

Metabolism and Cell Damage; Cytotoxicity and Regenerative Proliferation

Correlation of damage to urinary cells when reductive metabolism to toxic metabolite induced (D*)

D* in urine in animal studies increased at doses causing tumours (> than the *in vitro* LC50)

Supporting “Omic” Data

The spectrum of altered molecular pathways consistent - urothelial hyperplasia

Magnitude at the higher concentrations consistent

Genotoxicity

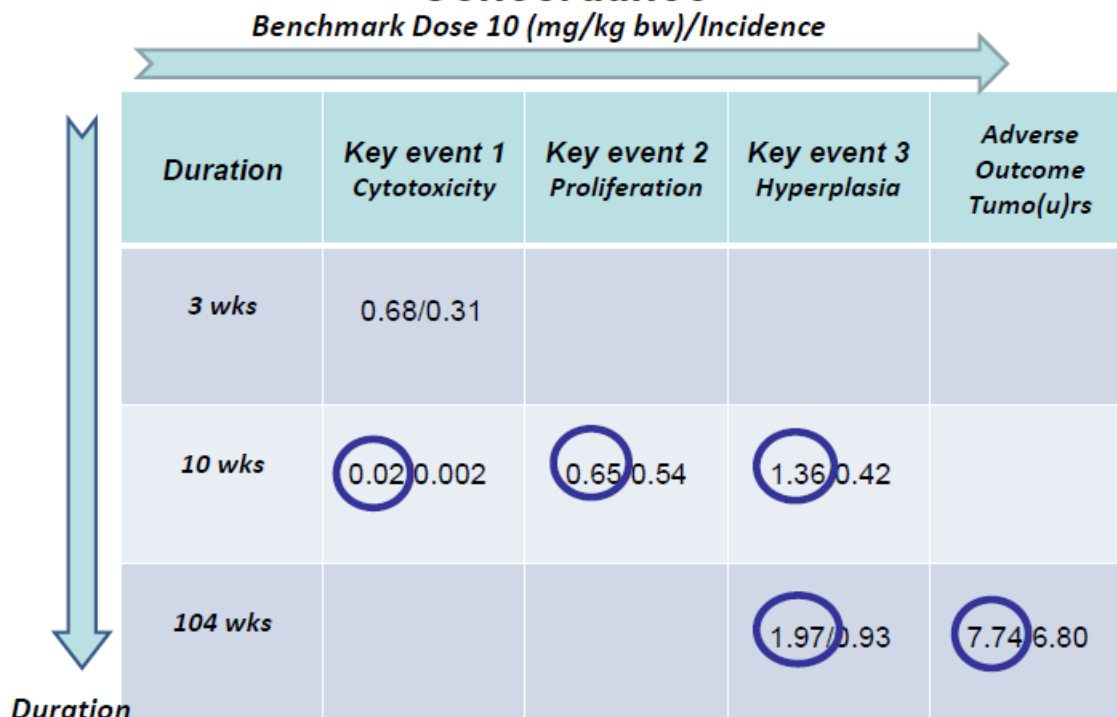
Data in a wide range of *in vitro* and *in vivo* assays; D & D* clastogenic at high doses

Human Data

Toxic metabolite in human urine

23

Case Example 2: Dose/Incidence/Temporal Concordance



Qualitative Confidence

- **Biological Plausibility – KERs**
 - Well accepted and documented path to cancer
 - i.e., metabolic activation – persistent cytotoxicity - tumours
- **Essentiality – KEs**
 - No adverse effects without relevant enzyme in null mice
 - Reversibility
 - Inhibition of reductive metabolism decreases toxicity; induction increases toxicity
- **Empirical Support – KERs**
 - Pattern of Temporal, Dose Response and Incidence Concordance supports hypothesized MOA based on extensive database (metabolic, omics, histopath)
 - a few explainable exceptions



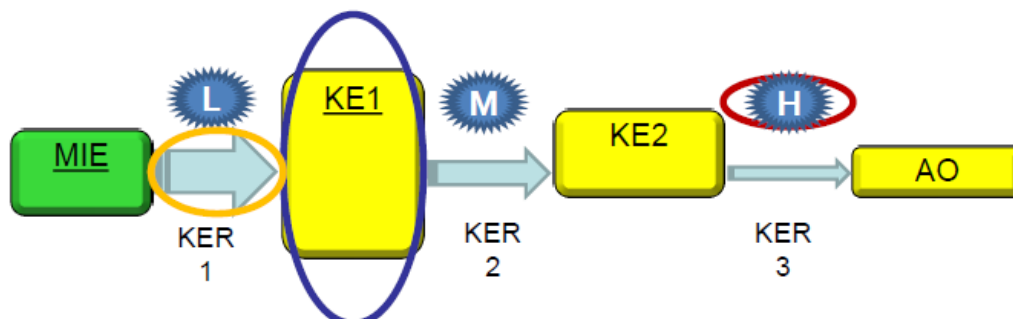
25

The Concordance Analysis Integrating TK Data

Key Event	Qualitative Concordance			Quantitative Concordance Humans
	Animals	Humans	Strength	
Reductive Metabolism / Active metabolite (D*) in urine	Yes	D* present in urine following exposure to analogue	Considerable In animals; limited but relevant to humans	PBPK model incorporating metabolic rates, enzyme affinities and distribution based on <i>in vitro</i> human data supported by <i>in vivo</i> data
Sustained cell damage and repair (cytotoxicity; proliferation)	In all cases at doses that induce tumours	Unknown Potential if sufficient D*	Considerable in animals, possible in humans but limited data	Genes upregulated in human & rat urinary bladder with rat cell line more sensitive
Bladder tumours	F 344 rats	Possible	Considerable in animals,; highly plausible in humans	No data

26

A "Snapshot View" to Facilitate Consideration of Context Specific Application



Confidence (Qualitative) Elements:

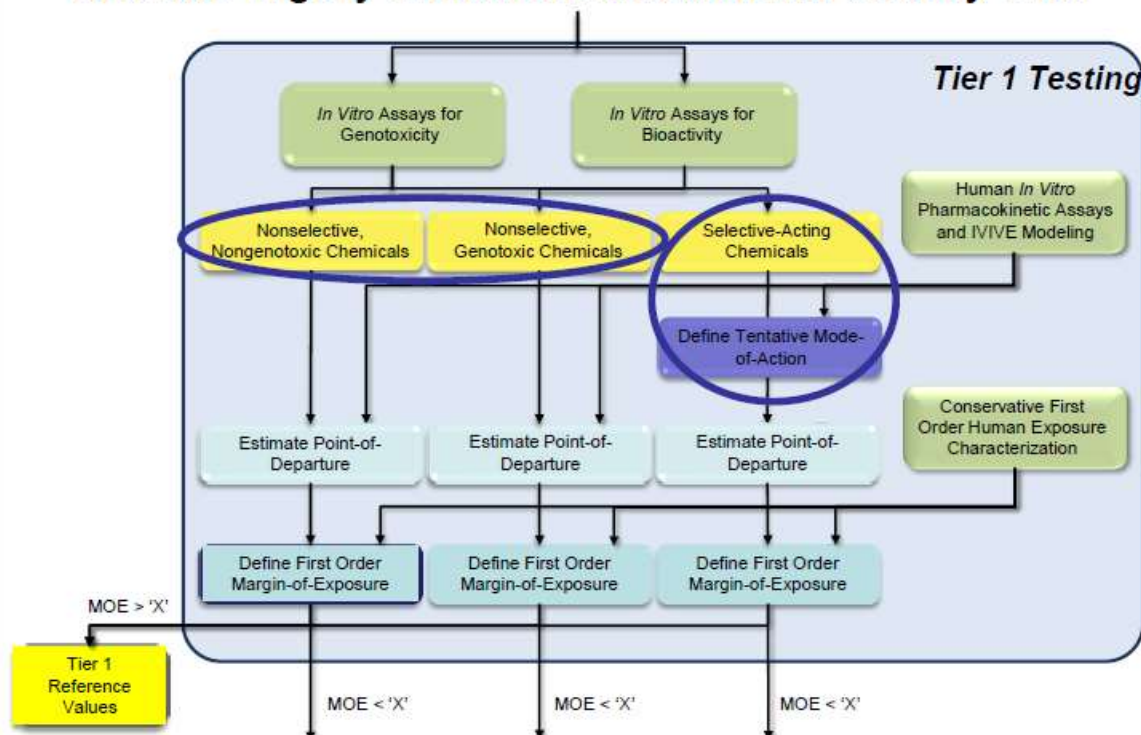
KEs – Essentiality (size of the node to represent H, M, L confidence?)

KERs – Biological Plausibility, Empirical Support

Degree of Quantitation of KERs (size of the arrow for H, M, L)

27

New and Legacy Chemicals with Minimal Toxicity Data



Thomas et al., (2013). Tox. Sci. 136(1), 4–18.

Documented MOAs

Mode of action	Case study	Reference
Tumors of various organs associated with mutagenic modes of action	Ethylene oxide 4-Aminobiphenyl	Meek et al. (2003) Cohen et al. (2006a)
Mammary tumors associated with suppression of luteinizing hormone	Atrazine	Meek et al. (2003)
Thyroid tumors associated with increased clearance of thyroxine	Phenobarbital Thiazopyr	Meek et al. (2003) Dellarco et al. (2006)
Bladder tumors associated with the formation of urinary tract calculi	Melamine	Meek et al. (2003)
Liver/kidney tumors associated with sustained cytotoxicity and regenerative proliferation	Chloroform	Meek et al. (2003)
Acute renal toxicity associated with precipitation of oxalate	Ethylene glycol	Seed et al. (2005)
Androgen receptor antagonism and developmental effects	Vinclozolin	Seed et al. (2005)
Nasal tumors associated with DNA reactivity and cytotoxicity	Formaldehyde	McGregor et al. (2006)

29

Recommendations/Conclusions

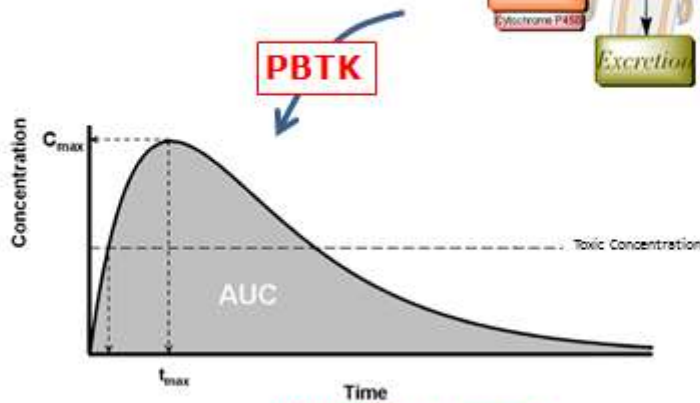
- MOA/AOPs builds on long standing regulatory experience
- Provides construct for coordinating input of the research community to address regulatory application
- Need to recognize/draw upon experience of a range of communities to facilitate simplicity, consistency, transparency
 - Need for simplified figurative output to facilitate consideration of application (the “Network” or “snapshot” view) based on research/regulatory input
- Stepwise process
- The need to document novel/data poor AOPs in systematic fashion from outset (frames next steps)

30

ANNEX 6 PRESENTATION: TK AND TD MODELLING APPROACHES RELEVANCE FOR USING MOA/AOP IN IATA

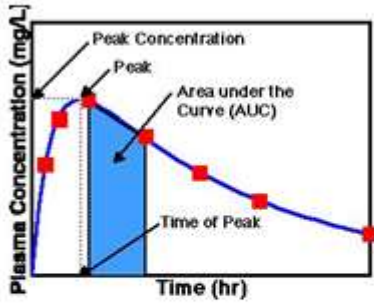


ToxicoKinetics
Basics + Simulation



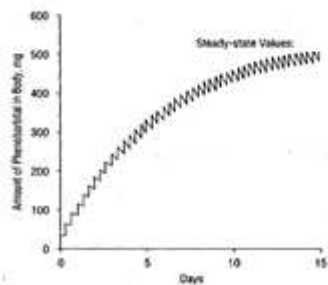
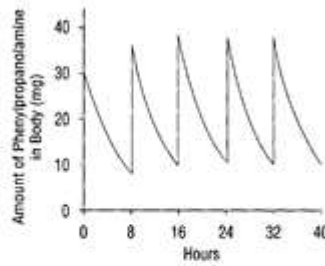
What is relevant in MoA?
 → Parent chemical ? Metabolite ?
 → AUC ? C_{max} ?
 → Rate of metabolite formed per minute ?

Single dose



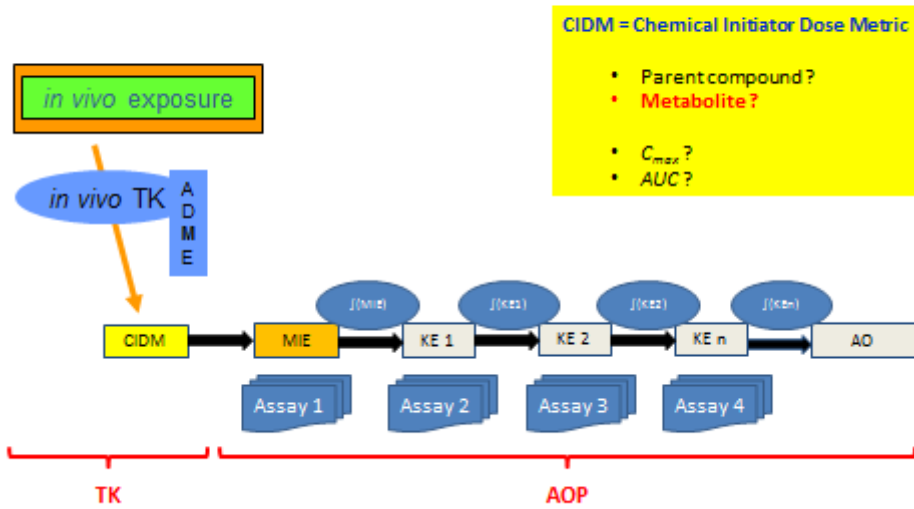
OECD 120 Workshop | US EPA, Washington DC, USA, 2014.11.17-19

Repeated dose



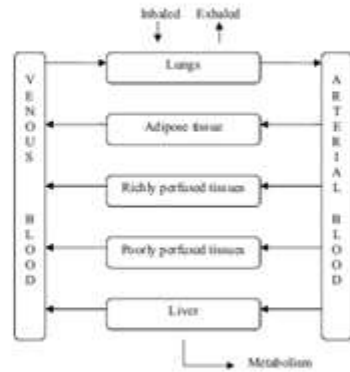
2

TK + Quantitative AOP = MOA



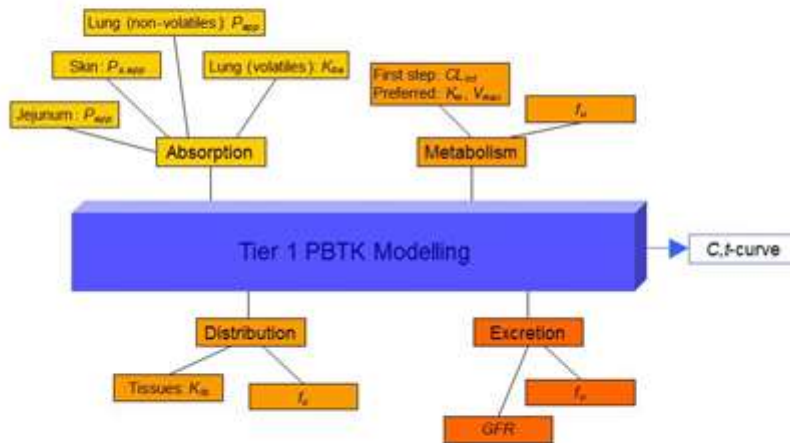
OECD 120 Workshop | US EPA, Washington DC, USA, 2014.11.17-19 | EPA Research Report | EPA 823-B-15-001, Office of Chemical Health and Consumer Protection

Schematic representation PBTK model



WHO/IPCS 2010 – Guidance on application and characterization of PBTK models in risk assessment

..... as simple as you want (Tier 1 PBTK)



Bessems et al. (2014), Reg Tox Pharmacol

..... as complex as needed, e.g. focus on metabolism

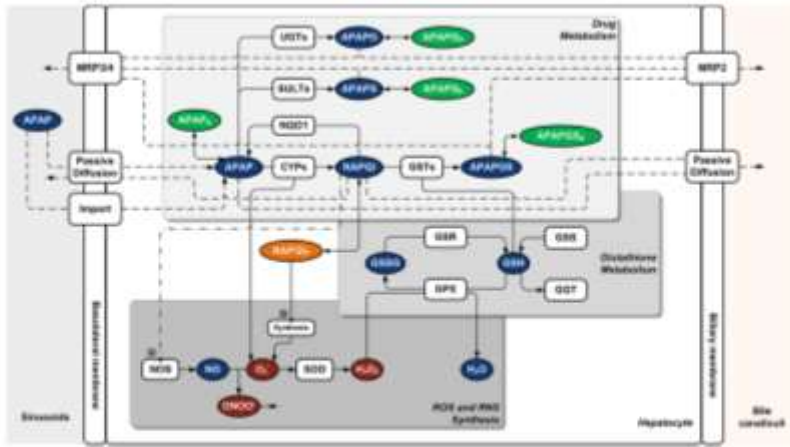


FIGURE 1 | Cellular metabolic network model for acetaminophen metabolism and toxicity. Abbreviations: APAP, acetaminophen; APAPG, acetaminophen-glucuronide; APAPGS, acetaminophen-glucuronide-sulfate; APAPS, acetaminophen sulfide; CYP, cytochrome P450 monooxygenase; GST, glutathione transferase; GPK, glutathione peroxidase; GSH, glutathione; GSR, glutathione reductase; GSS, glutathione synthase; GSSG, glutathione disulfide; GST, glutathione S-transferase; MRP120/4, multidrug resistance related protein; NQO1, N-acetyl-p-benzoquinone imine; NOS, nitric oxide synthase; NQO1, NAD(P)H quinone reductase; SOD, superoxide dismutase; SULT, sulfotransferase; UGT, UDP-glucosyltransferase; Index: "R", non-specifically bound protein/pep; index "P", non-specifically bound to protein.

Diaz Ochoa et al, *Toxicol Pharmacol*, 2015

APPENDIX A

A GENERAL REFERENCE ON PHYSIOLOGICALLY-BASED PHARMACOKINETIC/PHARMACODYNAMIC (PBPK/PD) MODELING

Yang, R.S.H., M.E. Anderson, J.B. Demison, Y.C. Gu, K.M. Uao and S. Rowford, 2004. Physiologically based pharmacokinetic and pharmacodynamic modeling. Chapter 23. In: *Molecular Models of Cancer*, S.C. Holland, Ed. Wiley Inc., New York, NY, p. 501-408.

[US EPA link](#)

Table 23.1. Dosing Schedule Dependence of Methotrexate Toxicity in Mice

Dose (mg/kg)	Schedule	Total Dose (mg/kg)	Peak Plasma Concentration (M)	Effect
350	Single dose	350	10 ⁻³	LD50
25	Twice daily	50	10 ⁻⁴	LD50
3	Every 3 h, 5 times, rest 8 h, then every 3 h, 3 times	24	10 ⁻⁵	>LD50 ^a
0.5	Every 3 h, 20 times	10	10 ⁻⁶	>LD50 ^a
0.8 µg/h	Infusion 96 h	3	10 ⁻⁸	>LD50 ^a

^aHigher toxicity than LD₅₀. Source: Morrison et al., 1987.

Table 23.1 shows the toxicity of methotrexate to mice under a variety of dosing schedules (Morrison et al., 1987). Obviously, toxicity does not directly correlate with total dose. Decreasing total dose by a factor of 117 (350 to 3) led to no increase, rather than a substantial decrease, in toxicity. In addition, the area under the plasma concentration-time curve (AUC), a frequently used pharmacokinetic parameter for bioavailability, did not correlate with toxicity. For instance, the AUC for a bolus dose of 350 mg/kg is about two orders of magnitude

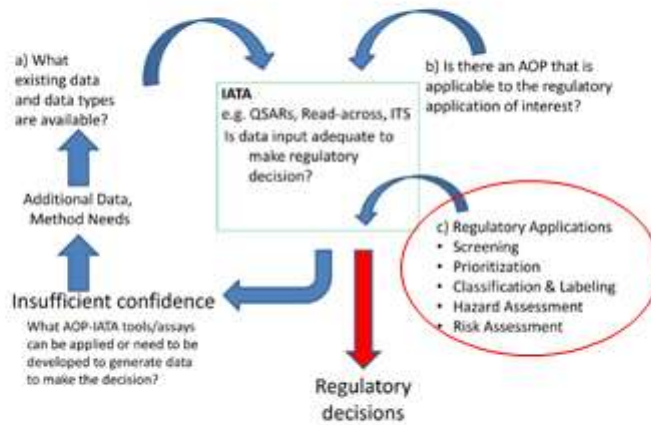
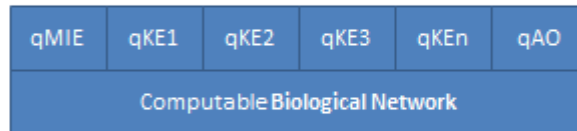
greater than that of the 96-h infusion at 0.8 µg/h, yet toxicity is higher with infusion.

The reason for the above phenomenon turned out to be intimately associated with the pharmacokinetics of methotrexate (and its polyglutamate metabolites) and the threshold and length of time that dihydrofolate reductase is inhibited (thus inhibition of DNA synthesis). At a bolus dose of 350 mg/kg, even though there is a short period of very high blood and tissue concentration, the inhibition of DNA synthesis did not persist long enough to cause lethality in at least some of the mice. At an infusion rate of 0.8 µg/h for 96 h, even though the blood and tissue levels were low, they were nevertheless high enough to cause sustained inhibition of DNA synthesis, which ultimately translated into higher lethality in the animals. These scientific discoveries eventually led to the revision of the methotrexate PBPK model by incorporating the inhibition of DNA synthesis into the model.

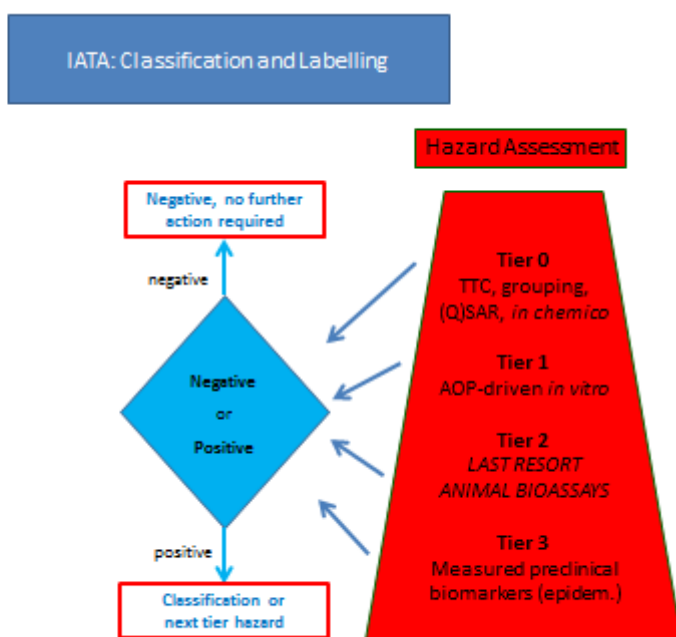
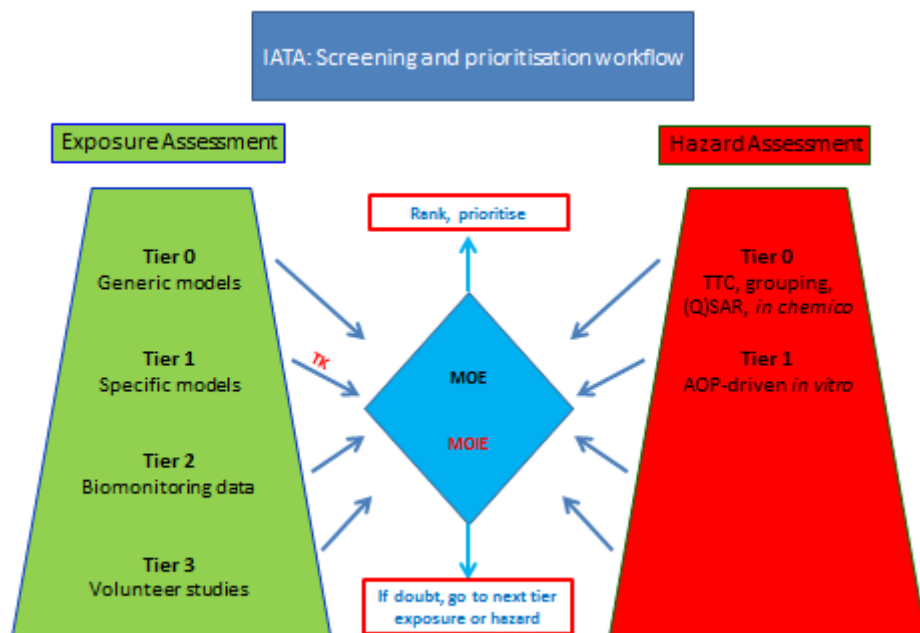




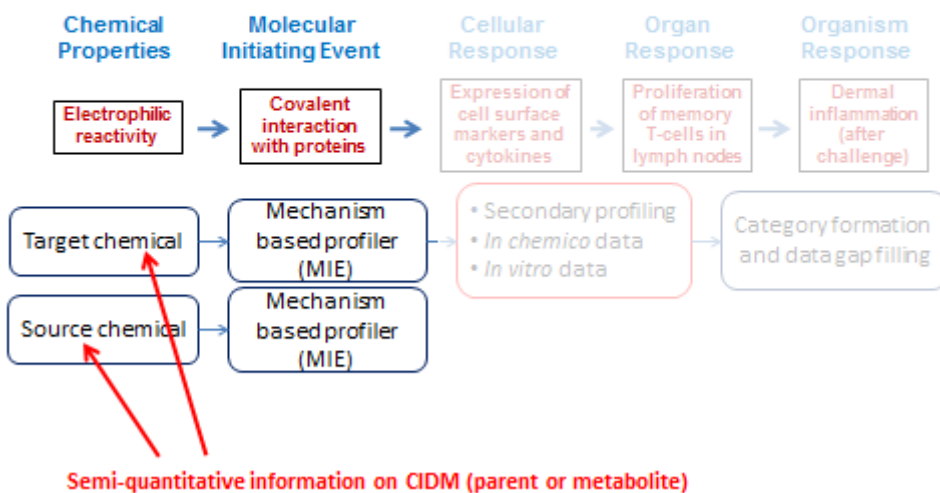
a.o. Sturla et al, Chem Res Tox, online 2013



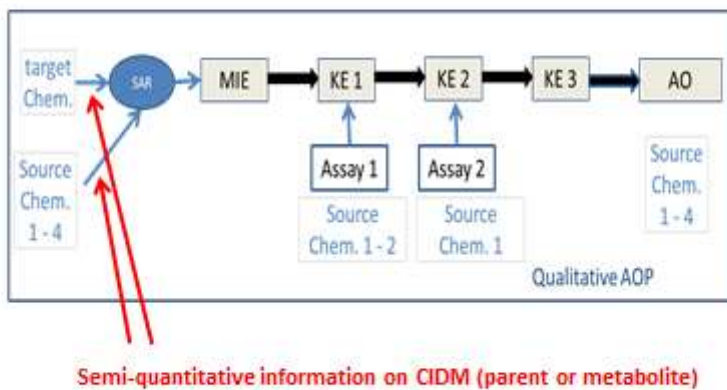
Tollenaar et al. (2014) 1177



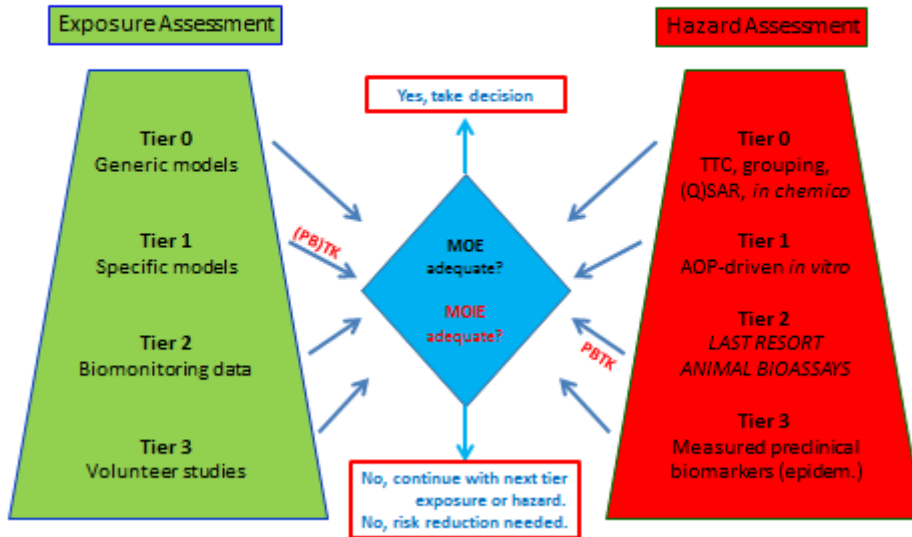
AOP: An example for skin sensitisation



Use of the AOP concept to categorise chemicals for a specific endpoint



IATA and probabilistic approaches: If insight in uncertainty and variability is needed, establish all values (exposure pillar + hazard pillar) as distributions






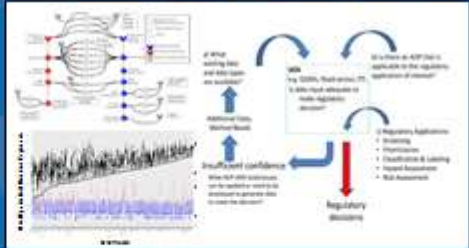
תודה
Dankie Gracías (شكر)
Спасибо Merci Takk
Köszönjük Terima kasih
Grazie Dziękujemy Děkujeme
Ďakujeme Vielen Dank Paldies
Kiitos Tänname teid 謝謝
Thank You Tak
感謝您 Obrigado Teşekkür Ederiz
Σας ευχαριστούμε 감사합니다
Bedankt Děkujeme vám
ありがとうございます
Tack

Joint Research Centre (JRC)
The European Commission's in-house science service
www.jrc.ec.europa.eu
Serving society - Stimulating Innovation - Supporting legislation

ANNEX 7 PRESENTATION CASE STUDY 1: THE USE OF THE AOP FOR AR / ER BINDING THYROID EFFECT TO PRIORITISE AND SCREEN CHEMICALS

 **The use of the AOP for AR / ER binding / thyroid effect to prioritise and screen chemicals**

*Kevin M. Crofton and Richard Judson
U.S. EPA, National Center for Computational Toxicology*



Office of Research and Development

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



Outline

- 1) Regulatory problem
 - EDSP Prioritization and Screening
- 2) Incorporating AOPs into IATA
 - Estrogen receptor driven adverse reproductive outcomes
 - Thyroid receptor driven adverse neurodevelopmental outcomes
- 3) Regulatory Outcomes
- 4) Lessons Learned



Legislative Mandates for Endocrine Testing

- **1996 Federal Food, Drug and Cosmetic Act**
 - Develop a screening program to test for endocrine activity
- **1996 Safe Drinking Water Act Amendments**
 - Test chemicals found in drinking water for endocrine activity

Office of Research and Development



Endocrine Screening Program

Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) Key Recommendations:

- Expand Protection to Include Human Health and Wildlife
- Include Estrogen, Androgen and Thyroid Pathways
- Develop a Two-Tiered Screening and Testing Program:

EDSTAC Conceptual Testing Framework:

- **Tier 1 Screening for *Potential* to Interact**
 - *Potential to interact* with the estrogen, androgen or thyroid hormone systems
- **Tier 2 Testing to determine *Interaction* with the endocrine system**
 - If endocrine-mediated adverse effects then *quantify dose-response relationship*

Office of Research and Development

Slide 3



The Problem

Too many chemicals to test with standard animal-based methods

– Cost, time, animal welfare

Need for more efficient methods that prioritize the 'worst first'

– Include HTS assay results – empirical data
– Incorporate AOP to increase confidence in use



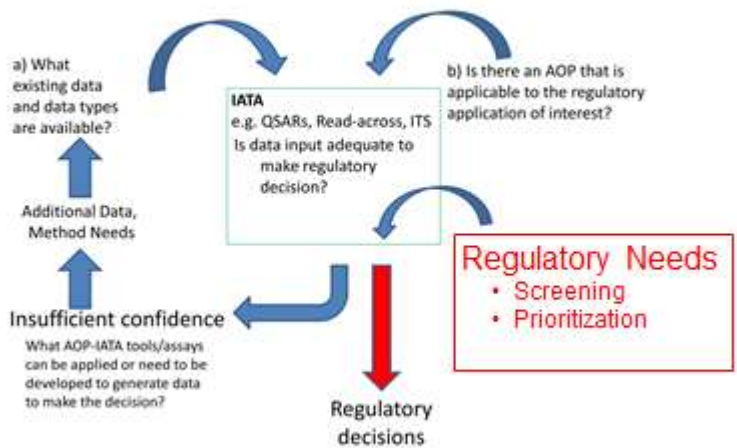
The Regulatory & Science Balance

- **The regulatory need is clear – 1000s of chemicals with no data and a need to prioritize them for further testing**
 - *Need to predict outcomes from disruption of estrogenic and thyroid systems*
- **The science needs to inform the regulatory process:**
 - With up-to-date science
 - Estimates of uncertainty in the data
- **This allows regulators to determine with the confidence (i.e., uncertainty) matches the decision needs**
- **IATA provides a framework to do this**

Office of Research and Development



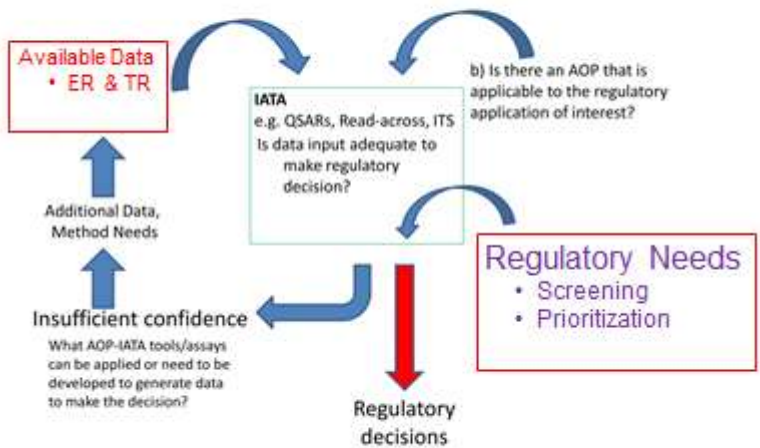
AOPs and IATA and EDSP



Office of Research and Development



AOPs and IATA and EDSP



Office of Research and Development



Available Data - ER

- 18 different assay-endpoints from ToxCast (1800 chemicals) and Tox21 (8500 chemicals)
- Different technologies and different points in the cellular ER pathway
 - 3 Radioligand Biochemical Assays (Novascreen)
 - 8 Protein Complementation Assays (Odyssey Thera)
 - 2 RNA Transcription Activation Assays (Attagene)
 - 4 Protein Transcription Activation Assays (Tox21)
 - 1 Proliferation Assay (ACEA)
- In vivo data sets from guideline and 'guideline-like' uterotrophic assays (29 active, 13 inactive chemicals)
 - Represents the best estimate of 'truth'.....

Office of Research and Development



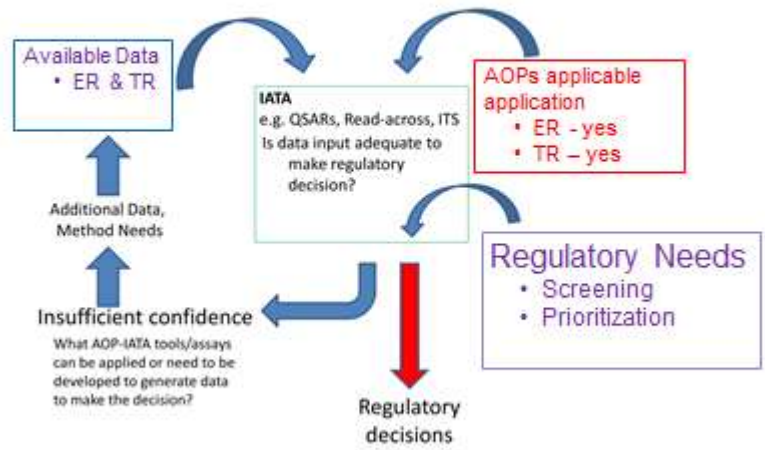
Available Data - TR

- 4 different assay-endpoints from ToxCast (1800 chemicals) and Tox21 (8500 chemicals)
- Different technologies and different points in the cellular ER pathway
 - 1 Radioligand Biochemical Assays (Novascreen)
 - 1 RNA Transcription Activation Assays (Attagene)
 - 2 Protein Transcription Activation Assays (Tox21)
 - 1 TRH Cell-free receptor binding GPCR (Novascreen)
- In vivo data sets from ToxRefDB including subchronic, chronic, pubertals and reproductive studies (xxx studies and xxx chemicals)
 - Represents the best estimate of 'truth' for thyroid disruption outcomes

Office of Research and Development



AOPs and IATA and EDSP

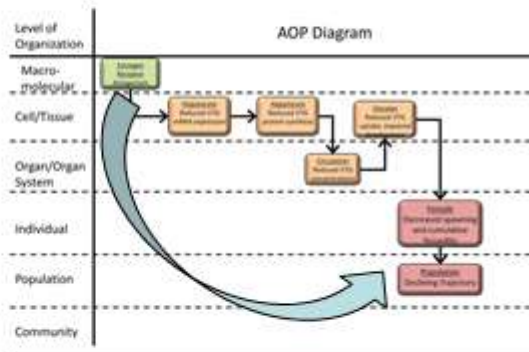


Office of Research and Development



ER AOP

- Well developed AOP for reproductive outcomes
 - Clear causative links, more limited quantitative links
- Currently used in the OECD QSAR Toolbox
- OECD AOPWiki entry being developed for fish reproductive outcomes for both agonists and antagonists
- High degree of confidence that alterations in the MIE will lead to adverse outcome (*qualitative*)



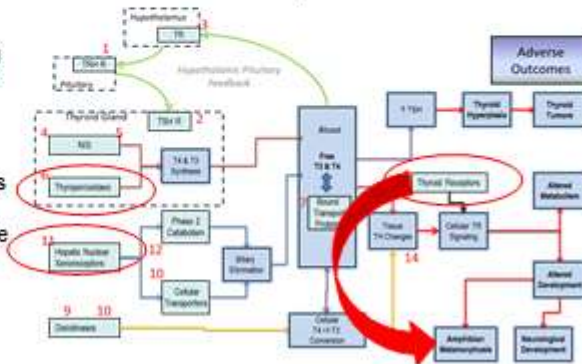
Office of Research and Development



TR AOPs

- Can we take the same approach as ER?
- No AOP available that is specific for a TR to AO pathway
- Two MOAs available that include TR as KE leading to adverse neurodevelopmental outcomes
- Similar status – clear causative links – limited quantitative

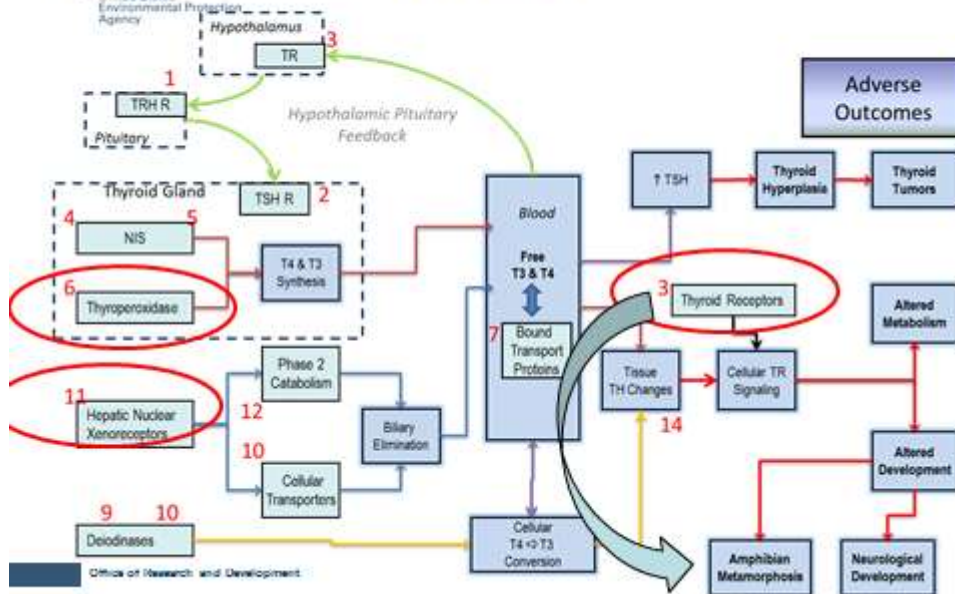
- Two MOAs demonstrating relevance to humans (Crit Rev Tox. 2005)
- OECD AOPWiki entries being developed for the two human relevant MOAs
- High degree of confidence that alterations in the MIE will lead to adverse outcome (*qualitative*)



Office of Research and Development



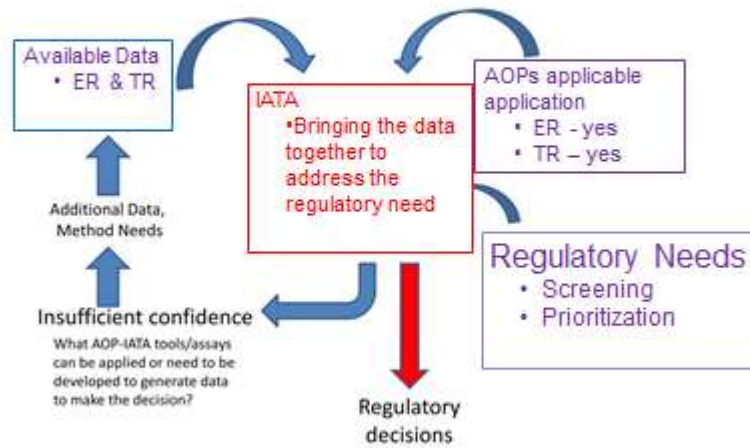
TR AOP



Office of Research and Development



AOPs and IATA and EDSP



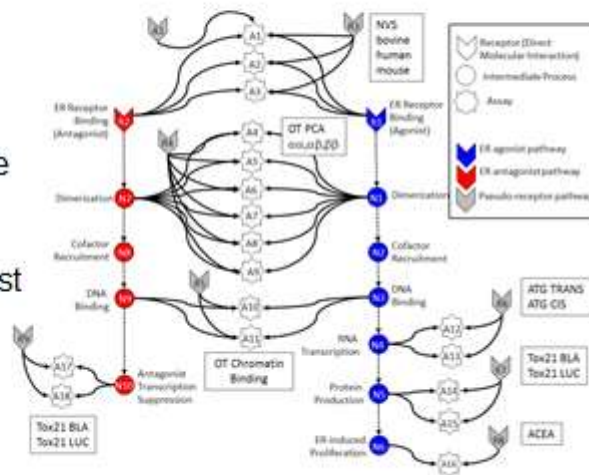
Office of Research and Development



ER as First Example

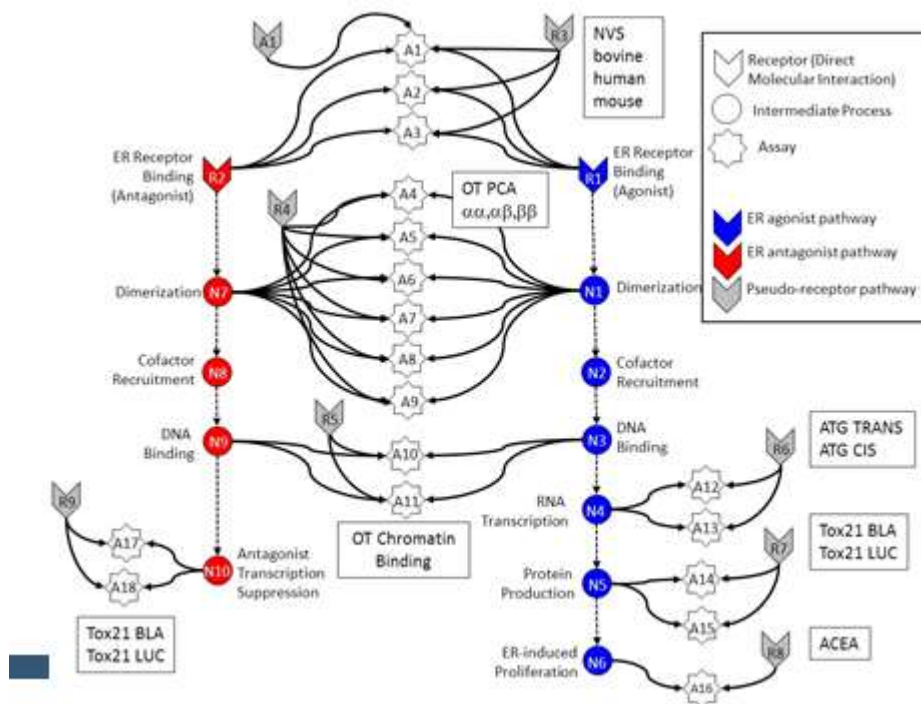
Combines results from multiple in vitro assays

- No assay is perfect
 - Assay Interference
 - Noise
- Use model to integrate assays
- Evaluate model against reference chemicals



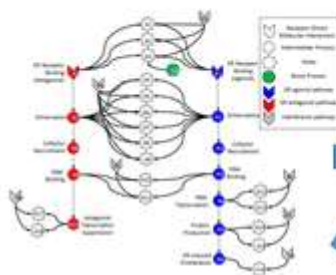
Office of Research and Development

Judson, unpublished



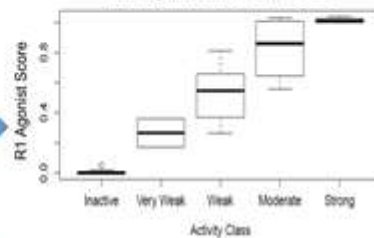
Checking Prediction and Ranking Chemicals

Checking Predictions

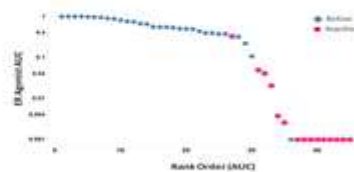


Predictive Consensus Model
(18 ToxCast assay endpoints for ER)

Model Agonist Scores vs Expert Calls

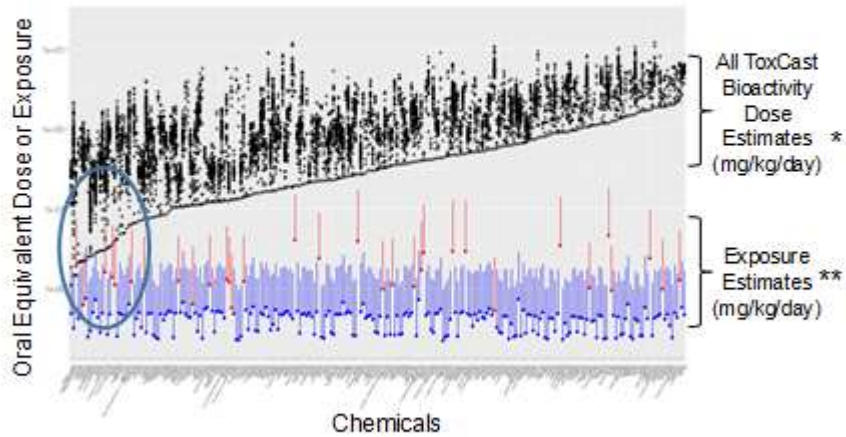


ER Agonist Calls vs In vivo Uterotopic Activity



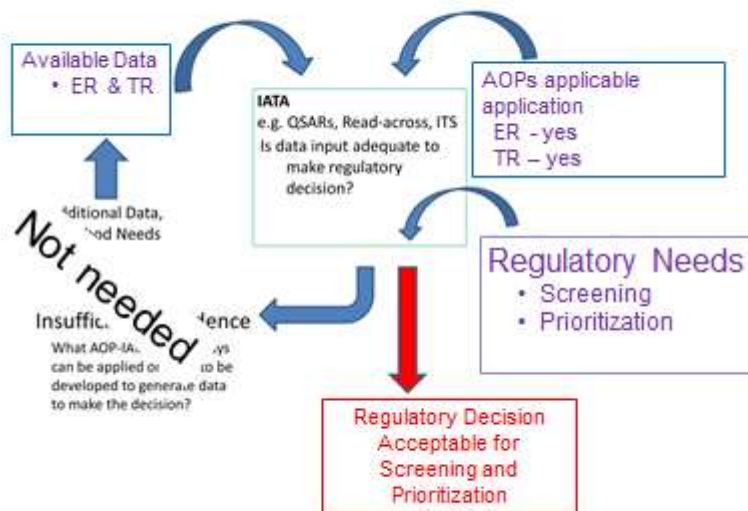
Integrating Bioactivity and Exposure

Prioritization = Hazard + Exposure



* Bioactivity converted to oral equivalent with rTK
 ** Daily exposure estimates from ExpoCast modeling

IATA and ER



ER IATA Summary

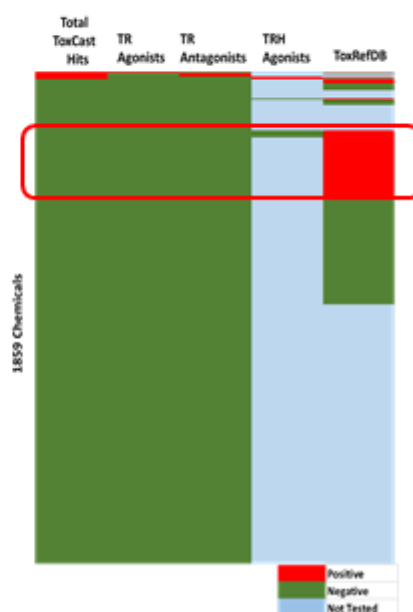
- **Consensus model provides good predictions**
- **Combined with exposure estimates provides a prioritization for follow-up testing of the 'worst first'....**
- **Remaining uncertainties**
 - Limited number of false negatives
 - Limited metabolism in HTS models
 - The estrogen receptor may not be only MIE

IATA for TR

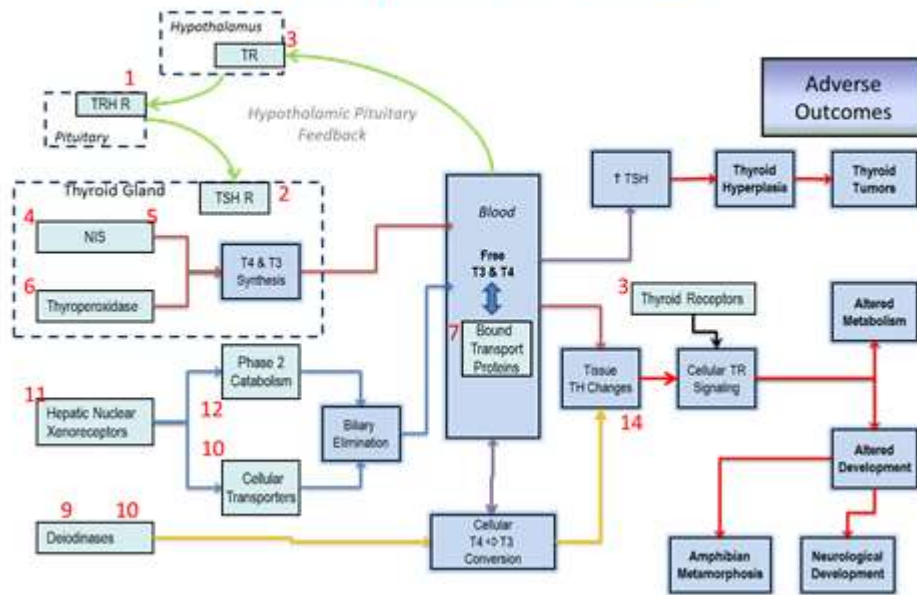
ToxCast Assay Endpoints vs ToxRefDB

Assay Endpoint Name	Component	Assay Type	Target	Species
Tox21_TR_LUC_GH3_Agonist		Cell based luciferase reporter gene	TR	rat
Tox21_TR_LUC_GH3_Antagonist		Luciferase Reporter Gene Assay	TR	rat
ATG_THRa1_TRANS		Cell based transcription factor activity	TR alpha	human
NVS_NR_hTRa		Cell free coregulator recruitment, antagonist mode	TR alpha	human
NVS_GPCR_rTRH		Cell free receptor binding	TRH	rat

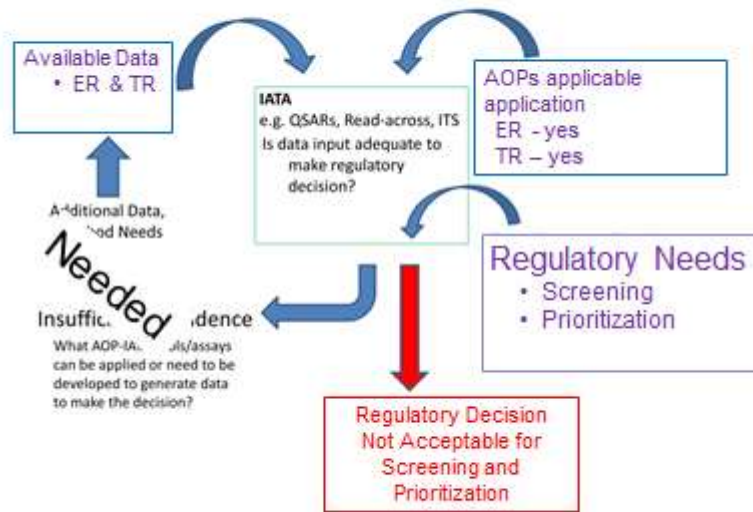
HTS Assays for TR do not predict adverse outcomes from thyroid system disruption



Thyroid System AOP & MIEs



IATA and TR



IATA and TR - Summary

- **No confidence** for use in regulatory decisions due to failure to predict majority of in vivo adverse effects
 - Not surprising based on multiple AOPs that lead to AO
- Data gaps
 - Need additional HTS data for other MIEs
 - Need methods development for some MIEs

Lessons Learned

- **Regulatory needs must be clearly articulated**
 - In the case the need was Screening and Prioritization
 - Drives the degree of confidence needed from the AOP
- **Use of HTS assay data for Screening and Prioritization**
 - Match the uncertainty in the AOP with regulatory decision
 - Multiple assays integrated into a consensus model provide more accurate predictions
 - Qualitative AOPs are sufficient if predictive nature is established
- **“Failure” to meet needs of regulatory decision**
 - IATA process provide unique opportunity to fine-tune data needs
- **Integration of bioactivity and exposure estimates within IATA is critical for prioritization process**

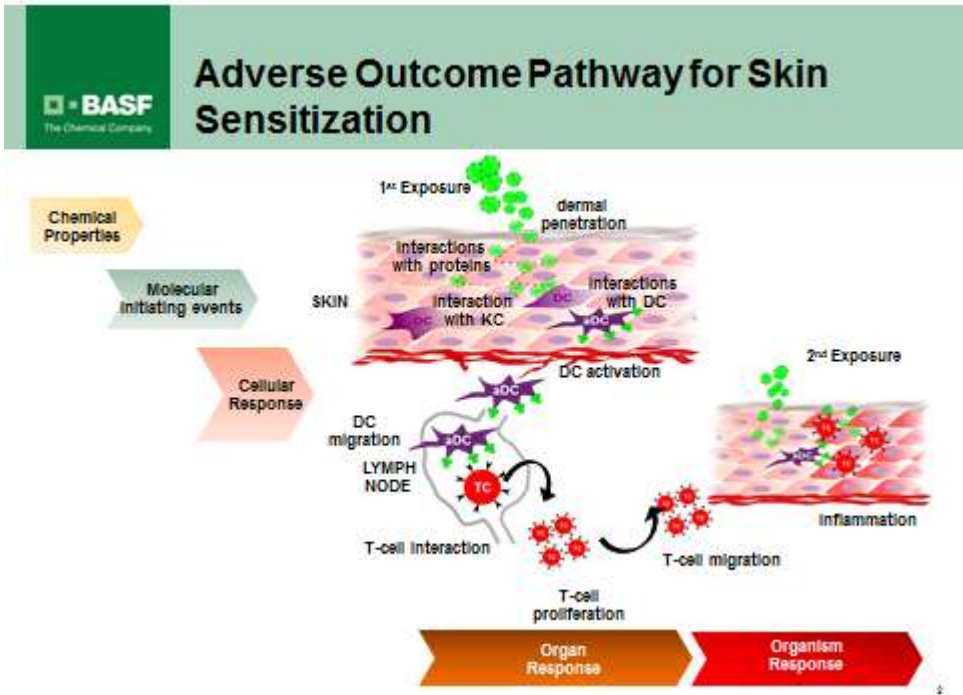
ANNEX 8 PRESENTATION CASE STUDY 2: ADVERSE OUTCOME PATHWAY FOR SKIN SENSITIZATION

BASF
The Chemical Company

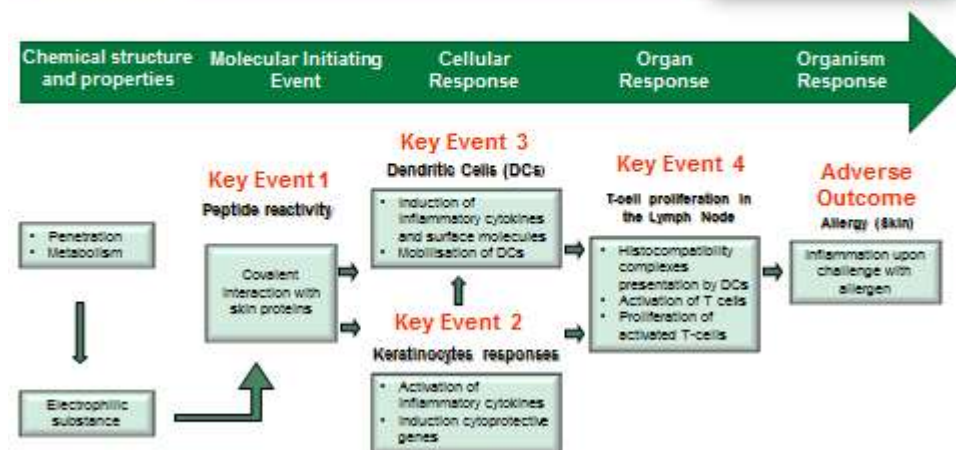
Adverse Outcome Pathway for Skin Sensitization

For Grouping of Substances, Read Across and Regulatory Application

Frank Faulhammer, Global Product Safety



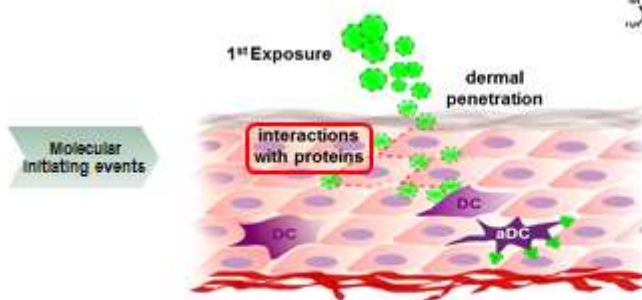
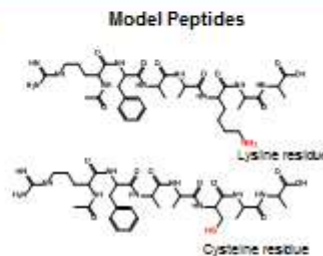
The adverse outcome pathway for skin sensitization (adapted from OECD; 2012)



4

Reactivity and Protein Binding
Direct Peptide Reactivity Assay (DPRA)

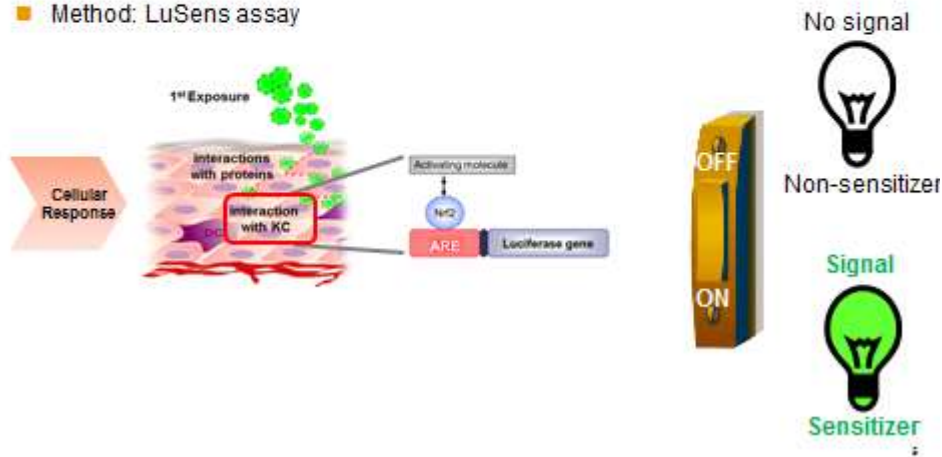
- Small molecules (haptens) must bind to proteins to attain their allergenic potential
- Method: Direct peptide reactivity assay (DPRA)



4

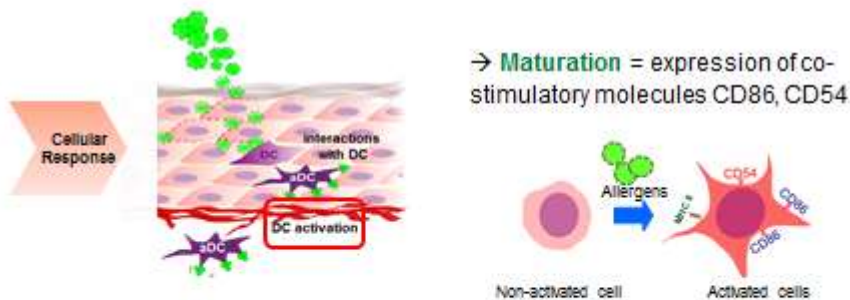
BASF The Chemical Company
Cellular responses
ARE-dependent reporter cell lines (LuSens)

- Cells (e.g. keratinocytes) must be activated to produce essential signaling molecules
- Method: LuSens assay



BASF The Chemical Company
Cellular responses
Dendritic cell activation assays

- Antigen presenting cells must upregulate cell surface markers to interact with T-cells.
- Method(s): MUSST and/or h-CLAT





- **Domain of applicability of the AOP**

- The key events for this AOP appear to be conserved across all mammals.

- **Relative level of confidence in the AOP**

- **Biological plausibility**

There is good agreement between the sequences of biochemical and physiological events leading to skin sensitization

- **Empirical Support for each of the Key Event Relationships**

While there is general agreement regarding the events, understanding of the underlying biology of some of the key events remains incomplete

- **Uncertainties/Inconsistencies**

There are uncertainties in the AOP; for example it is known that certain chemicals lead to T-cell proliferation in the LLNA without being skin sensitizers

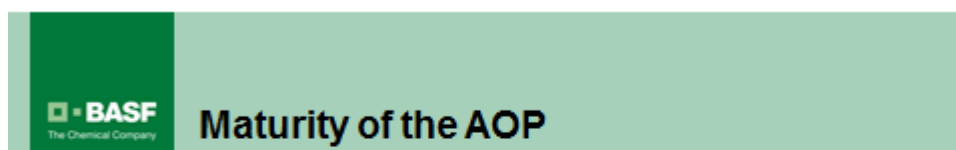
- **Essentiality of the Key Events**

The molecular initiating event (protein binding reactions) is based on long-standing, well-studied organic chemical mechanisms and reactions. Sensitization is causally linked to keratinocyte activity and T-cell proliferation and, to a lesser extent, dendritic cell activation/maturation.

- **Degree of Quantitative Understanding**

For skin sensitization, a major hurdle is moving from a qualitative AOP to a quantitative AOP. The qualitative AOP is clearly supported as a means to identify and characterize the potential for a chemical to be a sensitizer, the ability to consistently predict relative potency is currently lacking.

7



- **Correlative or incomplete AOPs / MoAs** have only qualitative or limited quantitative understanding of one or two cause and effect linkages between KEs or a KE and the AO. These pathways are often based on a few stressors tested in a limited number of assays with a low level of confidence in the AOP.
- **Qualitative AOPs / MoAs** have qualitative understanding of critical components of the AOP / MoA. These pathways are often based on one or a few well-studied stressors where there is experimental evidence for the most critical KEs and the AO. The level of confidence in the AOP is moderate.
- **Semi-Quantitative AOPs / MoAs** have, in addition to qualitative understanding of the entire AOP / MoA, **semi-quantitative understanding of some of the KEs**. These pathways are **based on multiple compounds** and/or stressors **evaluated at several KEs and the AO**. The level of confidence in the AOP is moderate to high.
- **Quantitative AOPs / MoAs** have in addition to quantitative understanding of critical components of the AOP, empirical data across the spectrum of KEs and AO. These pathways are based on many compounds evaluated for all KEs and the AO so **in vitro effects can be scaled to in vivo effects for risk assessment**. The level of confidence in the AOP is high.

8



Application of the AOP

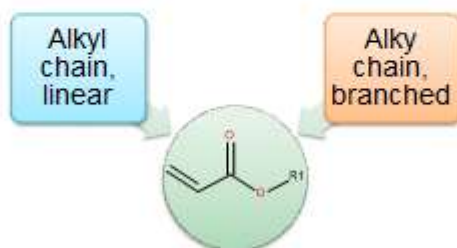
- Prioritize chemicals
- Grouping of chemicals
- Read Across to fill data gaps
- Replacement of the animal test



Building Categories Guided by the AOP

- Decide if target chemical falls into applicability domain
- Select similar chemicals based on common mechanism of action
- Check for common molecular initiating event as a first step in category building
- Divide into subcategories based on structural similarity

Common mechanism of Action: Acrylates



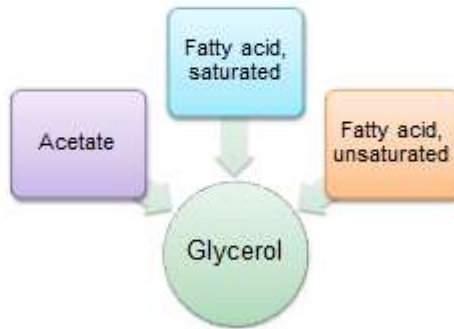
11

Data Matrix for Acrylates

Alkyl chain	Methyl-	Ethyl-	n-butyl	2-ethylhexyl-	2-propylheptyl
In vivo	Positive	GAP	Positive	Positive	GAP
OECD Toolbox PB alert	Michael acceptor	Michael acceptor	Michael acceptor	Michael acceptor	Michael acceptor
DPRA	Positive	Positive	Positive	Positive	Positive
LuSens				Positive	Positive
mMUSST				Positive	Positive

12

BASF The Chemical Company
Target Chemical in Existing Category: Glycerides



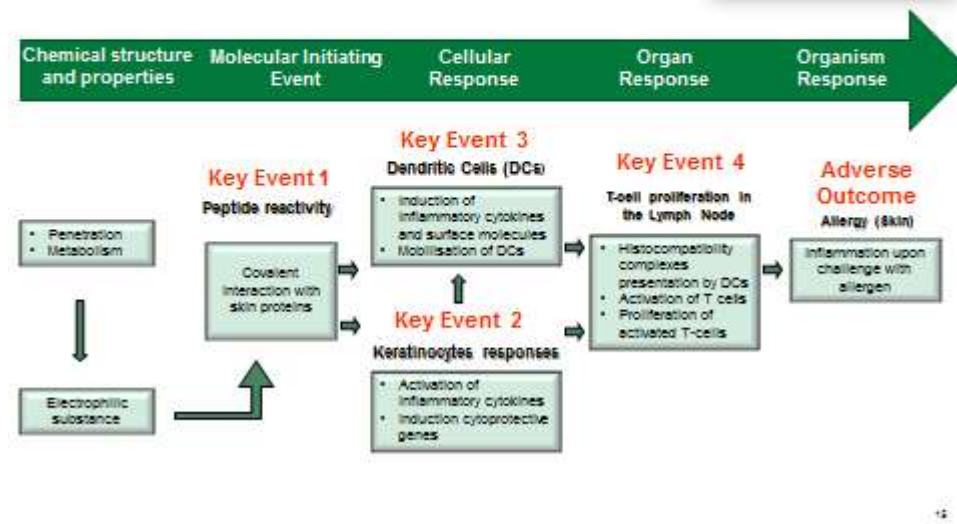
13

BASF The Chemical Company
Data Matrix for Some of the Glycerides

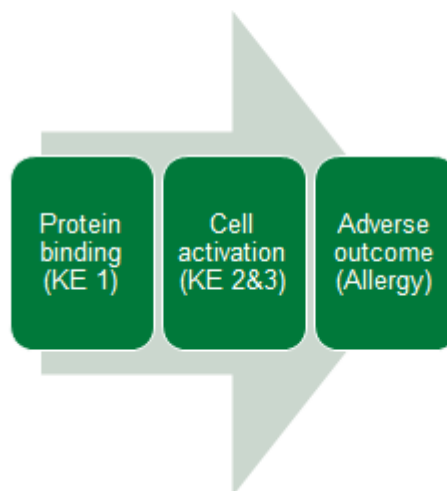
Fatty acid, saturated	Tri-C7	C8-18	C8-21	C14-18 mono and di	C18	C18
Fatty acid, unsat.		C18	C8-21	C16-18/C18 hydroxy		C18
Acetate			Mono and di			Mono and di
OECD Toolbox PB alert	No alert	No alert	No alert	No alert	No alert	No alert
In vivo	Negative	Negative	Negative	GAP	Negative	Negative
DPRA				Negative		

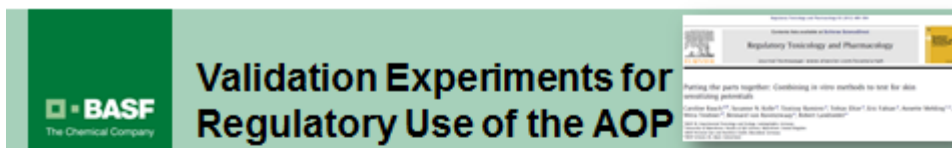
14

The adverse outcome pathway for skin sensitization (adapted from OECD; 2012)



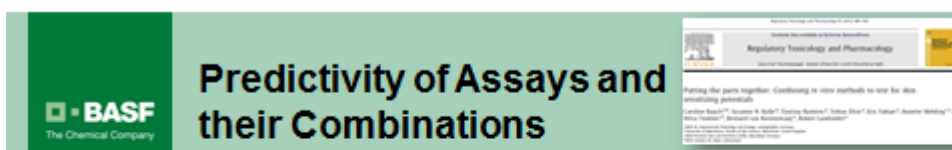
Can the AOP be used to predict the Apical Outcome?





- 59 test substances including LLNA performance standards
 - Additives/ stabilizers/ detergents 30%
 - Fragrances 24%
 - Cosmetic preservatives 22%
 - Cosmetic solvents 11%
 - Cosmetic dyes 7%
- 5/59 substances initially selected turned out not to be applicable due to technical reasons
- 54 substances with available LLNA and human skin sensitization information were evaluated in the 4 *in vitro/in chemico* assays in the validation process (DPRA, KeratinoSens, h-CLAT, mMUSST) along with the LuSens assay (similar to the KeratinoSens)

17



Compared to human		Positive predictive value	Negative predictive value	Accuracy
<i>In vivo</i> standard	LLNA	86 %	94 %	89 %
Individual assays	DPRA	88 %	86 %	87 %
	LuSens	83 %	81 %	82 %
	mMUSST	100 %	73 %	85 %
	h-CLAT	83 %	71 %	78 %
Combinations (one of two is positive)	DPRA and LuSens	80 %	100 %	85 %
	DPRA and mMUSST	100 %	69 %	81 %
	DPRA and h-CLAT	100 %	71 %	83 %
	LuSens and mMUSST	100 %	67 %	80 %
	LuSens and h-CLAT	79 %	88 %	82 %
Prediction model	DPRA, LuSens and mMUSST	97 %	91 %	94 %

18

BASF
The Chemical Company

Two Out of Three: Additional Data

A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation.

Assays tested: U937/CD86 Test (MUSST), Keratino Sens™ Assay, DPRA vs LLNA, WoE (2 of 3 tests) vs LLNA

	U937- CD86 Test (MUSST) vs LLNA	DPRA vs LLNA	Keratino Sens™ Assay vs LLNA	WoE (2 of 3 tests) vs LLNA
Sensitivity	71	82	79	82
Specificity	70	74	72	77
Accuracy	71	80	77	81
n	141	145	145	145

high sensitivity => minimization of false negatives; high specificity => minimization of false positives

- 43 non-sensitizers according to the LLNA, 33 weak, 39 moderate, 19 strong and 11 extreme sensitizers (Natsch et al., 2013)
- Molecular weight: majority ranged between 100 and 200 Da

19


BASF
The Chemical Company

Two out of Three: Predictivity Based on 54 and 145 Chemicals

Assay		Accuracy 54 chemicals (Bauch et al., 2012) compared to human data	Accuracy 54 chemicals (Bauch et al., 2012) compared to LLNA data	Accuracy 145 chemicals (Natsch et al., 2013) compared to LLNA data
Individual assays	DPRA	87%	79%	80%
	ARE reporter gene assay (Lu Sens or KeratinoSens)	82%	81%	77%
	U937/CD86 Test (MUSST-like Test)	85%	74%	71%
2 of 3	DPRA, ARE-based assay and U937/CD86 Test	94%	83%	81%

- Similar accuracy between both studies despite the extended data set
- Additional data from human studies were not available for all 145 substances; accuracy compared to human data was not determined

20




Two out of Three using „Real Life“ Chemicals

In vitro skin sensitization testing strategy – in-house post-validation with “real-life” compounds
 Sabatini-Galli, Susanne N. Kuhn, Rosalind Homan, Britta Werning, Annette Wenzel, Alexander von Stechow, Robert Landwehr
 10th Int. Conference on Toxicology, Cologne, Germany
 10th Int. Conference on Toxicology, Cologne, Germany
<http://www.woc9.com/poster/101-103>

Guth et al., WOC9, 2014. Poster-ID: IH1-103

- Real-life substances and formulations generally have a lower purity and contain some other byproducts
 - Tests were conducted in parallel to the in vivo tests (cosmetic ingredients prior to March 13, 2013)
 - Plant extracts and formulations were tested using gravimetric approaches instead of MW
- 24 sensitizers, 16 non-sensitizers (either LLNA or GPMT)
 - 7 isocyanates (acylating agents)
 - 5 acrylates (Michael acceptors)
 - 5 agrochemical formulations
 - 3 polyethylene imine polymers
 - 6 surfactants
 - 6 other cosmetic ingredients
 - 7 plant extracts
 - 1 peptide
 - no known pre/pro-haptens

21



Two out of Three using „Real Life“ Chemicals

In vitro skin sensitization testing strategy – in-house post-validation with “real-life” compounds
 Sabatini-Galli, Susanne N. Kuhn, Rosalind Homan, Britta Werning, Annette Wenzel, Alexander von Stechow, Robert Landwehr
 10th Int. Conference on Toxicology, Cologne, Germany
 10th Int. Conference on Toxicology, Cologne, Germany
<http://www.woc9.com/poster/101-103>

Guth et al., WOC9, 2014. Poster-ID: IH1-103

	in-house post-validation						Bauch, 2012			Natsch, 2013
	WoE I	WoE II	WoE I w/o PEI, AF	WoE II w/o PEI, AF	WoE I w/o PEI, AF, PE	WoE II w/o PEI, AF, PE	WoE I	WoE I	LLNA	WoE III
n	38	35	24	21	24	21	50	53	54	145
vs.	LLNA/GPMT	LLNA/GPMT	LLNA/GPMT	LLNA/GPMT	LLNA/GPMT	LLNA/GPMT	human	LLNA	human	LLNA
sensitivity	71	75	88	94	93	93	93	81	96	82
specificity	86	73	85	70	90	86	95	88	81	77
accuracy	76	74	87	85	92	90	94	83	89	81

WoE I: DPRA, LuSens, mMUSST; WoE II: DPRA, LuSens, h-CLAT; WoE III: DPRA, KeratinoSens, MUSST
 AF: agrochemical formulation; PEI: polyethylene imine; PE: plant extract
 high sensitivity => minimization of false negatives; high specificity => minimization of false positives

- The protocols for the test methods are intended for defined substances (e.g. require use of molar equivalents)
- Agrochemical formulations and polyethylene imine based polymers were not well predicted by the in vitro strategy indicating a need to adapt the methods

22

BASF
The Chemical Company

Two out of Three Using an in silico/in vitro Combination

Analysis of in Chemico and in Silico Methods to Predict Skin Sensitization Potentials
Ulrich et al., WCS, 2014, Poster-ID: 01-1103

Reference: LLNA data; Results in [%]	in chemico approach (QSAR Toolbox v3.2)				Ref.: LLNA & Human data; Results in [%]	'2 of 3 WoE' = KeratinoSens, (m)MUSST + ...			
	DPRA	OASIS Profiler for PB	OECD Profiler for PB	Overall in silico Result		DPRA		OECD Toolbox Profilers for PB	
Sensitivity	82	67	65	66	Sensitivity	80	86	79	82
Specificity	72	86	86	90	Specificity	80	96	85	96
Accuracy	79	73	71	73	Accuracy	80	89	81	86

high sensitivity => minimization of false negatives; high specificity => minimization of false positives

- 45 non-sensitizers and 113 sensitizers
- The OECD toolbox 3.2 offers a tool for protein reactivity
- When used alone, the overall accuracy is moderate
- When combined with other in vitro methods with an AOP based rationale (2 out of 3) good accuracies can be achieved

23

BASF
The Chemical Company

Nothing is Perfect – Some Limitations of the Two out of Three Strategy

Substances may be incorrectly predicted if they:

- Have a high cytotoxicity
- Have a low solubility in aqueous media (cell cultures)
- Are not stable at high pH (DPRA)
- Primarily react with lysine and not cysteine
- Are pre- or prohaptens (metabolism not covered in vitro!)

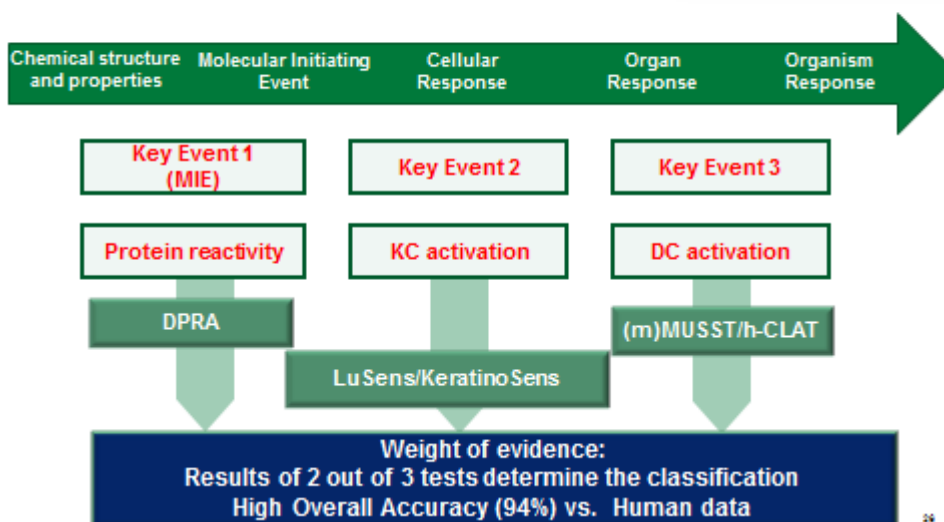
24

BASF The Chemical Company **Summary**

- The AOP for skin sensitization fulfills aspects of semi-quantitative and quantitative AOPs
- Single assays that cover key events of the AOP
 - may be used to prioritize chemicals for testing
 - add confidence to a read-across approach
 - help to group chemicals
- A2 out of 3 in vitro test strategy represents a viable ITS for a “yes or no answer” but is **not yet** applicable
 - to determine potency (e.g. strong, moderate and weak)
 - to sub-categorize according to GHS (e.g. Cat. 1A or 1B)
 - to assess complex mixtures/substances such as polymers and formulations

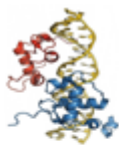
25

BASF The Chemical Company **Proposed test strategy „2 out of 3“ (majority vote, 2012)**



26

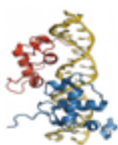
ANNEX 9 PRESENTATION CASE STUDY 3: DEVELOPMENT OF IATAS BASED ON THE ADVERSE OUTCOME PATHWAY (AOP) OF SUSTAINED AHR ACTIVATION LEADING TO RODENT LIVER TUMOR PROMOTION



Development of IATAs Based on the Adverse Outcome Pathway (AOP) of Sustained AHR Activation Leading to Rodent Liver Tumor Promotion

Katy O. Goyak, PhD, DABT
Senior Toxicologist
ExxonMobil Biomedical Sciences, Inc

November 18, 2014



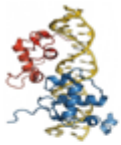
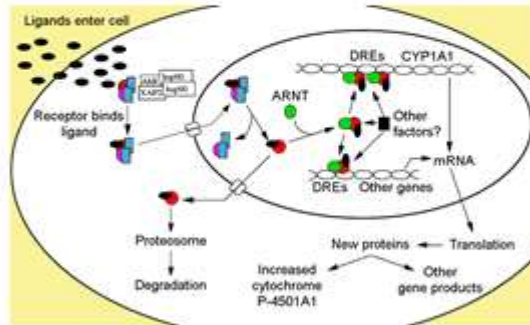
Road Map

- Biology of the Aryl Hydrocarbon Receptor and the Associated Tumor Response
- Description of the AOP
- Expressing the MIE in terms of both Dose and Time, i.e. Area-Under-the-Curve or AUC
- Quantitative Considerations of KE Occurrence and KE Relationships



The Aryl Hydrocarbon Receptor (AHR)

- The AHR is a ligand-activated transcription factor and part of the basic helix-loop-helix (bHLH) Per-Arnt-Sim (PAS) superfamily
- Activated by a variety of exogenous chemicals
 - ▣ Dioxins, PCBs, Dibenzofurans
 - ▣ Other planar polycyclic aromatic hydrocarbons
 - ▣ Natural phytochemicals, flavinoids and indoles
 - ▣ Multiple endogenous ligands proposed, e.g. FICZ
- Regulates a diverse array of genes
 - ▣ Phase I metabolic enzymes (e.g., Cyp1a1, Cyp1a2)
 - ▣ Phase II metabolic enzymes (e.g., Ugt1a2, Gsta1)
 - ▣ Others (e.g., Tiparp, p27Kip1, Bach2)

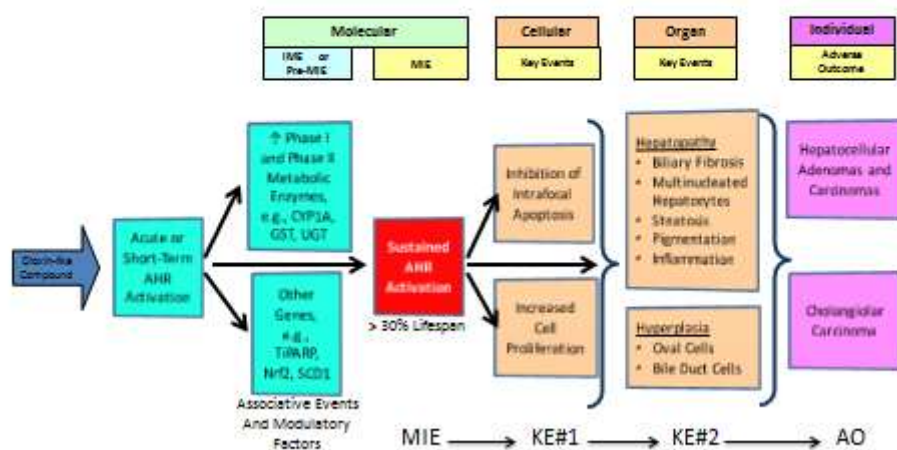


AHR mediated Liver Tumors

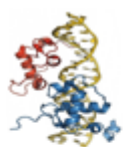
- The NTP cancer bioassay in Sprague-Dawley rats observed increased incidences of several cancers (Walker et al. 2007) including
 - hepatocellular adenoma
 - gingival squamous carcinoma (oral)
 - cholangiocarcinoma
 - cystic keratinizing epithelioma (lung)
- AHR activation is considered to be the initial key event for dioxin-induced tumorigenesis
- However, many ligands can activate the AHR and do not produce tumors, e.g. indole-3-carbinol in broccoli, omeprazole
- Thus acute or short-term AHR activation is the initial key event – can be termed “initial molecular event” or “pre-MIE” but is not the MIE
- Data clearly shows that it is the sustained activation of the AHR that is the MIE



The AOP: Sustained AHR Activation Leading To Rat Liver Tumor Promotion



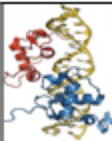
3



The AOP: Sustained AHR Activation Leading To Rat Liver Tumor Promotion

Key Event number	Key Event	Level of biological organization	Testing/Non-Testing approaches
Initial Molecular Event (IME)	Short-term AHR activation	Molecular level	Endogenous gene expression, reporter gene expression, receptor binding assay, (Q)SAR/radioligand
Key Event (KE) 1 (MIE)	Sustained AHR activation	Molecular level	Endogenous gene expression from sub-chronic in vivo studies ¹ Toxicokinetics
KE 2	Changes in Apoptosis, Proliferation and Cellular Homeostasis in vivo (initiation-promotion) and in vivo primary hepatocyte evidence	Cellular response	In vivo primary parenchymal and non-parenchymal cell studies in vivo histopathology and immunohistochemical staining from in vivo repeated dose studies and initiation-promotion studies
KE 3	Hepatopathy Constellation of histopathological changes, hyperplasia ↑fast-labelling, oval cells and bile duct hyperplasia	Organ response	In vivo primary liver cell studies; in vivo histopathology and immunohistochemical staining from in vivo repeated dose studies
Adverse Outcome (AO)	Liver Tumors ↑Hepatocellular adenomas, cholangiomas and cholangiocarcinomas	Organism response	Rat Liver Cancer Bioassay

5



Tailored BH Considerations for Weight of Evidence (WoE) of the Sustained AHR Activation RLTP AOP

1. Biological Plausibility of KERs	The biological plausibility of the overall AOP is High: The AOP is well supported by the key events, consistent with the biology of carcinogenesis and the events of tumor promotion.
2. Essentiality of KEs	The evidence in support of essentiality is High. There is direct evidence consisting of stop/reversibility studies, studies of non-persistent AHR activators and dose-response studies of persistent AHR activators showing that if sustained activation over a substantial portion of the lifespan is not achieved, rat liver tumors are not induced.
3. Empirical Support for KERs	The empirical support for the overall AOP (MIE of sustained AHR activation => promotion of hepatocellular and bile duct cell tumors in rats) is High.

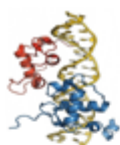
7



Dose - Time Concordance Table for the Sustained AHR Activation RLTP AOP

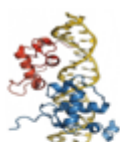
Increasing Dose as average liver conc. in ng/kg	Increasing Time →				
	Immediate	Days to weeks	Months	1 year	2 years
<100	IME	• Apoptosis decrease			
100 - 1000	IME	• Apoptosis decrease	• Apoptosis decrease • Proliferation / AHF volume increase • Proliferation / BrdU labeling		
1000 - 2000	IME	• Apoptosis decrease	• Apoptosis decrease • Proliferation / AHF volume increase • Proliferation / BrdU labeling • Bile duct hyperplasia		
2000 - 5000	IME	• Apoptosis decrease	• Apoptosis decrease • Increased proliferation (BrdU labeling) • Bile duct hyperplasia	• Bile duct hyperplasia	• cholangiocarcinomas
5000 - 10000	IME		• Apoptosis decrease • Proliferation / AHF volume increase • Proliferation / BrdU labeling • Bile duct hyperplasia • Multinucleated hepatocytes	• Proliferation / BrdU labeling • Bile duct hyperplasia • Multinucleated hepatocytes • Diffuse fatty change	• Hepatic adenomas, cholangiomas, cholangiocarcinomas
> 10000	IME	• Apoptosis decrease	• Apoptosis decrease • Proliferation / AHF volume increase • Proliferation / BrdU labeling • Bile duct hyperplasia • Oval cell hyperplasia • Multinucleated hepatocyte • Diffuse fatty change	• Proliferation / BrdU labeling • Bile duct hyperplasia • Multinucleated hepatocytes • Diffuse fatty change	• Hepatic adenomas, cholangiomas, cholangiocarcinomas

8



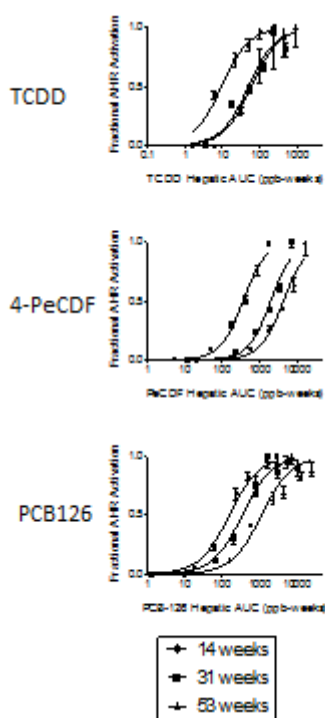
Molecular Initiating Event

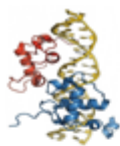
- Sustained AHR Activation
 - Substances that bind to AHR but exhibit rapid clearance (e.g. bergamottin in Earl Gray tea and grapefruit) do not produce rat liver tumors
 - AHR ligands that are poorly metabolized or persistent chemicals (e.g., TCDD) produce rat liver tumors
- We can quantify the MIE as an Area-Under-the Curve (AUC) for AHR activation



AUC Concept

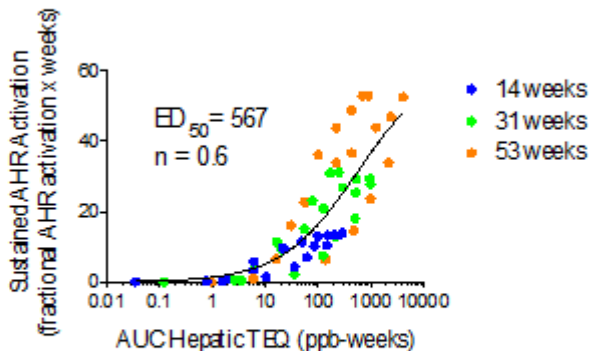
- The dose-response for AHR activation measured by EROD (CYP1A1 induction) using hepatic AUC of the dioxin-like compound (DLC) in ppb-weeks as the dose-term was similar at 14, 31 and 53 weeks in three NTP bioassays for TCDD, 4-PeCDF and PCB126.
- Expressing the response as the fractional AHR Activation (0-1 scale) shows the response is similar over the three time points (to the right).
- These graphs can be combined. The dose term will be the AUC of hepatic TEQ and the response will be sustained activation (SA) as the AUC of fractional AHR activation.





Relating the MIE (Sustained AHR Activation) to AUC for Dose

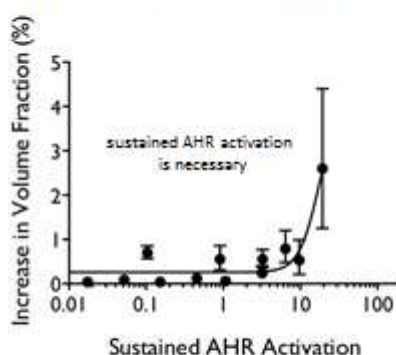
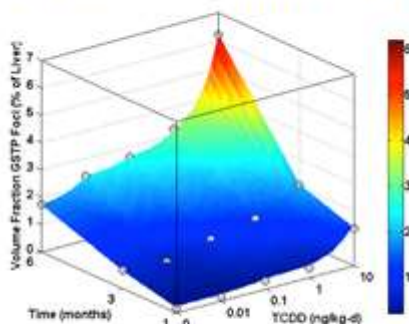
- Sustained Activation (SA) = AHR Activation Level x Time
- Fitting the dose response SA to TEQ Hepatic AUC is consistent with a Hill dose-response model
- The relationship of SA to AUC allows us to examine the “dose-response” of downstream events to SA in a quantitative fashion



11



KER: MIE → KE1 (SA) → Alteration of Cellular Growth Homeostasis



- 3D Dose-time Plot of Volume Fraction Increase of GSTP-positive Foci

- Volume Fraction Increase of ATPase-deficient Foci vs. SA
- ESA_{50} is a measure of the “potency” of the MIE

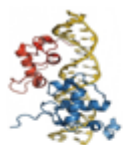
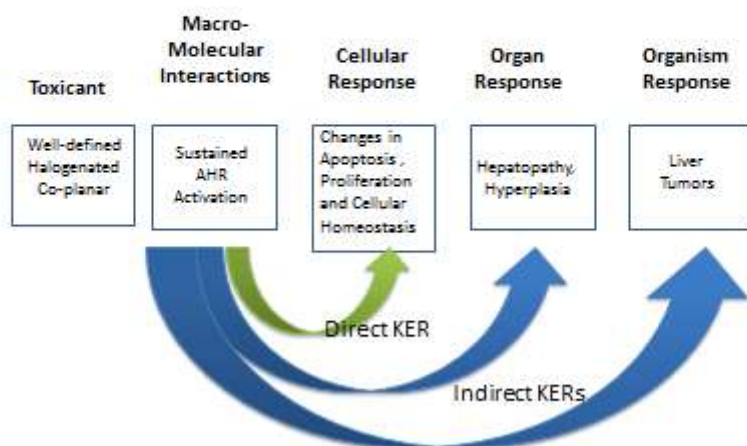
Budinsky et al., 2014, Crit Rev Toxicol 44(1):83

Data from initiation-promotion protocol: Teeguarden et al. 1999, Toxicol Sci 51(2):211

12

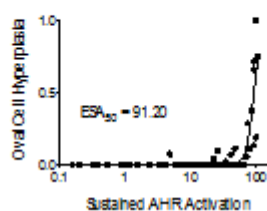
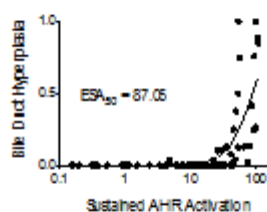
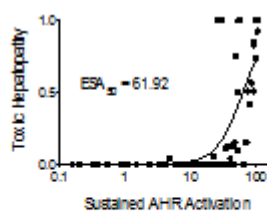


Basic AHR AOP for Rat Liver Tumor Promotion



KER: MIE → KE2, SA
→ Hepatotoxicity,
Hepatopathy

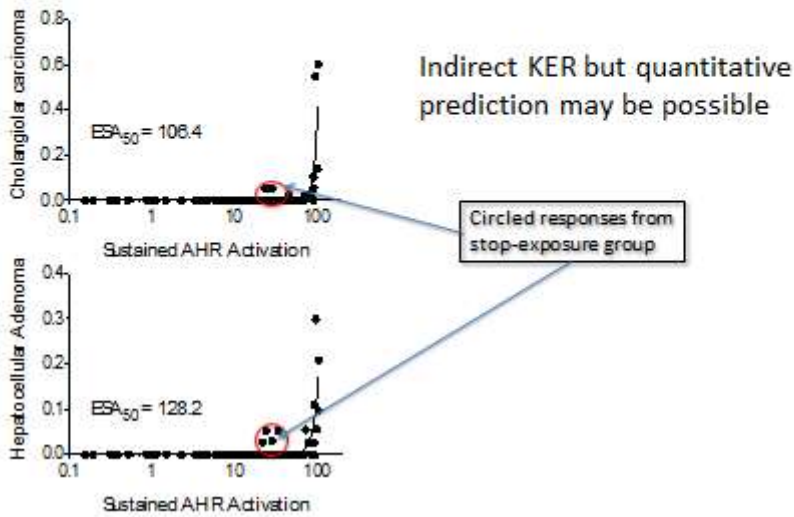
- Indirect KER between MIE and KE2
- Possibility of examining the direct relationship of KE1→KE2→AO because of many initiation-promotion studies for dioxin-like chemicals
- How do changes in cellular growth homeostasis leading to organ-level proliferation and tumors?



14



KER: KE2 → AO, MIE → AO
SA → Hepatotoxicity → Tumor Formation



15

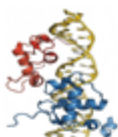


Dose-Temporality Concordance

The potential predictive value of SA emphasizes the importance of the dose-time concordance table (Budinsky et al., 2014, Crit Rev Toxicol 44(1):83; Meek et al., 2014, J Appl Toxicol 34(6):595; Simon et al. 2014, Crit Rev Toxicol 44 Supp 3:17)

Dose (ng/kg)	Temporal										Tumors	
	Key Event 1 (Immediate)	Key Event 2 (Days to Weeks)	Key Event 3 (Months)				Key Event 4 (Months)					
	AHR Activation/ Transcrip- -tion (XME) (1,2,3,4)	↓ Apoptosis (5,6,10)	Proliferation/Hyperplasia		Toxicity							
	AHF Vol. (5,8-12)	BrdU LI (1,7,12)	Bile duct (BDH) (1,11,12)	Oval Cell (OCH) (1)	Multi- nucleate Hepatocytes (MNH) (1,11) (weeks)		Diffuse Fatty Change (DFC) (1) (weeks)		Hepatic Adenoma (1)	Cholangio- carcinoma (1)		
<100		+										
100-1000	++++	+	+	+								
1000-2000	++++	+	++	+	+							
2000-5000	++++	+	+	+	+							
5000-10000	++++		++	+	++		+	+		+		+
>10000	++++	++	+++	++	++	+	+	++	+	+	+	+

16

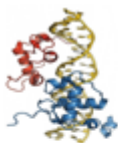


Scientific Confidence Framework for AOPs

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 <i>in vivo</i> 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. Whilst these processes have yet to be defined globally, in time, these should evolve to enable credible and transparent use of AOPs with sufficient scientific confidence by all stakeholders.

Described in Patlewicz et al Using a Scientific Confidence Framework to Support Application of Adverse Outcome Pathways for Regulatory Purposes submitted to Reg Pharm Toxicol

17



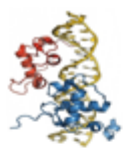
(1) Lessons Learned That Influence Development of IATAs from this AOP

- Receptor binding and acute transcriptional changes represent the Initial Molecular Event (IME or a Pre-MIE) but this may not be predictive of the AO

For IATA

- Assays that just measure AHR binding will have limited utility
 - May be used in decision tree as initial step to differentiate binders from non-binders
 - Cannot be used to predict other KEs or AO

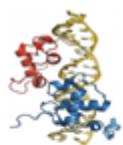
18



(2) Lessons Learned That Influence Development of IATAs from this AOP

- The Molecular Initiating Event is sustained AHR activation
 - Quantitative dose-response relationship of AUC (hepatic TEQ) to AHR Sustained Activation (SA) opportunity for inclusion of relevant assays in an IATA
 - The relationship of SA to 1) induction of hepatic foci, 2) inhibition of intrafocal apoptosis and 3) production of proliferative stimulus in the liver are all potential endpoints for assays in an IATA
 - short term *in vivo* rat liver initiation-promotion assay (e.g. <http://toxsci.oxfordjournals.org/content/16/3/525.full.pdf>)

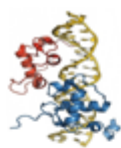
19



(3) Lessons Learned That Influence Development of IATAs from this AOP

- The Indirect KER of:
 - SA to hepatopathy
 - SA to bile duct hyperplasia and
 - SA to oval cell proliferation
- Hepatopathy, bile duct hyperplasia & oval cell proliferation may also all be potential endpoints for assays in an IATA
 - But these are *in vivo* responses that require considerable treatment time to be manifested

20



(4) Lessons Learned That Influence Development of IATAs from this AOP

- Evaluating activity in assays (e.g., AC50) in the context of human exposure improves interpretation and thus should be considered as an integral part of IATAs

– Examples of exposure:activity profiling:

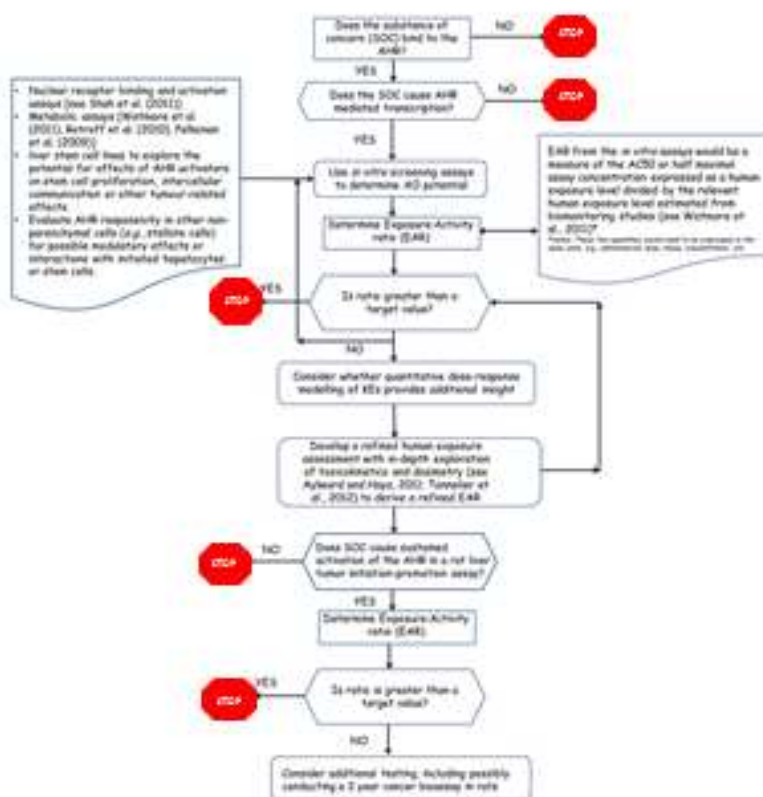
Becker et al. 2014. Interpreting Estrogen Screening Assays in the Context of Potency and Human Exposure Relative to Natural Exposures to Phytoestrogens. Birth. Defects. Res. B. Dev. Reprod. Toxicol. Birth 101:114-24

Wetmore et al. 2012. Integration of Dosimetry, Exposure and High-Throughput Screening Data in Chemical Toxicity Assessment. Toxicol. Sci. 125, 157-74

21



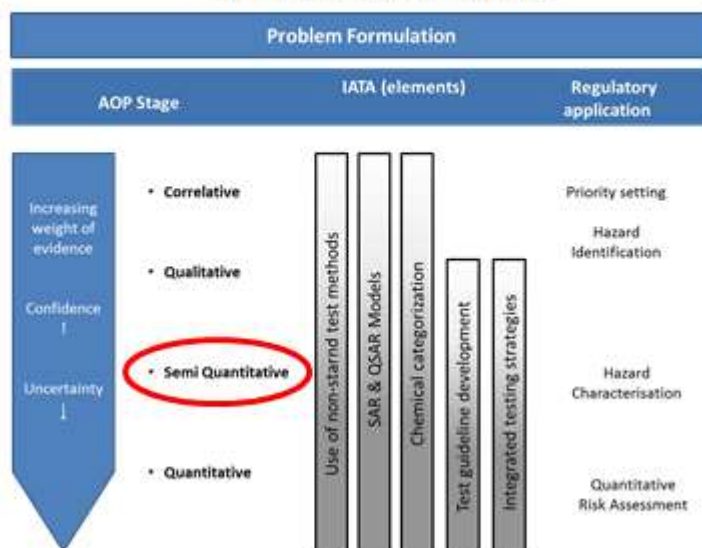
IATA for Sustained AHR activation



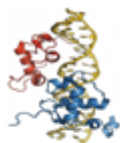
22



Relationship between different stages of AOP/MoA development and the informed types of IATA and its elements



23



Semi-Quantitative AOP

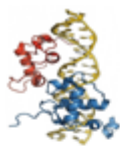
Has semi-quantitative understanding of the AOP- assessment of the experimental evidence and empirical data across the key events based on multiple compounds and/or stressors studied at the key events. The critical (early and late) key events are identified

Moderate to strong level of confidence in the AOP:

- Moderate to Strong support for the biological plausibility of some of the KERs (Good understanding of the KER based on previous documentation and broad acceptance; established mechanistic basis (e.g., mutation leading to tumours.)
- Moderate to strong support for the Essentiality of KE (Direct evidence from specifically designed experimental studies illustrating essentiality for at least one of the important key events)
- Moderate Empirical support for the KERs (Demonstrated dependent change in both events following exposure to a multiple number of specific stressors and some evidence inconsistent with expected pattern which can be explained by factors such as experimental design, technical considerations, differences among laboratories, etc.)
- Establishes quantitative understanding of some of the key events.

Not able to determining the response-to-response relationships required to scale in vitro effect to in vivo outcome.

24





AHR AOP for Rat Liver Tumor Promotion

Semi-Quantitative

Moderate to strong level of confidence in the AOP:

- Moderate to Strong support for the biological plausibility of some of the KERs
 - Moderate to strong support for the Essentiality of KE
 - Moderate Empirical support for the KERs
 - Establishes quantitative understanding of some of the key events
 - Not able to determining the response-to-response relationships required to scale in vitro effect to in vivo outcome
- Strong support for KERs (MIE-KE1 or KE2 and KE1-AO)
 - Moderate support for KER (KE2-AO)
 - Strong support for essentiality of KEs
 - Quantitative prediction potentially possible based on the Indirect KERs KE2-AO, MIE-AO
 - Exposure-Activity ratio calculations to facilitate risk assessment
 - SA (MIE) – AUC allows a Dose-response to be examined²⁵

ANNEX 10 PRESENTATION CASE STUDY 4: AROMATASE INHIBITION LEADING TO REPRODUCTIVE DYSFUNCTION (IN FISH). A QUANTITATIVE AOP CASE STUDY



[Aromatase inhibition leading to reproductive dysfunction \(in fish\)](#)
A quantitative AOP case study

Rory B. Conolly¹, Karen H. Watanabe², Wan-Yun Cheng¹, Edward J. Perkins³, Michael Mayo³, David H. Miller⁴, Daniel L. Villeneuve⁴

¹US EPA, Integrated Systems Toxicology Division, Research Triangle Park, NC, USA
²Oregon Health and Science Univ., Institute of Environmental Health, Portland, OR, USA
³US Army Engineer Research and Development Center, Vicksburg, MS, USA
⁴US EPA Mid-Continent Ecology Division, Duluth, MN, USA

** The contents of this presentation neither constitute nor, necessarily, reflect US EPA views or policies.*

Outline

- Background and purpose for AOP development
- Formal AOP description and weight of evidence evaluation
- A prototype quantitative AOP (Q-AOP)
- Considerations for Q-AOP development
- Q-AOP application(s)



Background

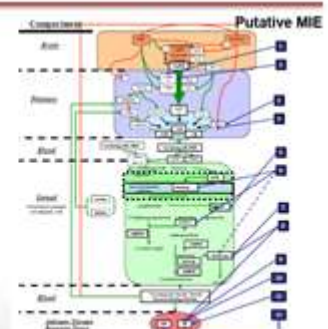
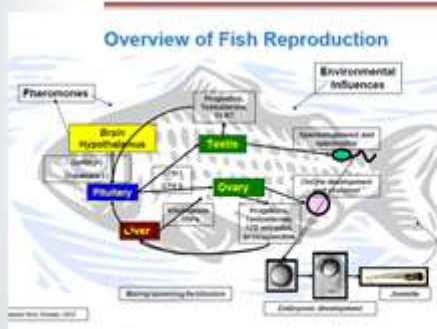
- Current US EPA, EDSP, Tier 1, includes a fish short term (21d) reproduction assay (OECD 229; OCSPP 890.1350)
- Screen for (anti)estrogens, (anti)androgens, steroidogenesis inhibitors
- Three week in vivo reproduction assay + 2 week acclimation (typically)
- Improving efficiency, reducing cost, and animal use would be desirable

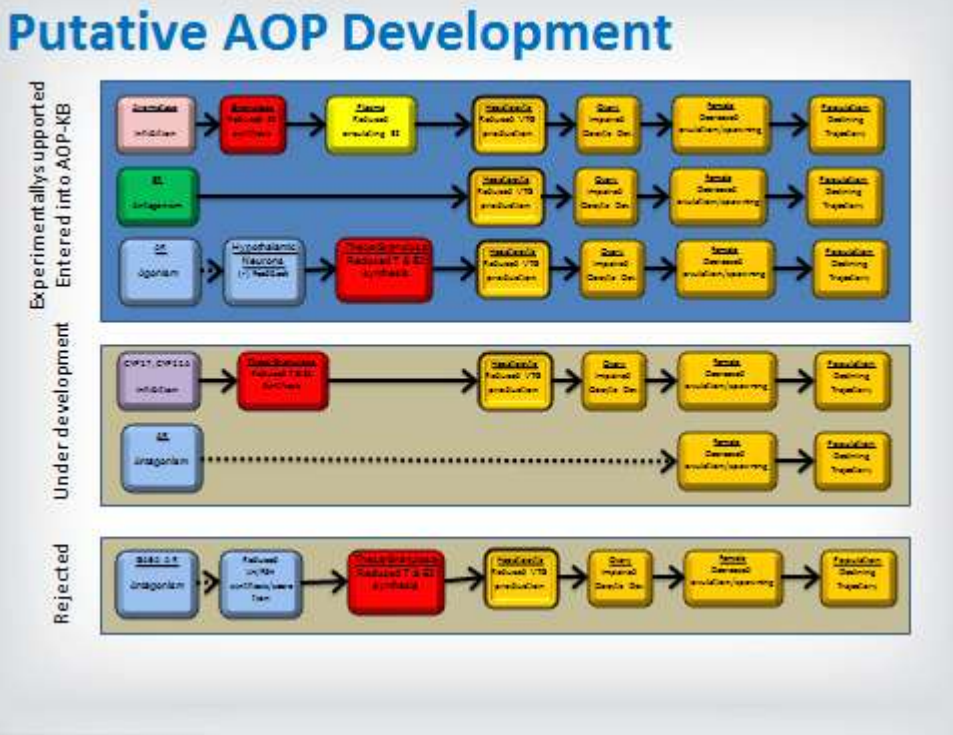


Background

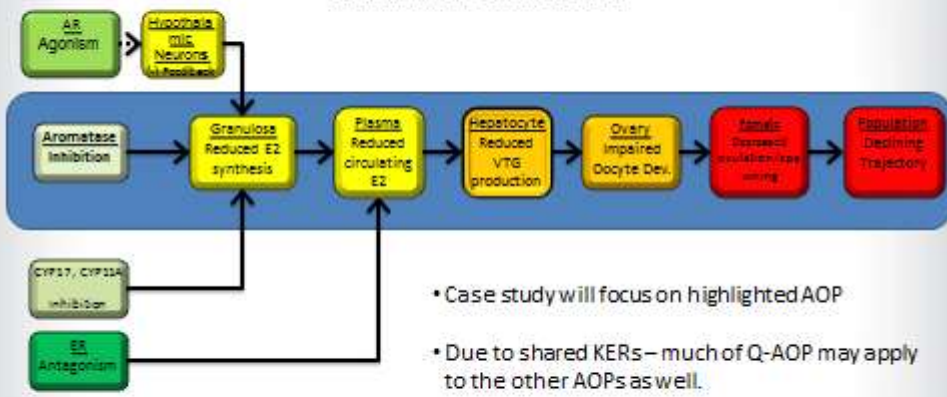


- Identify alternatives to EDSP Tier 1 fish short term reproduction assay (OECD 229)
- Support the use of high throughput in vitro as an alternative to FSTRA.






- Distinct MIEs, but shared KEs and KERs
- Form an AOP Network



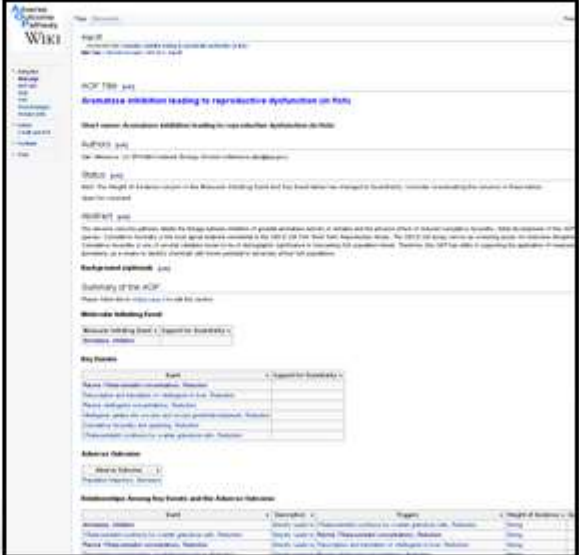
- Case study will focus on highlighted AOP
- Due to shared KERs – much of Q-AOP may apply to the other AOPs as well.

EPA Formal AOP Description & Weight of Evidence Evaluation



Aopkb.org

Users' handbook supplement to OECD guidance document for developing and assessing AOPs.

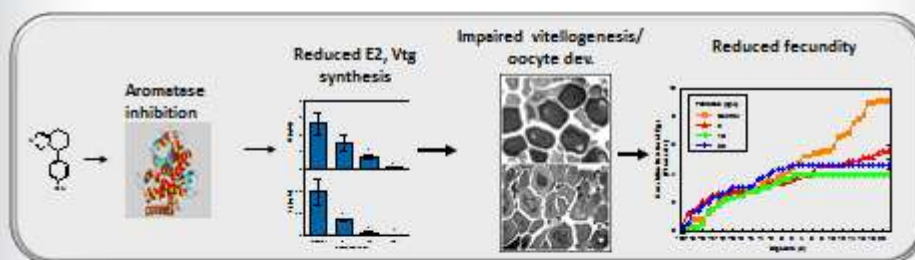


Formal AOP Development –Key Events

Section 5– Summary of the AOP & Key Event Descriptions



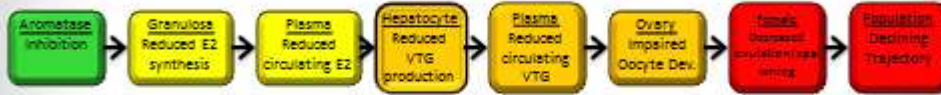
•AOP consists of 8 KE, 7 KER



- Each key event is observable
- We can describe its normal role in biological context
- We can describe how it can be measured directly or indirectly

Formal AOP Development-Key Events

Taxonomic Relevance



Aromatase first appears in common ancestor to amphioxus and vertebrates (Baker 2011).



MIE – likely applicable to most vertebrates

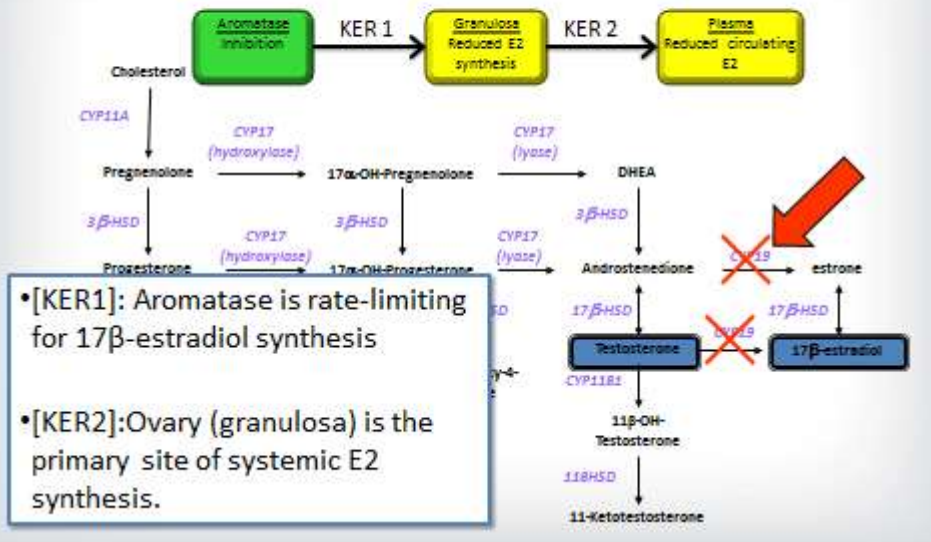
Vitellogenesis-related key events likely applicable to oviparous vertebrates

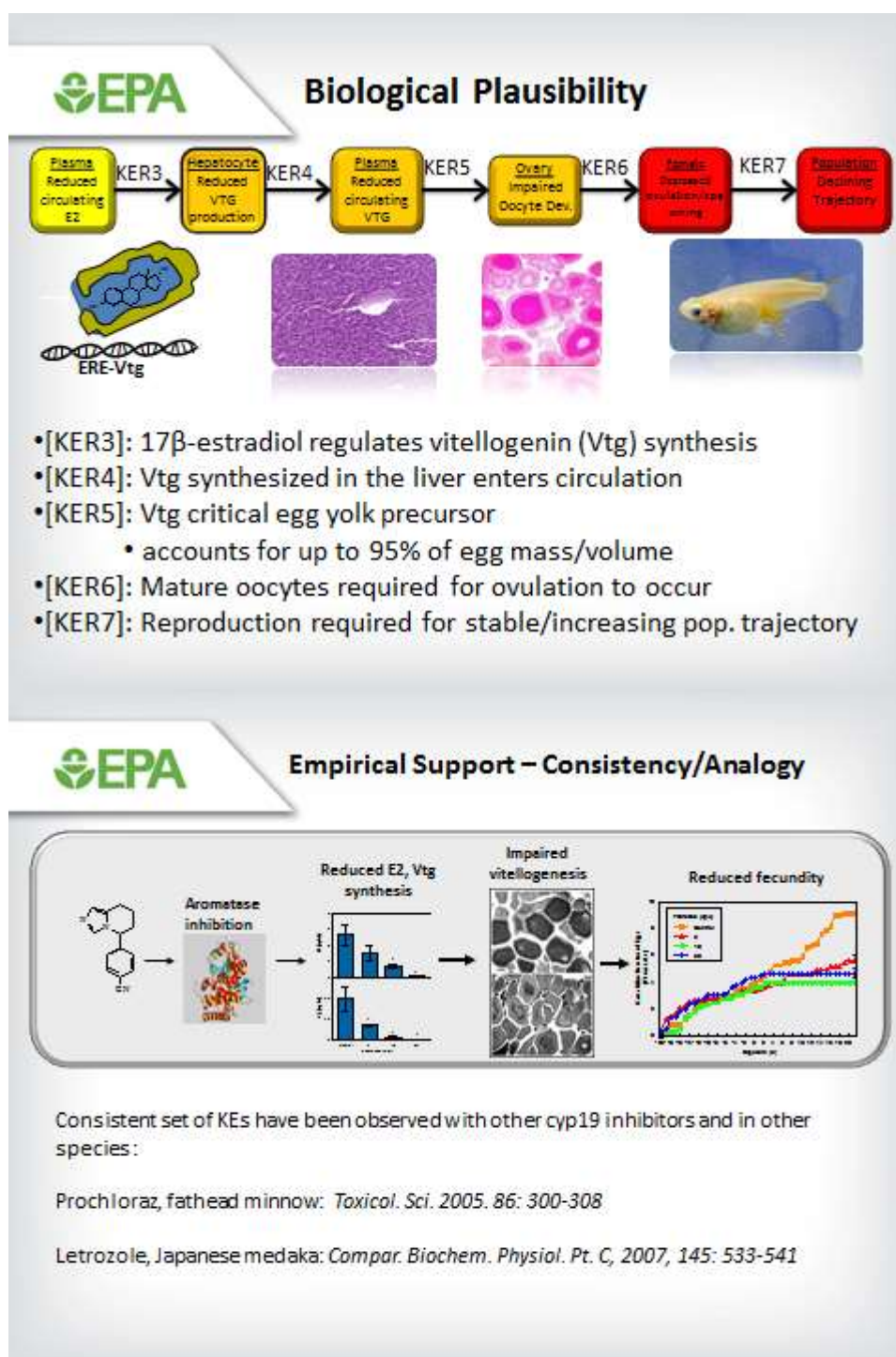


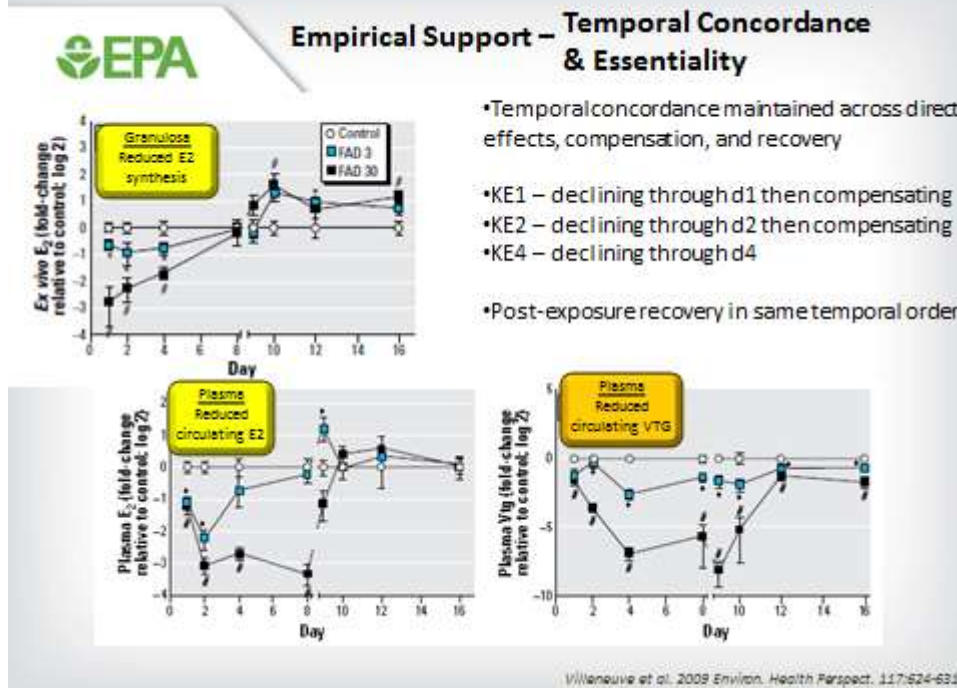
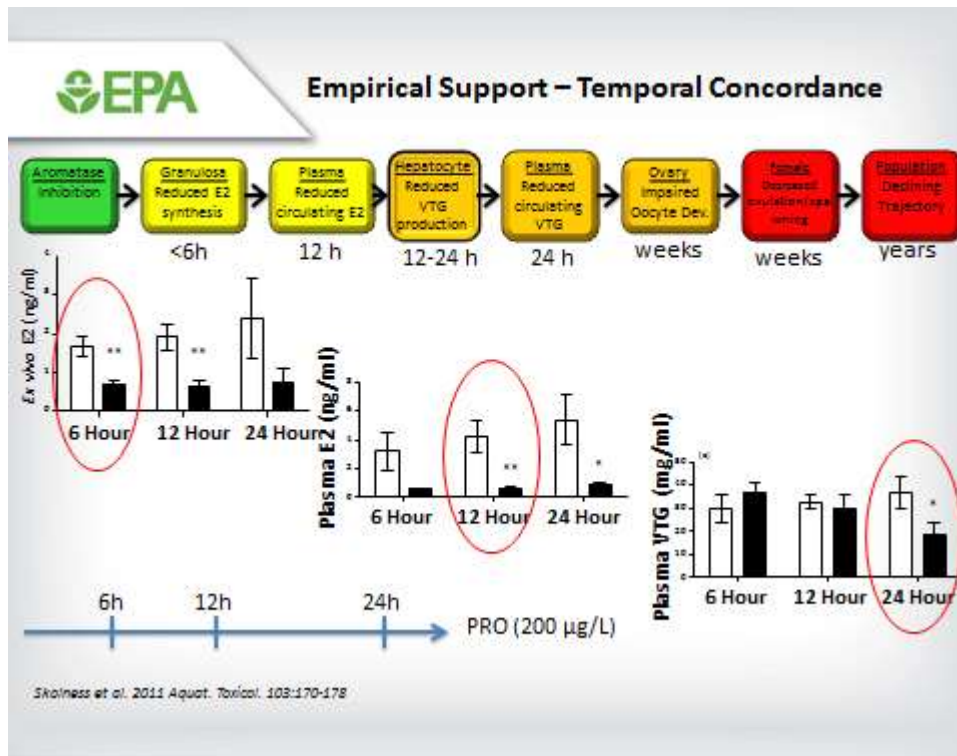
Baker ME. 2011. Molec. Cell. Endocrinol. 334:14-20



Biological Plausibility









Qualitative Confidence in AOP


Upstream Event	Description	Downstream Event	Weight of Evidence
Aromatase, Inhibition	Directly Leads to	17beta-estradiol synthesis by ovarian granulosa cells, Reduction	Strong
17beta-estradiol synthesis by ovarian granulosa cells, Reduction	Directly Leads to	Plasma 17beta-estradiol concentrations, Reduction	Strong
Plasma 17beta-estradiol concentrations, Reduction	Directly Leads to	Transcription and translation of vitellogenin in liver, Reduction	Strong
Transcription and translation of vitellogenin in liver, Reduction	Directly Leads to	Plasma vitellogenin concentrations, Reduction	Strong
Plasma vitellogenin concentrations, Reduction	Directly Leads to	Vitellogenin uptake into oocytes and oocyte growth/development, Reduction	Moderate
Vitellogenin uptake into oocytes and oocyte growth/development, Reduction	Directly Leads to	Cumulative fecundity and spawning, Reduction	Moderate
Cumulative fecundity and spawning, Reduction	Directly Leads to	Population trajectory, Decrease	Moderate

Overall, based on plausibility, essentiality, and empirical support we have **strong** confidence in the qualitative relationships depicted in the AOP.

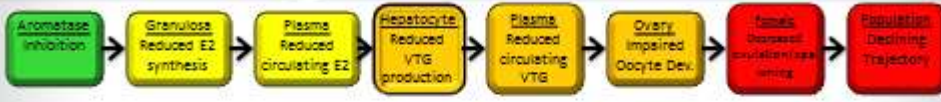


Qualitative Confidence in AOP


- Structural alerts and in vitro measures indicative of aromatase inhibition provide strong evidence of potential hazard as a reproductive toxicant (in fish).
 - Viable alternatives to FSTRA as a Tier 1 screen
- We have strong scientific confidence that in vivo observation of sustained reductions in ovarian aromatase activity, circulating E2, and circulating VTG are indicative of probably reproductive hazard (in fish).
 - Suitable as confirmatory endpoints that account for ADME
- AOP can support the development of a tiered or sequential testing strategy for aromatase inhibitors:
 - In vitro screening
 - Short-term in vivo (e.g., 24h) – focused on KE 1-5
 - Long-term in vivo (FSTRA) – focused on KE 1-7



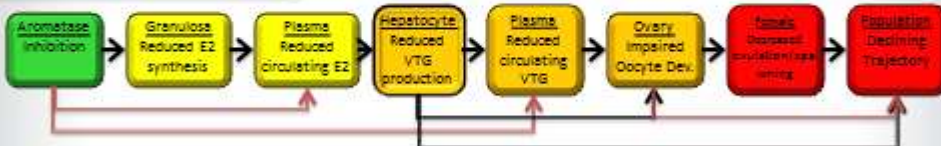
Quantitative Understanding



- Can in vivo measures of KE 1-5 be used to predict probability or severity of AO (KE 7, 8)
- Can the dose-response time-course behaviors of KE 1-5 be predicted in a manner that informs the design of confirmatory short-term in vivo tests?
- Can in silico structural alerts or in vitro surrogate measures of KE 1,2,4 be used to predict the probability or severity of the AO (KE 7, 8)?



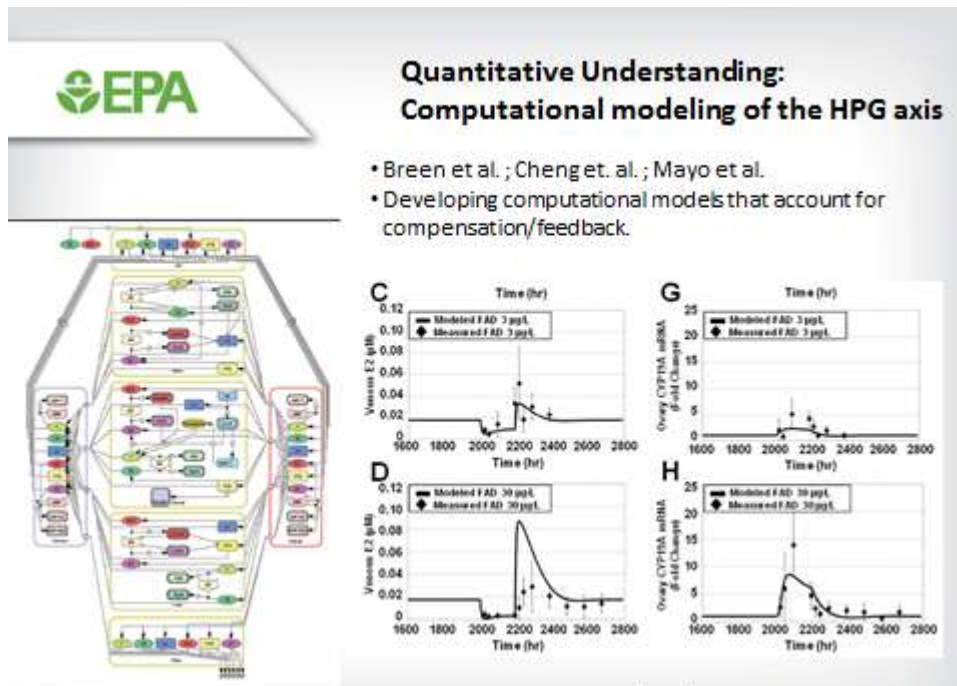
Quantitative Understanding



Computational Model	Input KE(s)	Output KE(s)	Type ¹
Breen et al. 2013	Aromatase inhibition	Circulating E2	BB
Cheng et al. (in development)	Aromatase inhibition	Circulating E2 and VTG	BB
Mayo et al. (in development)*	Aromatase inhibition	Circulating E2 and VTG	BB
Li et al. 2011	Plasma VTG	Cumulative fecundity and spawning	S/C
Miller and Ankley 2004	Cumulative fecundity	Population trajectory	BB
Miller et al. 2007	Plasma VTG	Population trajectory	BB/S/C

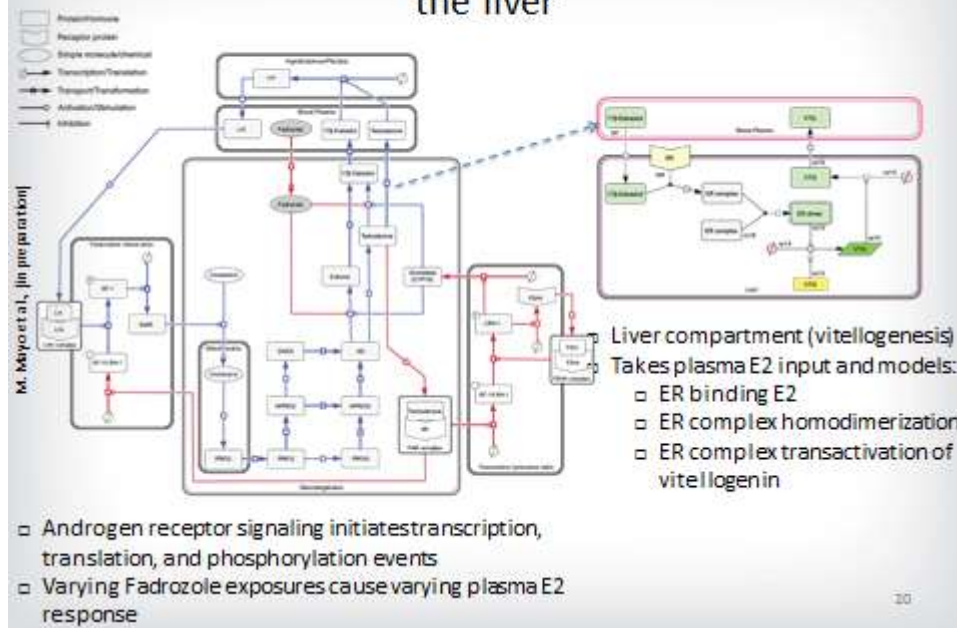
*employs a different model construct than Cheng/Breen

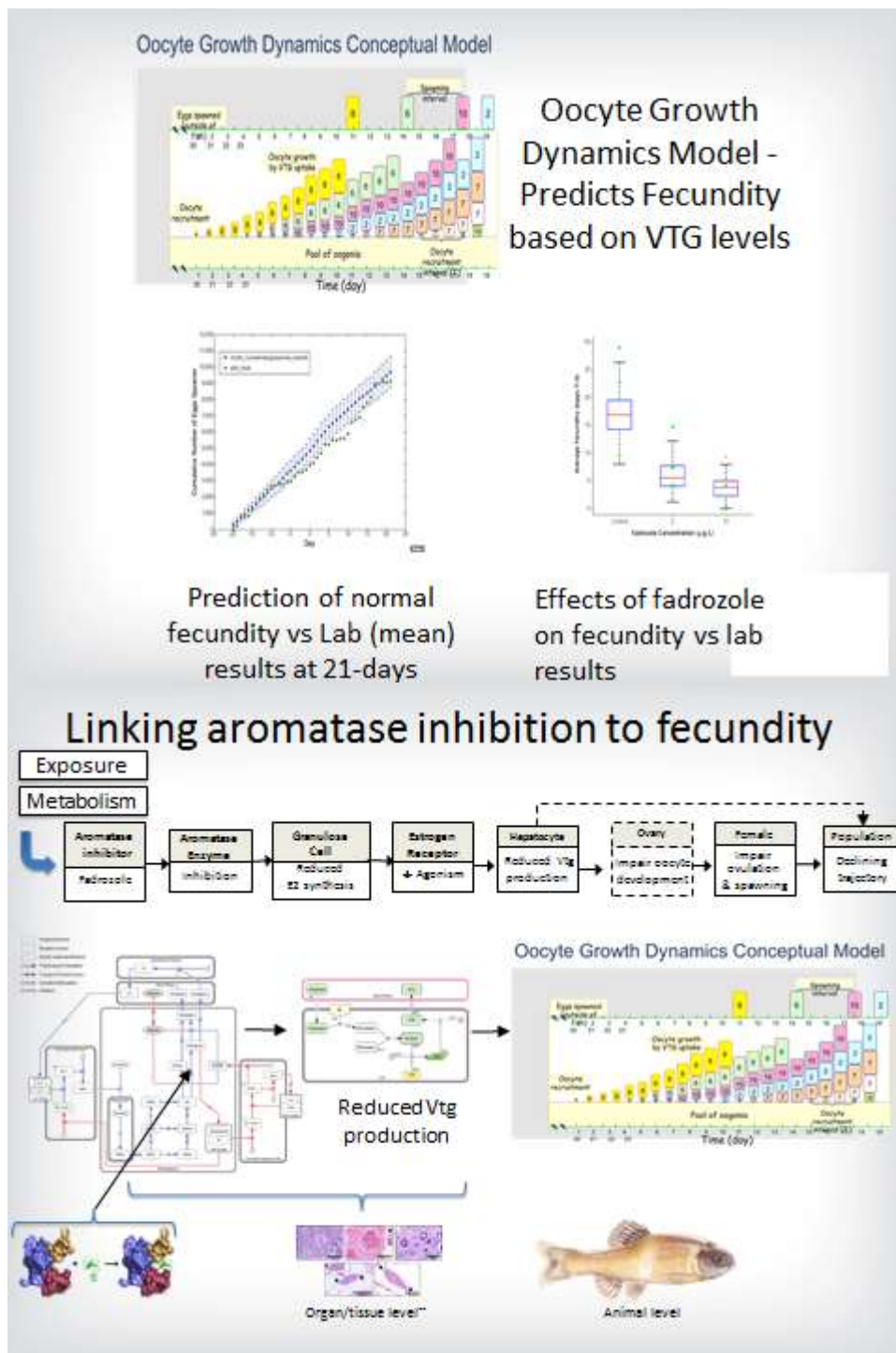
¹ BB = biologically-based (mechanistic); S/C = statistical/correlation-based

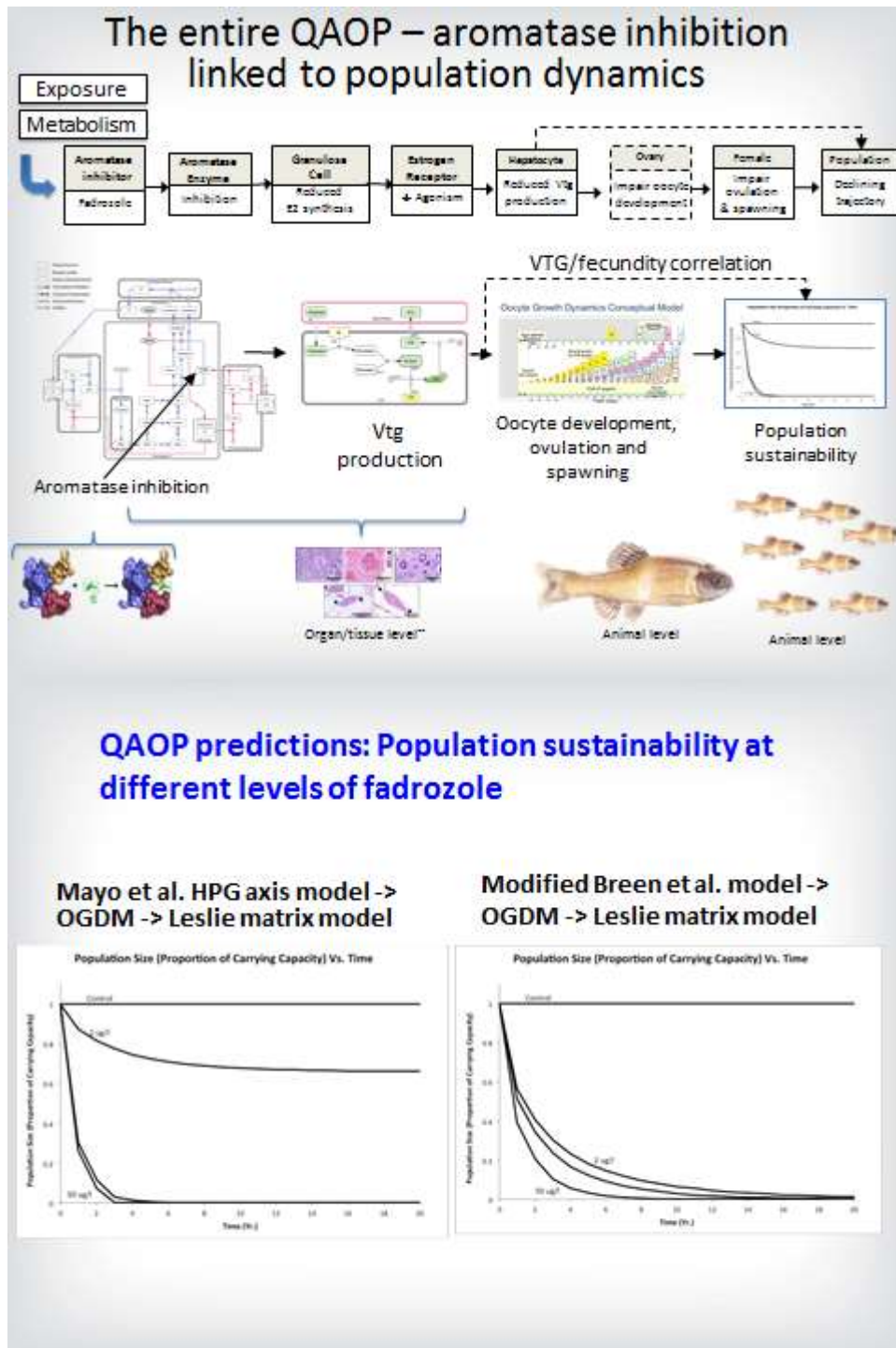


Breen et al. 2013 Toxicol. Sci. 133: 234-237

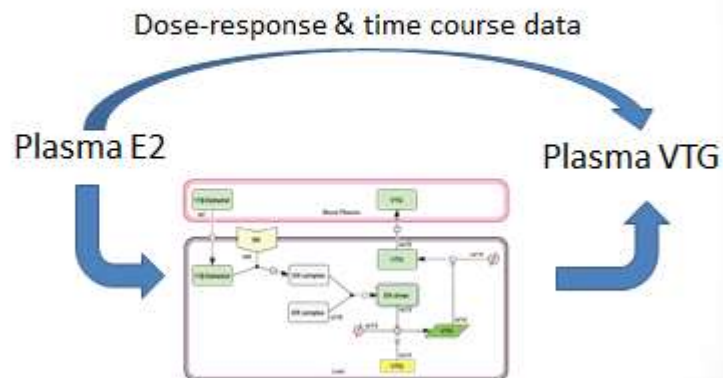
Linkage of HPG axis to vitellogenin synthesis in the liver







QAOP and key event relationships



Binary and higher-order mixtures of aromatase inhibitors

- Use “Toxicity Equivalent” approach with data from the EPA’s ToxCast® database.
- Specify the potency of other aromatase inhibitors relative to fadrozole.
 - chemical X is 50% as potent as fadrozole
 - Y molar X + Y molar fadrozole = 1.5*Y molar fadrozole
 - Assumes the individual dose-response curves are parallel.
- Computational modeling and lab work in progress!

Return on the investment

- A fully developed QAOP is powerful predictive tool.
 - Input exposure scenario of interest
 - Output prediction of change in adverse outcome
- But data needs are large
 - Expensive and time consuming
- Mature QAOP could serve as an “in silico” description of in vivo biology to aid in design of in vitro tests and interpretation of in vitro data
 - IVIVE

Confidence in the QAOP

- Evaluation of confidence in predictions provided by the QAOP will be a critical step.
 - Handbook contributing to consistency in considerations
- Regulatory decision-makers are likely to use the decision-support tools that are most reliable, least uncertain.
 - Degree of confidence application-dependent
- Need to plan for evaluation of not only of confidence in the QAOP, but of the confidence in the QAOP relative to that for the decision-support tools used in the absence of the QAOP.

Research Needs for Development of qAOPs

Computational modeling

- Experimental studies needed to establish better links between existing mechanistic/computational models at different levels of biological organization
- Determine the feasibility of developing a single computational model to represent a qAOP (from MIE to AO) versus linking existing models
- Development of a seamless modeling framework to facilitate implementation of qAOPs in risk assessment

**ANNEX 11 PRESENTATION OUTCOMES FROM THE SOMMA LOMBARDO WORKSHOP:
ADVANCING ADVERSE OUTCOME PATHWAYS FOR INTEGRATED TOXICOLOGY AND
REGULATORY APPLICATIONS**

**Outcomes from the Somma Lombardo
Workshop: *Advancing Adverse Outcome
Pathways for Integrated Toxicology and
Regulatory Applications***

Ed Perkins, US Army Corps Engineers

Chair: Natàlia Garcia-Reyero, Mississippi State University

Organizers: Rick Becker¹, Natàlia Garcia-Reyero², Ksenia Groh³,
Marlies Halder⁴, Sean Kennedy⁵, Teresa Lettieri⁴, Edward J
Perkins⁸, Knut Erik Tollefsen⁶, Bart Van der Burg⁷, Dan Villeneuve⁸,
Maurice Whelan⁴


US Army Corps Engineers Engineer Research and Development Center, Vicksburg,
MS, USA; ¹ American Chemical Society, USA; ² Mississippi State University, USA; ³
EAWAG, Switzerland; ⁴ Joint Research Center, Italy; ⁵ Environment Canada, Canada;
⁶ NIVA, Norway; ⁷ BioDetection Systems, The Netherlands;
⁸ Environmental Protection Agency, USA.

Purpose of the Workshop

Provide feedback, consensus opinion, and
recommendations concerning the practical
implementation of AOPs and the AOP Work
Process to both the OECD and to scientists and
regulators.

With the goal of....

***Advancing AOPs for Integrated Toxicology and
Regulatory Applications***



OECD
*Extended Advisory Group on Molecular Screening and
Toxicogenomics*

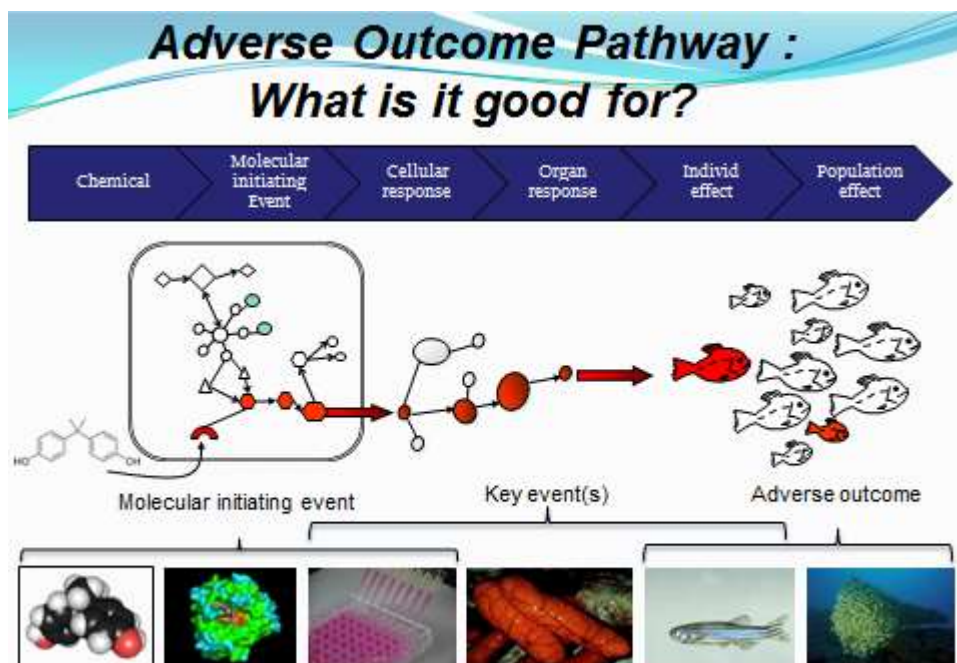
**WORKPLAN FOR DEVELOPMENT, ASSESSMENT, AND
USE OF ADVERSE OUTCOME PATHWAYS (AOPs)**

- **Inform Test Guidelines development**
- **Permit development of Integrated Approaches to Testing and Assessment (hypothesis-driven testing)**
- **Support the use of the QSAR for grouping chemicals**

Participants



- **Over 50 participants from 11 different countries**
 - Industry, academia, government, non-profit
- **Different perspectives/disciplines working together**
 - Human health & Eco: Together



Overview of OECD Work Process for developing AOPs

1. Proposal by stakeholder to develop AOP
2. Development of an AOP in the AOP-KB
3. Review by OECD expert Groups
4. Approval by Sub-bodies of the JM declassification and publication



Courtesy of Anne Gourmelon, OECD

Alignment of workgroup charges with OECD work process

1. Proposal by stakeholder to develop AOP	➔	WG1: Priorities for AOP development
2. Development of an AOP in the AOP-KB	➔	WG2: Strategic approaches to AOP development
3. Review by OECD expert Groups	➔	WG3: WOE evaluation of an AOP
4. Approval by Sub-bodies of the JM declassification and publication.	➔	WG4: Review and "acceptance" for regulatory application
5. Practical application in Testing and assessment	➔	WG5: Application to IATA – hypothesis driven testing

development in Ecotoxicology

Objectives

Identify priority areas for AOP development in ecotoxicology

Background:

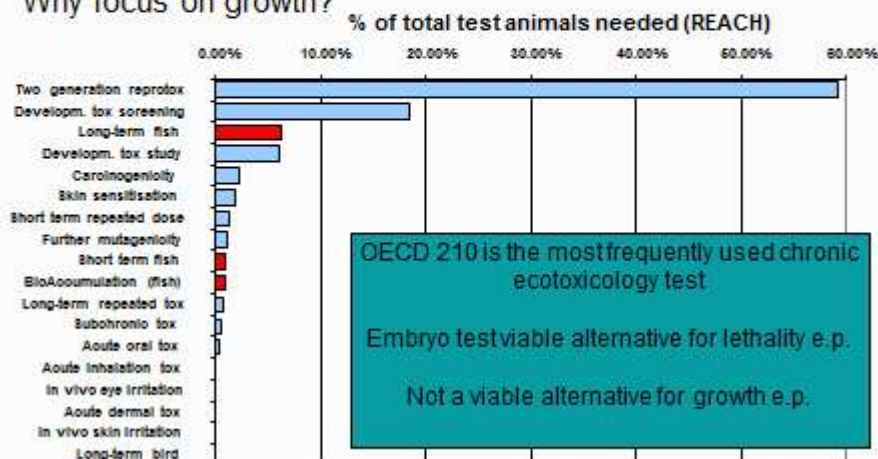
- A number of sublethal effects on organisms have high significance for potential population-level effects but are not efficiently characterized using current testing guidelines and strategies.
- E.g., effects on behavior, growth, immune function, etc.

Case studies

Growth impairment as an outcome of chronic toxicity in fish

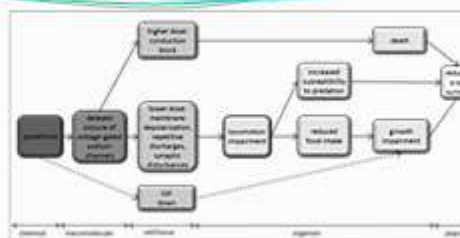
WG1: Priorities for AOP development in Ecotoxicology

Why focus on growth?



•Developed a conceptual model of growth regulation in fish

- E.g., role of growth hormone axis
- Nutrition
- Balancing energy allocation: survival, growth, reprod.



•Used conceptual model to identify relevant key events to use as a nucleus for AOP development.

•Prioritize research needs to fill gaps in AOPs related to the case study.



Groh et al 2014. Development and application of the adverse outcome pathway framework for understanding and predicting chronic toxicity: I. Challenges and research needs in ecotoxicology. Chemosphere. In press.
 Groh et al 2014. Development and application of the adverse outcome pathway framework for understanding and predicting chronic toxicity: II. A focus on growth impairment in fish. Chemosphere. In press.

WG2: Strategic approaches to AOP development

AOP development involves identifying:

- relevant molecular initiating event(s)
- key events
- adverse outcome(s)
- establishing the biological plausibility and evidence that define and support the predictive relationships between those events.

OECD guidance on developing and assessing adverse outcome pathways defines the type of information that should be included in an AOP description.

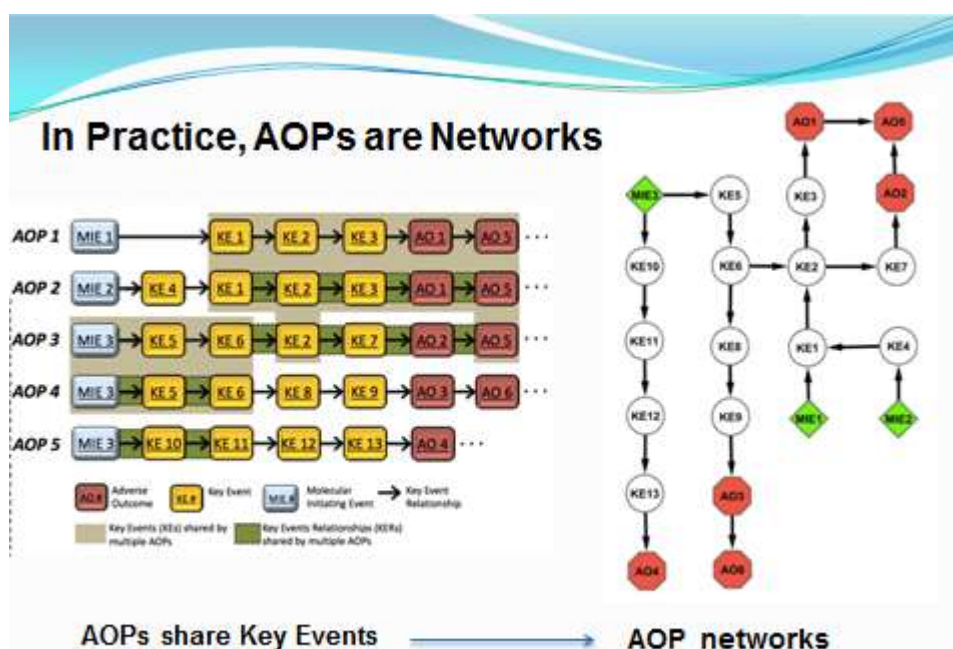
Challenge: common questions and challenges encountered by new AOP developers and practical strategies and best practices have been lacking.

WG2: Strategic approaches to AOP development

Core Principles of AOP Development

1. AOPs are not chemical-specific
2. AOPs are modular
 - Key Events – functional unit of observation – nodes
 - Key Event Relationships – functional unit of inference/extrapolation – edges
3. AOPs (linear) are a pragmatic functional unit of development and evaluation.
 - For a theoretical “pure ligand” they are a functional unit of prediction
4. For most real-world scenarios, AOP networks are the functional unit of prediction.
5. AOPs are living documents

Villeneuve et al. Adverse Outcome Pathway (AOP) Development I: Strategies and Principles. *Toxicol.* in press.



Villeneuve et al (submitted) Adverse Outcome Pathway (AOP) Development II: Best Practices

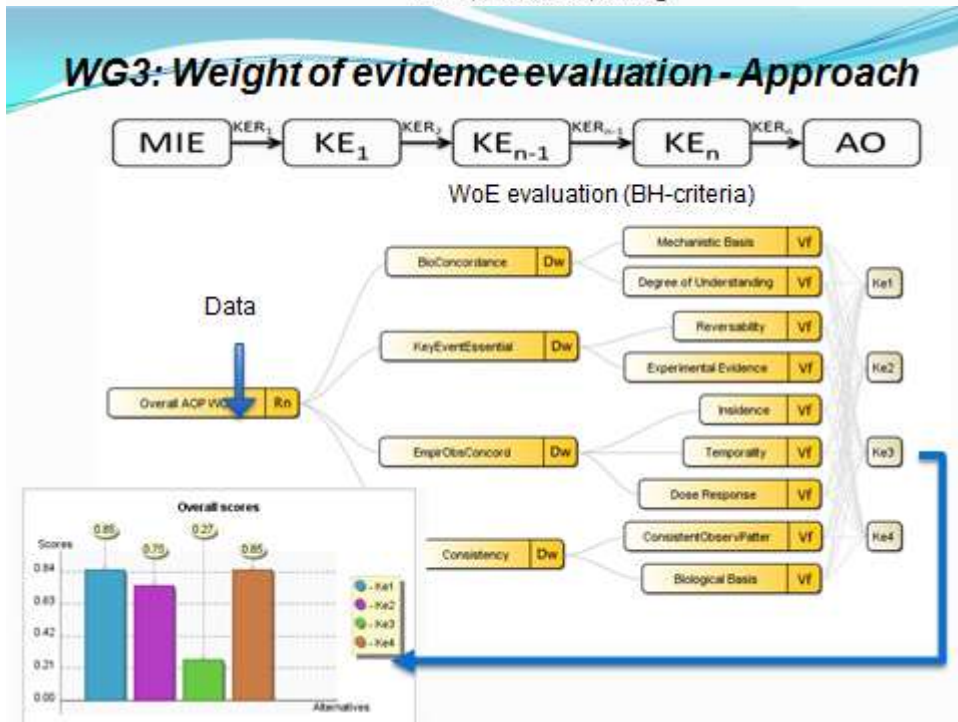
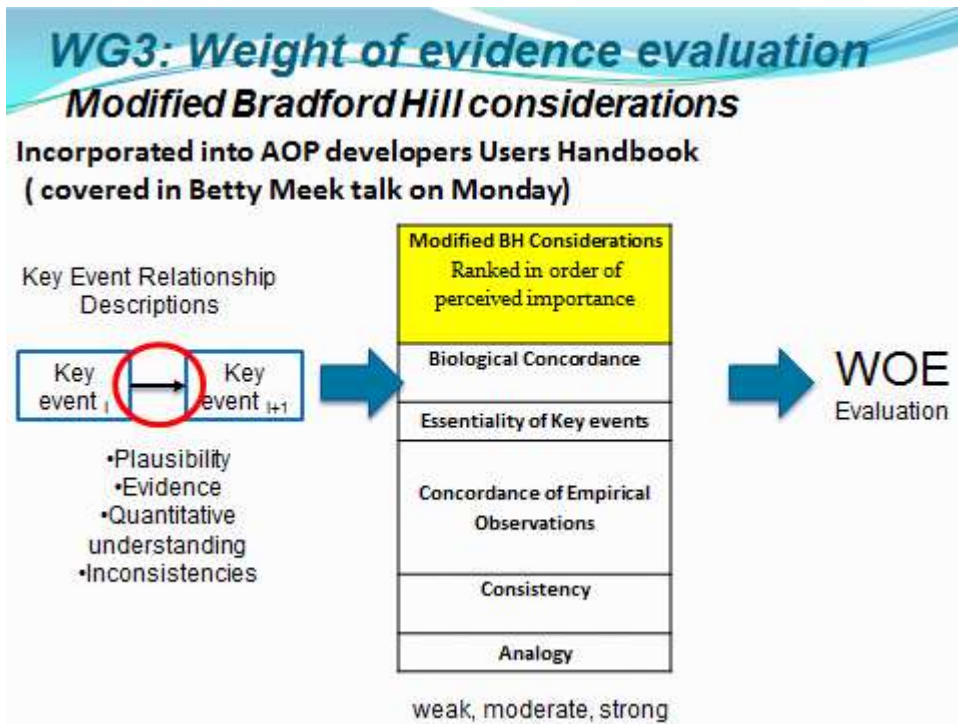
WG3: Weight of evidence evaluation of the quality of an AOP

Challenge: Data populating AOPs can be relied upon to different extents. Consistent use of a standardised approach to Evaluation the Evidence is key to successful use of AOP.

Objectives: Assess the suitability of (modified) Bradford Hill considerations for WoE evaluations related to AOP development and provide guidance/documentation to use in regulatory applications.

Case studies: Endocrine disruption (Estrogen/Androgen), Sustained AhR activation leading to toxicity...and more.

Becker et al. Increasing Scientific Confidence in AOPs: Application of Tailored Bradford Hill Considerations for Evaluating Weight of Evidence. In preparation.



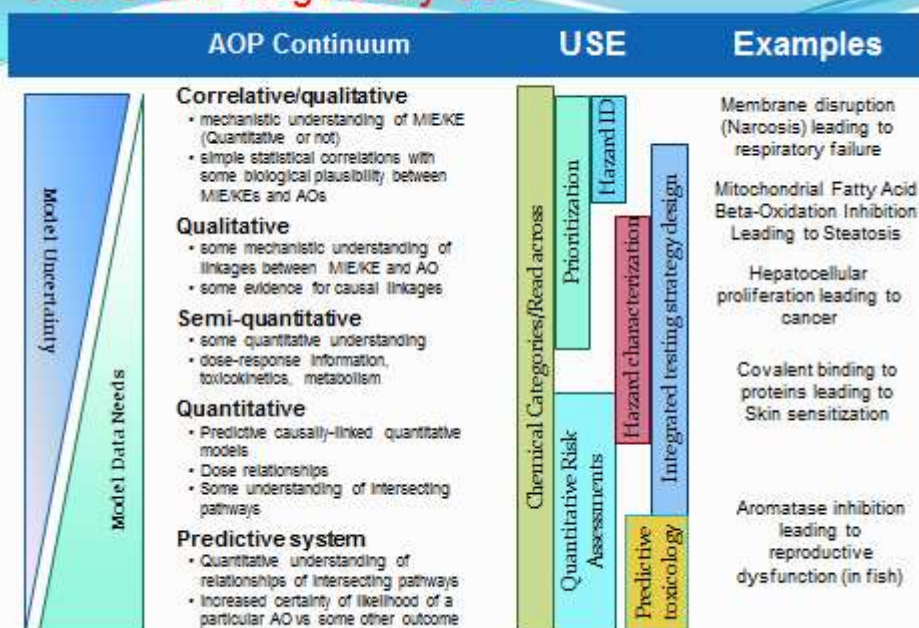
WG4: Using AOPs for regulatory applications

Challenges: AOPs offer a scientifically-credible foundation for hazard assessment and regulatory decision-making, but a roadmap/guide for practical use has not been established.

Objectives: develop a conceptual approach to weight AOP data to quantify and reduce uncertainty of different levels of the AOP.

Case studies: Reproductive toxicity via aromatase inhibition, Skin sensitization, hepatocellular proliferation leading to cancer, mitochondrial fatty acid beta-oxidation inhibition leading to steatosis, membrane disruption (Narcosis) leading to respiratory failure

AOPs and Regulatory use



Perkins et al 2014. Using Adverse Outcome Pathways for Regulatory Applications

Three different types of qAOPs

Scoring/ Weight of Evidence Quantitative Approach

- Elements are weighted value based on expert opinion and well-documented criteria.

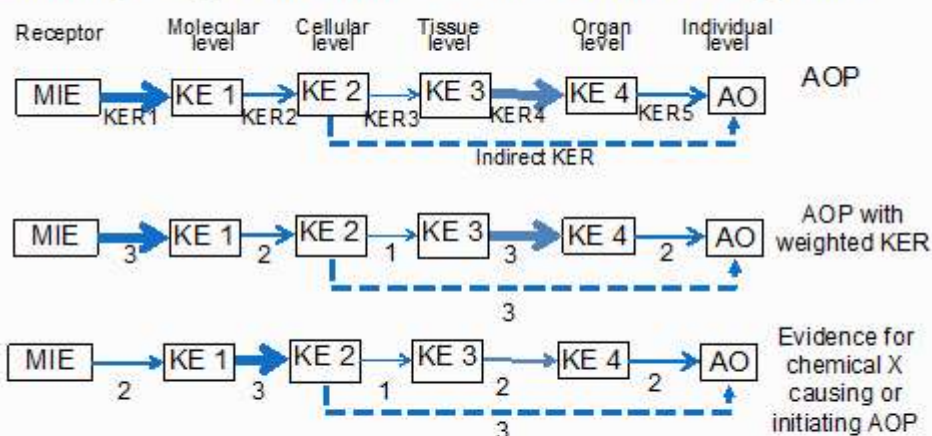
Probabilistic Quantitative Approach

- Statistical relationships exist that permit extrapolation between MI or KE and the AO

A Mechanistic Quantitative Approach

- Mathematical models of MIE, KE and KER are used to quantitatively predict AO in a dose-responsive manner

Scoring/ Weight of Evidence Semi-Quantitative Approach.

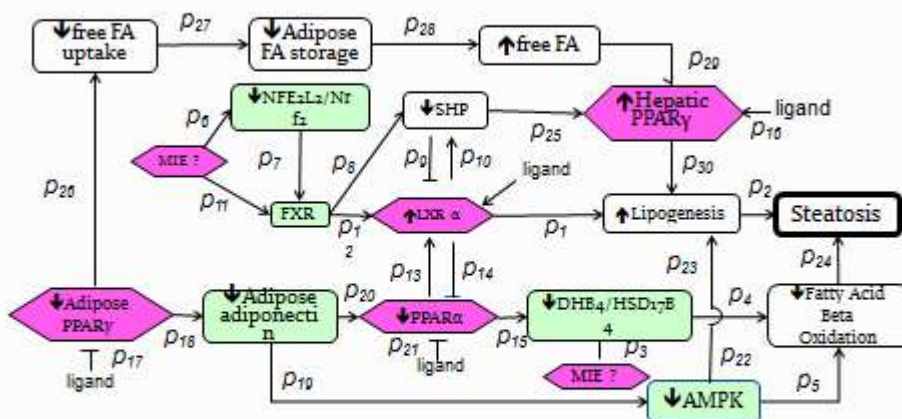


Based on Criteria established in modified Bradford Hill based weighting in AOP users handbook

https://aopkb.org/common/AOP_Handbook.pdf

Probabilistic Quantitative Approach:

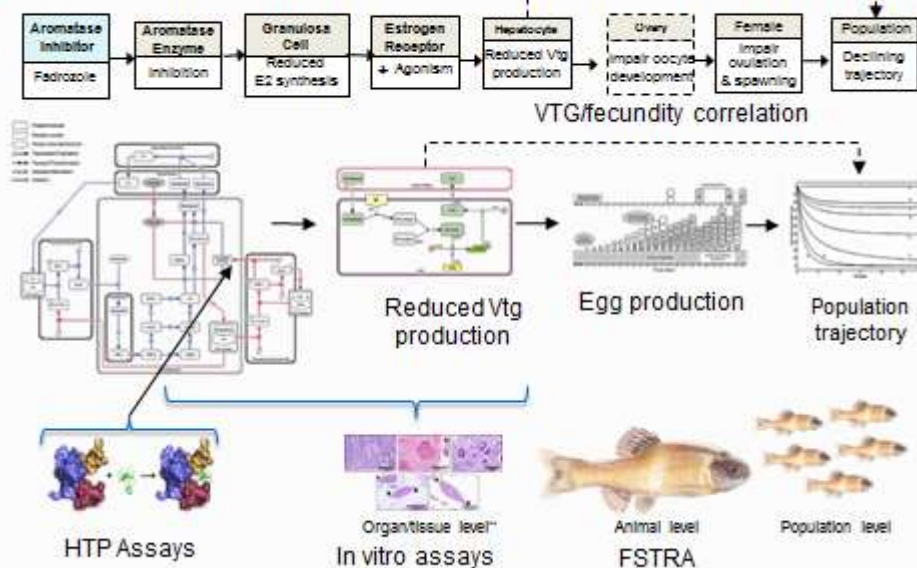
Network of AOPs leading to non alcoholic liver steatosis



Adverse Outcome Pathway (AOP) network for non alcoholic liver steatosis. Eight different AOPs are initiated by molecular initiating events (MIE, hexagon boxes) leading to the adverse outcome Steatosis. The probability (ρ_i) that one event leads to another is represented by an arrow. The probability an event will interfere with or inhibit another event is represented by a line and bar. The probability that any one AOP will result in steatosis is represented by the joint probability distribution across that AOP. Possible crosstalk between different AOPs are revealed in the network. The effect of complex mixtures could be assessed by examining the joint probability distribution across the entire network given the available data. Need to name acronyms still.

A Mechanistic Quantitative Approach:

Inhibition of aromatase leading to reproductive dysfunction



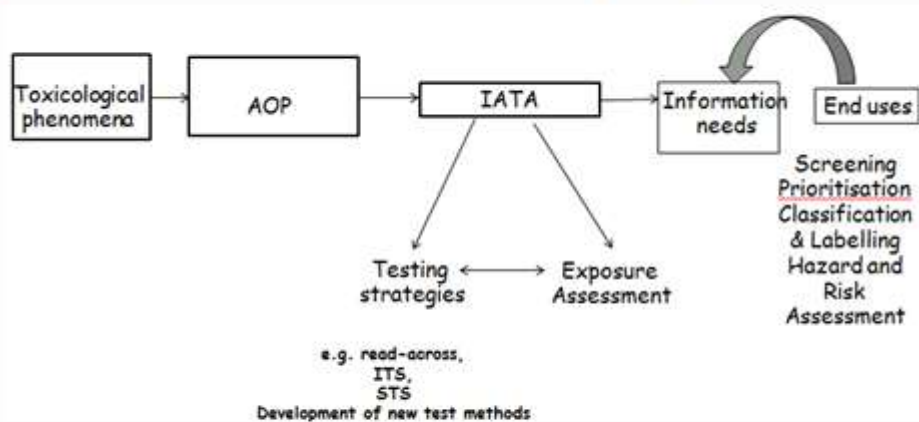
support *Integrated Approaches to Testing and Assessment (IATA)*

Challenges: AOPs offer conceptual approaches to the assembly and integration of data into knowledgebases, but are not, in and of themselves, sufficient for regulatory decision-making.

Objectives: Develop strategies for how AOPs may inform/assist practical applications for potential regulatory use.

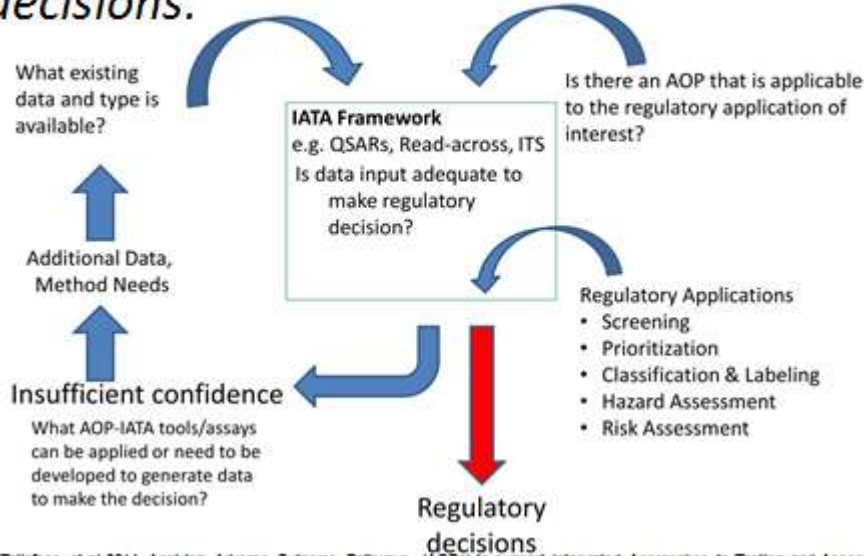
Case studies: Skin sensitisation, Endocrine disruption (EAT), AChE inhibition leading to Lethality.

AOPs – source for developing IATAs



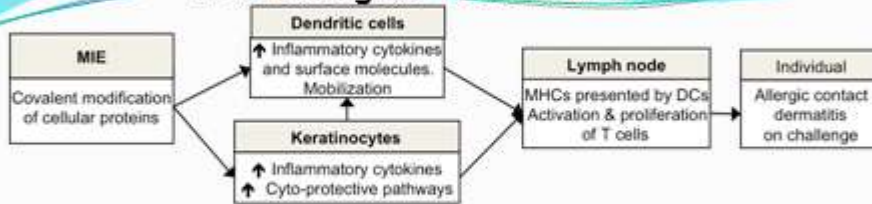
Tollefsen et al 2014. Applying Adverse Outcome Pathways (AOPs) to support Integrated Approaches to Testing and Assessment (IATA)

Conceptual framework for an AOP-informed IATA to support regulatory decisions.



Tollefsen et al 2014. Applying Adverse Outcome Pathways (AOPs) to support Integrated Approaches to Testing and Assessment

Covalent binding to proteins leading to Skin Sensitization AOP biological context for IATA



Example: Bayesian integrated testing strategy

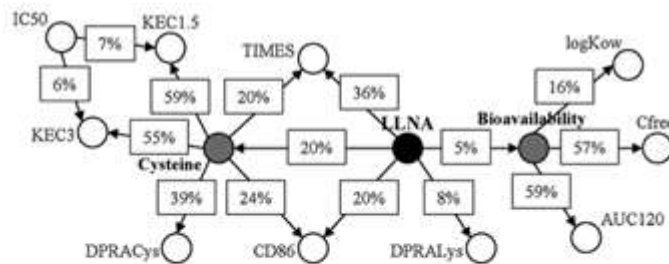
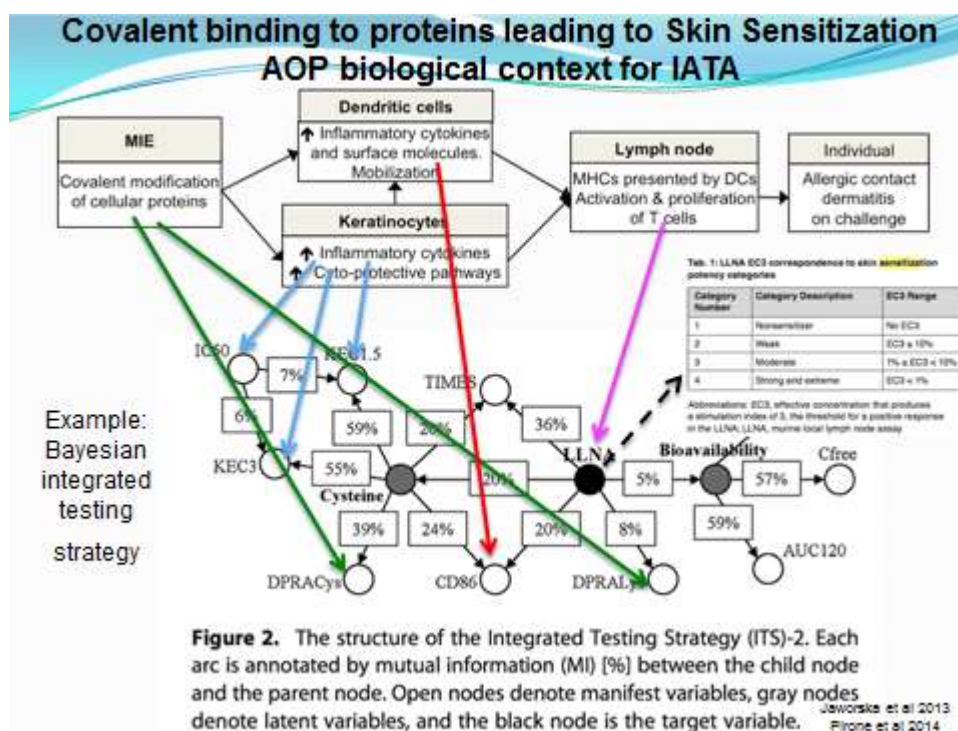


Figure 2. The structure of the Integrated Testing Strategy (ITS)-2. Each arc is annotated with mutual information (MI) [%] between the child node and the parent node. Open nodes denote manifest variables, gray nodes denote latent variables, and the black node is the target variable.

Jaworska et al 2013
Pirone et al 2014



Principles for AOP-IATAs

The initial principles proposed are framed by a clear identification of the regulatory requirement as well as the applicability domain of the IATA itself:

- define the endpoint of regulatory concern being assessed;
- define the purpose/application for which the IATA is proposed;
- describe the rationale, including mechanistic basis (e.g. AOP), according to which the IATA is constructed;
- describe the individual information sources constituting the IATA;
- characterize the predictive performance and applicability domain of the IATA, or IATA subcomponent(s) that can be expressed as a prediction model(s).

Summary

- ❖ Addressed key questions and challenges for practical application and implementation of the OECD work process.
- ❖ AOPs and elements of AOPS can be quantified
- ❖ New guidance produced for how to develop AOPs and how they could be used.
- ❖ Set the stage for additional issues to explore in future AOP workshops

- ❖ Products:
- ❖ Report highlighting major outcomes presented to OECD (May 2014)
- ❖ AOP WoE tables adapted and incorporated into the OECD AOP Users Guidance (AOPKB.org)
- ❖ 9 manuscripts providing detailed conclusions and case studies (5 already published).

<https://aopkb.org/saop/workshops/somma.html>

Acknowledgements

Organizing committee

- Natália Garcia-Reyero (MSU, USA)
- Rick Becker (ACS, USA)
- Ksenia Groh (EAWAG, Switzerland)
- Marlies Halder (JRC, EC)
- Sean Kennedy (Env. Canada, Canada)
- Teresa Lettieri (JRC, EC)
- Edward J. Perkins (ERDC, USA)
- Knut Erik Tollefsen (NIVA, NO)
- Bart van der Burg (BDS, The Netherlands)
- Dan Villeneuve (USEPA, USA)
- Maurice Whelan (JRC, EC)



ANNEX 12 PRESENTATION OUTCOMES FROM THE WORKSHOP ON THE ADVERSE OUTCOME PATHWAYS: FROM RESEARCH TO REGULATION



Adverse Outcome Pathways: From Research to Regulation

September 3-5, 2014

William H. Natcher Conference Center
National Institutes of Health
Bethesda, Maryland, USA

Workshop Summary
Michelle Embry (ILSI HESI)
OECD IATA Workshop
18 November 2014



AOP workshop was co-sponsored by:
NICEATM & PCRM
~120 in-person
>350 webcast



- Attendees/Speakers:
 - Research scientists
 - Regulatory decision-makers
 - Industry stakeholders
 - Nonprofit groups
 - Test method developers
 - Computational modelers
 - Epidemiologists
 - Informaticians
- Format:
 - Symposium talks
 - Discussion forums
 - Poster sessions
 - Junior investigator awards
 - Hands-on demonstrations
 - AOP Wiki/Effectopedia
 - Rotating breakout groups
 - Case study presentations
 - Charge questions

Workshop Steering Team

- Nicole Kleinstreuer (NIEHS)
- Kristie Sullivan (PCRM)
- Warren Casey (NIEHS)
- Craig Rowlands (Dow Chemical)
- George Daston (Procter & Gamble)
- Donna Mendrick (USFDA)
- Joanna Matheson (CPSC)
- Elizabeth Maul (NIEHS)
- Sharon Munn (EC JRC)
- Stephen Edwards (USEPA)
- Michelle Embry (ILSI HESI)



Workshop Topics

- **Workshop sessions:**
 - Building Upon Other Efforts
 - AOPs Under Development
 - Case Studies: Regulatory Uses for Well-Identified AOPs
 - The Risk Context
- **Breakout group topics**
 - The Process of Regulatory Acceptance
 - Using AOPs for Regulatory Decisions: Confidence and Criteria
 - Taking Qualitative AOPs to the Next (Quantitative) Level



Breakout Groups Conclusions:



- Need to incorporate variability and uncertainty around exposure, species differences, kinetics, dynamics, and quantification of AOPs
- Develop systematic, transparent frameworks for creating confidence in AOPs across all stakeholders, based on the application (prioritization, risk assessment, test method alternatives, etc.)
- OECD offers a path for international cooperation in the development, evaluation, and application of AOPs, supported by tools such as the Wiki Knowledge Base and Effectopedia

Breakout Group Conclusions: (cont'd)



- Weight of evidence approaches using the Bradford-Hill criteria and reproducibility analyses, combined with databases of validated assays, decision strategies (including assumptions and applicability domains) and AOP networks, will allow fit-for-purpose AOP validation
- Some priority pathways were identified based on public health concerns (e.g. cardiovascular, respiratory sensitization, diabetes, developmental toxicity)

Key Messages:
People, Process,
Priorities,
Partnering



Key Messages:
People



- Engage a broader community
 - Expand education and outreach
 - Integrate disciplines beyond toxicology (e.g., medical, IT)
- Help biologists become more computational
- Ensure that communication/momentum maintained

Key Messages: Process



- Needs to be systematic/transparent
- Many aren't aware of how to engage in the OECD process
- Distinguish development of AOPs from application of AOPs
- AOPs are useful even if they are not complete, but should be applied with caution
- Establish what is the minimum info (qualitative vs. quantitative) needed to develop a confidence framework

Key Messages: Priorities



- Determine priority AOPs to move forward, focus efforts on those first
- Facilitate communication between groups (**NICEATM AOP listserve established**)

Key Messages: Partnering



- Determine how best to leverage resources to build AOPs and facilitate regulatory use
- Need to ensure that industry is engaged
- How sustainable is the current mechanism for getting AOPs done? (currently constructed based on "volunteer" efforts)
- Could establish working groups that could develop AOPs rather than the ad hoc mechanism as currently done.

Next Steps / Follow-up



- Manuscript in preparation; will be reviewed by workshop steering team
- NIH AOP Listserv Established:
 - <https://list.nih.gov/cgi-bin/wa.exe?SUBED1=AOP&X=0C98C98B03C7F721C4>

ANNEX 13

**PARTICIPANTS LIST FOR WORKSHOP ON A FRAMEWORK FOR THE DEVELOPMENT
AND USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT**

Canada

Dr. Tara S. BARTON-MACLAREN
*Manager, Hazard Methodology Division
Existing Substances Risk Assessment Bureau
Health Canada*

*** Chair**

Ms. Kathy HUGHES Chief, Hazard Methodology Division
*Research Manager, Hazard Identification
Environment Nutrition and health Science Directorate
Health Canada*

*** Rapporteur**

Dr. Bette MEEK
*Associate Director, Chemical Risk Assessment
University of Ottawa
McLaughlin Centre for Risk Science*

Denmark/Danemark

Ms. Marie Louise HOLMER
*Danish Environmental Protection Agency, Chemicals
Division
Danish Environmental Protection Agency*

Dr. Susanne HOUGAARD - BENNEKOU
*Ph.D. Toxicologist
Pesticides and Genetechnology
Ministry of the Environment (MIM)*

Mr. Magnus LØFSTEDT
*Ecotoxicologist
Danish Ministry of the Environment*

France

Ms. Sandrine ANDRES
*Responsable de l'Unité
Direction des risques chroniques - unité ETES
Institut National de l'Environnement Industriel et des
Risques (INERIS)*

Professor Barbara DEMENEIX
*Directeur CNRS UMR 7221/MNHN USM 501
Lab Evolution des Régulations Endocriniennes
National Museum of Natural History*

Germany/Allemagne

Dr. Matthias HERZLER
*Department Chemicals Safety, Unit Toxicology of
Chemicals*

Federal Institute for Risk Assessment (BfR)

Professor Gerrit SCHÜÜRMAN
*Department of Ecological Chemistry
UFZ Centre for Environmental Research*

Mr. Gerd MAACK
*Section IV 2.2
Federal Environmental Agency*

Japan/Japon

Dr. Takashi YAMADA
*Safety Assessment Division, Chemical Management Center
National Institute of Technology and Evaluation*

Dr. Hajime KOJIMA
*Secretary General
Japanese Center for the Validation of Alternative Methods
(JaCVAM), Division of Pharmacology, Biological Safety and
Research Centre
National Institute of Health Sciences (NIHS)*

Dr. Taisen IGUCHI
*Okazaki Institute for Integrative Bioscience
National Institute for Basic Biology, National Institutes of
Natural Sciences*

Netherlands /Pays-Bas

Mrs. Jeannette GOMEZ
*Risk Assessment Officer, Regulatory Toxicity
Centre for Safety of Substances and Products
National Institute of Public Health and the Environment
(RIVM)*

Professor Aldert PIERSMA
*Laboratory for Health Protection Research
National Institute for Public Health and the Environment
RIVM*

Mrs. Lidka WIJKHUIZEN MASLANKIEWICZ
*Department for Consumers and Product Safety
National Institute of Public Health and the Environment
(RIVM)*

Switzerland/Suisse

Dr. Andreas BUSER
*PhD, Scientific Advisor
Transport, Energy and Communications DETEC
Federal Office for the Environment (FOEN)*

Ms. Sabine FREY
Federal Office of Public Health (FOPH)

United Kingdom/

Dr. Miriam JACOBS

Royaume-Uni

*Toxicology Department
Public Health England*

Dr. Gavin MAXWELL
*Safety & Environmental Assurance Centre (SEAC)
Unilever*

United States /États-Unis

Dr. Rory CONOLLY
*Senior Research Biologist
ORD/NHEERL/ISTD
U.S. Environmental Protection Agency*

Dr. Kevin CROFTON
*Acting Deputy Director
National Centre of Computational Toxicology
U.S. Environmental Protection Agency*

Dr. Steve EDWARDS
*Systems Biologist
ORD/NHEERL/ISTD
U.S. Environmental Protection Agency*

Suzanne FITZPATRICK
*Senior Advisor for Toxicology
Center for Food Safety & Applied Nutrition / US Food and
Drug Administration
DHHS/FDA/OC/CFSAN/OCD*

Dr. Abigail JACOBS
*ODE Assoc. Dir. Pharm/Tox ONDIO/CDER/FDA
Pharmacology/Toxicology for Several Offices of Drug Evaluation
U.S. Food and Drug Administration*

Ann LOWIT
*Co-chair of ICCVAM
Office of Pesticides
US Environmental Protection Agency*

Dr. Edwards PERKINS
*Senior Scientist
Environmental Laboratory
US Army Engineer Research and Development Center*

*** Chair**

Dr. Jennifer SEED
*Risk Assessment Division
U.S. Environmental Protection Agency (EPA)*

EU /UE

Dr. Jos BESSEMS
*Systems Toxicology Unit - EURL ECVAM
JRC - Institute for Health and Consumer Protection (IHCP)*

*** Chair**

Dr. Jean-Lou DORNE
Emerging Risks
European Food Safety Authority (EFSA)

Dr. George FOTAKIS
Scientific Officer
ECHA
European Chemicals Agency (ECHA)

Professor Maurice WHELAN
Head of Unit - Systems Toxicology
Institute for Health and Consumer Protection
European Commission Joint Research Centre

Business and Industry
Advisory Committee
(BIAC)/Comité consultatif
économique et industriel
(BIAC)'

Dr. Klaas DEN HAAN
Petroleum Products
CONCAWE

Dr. Frank FAULHAMMER
Regulatory Toxicology Chemicals II
BASF - The Chemical Company

Dr. Katy GOYAK
Senior Toxicologist
Toxicology
ExxonMobil Biomedical Sciences, Inc.

*** Rapporteur**

Dr. Grace PATLEWICZ (TIER)
DuPont Haskell Global Centers

Environmental NGO

Dr. Jennifer MCPARTLAND
Environmental Defense Fund (EDF)

International Council
on Animal Protection
in OECD Programmes
*** Rapporteur**

Dr. Katy TAYLOR
British Union for the Abolition of Vivisection

Dr. Kate WILLETT *Director*
Regulatory Toxicology, Risk Assessment and Alternatives
Humane Society of the U.S.

International Life
Sciences Institute (ILSI)

Michelle R. EMBRY
Senior Scientific Program Manager
ILSI Health and Environmental Sciences Institute (HESI)

OECD/OCDE

Dr. Joop DE KNECHT
Principal administrator
ENV/EHS

*** Chair**

Professor Terry SCHULTZ
University of Tennessee

Lhasa Limited

Mr. Mukesh PATEL
Principal Scientist

***Liverpool John Moores
University
* Chair***

Dr. Mark CRONIN
School of Pharmacy and Chemistry