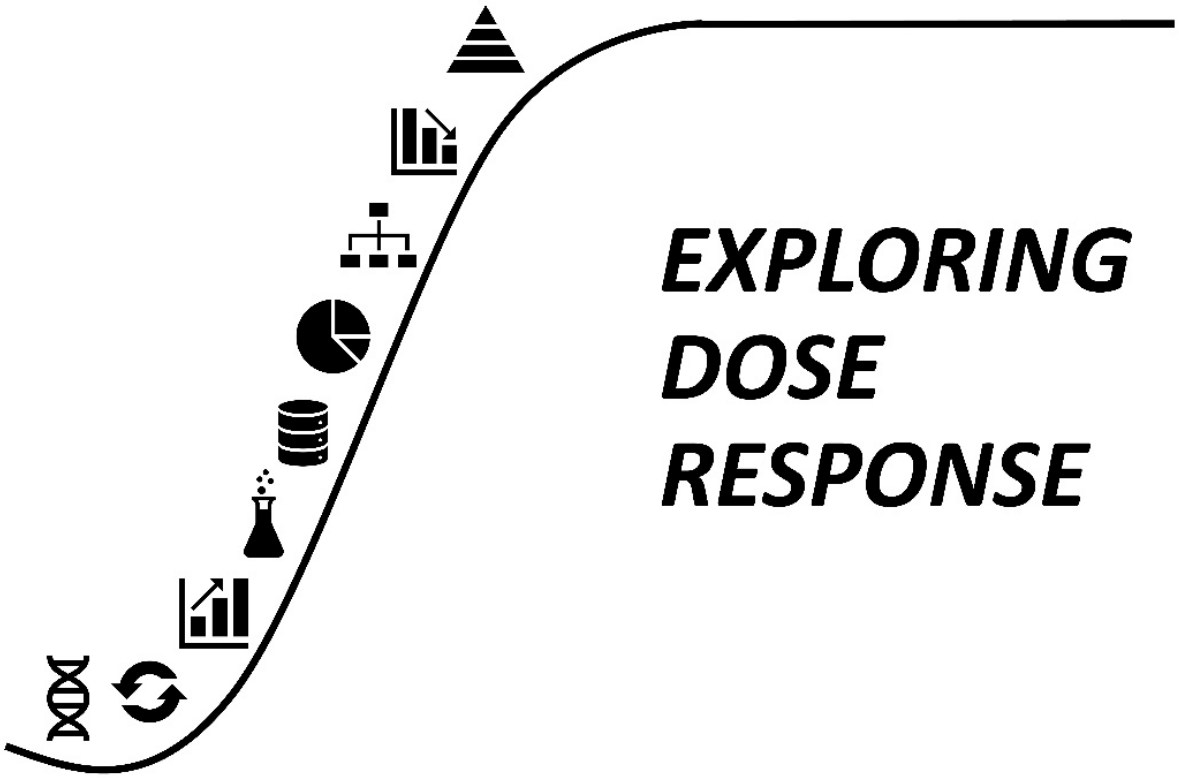
Workshop Report

New Approach Methodologies (NAMs) In Regulatory Risk Assessment



Manchester

March 2020

Committee on Toxicity

Food Standards Agency

Contents **Page Number**

Abstract 3
About the COT 4
About the FSA 4
Executive Summary

Introduction and Background 7
Chemical Landscape 7
Potency Estimation 8
Overview of in silico toxicology 8
Integrated Approaches to Testing and Assessment 9
Previously 9
Present-Why now? 10
FSA requirement for potency estimation 10
Background work 10

Mission and Vision: Objectives and outline of workshop 11
Objectives of the workshop 11
Outline of the workshop 11
Brief overview of case studies 12

Presentations and panel discussions 13
Presentations 14
Panel Discussion Session Outputs 18

Case Studies 23
Tropane Alkaloids 23
Polymers 26
Selective Androgen Receptor Modulators 30
Vinyl Acetate Monomer 35

Future Steps 37
Driving future research on point of departures (PODs) 37
Research priorities & recommendations 41

Overarching Conclusions and Recommendations 42
Moving from research to risk assessment to regulatory setting and beyond 43

Abbreviations 46
References 48
Technical Terms 52

Workshop Organizing Committee 54
COT Members 54
COT Secretariat 55

Abstract

The UK Food Standards Agency (FSA) and the Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment (COT) held an “Exploring Dose Response” workshop in a multidisciplinary setting inviting regulatory agencies, government bodies, academia and industry. The workshop provided a platform from which to address and enable expert discussions on the latest in silico prediction models, new approach methodologies (NAMs), physiologically based pharmacokinetics (PBPK), future methodologies, integrated approaches to testing and assessment (IATA) as well as methodology validation.

Using a series of presentations from external experts and case study (plastic particles, polymers, tropane alkaloids, selective androgen receptor modulators) discussions, the workshop outlined and explored an approach that is fit for purpose applied to future human health risk assessment in the context of food safety. Furthermore, possible future research opportunities were explored to establish points of departure (PODs) using non-animal alternative models and to improve the use of exposure metrics in risk assessment.

About the Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment (COT)

The Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment (COT) is an independent scientific committee that provides advice to the Food Standards Agency, the Department of Health and Social Care, and other Government Departments and Agencies on matters concerning the toxicity of chemicals.

About the FSA

The UK Food Standards Agency is an independent Government department working across England, Wales and Northern Ireland to protect public dietary health and consumers' wider interests in food. The FSA uses expertise and influence so that people can trust that the food they consume is safe and is what it says it is, and food is healthier and more sustainable.

The Science, Evidence and Research Division (SERD) of the FSA provides strategic analysis, insight and evidence across the FSA's remit to underpin the development of policies, guidance and advice on food safety.

SERD is a multi-disciplinary team of scientists, risk assessors, economists, statisticians, social scientists and operational researchers which provides high quality, timely and robust evidence. We strengthen our knowledge base using a range of external science capabilities, such as our independent Scientific Advisory Committees (independent groups of experts that advise the FSA on various aspects of food safety), by commissioning research and surveys, and engaging with academia, research councils through sponsoring PhDs and post-doctorate fellowships.

Executive Summary

Advances in biology, computer science, and other fields are paving the way for major improvements in how we evaluate environmental and public health risks posed by potentially toxic chemicals. The combined advances in discovery and clinical sciences, data science and technology have resulted in toxicity testing which has reached a pivotal transformation point known as part of the 4th industrial revolution (4IR). One of the major recent scientific advancements is the development of alternative toxicity testing and computer modelling strategies for the evaluation of hazard and exposure.

The UK Food Standards Agency (UK FSA) and the Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment (COT) held an "Exploring Dose Response" workshop in a multidisciplinary setting inviting attendees from regulatory agencies, government bodies, academia and industry. The workshop provided a platform from which to address and enable expert discussions on the latest in silico prediction models, new approach methodologies (NAMs), physiologically based pharmacokinetics (PBPK), future methodologies, integrated approaches to testing and assessment (IATA) as well as methodology validation.

Using a series of presentations from external experts and case study (including plastic particles, polymers, tropane alkaloids, selective androgen receptor modulators) discussions, the workshop outlined and explored an approach that is fit for purpose applied to future human health risk assessment in the context of food safety. Furthermore, we explored possible future research opportunities to establish points of departure (PODs) using non-animal alternative models and to improve the use of exposure metrics in risk assessment.

The overall conclusions and recommendations were as follows:

1. The use of pragmatic guidelines/framework for incorporating these models into risk assessment.

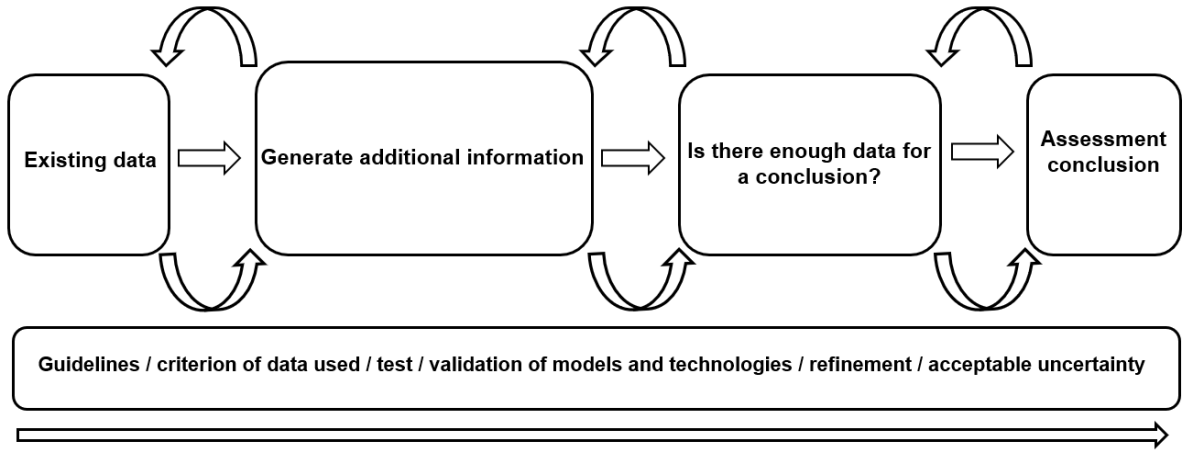
2. Case studies, such as those outlined in the workshop, should be used to determine applicability, and provide confidence in the models.

3. Human biomonitoring data will be key to identify a realistic snapshot of exposure scenarios as well as ‘big data’, which need to be linked to human clinical data.

4. Exposure data and exposure science will be key in developing in silico modelling in risk assessment and to explore the use of exposomics.

5. There should be transparency throughout the process i.e., Consumer facing engagement of new approach methods.

Ultimately, it was collectively agreed by attendees, that integration of these new technologies, as part of our risk assessment methodologies with a validation process throughout, will be key in the acceptance of the models (by regulatory bodies) and will be fundamental in the future of human and environmental safety.



Introduction and Background

Advances in biology, computer science and other related fields are paving the way for major improvements in how we evaluate environmental and public health risks posed by potentially toxic chemicals. The combined advances in discovery and clinical sciences, data science and technology have resulted in toxicity testing which has reached a pivotal transformation point known as part of the 4th industrial revolution (4IR). One of the major recent scientific advancements is the development of alternative toxicity testing and computer modelling strategies for the evaluation of hazard and exposure.

The volume of data produced in the world is growing ever more rapidly, from 33 zettabytes in 2018 to an expected 175 zettabytes in 2025 (IDC, 2018). The Department for Business, Energy and Industrial Strategy (BEIS) white paper on 'Regulation for the Fourth Industrial Revolution' notes that changes in technology are occurring at a "scale, speed and complexity that is unprecedented". The use of these new and changing technologies can help improve regulatory processes in several ways such as to improve the efficiency of data collection and exploit data already held by agencies to support better analysis and risk assessment.

Chemical Landscape

Over 350,000 chemicals and mixtures of chemicals have been registered for production and use worldwide. This is up to three times as many as previously estimated and with substantial differences across countries/regions. A noteworthy finding is that the identities of many chemicals remain publicly unknown because they are claimed as confidential (over 50,000) or ambiguously described (up to,70 000) (Wang et al., 2020).

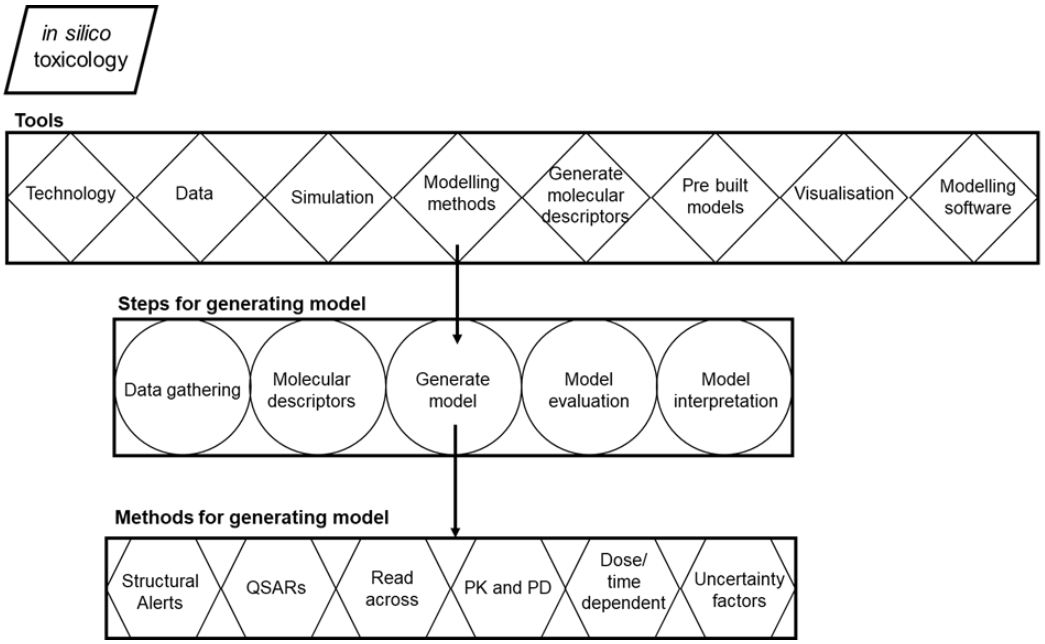
As a result, thousands of chemicals are in common use, but only a portion of them have undergone significant toxicologic evaluation, and as more emerge it is important to prioritize the remainder for targeted testing (Judson et al., 2009). This is especially important for chemicals (found in food and in the environment) where sometimes little or no toxicological information is available.

Potency Estimation

Potency measures can be applied to chemicals for rapid identification of pharmacoactive hits or toxicological assessment and used as input data for prediction modelling or association mapping.

Overview of in silico toxicology

In silico toxicology encompasses a wide variety of computational tools (Figure 1): databases for storing data about chemicals, their toxicity, and chemical properties; software for generating molecular descriptors; simulation tools for systems biology and molecular dynamics; modelling methods for toxicity prediction; modelling tools such as statistical packages and software for generating prediction models; expert systems that include pre-built models in web servers or standalone applications for predicting toxicity; and visualization tools. In general, methods include the following steps while developing prediction models (Figure 1): gathering biological data that contain associations between chemicals and toxicity endpoints, calculating molecular descriptors of the chemicals, generating a prediction model, evaluating the accuracy of the model, and validation of the model (Patterson et al., 2020).



Key
QSARs: Quantitative Structure Activity Relationships
PK: Pharmacokinetic
PD: Pharmacodynamic

Figure 1. Overview of *in silico* toxicology. Tools, steps for generating model and methods for generating model (Figure adopted from Raies and Bajic 2016).

Integrated Approaches to Testing and Assessment

Integrated approaches to testing and assessment (IATAs) provide a means by which all relevant and reliable existing information about a chemical can be used to answer a defined hazard characterization question. Information considered, can include toxicity data, exposure routes, use cases, and production volumes. This information is used to characterize outcomes that can inform regulatory decision-making.

The drawbacks of traditional toxicity testing approaches using laboratory animals may be overcome, by the use of human cell-based, biochemical, and/or computational methods to predict chemical toxicity. Due to the complexity of toxicity mechanisms, data from several methods usually need to be considered in combination to adequately predict toxic effects. IATAs provide a means by which these data can be considered in combination. When necessary, IATAs can guide generation of new data, preferably using non-animal approaches, to inform regulatory decision-making.

Previously

In 2009, the COT held a workshop on 21st century toxicology. The workshop addressed the United States (US) National Academy report called Toxicity Testing in the 21st Century: A Vision and a Strategy. The report called for accelerated development and adoption of human cell *in vitro* and *in silico* methods for the prediction of hazards, the determination of mechanistic information, and the integration of data.

Present - Why now?

As it is now halfway through the strategy period (10 years) it would be pertinent to review the current methodologies available whilst holding a workshop and to discuss their applicability in risk assessment including the current regulatory landscape.

FSA requirement for potency estimation / exploring dose response/ PBPK

The UK FSA have identified a need for potency estimation to aid in risk assessment.

This will be fundamental in risk assessment scenarios where limited to no information is available on the toxicity of a chemical.

When responding to food incidents the UK FSA regularly assess chemicals, particularly unauthorised novel food ingredients and sports/dietary supplements

where there is very little toxicological information available, and it is not possible to provide meaningful valid risk advice to FSA Policy colleagues.

Background work

Background work was undertaken to help set the scene for the workshop and output. A scoping paper on Environmental, health and safety alternative testing strategies: Development of methods for potency estimation was presented to the COT in December 2019. The COT were provided with a concise review of currently available methods, which included databases, different kinds of QSAR methods, adverse outcome pathways (AOPs), High Throughput Screening (HTS), read-across models, molecular modelling approaches, machine learning, data mining, network analysis tools, and data analysis tools using artificial intelligence (AI) to inform the objectives of the workshop.

Mission and Vision: Objectives and outline of workshop

Objectives of the workshop

The application of these alternative strategies to human health risk assessment requires effective collaboration between scientists including chemists, toxicologists, informaticians and risk assessors. As such, this multi-disciplinary workshop drew upon delegates and speakers from industry, academia, and regulatory agencies with a diverse range of experience.

The workshop provided a platform (Figure 2A) from which to address the latest in silico prediction models and PBPK modelling techniques. It provided speakers with the opportunity to share their knowledge and experience through case studies and roundtable discussions on approaches fit for purpose, including consideration of their validation and integration into current health risk assessment practices (Figure 2B).

Outline of the workshop

The workshop was divided into different area sessions; New Approach Methodologies & Special Scenarios, Approaches fit for purpose: Validation of methodologies, PBPK modelling, Future Methodologies (micro-physiological environment) and consisted of presentations accompanied by roundtable discussions of case studies with feedback and a discussion about future research needs. The presentations were delivered by invited experts (Figure 3) and had been designed to provide relevant information to inform the later discussions and case studies.

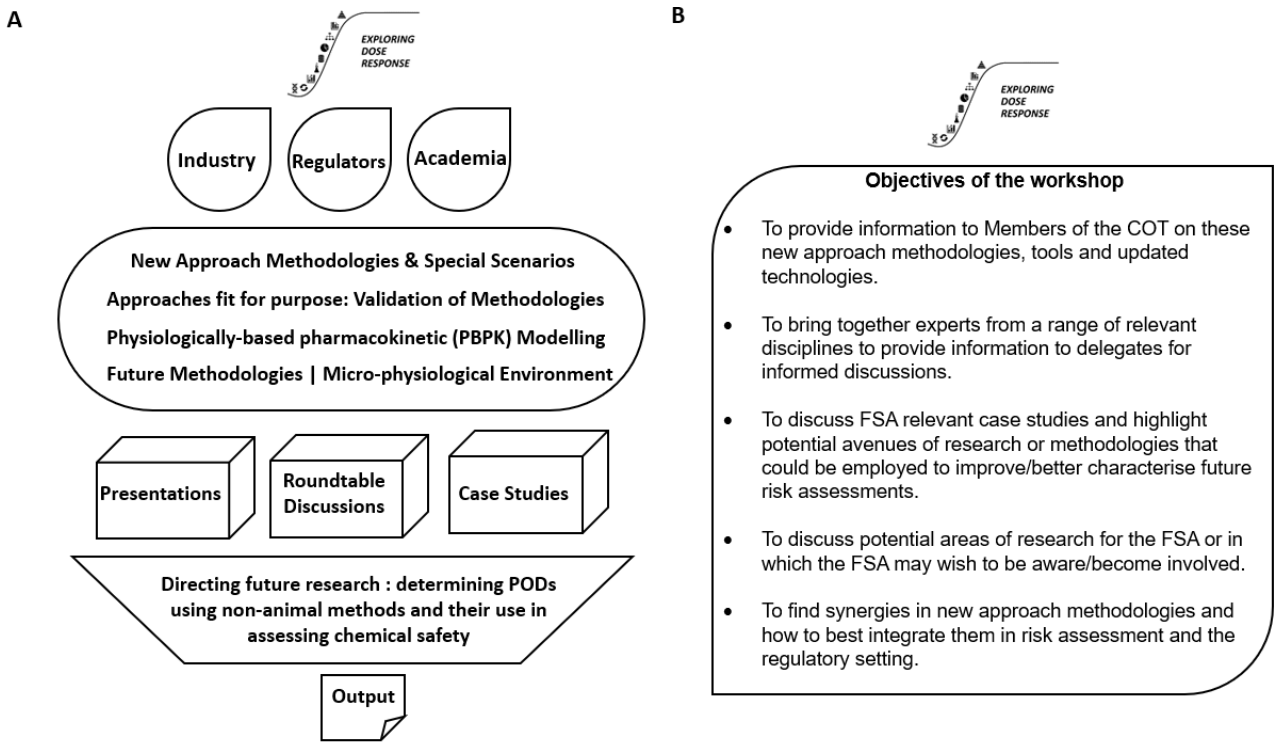


Figure 2. Diagrams representing outline and objectives of workshop. Overview and outline (A) and objectives (B) of the Exploring Dose Response Workshop.

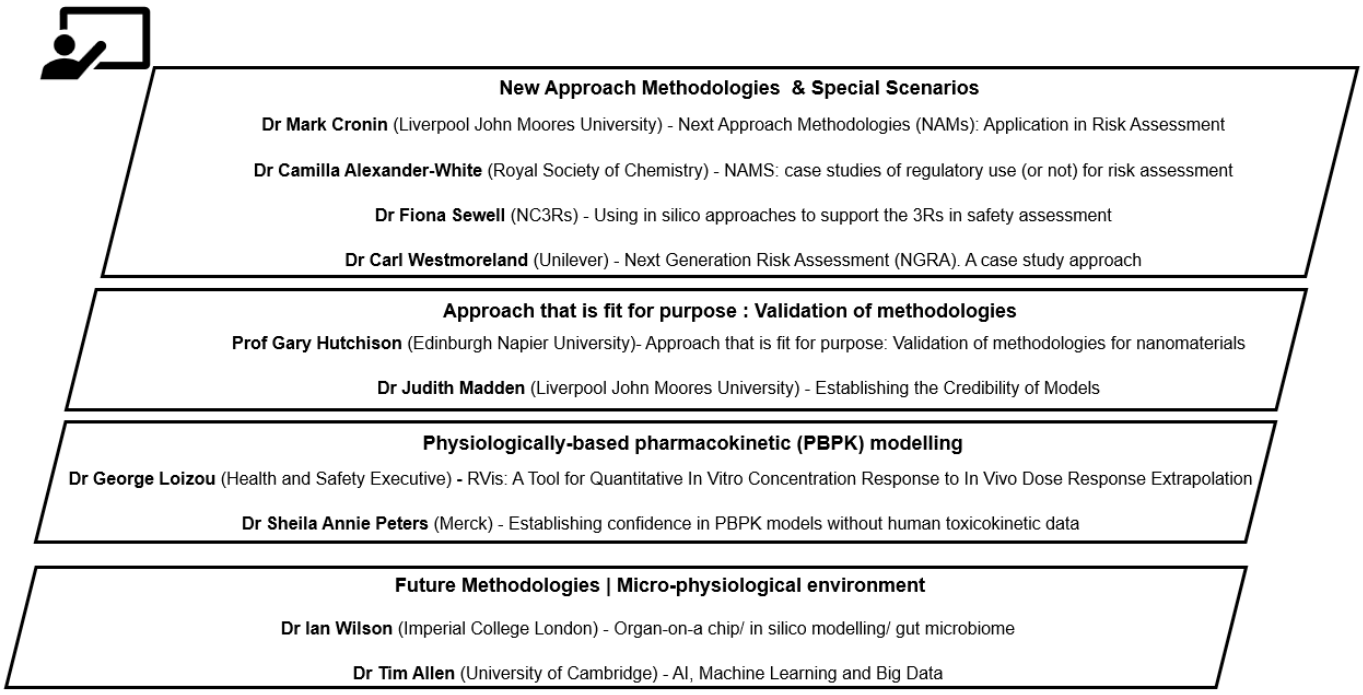


Figure 3. Diagram overview of sessions and presentations of the Exploring Dose Response Workshop.

Brief overview of case studies

The case studies (Figure 4) were based on a range of chemical areas (man-made and environmental) relevant to the FSA, where limited information was available for the purposes of risk assessment: contaminants (tropane alkaloids (TAs)), food supplements (selective androgen receptor modulators (SARMs)), food contact materials (vinyl acetate monomer (VAM)) and polymers (plastic particles). Background on the chemical groups was provided and questions were asked to prompt discussion. Each of the case study discussion groups included invited experts, Members of the COT, COT Secretariat and other attendees to try and ensure consistent expertise across groups.

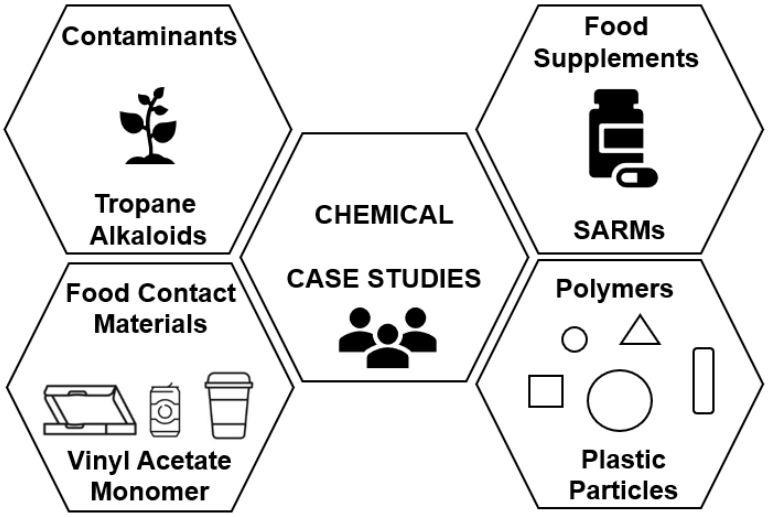


Figure 4. Case studies overview. The case studies were based on a range of chemical areas (man-made and environmental) relevant to the FSA, where limited information was available for the purposes of risk assessment: contaminants (tropane alkaloids (TAs)), food supplements (selective androgen receptor modulators (SARMs)), food contact materials (vinyl acetate monomer (VAM)) and polymers (plastic particles).

Presentations and panel discussions

The workshop was divided into different area sessions: New Approach Methodologies & Special Scenarios; Approaches fit for purpose: Validation of methodologies; PBPK modelling; and Future Methodologies (micro-physiological environment) and consisted of presentations accompanied by roundtable discussions of case studies with feedback and a further discussion about future research needs. The presentations were delivered by invited experts and had been designed to provide relevant information to inform the later discussions and case studies.

Presentations

In the session **New Approach Methodologies & Special Scenarios**, Professor Mark Cronin (Liverpool John Moores University) introduced the topic of "New Approach Methodologies (NAMs): Application in Risk Assessment". The future short-term aims of NAMs in risk assessment include filling in data gaps and provision of relevant information for regulatory submissions. Longer term, NAMs aim to support the mechanistic and exposure profiling of chemicals and will provide the data to support the new paradigm in non-animal safety assessment of chemicals. Highlighting the challenges faced when using NAMs in risk assessment, are their translation from theory to practice; the development of robust, reliable, and reproducible methods; integration into schemes for risk/safety assessment; and their global harmonisation with regard to regulatory acceptance.

Dr Camilla Alexander-White (Royal Society of Chemistry) discussed recent case studies of regulatory use (or not) for risk assessment. This included chemical grouping; human biomonitoring 4 EU programme (HBM4EU) in Europe); an example of how a PBPK model was accepted by EU regulators built on a plethora of *in vivo* (Bernauer et al., 2016) data and an example of a case study using quantitative *in vitro* to *in vivo* extrapolation for environmental esters (Campbell et al., 2015).

Dr Fiona Sewell (NC3Rs) presented on using in silico approaches to support Replacement, Reduction and Refinement (3Rs) in safety assessment by the UK National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs). NC3Rs is a science-led and evidence-based organisation established in 2004 to accelerate the development and uptake of new models and tools that replace, reduce, or refine the use of animals in research. The presentation described how these new approaches can be incorporated to improve decision-making. Though the ultimate aim is to work towards replacement, 'alternatives' are unlikely to offer a direct 1:1 solution and a tiered/combinatory approach may be necessary.

Dr Carl Westmoreland (Unilever) highlighted recent publications in the area of *in vitro* and *in silico* risk assessment, a tiered approach to be used (highlighting the

processes and current methodologies available) and the principles of Next Generation Risk Assessment (NGRA) according to the International Cooperation on Cosmetics Regulation. A case study method was presented, which was used to test the NGRA tiered approach assuming that there were no traditional toxicology data for a commonly used ingredient (coumarin).

In the session for **Approach that is fit for purpose: Validation of methodologies** Professor Gary Hutchison (Edinburgh Napier University) presented on alternative testing and exposure strategies for nanomaterials (NM) outlining the various Horizon 2020 projects such as Grouping, Read-across, Characterisation and classification framework for regulatory risk assessment of manufactured nanomaterials and Safer design of nano (GRACIOUS). It also considered the investigation and verification of current testing methods for engineered NMs and their applicability for use with nanobiomaterials (NBMs) and how these could be used in proposed Integrated Approaches to Testing and Assessment (IATA) for Developmental and Reproductive Toxicology (DaRT). Finally, the talk outlined key challenges for the future, highlighting the move to complex 3D cell models and microfluidic systems and how we ascertain dose may be challenging; agreement on definitions and measurement of dose (mass, surface area); stability of (Bio) nanomaterials in solution; corona assessment; assessment of complex 3rd generation bio nanomaterial, within possible matrices, will challenge traditional approaches and understanding the implications of endotoxin contamination in production lines are all key areas that need to be worked through to support the safe development of the technology.

Dr Judith Madden (Liverpool John Moores University) presented on establishing the Credibility of Model including using credibility criteria. These alternative methods include leveraging existing data, in silico modelling and the use of (human relevant) *in vitro* models. The Joint Research Centre (JRC) EU Reference Laboratory for alternatives to animal testing (EURL ECVAM) report (2017) established four criteria for achieving model credibility: (i) understanding the model; (ii) understanding the data underpinning the model; (iii) clearly stating assumptions and hypothesis encoded and; (iv) considering the gap between the model and reality. Credibility of PBK models can be visualised using a matrix that characterises the degree of confidence in the components of the model: i.e., its biological plausibility, how well it

simulates known data and its overall reliability considering uncertainty and sensitivity. *In vitro* assays can be used to generate new data, these should be conducted in accordance with Organisation for Economic Co-operation and Development (OECD) Technical Guidance documents or the OECD Guidance Document on Good *In Vitro* Method Practices (GIVIMP, 2018). Model reporting needs to adequately justify and document both the model structure and the parameters used, to ensure reproducibility and confidence in the model. It was concluded that to advance PBK modelling, in the context of supporting chemical safety assessment, it is essential that there was an ongoing dialogue between model developers and (regulatory) users.

In the session **PBPK modelling**, Dr George Loizou (Health and Safety Executive) presented a software tool called RVis which is a prototype application for the analysis of structure and performance of physiologically PBPK and other models. The input parameters comprise anatomical, physiological, metabolic and physicochemical values and the calculated outputs are the rates of uptake, elimination and organ and tissue concentrations (i.e., the internal dose). The advantages of utilising RVis as a tool for probabilistic PBPK is that it accounts for human inter-individual variability, has the ability to determine a credible interval for BMD lower bound values, and also offers a fully quantified measure of uncertainty for quantitative *in vitro* to *in vivo* extrapolation.

Dr Sheila Annie Peters (Merck) discussed establishing confidence in PBPK models without human toxicokinetic data. This was done by introducing the barriers to establishing mechanistic credibility of PBPK models in bottom-up and top-down approaches. A workflow to verify and validate the predictive performance of a PBPK model was presented, in addition to the utility and role of sensitivity analysis. Food and Drug Administration (FDA) recent white paper published a framework that can be used by industry and regulatory agencies to assess the credibility of computational models. There are five key concepts that can be used to establish model credibility (namely, the question of interest, defining the context of use, assessing model risk, establishing risk-informed credibility, and assessing model credibility). However, it was noted that there is a lack of consensus on best practices for determining if a model is fit-for purpose (with reference to validation,

performance/sensitivity metrics, and platform independence). It was concluded that knowledge gaps and uncertainties in predicted human pharmacokinetics cannot be overcome by any level of sophistication in pharmacokinetic modelling. Furthermore, the number of model assumptions tends to be proportionate to model complexity to a point that a complex model could become too distant to whatever is being modelled. Transparent communication of underlying assumptions and knowledge gaps is needed. Although PBPK offers valuable opportunities for data integration, mechanistic basis and route extrapolation, value addition of PBPK needs to be objectively evaluated, demonstrated, and understood before it is adopted.

In the session **Future Methodologies: Micro-physiological environment**, Professor Ian Wilson (Imperial College London) addressed the potential use of organ-on-a-chip technology, *in silico* modelling and the gut microbiome in exploring dose response. A particular challenge for the future highlighted is the role of the gut microbiota. The micro-organisms resident in the gut represent a major and highly variable component of metabolism and prospects for the use of *in vitro* systems to aid in its modelling were detailed. However, variability in the composition of the gut microflora complicates modelling as it results in, sometimes significant, interindividual differences in the metabolism, pharmacology and toxicity of dietary components and xenobiotics. gut microflora can have an array of effects on the following: drugs and their metabolites, bioavailability of dietary constituents, expression of host drug metabolising enzymes, and toxicity. Ultimately, future *in vitro* and *in silico* models will have to take into account gut wall metabolism for oral exposures. In addition, such models should benefit from the increased *in vitro* assessment of gut microbial activity and the highly targeted use of both gut microflora and organ-based humanized *in vivo* models.

Dr Tim Allen (University of Cambridge) discussed AI, machine learning and big data in risk assessment. The talk first gave an overview on the AOP, after which structural alerts was discussed. Dr Allen presented on one his projects using 2D structural alerts to define chemical categories for molecular initiating events (Allen et al., 2018) It was discussed how molecular-initiating events (MIEs) are important concepts for *in silico* predictions. They can be used to link chemical characteristics to biological activity through an AOP. Furthermore, the project explained how the tool provides

the first step in an AOP-based risk assessment, linking chemical structure to toxicity endpoint. Neural networks and quantitative predictions were also introduced. In biologically inspired neural networks, mathematical relationships link artificial neurons in layers leading to a prediction in the output layer. There was also description of another project by Wedlake et al using 90 biological targets representing important human MIEs, structural alert-based models which have been constructed with an automated procedure that uses Bayesian statistics to iteratively select substructures. These networks can be used as both binary predictors and quantitative predictors, which are more suitable for a risk assessment procedure.

Panel Discussion Sessions Outputs

New Approach Methodologies & Special Scenarios

- Cost comparison vs traditional methodologies i.e., NAMs approaches to risk assessment may seem to be relatively inexpensive on a per assay basis, but as a number of approaches may need to be used as part of a tiered toxicity testing framework to give confidence in the results, costs and time can escalate and become expensive.
- For higher level exposures, greater uncertainty factors or more conservatism may be needed in the risk assessment as applied through a rigorous uncertainty assessment.
- Different tools and standards could be brought into the tiered approach and uncertainty assessment utilised for both the estimation of systemic exposure and ingredient bioactivity.
- Bespoke investigations can be designed to explore effects of chemicals as they are progressed through the tiers of a NAMs approach.
- High throughput transcriptomic (HTTr) data could be used, from the perspective of potentially establishing PODs based on a No Observed Transcription Effect Level (NOTEL), which are more conservative than No-observed-adverse-effect levels (NOAELs) derived from animal studies.

- Internal dose: dosimetry and *in vitro* kinetics are imperative to define/predict what concentration of chemical went into the cell rather than what was added to the well in an *in vitro* assay. This is important so that doses of effects can be more reliably translated, and more accurate predictions made.
- It needs to be established how good the strategy is for computational methods, since the models are only as good as the data going in i.e., if the data are not available, a model cannot be produced.
- Not all biological effects and complex stress responses are picked up with computational methods. Therefore, 'missing information' needs to be covered using biology assays and adverse outcome pathways. This could be achieved by transcriptomics.
- 'Big data' approaches need to be linked to human clinical data, biobanks and biomonitoring data, including the analysis of biofluids to tissues and organs.

Approach that is fit for purpose: Validation of methodologies

- Alternative testing and exposure strategies for nanomaterials was discussed, outlining the various Horizon 2020 projects such as:
    - GRACIOUS- Grouping, Read-across, Characterisation and classification framework for regulatory risk assessment of manufactured nanomaterials and Safer design of nano.
    - PATROLS - Physiologically Anchored Tools for Realistic nanOmaterial hazard aSsessment is establishing a battery of innovative, next generation safety testing tools to more accurately predict the adverse effects caused by long-term engineered nanomaterial (ENM) exposure in humans and the environment. (The ambition is to accurately predict adverse effects caused by long term (chronic), low dose engineered nanomaterial exposure in humans and environmental systems to support regulatory risk decision making).
    - Risk Management of Biomaterials (BIORIMA)- To adapt and validate current test methods and or develop new test methods to detect

adverse effects of nanobiomaterials (NBM) (*in vitro* and *in vivo*), as well as contribute to integrated testing strategies to support QSAR and PBPK/PD. This work supports the standardisation of NBM and methods for their eventual use in advanced therapy medical products (ATMP) and medical devices (MD). This includes benchmarking reference materials.

- Challenges for the future: highlighting the move to complex 3D cell models and microfluidic systems and how we ascertain dose may be challenging i.e. internal dose; agreement on definitions and measurement of dose (mass, surface area); stability of NBM in solution; corona assessment; assessment of complex 3rd generation NBM, within possible matrices, will challenge traditional approaches and understanding the implications of endotoxin contamination in production lines are all key areas that need to be worked through to support the safe development of the technology.
- Credibility of physiologically-based kinetic (PBK) models can be visualised using a matrix that characterises the degree of confidence in the components of the model: i.e., its biological plausibility, how well it simulates known data and its overall reliability considering uncertainty and sensitivity
- To ensure credibility of the input parameters for any model, consideration should be given to their origin.
- Model reporting needs to adequately justify and document both the model structure and the parameters used, to ensure reproducibility and confidence in the model.
- To advance the acceptance of PBK modelling, in the context of supporting chemical safety assessment, it is essential that there is an ongoing dialogue between model developers and (regulatory) users. Further uptake of PBK models is being facilitated by development of additional guidance documents, generation of case studies and improved resources for the generation of input parameters and models.

Physiologically-based pharmacokinetic (PBPK) Modelling

- For model reproducibility, generally, there is insufficient information in the documents (peer-reviewed literature) to allow reproduction for the same chemical, let alone other chemicals. One of the benefits, of the available PBPK software models, is that the user can put their own distributions into models. However, it is important to note that they should still have access to appropriate expertise. In the discussions, it was raised that dealing with contaminants is different to dealing with pharmaceuticals i.e., Model credibility depends on the intended purpose and must be taken into account in the risk assessment process.
- PBPK models are versatile but also need to be reliable. It was stated that it would be difficult to validate a model per se because it is dependent on how the model will be used. However, there have been on-going efforts to make reporting of models more consistent. Guidance is under development at OECD and Tan et al. (2020) published a reporting template.
- There is now much more available information on parameters. However, for contaminants it is not possible to get an understanding of unknown unknowns. It was stated that when sampling a population, you have to co-variate to get correlated sampling.
- At what point, if at all, should FSA consider consumer-facing transparency re: NAMs when used in risk assessment? That's the very reason that these methods are not being rushed into in risk assessment. The risk assessment will be clear and transparent about methods and uncertainties.
- Are the available microdosing data relevant, given the dose is below the saturation kinetics and how can we ensure the system is not overly saturated or exposure significantly underestimated? Polyethylene terephthalate (PET) was used as a good example for microdosing because there is not a dose that will cause anomalies.

Future Methodologies and Micro-physiological Environment

- Neural Networks are a class of Machine Learning Algorithms that can provide both binary and quantitative predictions.

- Structural Alerts, Random Forests and Neural Networks have been used to try and predict binary activity at Human MIEs.

- A combination of these models (e.g., structural alerts, random forests and neural networks) and understanding of their workings is key to highest performance and model use in toxicology decision making.

- Dose response relationships and risk assessment procedures ideally require quantitative information, but qualitative risk assessments can be carried out too.

- Quantitative predictions help push this methodology closer to use in risk assessment, rather than just hazard identification.

- The power of the machine learning algorithm is that it works in a similar way across the board. Models don't have to be built in a bespoke way every time, but it was stated that applicability is bespoke. The applicability domain is acceptable but perhaps there should be degrees of certainty in different areas of space. The initial cases and training data (used for validation) also need to be considered.

- Bayesian probability offers the opportunity to update the probability for a hypothesis as more evidence or information becomes available. It can look and filter the probability of accuracy and conditional probability. Therefore, you can relate the actual probability to the measured test probability. Alternative ways of doing dose response modelling are required to correct for errors. Data are not necessarily information; interpretation is required to achieve that transition.

- Discussion on the questions: When to adopt new schemes? how many failures are you prepared to have? When are there enough *in silico* predictions that a physical experiment does not have to be performed? It was debated whether *in silico* and *in vitro* methods are actually cheaper than *in vivo* studies. There is increasing confidence in computational approaches, but

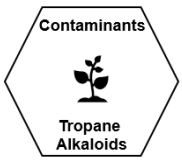
they may need additional approaches and Weight of Evidence (WoE) would still be used initially, which increase the cost.

- The Safety & Environmental Assurance Centre (SEAC) coumarin case study is a good example of building models. Increased confidence in the tiered/NAMs/PBPK approach are likely to predominantly come through case studies.

- When considering the biotransformation of bioactive compounds in food it needs to be accepted that the gut, including its microflora, should be considered as well as the liver. There are >2000 species of microflora in the gut. Some are essential, some not, and they represent a huge metabolic capability. It was discussed that the microbiome changes with environment, diet, age, sex, pharmaceutical use etc., how the information from gut microflora should/would be used in PBPK modelling might prove somewhat challenging. It should also be remembered that the gut microflora-derived metabolites across various cultures/countries will vary.

Case studies

Tropane Alkaloids Contaminants (Natural)



Tropane alkaloids (TAs) are plant toxins that are naturally produced in several families including Brassicaceae, Solanaceae (e.g., mandrake, henbane, deadly nightshade, Jimson weed) and Erythroxylaceae (including coca). TAs can occur in cereal-based foods through the contamination of cereals with seeds from deadly nightshade and henbane. Although more than 500 different TAs have been identified in various plants, respective data on toxicity and occurrence in food and feed are limited (EFSA, 2013). The COT has reviewed TAs and in 2017, the FSA commissioned a survey on the monitoring of TAs in food.

Attendees were asked to consider the following:

- A number of other TAs of unknown potency were present at higher concentrations than (-)-hyoscyamine and (-)-scopolamine, with some of these reported at detectable levels in up to 26% of cereal-based samples. Syndicate groups were asked to consider this group of compounds and explore ways of ascertaining the potency of similar molecules in the group, given that data are available on only a limited number of TA's.

- As it is thought that the effects of a combination of TAs would be different from those of exposure to a single TA, groups are asked to explore possible methods of quantifying this difference.

Discussion output points:

- With regards to potency, it would be prudent to first look into the known potencies of TAs. If it is assumed that all TAs are equipotent, then this would be the most conservative approach. However, the potency of most TAs is unknown but if there were standards used for their analysis, could potency be determined from these? The relationship between potency and antimuscarinic effects should then be explored. If this is not possible then an assumption could be made that their potency is equal to that of hyoscyamine and/or scopolamine. If they are equipotent then an assessment needs to be made as to the level of risk. It is important to note that if an assumption is being made on potency, then it cannot be ruled out that the potency of the TAs mentioned is more than that of hyoscyamine and/or scopolamine. For quantification, relative potency could be used taking advantage of data on effects on muscarinic receptors.

- There is potential exposure to various TAs from eating cereal-based products. Therefore, the risk assessment would have to consider different combinations. When looking at the effects in combination, it is important to consider all of the TAs detected and the potency, if we assume synergistic effects. It is possible that the effects may be geometric or have antagonistic actions. It is possible that when in combination, less potent compounds may bind the receptors and

prevent the more potent compounds from docking.

- When exploring antimuscarinic effects, *in vitro* tests should be conducted, and other endpoints investigated to check if TAs are all antimuscarinic. If it is assumed that all TAs are antimuscarinic then presumably combinations of TAs will have an addictive effect. Muscarinic receptors have known potency for these compounds. However, there are some limitations such as receptor ligand binding, receptor ligand responses. It would be worth exploring different HTS methods for TAs (binding assay) then using expert opinion to rank the data.

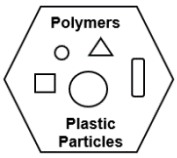
- It is important to consider whether TAs all have the same toxicokinetics. It would be desirable to measure bioavailability by looking at the metabolism and pharmacokinetics of TAs of known potency and then ranking potency levels of TAs and look into exposure of these chemicals. The structures could then be run through a QSAR programme to see if data gaps can be filled. It would be useful to look critically at the structures, such as substituents on the molecule and the variety of sidechains, for changes in the receptor. Questions arose such as:
    - Is there a way that the potency of TAs can be ranked using QSAR?
    - Could we use the acute reference dose (ArFD)?

- Structural differences in TAs could have different effects on a receptor. If the TA is structurally different it may hit a different site of the same receptor and modulate other TAs which may lead to competition. Read-across may still be the best estimate but there is always uncertainty because the substances are not the same. The limited data on TAs reduces the reliability of read-across. It was noted that there are structural alerts present for genotoxicity in some TAs. Therefore, one would characterise using genotoxicity and then TTC, giving the worst-case scenario. No exposure data is provided and there is no information on LOQ or LOD, but as there are alerts for genotoxicity this would suggest that any exposure is unacceptable. It would need to be investigated whether there are any common chemical groups throughout the TA structures

which trigger the antimuscarinic effect. Is a QSAR method able to differentiate between different effects? It is likely that a tiered approach will be required.

- Finally, it was noted that there has only been detection of 24 TAs in the cereal-based samples because these are what the samples were analysed for. However, there are more than 500 TAs, any of which could also be present. It was suggested that better agricultural processes could be used to mitigate and reduce the risk by reducing the presence of TAs in cereals. Additionally, analytical methodology could be applied to detect more TAs.

Polymers/Mixtures (man-made/ environmental): Plastic particles



Plastic particles (micro/nano plastics) are intentionally added to products (e.g., in cosmetics as exfoliants) or result from fragmentation of macroplastics into smaller sizes by natural processes (e.g., weathering, corrosion etc.). These particles can come in different sizes; nano (1 – ≤100 nm), micro (1 – 5 mm), and macro (> 5 mm). The occurrence of microplastics has been reported in seafood, honey, beer and salt, with most of the data being on occurrence in seafood. A full risk assessment on the potential toxic effects of micro and/or nanoplastics could not be carried out due to the lack of comparative data available for baseline levels of both compounds.

Furthermore, there is no established NOAEL for each polymer type. The European Food Safety Authority (EFSA) Scientific Panel on Contaminants in the Food chain (CONTAM) concluded that the risks of toxicity from micro and nanoplastics themselves, from oral exposure could not be assessed due to the lack of data, especially with regards to metabolism and excretion (EFSA, 2016). The COT is currently reviewing the potential risk of microplastics in food.

Attendees were asked to consider the following:

- Do you envision the AOP methodology to be able to assist in prioritising the potencies of the different types of plastic particles? If yes, how so?

- Do you agree with the read-across of plastic particles to tyre and road wear particles?

Discussion output points:

AOP methodology:

- AOP methodology would assist in prioritising the potencies of plastic particles but it is not ready yet.
- There is still a need for internal and external exposure data.
- AOP needs a single chemical, but plastic particles may well have mixtures of chemicals.
- Read-across might be challenging between different plastics as the composition of plastics will differ. For most particles it would depend on what the particles are made of in order to determine what effects they might have.
- There should be a criterion for inclusion of a certain adverse effect/pathway.
- There should also be standardisation for the data used in read-across.
- Testing against key events would tell us what the chemical does but not what it is i.e. Do we even know a key event that actually takes place at this stage?
- Different exposure routes will lead to a wide variety of adverse effects. The route of exposure currently includes inhalation, dermal and ingestion which will then have different effects on internal dose.
- When considering how to use the AOP diagram it needs to be borne in mind that there is a battery of processes to go through some of which are known, whilst others are unknown. It needs to be considered whether an OECD approach for AOPs should be followed.

Tyre and road wear particles:

- Read-across from other particles is very limited. Read-across will therefore be challenging as there are limited or no data on plastic particles. The use of

read-across of tyre and road wear to plastic particles was not currently considered useful.

- With regard to fibrotic response to accumulation it will be challenging to pinpoint e.g., if it is adverse or is nano-clumping occurring?
- It was noted that there is limited analytical methodology available for microplastics and even more so for nanoparticles which affects particle matter (PM)10 (for both tyre wear and atmospheric fibres). This is further complicated by the organic sample matter (food/tissue). However, it is possible that migration data from manufacturer's could be obtained, and a risk assessment potentially be performed on the leachates.

Other discussion points:

- Particle morphology (size and shape) plays an important role on the toxicity profiles. This should be considered.
- Formation of protein coronas.
- What different types of polymers are we exposed to?
- What toxicology has been done to date? Any pointers for potential hazards?
- Do polymers have systemic access?
- There is information available on particle matter $(PM)_{2.5,5,10}$ *etc*. Can this be extrapolated?
- There are currently analytical and sampling challenges with measuring plastic particles such as how to analyse them in food/tissues. Consideration needs to be given to what chemicals are potentially stuck on the surface. The analysis is technically very challenging, and it is currently not possible to detect plastic particles below 1 µm in complex samples. The sampling size/method would be different for the environment/food and the different particles. Do we have sufficient particles in samples to analyse for the particle chemical effect?
- Persistent organic pollutants or weathered particles may lose some inherent characteristics.
- The potential presence of biofilms needs to be considered as do microbial effects.

- The physical aspects of the particles are responsible for the effects. How do the particles break down and is the size we see in the food/environment the starting size or subsequent from break down?
- Certain polymer particles may be converted to Environmentally-Persistent Free Radicals (EPFRs) following UV photolysis.
- There is uncertainty around particle composition/size. There are various distributions.
- Analytical methods are needed to extract particles from the environment.
- It is not certain how reliable older data are. There are not many labs which have the technology/possibility to generate the data required.
- There is a need to consider the possibility of microplastics accumulating other toxic chemicals within themselves.
- It needs to be determined whether plastic should be analysed in its original form or whether the polymer should be considered; some of the components would have been assessed toxicologically but only for the chemicals and not for micro particles. However, this would still only provide a snapshot of that time/place.
- Animal/toxicology studies are carried out on the pure plastic not on weathered particles which are what the population are generally exposed to.
- Nano-particles and micro-particles will behave differently, therefore having different effects.
- More clarity is required on the routes of exposure to plastic particles.
- Limited human data have demonstrated that (micron size) particles are able to pass through the gut. However, it has been demonstrated that in the nano range (nanoplastics) are sufficiently small to be able to cross and interact with biological components i.e., nano bio interface.
- The model would need to take into account the implications/long-term health effect of particles being retained in the lung/gut.
- How would the AOP pathway take into account chemicals that are stuck to a particle surface and released?

Food supplement (man-made): Selective Androgen Receptor Modulators (SARMs)

These can be found in bodybuilding/gym-based supplements and are designed to have a similar effect to anabolic steroids, but without many of the unwanted side effects. Toxicological information for SARMs is scarce and, where available, the dosage used in supplements is usually at higher doses than was tested in clinical trials. Since the mechanisms by which tissue selectivity is achieved have not been clearly elucidated, there is poor understanding of the potential side effects associated with exposure to SARMs through supplements. Moreover, structural modifications could affect the binding affinity, specificity and potentially affect the potency of different SARMs. Understanding of the structure-activity relationships (SAR), molecular pathways involved as well as the potency of the various molecules is needed for the development of a risk assessment strategy.

Attendees were asked to consider the following:

- (Q1) What criteria could be used for the development of AOP methodologies for the risk assessment of SARMs?
- (Q2) Could read-across be used for risk assessment of SARMs with limited toxicological information? If yes, what criteria should be used and are there any classes of chemicals that are appropriate for read-across based on the information provided?
- (Q3) Is it prudent to attempt to extrapolate from the levels used in clinical trials to the levels used in supplements?
- (Q4) Could PBPK used for understanding distribution of SARMs in the body and would this approach be appropriate for determining potential side effects?

Discussion output points:

Q1. What criteria could be used for the development of AOP methodologies for the risk assessment of SARMs?

- Biologically relevant key events i.e., anabolic effects or tissue specific effects, antagonistic or agonistic effects should be used. Searches could be undertaken for tissue specific effects and androgenic effects.

- The criteria used should be biologically relevant and of key events leading to a specific outcome. Example: trying to build an AOP on suitable skeletal muscle system *in vitro*.

- Utilisation of *in vitro* assays to screen the responses, using the chemical structure as a starting point. However, it must be noted that the AOP is not chemical specific which could be a limitation. It should be testing potencies for androgenic effects. The criteria need to be biologically relevant and related to specific key events, then adverse outcomes i.e., Use tissue relevant *in vitro* assays to aid development of AOP for SARMs. For example: Skeletal muscle system *in vitro* then the development of the AOP. AOP would work with SARMs as classic mechanistic intervention event. MOA will be the key interaction.

- When looking at structures, the read-across will be challenging. Look at analogues within the groups rather than across groups. Use parent compounds to scope out how compounds act and compare to other compounds. SARMs have small structural changes.

- It is important to note that the amounts of SARMs used in supplements are higher than the clinical dose, therefore the levels are not comparable. It should also be noted that the toxicity might be extension of the pharmacology. Comparisons should be made with others in the androgen receptor (AR) space and compounds may be tested at higher doses. It should be

determined whether levels can be extrapolated. The pharmaceutical industry is selecting compounds for tissue specificity.

- It would be useful to assess potency first, such as the biospider approach, using the androgen receptor model system and classic initiating event.

- A bespoke strategy might be needed, depending on definition, initiating effect and mechanism.

- Transcriptomics could be used and an AOP would be written for androgen receptors.

Q2. Could read-across be used for risk assessment of SARMs with limited toxicological information? If yes, what criteria should be used and are there any classes of chemicals that are appropriate for read-across based on the information provided?

- Using read-across for risk assessment of SARMs may not yet be possible, although AOPs could be used for similar compounds to allow possibility of read-across. Read-across is unlikely to be useful in this instance as small structural changes will potentially lead to large conformational ones. Read-across would be limited to binding, gene activity and transcription. However, it may be possible to use *in vitro* and structure via read-across. Read across could only be used if the new compound was similar in structure and end points to chemicals already considered i.e., if it causes a similar biological effect and it has a related structure.

- It is possible to do a risk assessment for androgenic effects, and that may raise a concern. If not, that doesn't necessarily mean that there aren't other effects, i.e., read-across from other substances affecting the androgen receptor is useful if it indicates a concern, perhaps less so if it doesn't. The challenge is that there is no database of toxicological data, so the focus is on the androgen receptor. Do we know enough about AR-mediated effects?

Q3. Is it prudent to attempt to extrapolate from the levels used in clinical trials to the levels used in supplements?

- Benzimidazoles are from multiple origins and from different sources in the food chain. It becomes a risk-benefit equation and a co-intake issue.

- The higher doses being taken are not comparable to those tested in clinical trials; at high doses, receptors may be saturated, etc. There are limits in doses in phase 2 trials.

- It is not considered prudent to extrapolate the levels used in clinical trials to supplement use, as levels in supplements are higher than those used in clinical trials. Although, the dose level selection in clinical trials, may indicate what a suitable risk/benefit ratio is.

- Things to consider:
    1. Increased concentration via nanoencapsulation.
    2. Co-intake/poly-supplement use.
    3. Key ingredients have multiple origins and,
    4. Clinical data may indicate a risk/benefit ratio, but it is not prudent to extrapolate from this for supplement use.

Q4. Could PBPK used for understanding distribution of SARMs in the body and would this approach be appropriate for determining potential side effects?

- PBPK could be used for internal dose, but risk assessment approaches use external doses. However, this may not help as the tissue distribution is only a hypothesis and in order to run a PBPK model the tissue concentration is needed but is currently unknown. The effect of high doses on the pharmacokinetics are unknown. Once a PBPK calculation is achieved inside cells it may make a decision easier. PBPK modelling might be possible with clinical trial data but may need more than 1 model. Therefore, PBPK modelling would be a good start but is unlikely to be sufficient by itself.

- Is there a consistent chemical communality between the different SARMs? What does the structure do to the toxicity? The diverse chemistry may affect read-across.

- Enough is known about the effects of other androgens to perhaps predict what PODs we might expect for androgenic effects. Therefore, a risk assessment can be done for the androgen part, however we don't know what other effects could arise from exposure as it is currently unknown. Computationally, it could be anticipated what the adverse effects would be. AOPs do exist for effects on the AR. However, other potential aspects/effects are unknown.

- It may provide insight if it was known how the pharmaceutical industry selects SARMs for tissue selectivity, whether there is a specific method. It would be useful to know what reason they have for selecting certain SARMs and not others. It would be interesting to know why not all SARMs go on to phase 2 in clinical trials. PBPK modelling might be possible if the clinical trial data was made available, however, more than one model may be needed e.g., transport specific information, structural similarity might also be useful.

Other points raised to consider:

- Internal doses of supplements should be considered and compared to medicines. PBPK modelling could be used for this.
- There are currently no biomarkers and there is only an idea about the variability as there are only small numbers of volunteers in the studies.
- Regulatory assessment tends to model the hazard so historical data could be used.
- Are supplements really foods? It would be useful to revisit the definition of foods.
- What goes into supplements? Is the labelling correct?
- Comparisons have to be done carefully for selectivity activity across different targets/off-target effects.

- Different mechanisms will result in different side effects.
- Biomonitoring can be useful but is unlikely to be available.
- The habits of consumers should be considered:
    - Do people take supplements separately or in combination?
    - Phase 1 trials mostly involve men therefore, the reported effects are in men. However, women take these supplements as well. What is known about the effects in women?
    - Are they being used by men and women? The general consensus was that they were more targeted towards men.
    - Do users take combinations?
    - Do they cycle through different SARMs?

Food Contact Material (man-made): Vinyl Acetate monomer (VAM)

Vinyl acetate monomer (VAM) is solely used as an intermediate in the chemical industry for manufacturing (polymerisation) of vinyl acetate (co)polymers. Hence it is concluded that the entire production volume of VAM is used up for the manufacture of various (co)polymers, mainly polyvinyl acetate. Polymers manufactured from VAM are used in a broad spectrum of products, including adhesives (e.g., film and surface adhesives) for packaging products and contain traces of vinyl acetate as a residual monomer. Human data on the acute toxicity of vinyl acetate are not available, however there are some rat studies. Therefore, by applying PBPK modelling various risk assessments have been proposed and this could potentially be used in future.

Discussion output points

- Supplementary analysis (uncertainty and sensitivity analyses) should be conducted as part of the model building phase (and not afterwards, as implied in the guidance from WHO 2010).

- It was noted that although guidance from the WHO states that “the plausibility of a particular dose metric (that is to be simulated) is determined by its consistency with available information on the chemical’s MOA as well as

dose-response information for the toxicological endpoint of concern", there is no dose-response information in the case of vinyl acetate, only information on MOA (i.e., only one side of the equation). Therefore, there was disagreement with the "medium" level of confidence placed on the model for vinyl acetate by the WHO.

- A delegate is involved in preparing OECD on guidance on the issue/validation of human PBPK models without human pharmacokinetic data. It was noted that with lipophilic chemicals, there is increased potential for lymphatic uptake from the GI tract, an absorption pathway that is not always included in PBPK models. There is a need to establish computer modelling processes including read across to predict this uptake from logP values. Furthermore, there is a need for regulators to do read-across.

- In the case of paraquat (herbicide), there is significant binding of this chemical to cartilage. This is an example of where the underlying biological interactions need to be understood before a PBPK model can be built to accurately reflect these exemplar mechanisms.

- Read-across may be used to predict physico-chemical properties but accurate prediction of the pharmacokinetics is more challenging.

- It was agreed that the values of the PBPK parameters would change between a microdose and a larger occupational or domestic exposure dose. The extent of the change depends on the pharmacology of the molecule in question. The use of microdose data is only valid for linear behaviour and subsequently a narrow range of exposures and applicability. They may not therefore relate to higher levels of occupational exposure where saturation effects may occur; this has certainty been the case as seen in the pharmaceutical industry. There are also human ethical considerations that remain with the use of microdosing. Furthermore, the radiolabel may change the *in vivo* behaviour of the chemical.

- Use and test known case studies as if the known is unknown.

- For a conservative approach: look at Monte Carlo simulation and Bayesian methods and see if they match. It is possible that you could apply this methodology to PBPK, select a concentration range and use distribution around vulnerable groups.

- Animal to human PBPK prediction is possible. Inhalation/deposition pre-systemic exposure could be modelled, although the anatomy is different, so it would only work in limited circumstances.

Future Steps

Driving future research on point of departures (PODs)

At the end of the workshop, a collective roundtable discussion was held: Directing future research – determining PODs using non-animal methods and their use in assessing chemical safety.

The following is a summary of the discussion points about how we can combine experts, themes and knowledge gaps in a multidisciplinary setting to drive priorities forward.

- There is a need for pragmatic guidelines in how to develop and implement quality assured NAMs for safety evaluations. These new methods generate complex new types of data and there are always likely to be gaps in understanding. The scientific uncertainties of NAMs can and should be described alongside the data. Confidence needs to be increased in the predictions from new methods/models i.e., there is not necessarily a need for a full validation of a NAM approach vs the outcomes of animal data because the aim is to use the data afresh in a different way to try to decide whether there is a risk of an adverse outcome in humans or not. A decision needs to

be reached, based on scientific evidence and the uncertainty around the prediction needs to be explained. This needs to be described as clearly and rigorously as possible, such that decision-makers can decide whether predictions from NAMs indicate the risk is acceptable or not? This requires a complete framework to be developed, which could be tiered in terms of data generation and requirements.

- The use of probabilistic approaches (statistics) and in particular Bayesian approaches in machine learning could be used to explore uncertainty in combining data types. In Bayesian learning, everything is a distribution, therefore a mean of variance is produced. These algorithms can provide true uncertainties for every case giving an output and probability of that output being correct. There is also the possibility of combining two methods: distribution before and after generating a new piece of data.

- There were discussions on the use of benchmarking the output of NAMs with the use of *in vivo* animal data estimation of PODs in risk assessment.

- Toxicodynamic modelling verification: Test the impacts on potency of receptor-based mechanisms in AOPs by using PBPK models.

- Consumer facing engagement on new approach methods: There should be planning to take NAMs forward using social sciences research and technical research for integration, such that the public have confidence that NAMs can be used equally as effectively to keep them as safe as using traditional methods.

- Case studies can be used to evaluate NAMs and how they perform for safety decision-making of the assessment of the risks of contaminants in food. The FSA need to define scenarios and substances (through case studies) that would be evaluated and see what outputs occur. It would be useful if the FSA could gather and articulate current science issues in toxicology Limit of Detection (LoD) / Limit of Quantitation (LoQ), application of uncertainty and

reference materials methods i.e., Measurement issues: reliable and validated? reference methods materials and LOQs should be validated along the way.

- A challenge led approach should be defined, with case studies and the models and their basis should be evaluated. Straightforward case studies could be used to start with as an initial step in the process. It was suggested that FSA should define these technical challenges where solutions are needed.

- Validation and acceptance: Use the coumarin in cosmetics case study to show how NAMs could be used in principle in safety evaluation for low level exposures.

- Provocative questions were put forward such as: How are animal models relevant to humans? And when did we decide that animals were good models for human and that we were happy with the data? It has become a matter of social acceptance that using data from animal models in our traditional methods are protective for human consumers.

- The use of exposomics and the use of exposomics data alongside both untargeted and targeted metabolomics profiling. This may generate useful information on kinetic behaviour in the body for chemicals already in use in products to learn more about human exposure modelling.

- Computational methods such as QSAR and molecular docking could be used for potency estimation if the known molecular targets could be used in a dose response. However, absolute potency needs to be evaluated objectively, to understand the relationship between potential activity at a molecular target and *in vivo* response in a range of organisms with differing pharmacokinetic attributes.

- Chemicals are processed in the body by bacteria as well as our cells/tissues. Can we incorporate microbiome in the models using *in vitro* methods to reflect physiology? We could use learnings from the pharmaceutical industry to guide the food industry. However, it is likely to be extremely challenging to do this.

- With regards to PBPK modelling, the WHO have developed guidance on how to develop a scientifically robust model. The onus is on the modeller to assess the validation of model/regulatory acceptance according to the WHO criteria. Further guidance on validation of models (for a given purpose) has also recently been published (Parish et al., 2020). Questions arose such as: Are there any circumstances where we can use simpler *in silico* compartmental models vs PBPK?

- There is generally no United Kingdom (UK) biomonitoring data for chemicals exposure in human populations (akin to that from the National Health and Nutrition Examination Survey (NHANES) programme in the USA or in the human biomonitoring 4 EU programme (HBM4EU) in Europe). It would be helpful to have UK human data for priority chemicals of interest or understand how and when EU data could be used and interpreted as being relevant or not for the UK population. It may then be possible to develop human relevant PBPK models for some classes of chemicals using human data, to learn more about human kinetics.

- Human clinical metabolomics could be used i.e., to relate *in vitro* metabolite signatures to those *in vivo.* Leverage human metabolomic data and human exposure assessment i.e., to evaluate the relevance of dose metrics in *in vitro* systems.

- How can we use and combine data from new technologies going forward, using data from *in silico* and *in vitro* technologies and human clinical data types and integrate all these new types of data as part of the risk assessment process to arrive at probabilistic rather than deterministic conclusions? Integration of multiple data types in clear risk-based frameworks will be key.

Research priorities & recommendations:

1. Incorporate microbiome in the models using *in vitro* methods to reflect physiology.

2. Exploration into intracellular dosing. It is important to define/predict what went into the cell rather than what was added to the well. The objective is to try and get close to the free concentration in the tissue.

3. Assay applicability: assay model validation and applicability for toxicity testing in a regulatory setting.

4. Explore the use of AI algorithms to prove uncertainties throughout the process step by step.

5. Exposure science need to develop formal criteria and processes for validation.

Overarching conclusions and recommendations

- Pragmatic guidelines / frameworks are needed for incorporating these models into risk assessment.

- We need to describe the uncertainty of these methods. There needs to be confidence in the prediction from these methods/models i.e., there is not necessarily a need for a full validation.

- Use case studies like the ones outlined in the workshop to move towards applicability and confidence in the models.

- Human biomonitoring data will be key to identifying a realistic snapshot of exposure scenarios. Incorporating this data in the *in silico* models could be provided to enhance accuracy in exposure scenarios.

- PBPK models could be used to provide relevant substances to benchmark against known human biomonitoring data.

- Big data need to be linked to human clinical data and biomonitoring data including the analysis of biofluids.

- Exposure data and exposure science will be key in developing *in silico* models in risk assessment.

- Explore the use of exposomics and the use of exposomics data alongside metabolomics.

- NAMs approaches used for the cosmetics could be applied in the same way for food ingredients/contaminants specifically for higher level exposures through Uncertainty Assessment.

- Transparency throughout the process i.e., Consumer facing engagement on new approach methods. There should be planning to take forward these new methods using social sciences research and technical research for integration.

- Finding a synergy to use / combine these new technologies and integrate these as part of our risk assessment methodologies with a validation process throughout.

Moving from research to risk assessment to regulatory setting and beyond

Food authorities should strive to incorporate the best scientific methods available (Kavlock et al., 2018).

In the recent EU Farm to Fork strategy and the EU Green Deal Food 2030 Pathways for Action (Food systems and data) it states that value should be placed on emerging technologies, tools, standards and infrastructure for use in food systems.

The Joint Research Centre (JRC) EU Reference Laboratory for alternatives to animal testing (EURL ECVAM) published its Status Report 2019 on the development, validation and regulatory acceptance of alternative methods used for scientific purposes stating: “Innovation, collaboration and education initiatives drive progress in alternatives to animal testing.”

NAMs and IATAs are rarely accepted by regulatory bodies and the key is how can these approaches be facilitated in a regulatory setting and supporting the technology available. However, the potential use through various case studies as a proof of principle concept is becoming apparent.

The future direction of safety assessment science will depend heavily on the evolution of the regulatory landscape. A key challenge, though, is whether the regulatory framework can keep pace with the increasing speed of scientific and technological developments (Worth et al., 2019).

This will need close collaboration between chemists, toxicologists, informaticians and risk assessors to develop, maintain and utilise appropriate models. Not only must the different disciplines come together, but also those scientists from industry, academia and regulatory agencies must recognise the commonalities (Cronin et al., 2018). The challenge is to respond to the growing need for adaptable, flexible and even bespoke computational workflows that meet the demands of industry and regulators, by exploiting the emerging methodologies of Tox21 and risk assessment.

The focus of the 7th annual Global Summit on Regulatory Science (GSRS17) was Emerging Technologies for Food and Drug Safety. In the GSRS17 meeting, it was said that “moving forward toward greater integration of emerging data and novel

methodologies for chemicals risk assessment will need continuous efforts on capacity building. This will be accomplished through increased data accessibility and sharing, the maintenance and establishment of key partnerships, technical workshops and training sessions with international experts, and ongoing focus on data analysis tools development to address regulatory questions. It is also important to demonstrate proof of concept through various case studies and work collaboratively on the interpretation and application of new data for use in regulatory applications." This is currently being done at an international level under the OECD and as the focus of the Accelerating the Pace for Chemical Risk Assessment initiative co-lead by the US EPA, the ECHA and Health Canada (Kavlock et al., 2018).

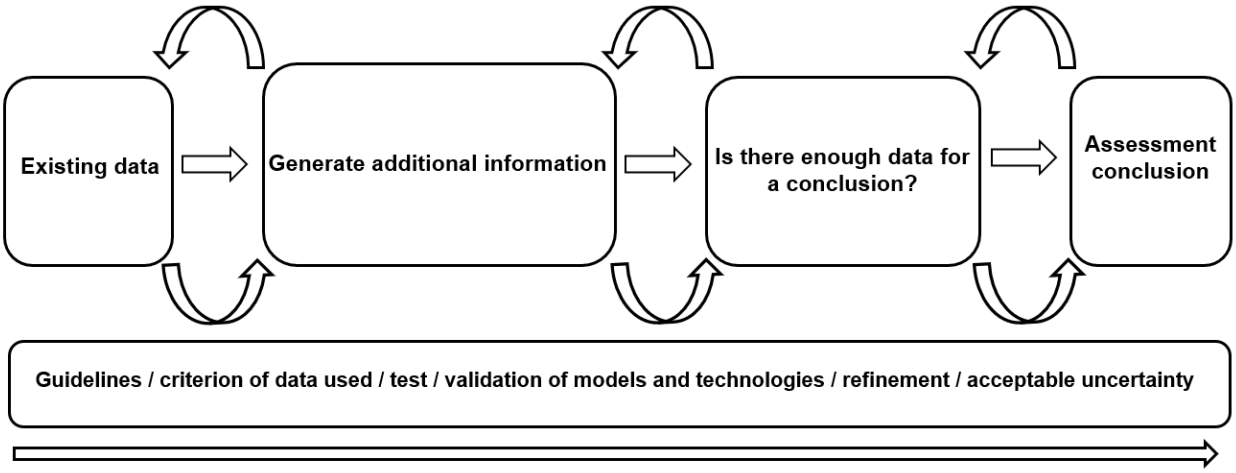


Figure 5. Diagram of concluding thoughts of workflow discussions around NAMs and IATA of the Exploring Dose Response workshop.

Ultimately, innovative technologies should be reviewed and evaluated once developed to be integrated as part of the risk assessment strategies for chemical testing for human health and the environment. Using a validation process via a science and evidence driven approach, to address the data gaps in the risk assessment process, will facilitate the acceptance and validity of these NAMs as well as pave the way for alternatives testing strategies with confidence (Figure 5). Furthermore, integration of these technologies as part of our risk assessment process to streamline our probabilistic rather than deterministic conclusions will be fundamental in the future of human and environmental safety.

Abbreviations

3Rs	Replacement, Reduction and Refinement
AI	Artificial intelligence
AOPs	Adverse outcome pathways
ATMP	Advanced therapy medical products
BEIS	Department for Business, Energy and Industrial Strategy
CONTAM	Scientific Panel on Contaminants in the Food chain
COT	Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment
EFSA	European Food Safety Authority
EURL ECVAM	EU Reference Laboratory for alternatives to animal testing
ENM	Engineered nanomaterial
FDA	Food and Drug Administration
HBM4EU	Human biomonitoring 4 EU programme
HTS	High Throughput Screening
HTTr	High throughput transcriptomic
IATAs	Integrated approaches to testing and assessment
JRC	Joint Research Centre
LoD	Limit of Detection
LoQ	Limit of Quantitation
MIEs	Molecular-initiating events
NAMs	New Approach Methodologies
NBM	Nanobiomaterials

NC3Rs	National Centre for the Replacement, Refinement and Reduction of Animals in Research
NHANES	National Health and Nutrition Examination Survey
NGRA	Next Generation Risk Assessment
NOTEL	No Observed Transcription Effect Level
OECD	Organisation for Economic Co-operation and Development
PET	Polyethylene terephthalate
PBPK	Physiologically based pharmacokinetics
PK	Pharmacokinetic
PD	Pharmacodynamic
POD	Points of departures
SAR	Structure-activity relationships
SARMs	Selective androgen receptor modulators
SEAC	Safety & Environmental Assurance Centre
SERD	Science, Evidence and Research Division
TAs	Tropane alkaloids
UK	United Kingdom
UK FSA	UK Food Standards Agency
US	United States
WoE	Weight of Evidence
VAM	Vinyl acetate monomer
QSARs	Quantitative Structure Activity Relationships

References

Allen, T.E., Goodman, J.M., Gutsell, S. and Russell, P.J., 2018. Using 2D structural alerts to define chemical categories for molecular initiating events. **Toxicological Sciences**, 165(1), pp.213-223.

Bernauer, U., Bodin, L., Celleno, L., Chaudhry, Q., Coenraads, P.J., Dusinska, M., Duus-Johansen, J., Ezendam, J., Gaffet, E., Galli, C.L. and Granum, B., 2016. SCCS OPINION ON decamethylcyclopentasiloxane (cyclopentasiloxane, D5) in cosmetic products. Available at: Opinion of the Scientific Committee on Consumer Safety on o-aminophenol (A14) (europa.eu)

Campbell, J.L., Yoon, M. and Clewell, H.J., 2015. A case study on quantitative in vitro to in vivo extrapolation for environmental esters: Methyl-, propyl-and butylparaben. **Toxicology**, 332, pp.67-76.

Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment. Environmental, health and safety alternative testing strategies: Development of methods for potency estimation (TOX/2019/70)
Available at: Developing methods for potency estimation (food.gov.uk)

Cronin, M.T., Madden, J.C., Yang, C. and Worth, A.P., 2019. Unlocking the potential of in silico chemical safety assessment–A report on a cross-sector symposium on current opportunities and future challenges. **Computational Toxicology,** 10, pp.38-43.

Department for Business, Energy and Industrial Strategy (BEIS) white paper on 'Regulation for the Fourth Industrial Revolution' Available at: The Use of Emerging Technologies for Regulation (publishing.service.gov.uk)

EFSA Panel on Contaminants in the Food Chain (CONTAM)
Scientific Opinion on Tropane alkaloids in food and feed. Available at: Scientific Opinion on Tropane alkaloids in food and feed - - 2013 - EFSA Journal - Wiley Online Library

EFSA Panel on Contaminants in the Food Chain (CONTAM)
Presence of microplastics and nanoplastics in food, with particular focus on seafood
Available at: Presence of microplastics and nanoplastics in food, with particular focus on seafood - - 2016 - EFSA Journal - Wiley Online Library

European Commission - European Strategy for Data. Available at: IMMC.COM%282020%2966%20final.ENG.xhtml.2_EN_ACT_part1_v11.docx (europa.eu)

EURL ECVAM Status Report 2017 on alternative methods. Available at: EURL ECVAM Status Report 2017 - European Commission (europa.eu)

EURL ECVAM Status Report 2019 on alternative methods. Available at: EURL ECVAM status report on the development, validation and regulatory acceptance of alternative methods and approaches (2019) - Publications Office of the EU (europa.eu)

HM Government Regulation for the Fourth Industrial Revolution: White Paper (2019)
Available at: Regulation for the Fourth Industrial Revolution: White Paper (publishing.service.gov.uk)

Horizon 2020. Available at: Horizon 2020 - European Commission (europa.eu)

Human biomonitoring 4 EU programme. Available at: HBM4EU – science and policy for a healthy future

Judson, R., Richard, A., Dix, D.J., Houck, K., Martin, M., Kavlock, R., Dellarco, V., Henry, T., Holderman, T., Sayre, P. and Tan, S., 2008. The toxicity data landscape for environmental chemicals. **Environmental health perspectives**, 117(5), pp.685-695.

Kavlock, R.J., Bahadori, T., Barton-Maclaren, T.S., Gwinn, M.R., Rasenberg, M. and Thomas, R.S., 2018. Accelerating the pace of chemical risk assessment. **Chemical research in toxicology, 31**(5), pp.287-290.

National Toxicology Program- Integrated Approaches to Testing and Assessment
Available at: Integrated Approaches to Testing and Assessment (nih.gov)

National Academies Toxicity Testing in the 21st Century, A Vision and a Strategy (2007). Available at: Toxicity Testing in the 21st Century: A Vision and a Strategy | The National Academies Press

OECD Guidance Document on Good In Vitro Method Practices (GIVIMP, 2018) Available at: Guidance Document on Good In Vitro Method Practices (GIVIMP) | en | OECD

Parish, S.T., Aschner, M., Casey, W., Corvaro, M., Embry, M.R., Fitzpatrick, S., Kidd, D., Kleinstreuer, N.C., Lima, B.S., Settivari, R.S. and Wolf, D.C., 2020. An evaluation framework for new approach methodologies (NAMs) for human health safety assessment. **Regulatory Toxicology and Pharmacology, 112**, p.104592.

Patterson, E.A., Whelan, M.P. and Worth, A.P., 2020. The role of validation in establishing the scientific credibility of predictive toxicology approaches intended for regulatory application. **Computational Toxicology,** p.100144.

Raies, A.B. and Bajic, V.B., 2016. In silico toxicology: computational methods for the prediction of chemical toxicity. **Wiley Interdisciplinary Reviews: Computational Molecular Science,** *6*(2), pp.147-172.

Safety & Environmental Assurance Centre: A Next Generation Risk Assessment Case Study for Coumarin in Hypothetical Cosmetic Products. Available at: A Next Generation Risk Assessment Case Study for Coumarin in Hypothetical Cosmetic Products (unilever.com)

Slikker Jr, W., de Souza Lima, T.A., Archella, D., de Silva Junior, J.B., Barton-Maclaren, T., Bo, L., Buvinich, D., Chaudhry, Q., Chuan, P., Deluyker, H. and Domselaar, G., 2018. Emerging technologies for food and drug safety. **Regulatory Toxicology and Pharmacology**, 98, pp.115-128.

Tan, Y.M., Chan, M., Chukwudebe, A., Domoradzki, J., Fisher, J., Hack, C.E., Hinderliter, P., Hirasawa, K., Leonard, J., Lumen, A. and Paini, A., 2020. PBPK model reporting template for chemical risk assessment applications. **Regulatory Toxicology and Pharmacology**, p.104691.

Wang, Z., Walker, G.W., Muir, D.C. and Nagatani-Yoshida, K., 2020. Toward a global understanding of chemical pollution: a first comprehensive analysis of national and regional chemical inventories. **Environmental Science & Technology, 54**(5),

pp.2575-2584.

Wedlake, A.J., Folia, M., Piechota, S., Allen, T.E.H., Goodman, J.M., Gutsell, S. and Russell, P.J., 2019. Structural Alerts and Random Forest Models in a Consensus Approach for Receptor Binding Molecular Initiating Events. **Chemical Research in Toxicology**.

Worth, A.P., 2019. The future of in silico chemical safety… and beyond. **Computational Toxicology**, (10) pp 60-62.

Technical Terms

Bayesian statistics	Bayesian statistics is a theory in the field of statistics based on the Bayesian interpretation of probability where probability expresses a degree of belief in an event.
Exposomics	Exposomics is the study of the exposome and relies on the application of internal and external exposure assessment methods. Internal exposure relies on fields of study such as genomics, metabonomics, lipidomics, transcriptomics and proteomics.
Internal Dose	Internal dose: the amount of a substance that entered the body through the skin, eyes, lungs, or digestive tract and was taken up by organs or particular tissues).
Molecular Initiating Event (MIE)	A Molecular Initiating Event is the initial interaction between a molecule and a biomolecule or biosystem that can be causally linked to an outcome via a pathway.
muscarinic receptor antagonist	A muscarinic receptor antagonist is a type of anticholinergic agent that blocks the activity of the muscarinic acetylcholine receptor
Protein Corona	The term 'protein corona' was proposed in 2007 to describe such spontaneous self-assembly and layering of proteins onto NP surfaces (Cedervall, T., Lynch, I., Lindman, S., Berggård, T., Thulin, E.,

	Nilsson, H., Dawson, K.A. and Linse, S., 2007. Understanding the nanoparticle–protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. Proceedings of the National Academy of Sciences, 104(7), pp.2050-2055.)
Toxicity tiered testing	Toxicity tiered testing is a set of biologically based toxicity testing decision triggers, developed and analysed within a tiered testing and decision-making framework for evaluating potential human health hazards and risks associated with chemical exposures.
Transcriptome	The transcriptome is the set of all RNA transcripts, including coding and non-coding, in an individual or a population of cells.
Zettabytes	Zettabytes are 1,000,000,000,000,000,000,000 bytes. Zettabyte is approximately equal to a thousand Exabytes, a billion Terabytes, or a trillion Gigabytes.

Workshop Organizing Committee

Dr Olivia Osborne

Ms Claire Potter

Ms Cath Mulholland

Mr Freddie Lachman

COT Members

Professor Alan Boobis (Chair)

Ms Jane Case

Dr Stella Cochrane

Dr James Coulson

Dr Rene Crevel

Dr Phil Botham

Professor John Foster

Dr Caroline Harris

Professor Gary Hutchison

Dr Sarah Judge

Professor Gunter Kuhnle

Dr David Lovell

Dr Mac Provan

Ms Juliet Rix

Professor Faith Williams

Dr Michael Routledge

Dr Cheryl Scudamore

Dr Natalie Thatcher

Professor Matthew Wright

Professor Maged Younes

Dr John Thompson

Professor Mireille Toledano

COT Secretariat

Ms Cath Mulholland

Dr Alex Cooper

Dr Barbara Doerr

Mr Barry Maycock

Dr Douglas Hedley

Ms Cleanncy Hoppie

Dr Olivia Osborne

Ms Claire Potter

Dr Joseph Shavila

Ms Chloe Thomas

Ms Sabrina Thomas

Ms Chara Tsoulli

Ms Frederique Uy

Ms Frances Hill

Food Standards Agency

2020