Organisation for Economic Co-operation and Development

ENV/JM/MONO(2020)24

Unclassified English - Or. English

8 December 2020

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

Cancels & replaces the same document of 24 September 2020

REPORT ON CONSIDERATIONS FROM CASE STUDIES ON INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT (IATA)

Fifth Review Cycle (2019)

Series on Testing and Assessment
No. 328

JT03469520

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 328

REPORT ON CONSIDERATIONS FROM CASE STUDIES ON INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT (IATA) - FIFTH REVIEW CYCLE (2019)

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 2020

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 37 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 twelve 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; Safety of Manufactured Nanomaterials;** and **Adverse Outcome Pathways**. More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (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

Forward

OECD member countries have been making efforts to expand the use of alternative methods in assessing chemicals. The OECD has been developing guidance documents and tools for the use of alternative methods such as (Quantitative) Structure-Activity Relationships (Q)SAR), chemical categories and Adverse Outcome Pathways (AOPs) as a part of Integrated Approaches to Testing and Assessment (IATA). There is a need for the investigation of the practical applicability of these methods/tools for different aspects of regulatory decision-making, and to build upon case studies and assessment experience across jurisdictions.

The objective of the IATA Case Studies Project is to increase experience with the use of IATA by developing case studies, which constitute examples of predictions that are fit for regulatory use. The aim is to create common understanding of using novel methodologies and the generation of considerations/guidance stemming from these case studies.

This document reports the learnings and lessons obtained from the review experience of the eight case studies, listed below, submitted to the 2019 review cycle of the IATA Case Studies Project. The topics discussed in this document include the strongest aspects and uncertainties of each case study. The IATA Case Studies Project has also identified a variety of areas for developing further guidance on IATA.

1. CASE STUDY ON USE OF AN INTEGRATED APPROACH TO TESTING AND ASSESSMENT (IATA) AND NEW APPROACH METHODS TO INFORM A THEORETICAL READ-ACROSS FOR DERMAL EXPOSURE TO PROPYLPARABEN FROM COSMETICS, ENV/JM/MONO(2020)16.
2. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR SYSTEMIC TOXICITY ARISING FROM COSMETIC EXPOSURE TO CAFFEINE, ENV/JM/MONO(2020)17.
3. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR 90-DAY RAT ORAL REPEATED-DOSE TOXICITY OF CHLOROBENZENE-RELATED CHEMICALS, ENV/JM/MONO(2020)18.
4. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT TO INFORM READ-ACROSS OF P-ALKYLPHENOLS: REPEATED-DOSE TOXICITY, ENV/JM/MONO(2020)19.
5. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR PREDICTION OF A 90 DAY REPEATED DOSE TOXICITY STUDY (OECD 408) FOR 2-ETHYLBUTYRIC ACID USING A READ-ACROSS APPROACH FROM OTHER BRANCHED CARBOXYLIC ACIDS, ENV/JM/MONO(2020)20.
6. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR READ-ACROSS BASED FILLING OF DEVELOPMENTAL TOXICITY DATA GAP FOR METHYL HEXANOIC ACID, ENV/JM/MONO(2020)21.
7. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR IDENTIFICATION AND CHARACTERISATION

OF PARKINSONIAN HAZARD LIABILITY OF DEGUELIN BY AN AOP-BASED TESTING AND READ ACROSS APPROACH, ENV/JM/MONO(2020)22.

8. CASE STUDY ON THE USE OF INTEGRATED APPROACHES TO TESTING AND ASSESSMENT FOR MITOCHONDRIAL COMPLEX-III-MEDIATED NEUROTOXICITY OF AZOXYSTROBIN - READ-ACROSS TO OTHER STROBILURINS, ENV/JM/MONO(2020)23.

These case studies are illustrative examples, and their publication as OECD monographs does not translate into direct acceptance of the methodologies for regulatory purposes across OECD countries. In addition, these cases studies should not be interpreted as official regulatory decisions made by the authoring member countries.

This document has been prepared by a project team of the Working Party on Hazard Assessment and was endorsed at the 4th meeting of the Working Party on Hazard Assessment in June 2020.

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

Table of contents

Tables

1. INTRODUCTION

OECD member countries have been making efforts to expand the use of alternative methods in assessing chemicals. The OECD has been developing guidance documents and tools for the use of alternative methods such as (Quantitative) Structure-Activity Relationships ((Q)SAR), chemical categories, Adverse Outcome Pathways (AOPs) and *in vitro* testing as a part of Integrated Approaches for Testing and Assessment (IATA). There is a need for the investigation of the practical applicability of these methods/tools for different aspects of regulatory decision-making, and to build upon case studies and assessment experience across OECD member countries.

The Cooperative Chemicals Assessment Programme (CoCAP)[1] was revised in 2014 to enhance the activity of the development and the application of IATA. This programme provides a forum for scientific exchange of approaches on how novel methods are applied to assess the hazard of chemicals, and establish common and best practices for the use of these methods for assessing different types of chemicals. The IATA Case Studies Project[2] was launched in 2015 under the revised CoCAP. The objective of the project is to increase experience with the use of IATA by developing case studies, which constitute examples of predictions that are fit for regulatory use. The aim is to create common understanding of using novel methodologies and the generation of considerations/guidance stemming from these case studies.

This project reviews case studies submitted from member countries every year. The review results are discussed in a project meeting. The discussion includes the topics of strongest aspects of case study, uncertainty of case study, areas for further developing guidance and possible use of each case study in a regulatory context. In every review cycle, the case studies approved will be published with a considerations document capturing the learnings and lessons stemming from case studies. The outcomes of the past four review cycles of the project (2015 – 2018) included fifteen case studies and four considerations documents, which have all been published (OECD, 2016a; 2016b; 2016c; 2016d; 2016e; 2017a; 2017b; 2017c; 2017d; 2017e; 2017f; 2018a; 2018b; 2018c; 2018d; 2018e; 2019a; 2019b; 2019c).

In the fifth review cycle (2019), the eight case studies shown in Table 1 were reviewed. The final case studies are published [ENV/JM/MONO(2020)16 -23, Series on Testing and Assessment No. 320-327]. These case studies are illustrative examples, and their publication as OECD monographs does not translate into direct acceptance of the methodologies for regulatory purposes across OECD member countries. In addition, these cases studies should not be interpreted as official regulatory decisions made by the authoring member countries. This document describes the review results of each of the eight case studies and summarises the learnings and lessons stemming from the case studies reviewed in the fifth review cycles.

[1] OECD, Cooperative Chemicals Assessment Programme (CoCAP). http://www.oecd.org/chemicalsafety/risk-assessment/oecdcooperativechemicalsassessmentprogramme.htm

[2] OECD, IATA Case Studies Project. http://www.oecd.org/chemicalsafety/risk-assessment/iata-integrated-approaches-to-testing-and-assessment.htm

Table 1. Case Studies Reviewed in the Fifth Review Cycle (2019)

No.	Title	Lead	Purpose of Use	References
1	Inform a Theoretical Read-Across for Dermal Exposure to Propylparaben from Cosmetics	BIAC	Fill a theoretical data gap for the reproductive toxicity of Propylparaben based on read-across and demonstrate the safety (i.e. a sufficient margin of exposure or MOE) of Propylparaben as used in cosmetics. Use data from *in silico*, *in vitro*, and toxicogenomics, and bioactivity data to support the biological similarity of analogues and establish potency trends to inform selection of the best source chemical from within a category.	OECD, 2020a
2	Systemic Toxicity Arising from Cosmetic Exposure to Caffeine	BIAC	Demonstrate how read-across can be applied in order to fill data gaps in an assessment of the potential risk for the consumer from exposure to caffeine by using *in silico*/*in vitro* biokinetic and toxicodynamic data as well as a physiologically-based biokinetic (PBBK) model.	OECD, 2020b
3	90-Day Rat Oral Repeated-Dose Toxicity of Chlorobenzene-Related Chemicals	BIAC	Demonstrate the use of QSAR, grouping approach, read-across approach and other *in vitro* methods to estimate 90-day rat oral repeated-dose toxicity data for some chlorobenzene-related chemicals	OECD, 2020c
4	Inform Read-Across of p-Alkylphenols: Repeated-Dose Toxicity	BIAC	Fill data gap of for the sub-chronic rat oral repeated dose toxicity for p-alkylphenols based on a grouping approach combined with use of MOA-based *in vitro* methods	OECD, 2020d
5	Prediction of a 90 Day Repeated Dose Toxicity Study (OECD 408) for 2-Ethylbutyric acid Using a Read-Across Approach to Other Branched Carboxylic acids	BIAC	Characterise the hazard of 2-ethylbutyric acid after 90-day oral exposure by using a read-across approach based on an AOP-based *in vitro* testing battery.	OECD, 2020e
6	Read-Across Based Filling of Developmental Toxicity Data Gap for Methyl Hexanoic acid	BIAC	Demonstrate that new approach methodologies (NAMs) can effectively support read across to fill the developmental toxicity data gap for 2-methylhexanoic acid by using a battery of *in vitro* tests with clear relevance to developmental toxicity.	OECD, 2020f
7	Identification and Characterisation of Parkinsonian Hazard Liability of Deguelin by an AOP-based Testing and Read Across Approach	BIAC	Assessment of an AOP-based testing strategy for hazard identification of Parkinson's disease-associated neurological effects caused by chemicals. The goal is to provide an IATA test battery for risk assessment of complex I inhibition and liability for parkinsonian motor deficits with a similar MIE as rotenoids. It is also to demonstrate that deguelin would give similar biological interactions and activation of key events as rotenone in the selected test battery that represents the key events of the AOP but with differences in potency.	OECD, 2020g
8	Mitochondrial Complex-III-Mediated Neurotoxicity of Azoxystrobin - Read-Across to Other Strobilurins	BIAC	Show that a neurotoxicity study (OECD TG 424) triggered by the potential neurotoxicity based on *in vitro* data would not be warranted based on other substances of the same class and their *in vivo* neurotoxicity data and that low neurotoxic potential of azoxystrobin mediated by inhibition of Complex III of the mitochondria be predicted based on read-across by toxicodynamic and toxicokinetic NAM data	OECD, 2020h

2. PROCESS FOR REVIEWING THE CASE STUDIES

The following 16 countries/organisations participated in the fifth review cycle: Australia, Canada, Denmark, Germany, Italy, Japan, the Netherlands, Norway, Sweden, the United Kingdom (UK), the United States (US), EU/Joint Research Centre (EU/JRC), EU/European Chemicals Agency (EU/ECHA), Business and Industry Advisory Committee to the OECD (BIAC), International Council for Animal Protection in OECD Programmes (ICAPO), Lhasa Limited.

The authors were requested to consider the templates provided in Annex 1 and Annex 2 for the documentation of the case studies. The template for the case studies on read-across was developed based on the reporting format in the OECD Guidance on Grouping of Chemicals (OECD, 2014a) and a case study document (OECD, 2014b). The general template for IATA case studies, based on building blocks, was developed in order to fit the case studies using multiple IATA components, such as AOPs / Mode of Action (MOA), Defined Approaches (DAs), Workflows, Grouping/Read-Across. Both of the templates have been continuously updated based on the review experience of the case studies in the past review cycles.

Reviewers were requested to answer the following guided questions when reviewing the case studies:

1. Is the purpose of the case study clear?
2. Are the justifications presented in the different sections sound? (e.g. hypothesis; analogue selection; justification for data gap filling; integrated conclusion; uncertainty discussion; other). If not, suggest how to improve it.
3. Are there specific topic areas in the case study that could benefit from the development of further guidance for application or interpretation? (e.g. building the hypothesis; identifying important IATA elements for the endpoint; selecting analogues; deriving integrated conclusion; uncertainty communication. etc.)
4. What are the strongest aspects of the case study?
5. What are the dominant and most relevant areas of uncertainty and how do you think they could be reduced? Could their reduction lead to a different conclusion of the case study?
6. Would you use the results of such a case study in your regulatory context? If no, why not (legislative/policy/scientific reasons)?
7. Does the template work well?
8. Other?

In addition, case study authors were requested to also answer the following guided questions:

1. Which areas of the case study was the most difficult to justify and why?
2. What information would have helped you in developing the case study?

3. Would the availability of guidance or tools in a particular area have helped you in developing the case study?
4. Would you use the results of such a case study in your regulatory context? If no, why not (legislative/policy/scientific reasons)?
5. Does the template work well?
6. Other?

The reviewer's comments and the revised case studies were discussed at the fifth meeting of the IATA Case Studies Project (19-20 November 2019), and follow-up written commenting rounds for some of the case studies, in order to finalise the case studies and summarise the learnings and lessons.

3. SUMMARY OF REVIEW RESULTS

3.1. Case Study 1: Case Study on use of an Integrated Approach to Testing and Assessment (IATA) and New Approach Methods to Inform a Theoretical Read-Across for Dermal Exposure to Propylparaben from Cosmetics

This case study aims at establishing a proof-of-concept for the value added of New Approach Methodologies (NAMs) in read across. The safety assessment for propylparaben (PP) was conducted according to the *ab initio* framework/workflow described in Case Study 2016-5 (OECD, 2017f). This approach demonstrates the use of read-across to fill a theoretical data gap for reproductive toxicity of PP and demonstrate the safety of PP as used in cosmetics.

Tier 0 involves consideration of the chemical structure of parabens and the collection of available *in silico*, *in vitro* and *in vivo* data as well as the relevant exposure information from the use of parabens in cosmetic products. Overall, docking simulations for nuclear receptors indicated a homogenous profile of weak endocrine activity. The level of external exposure to parabens was calculated to be higher than the Threshold of Toxicological Concern (TTC). Therefore, the safety assessment progressed to Tier 1 relying on a read-across approach.

In Tier 1, the chemical category was formed for read-across: the short linear chain parabens. Several estrogenicity studies evidenced low activity for parabens, and an *in vivo* reproductive toxicity study (uterotrophic assay) on Butylparaben (BP) resulted in weak estrogenic activity. The identified Point of Departure (PoD) in this study was a NOEL of 2 mg/kg/day, which was considered very conservative. Based on Tanimoto similarity only, BP was identified as the closest analogue related to the PP target compound.

In Tier 2, NAMs were considered for this case study by comparing the biological profile of the target with that of the source compounds. Available existing *in vitro* and *in silico* data were collected as well as new NAM data were generated including *in vitro* toxicodynamic and toxicokinetic data, toxicogenomics and bioactivity data. These data were used to evaluate similarity in metabolism, potential modes of action (MOA), bioactivity across the short-chain parabens, as well as to establish potency trends across the category. In addition, physiologically-based biokinetic modelling (PBBK) was used to estimate internal exposure concentrations from the read-across animal PoD and from the consumer exposure estimates to PP. Margins of Internal Exposure (MoIE) were calculated to be 258 using the Scientific Committee on Consumer Safety (SCCS) deterministic values, and 315 and 9459 for Tier 1 (deterministic) and Tier 2 (probabilistic) consumer exposure estimates, respectively.

Finally, the overall read-across justification and integrated conclusion is provided, as well as a final conclusion on the case study as an example safety assessment, informed by NAMs, for low toxicity chemicals.

Please refer to ENV/JM/MONO(2020)16 Series on Testing & Assessment No. 320 for the case study to put the following points into context.

The strongest aspects of the case study were identified as follows:

- The application of a combination of different NAMs to evaluate the absorption, distribution, metabolism, and excretion (ADME)/toxicokinetic, toxicodynamic and toxicogenomics profiles of the category members increased the weight of evidence, and further supported the read-across hypothesis justification.
- The use of a tiered exposure assessment assisted to conclude on safe use of propylparaben based on dermal risk assessment.
- The introduction, problem formulation, hypothesis and objective for the case study are well explained and the explanation of how the NAMs are used in the risk assessment is good.
- The analysis was well thought-out with reasonable assumptions, and the results and conclusions were well-described.
- The category of short chain parabens is very clearly defined so that the different compounds show rather similar properties. The selection of methods is comprehensive and the selection of compounds sound.

The main points discussed for revising the case study were as follows:

- The case study was restructured to align more directly with the tiered assessment workflow presented in the case study and moving most of the data tables and all method description to appendixes in order to make this case study clearer and make it easier to follow the IATA workflow. In addition, the case study was made more understandable by adding a summary on the supporting evidence for the hypothesis at each tier of the IATA.
- Toxicogenomics data on para-hydroxybenzoic acid (pHBA) as common metabolite of parabens has been added to show much lower effect of this metabolite on gene expression changes in breast cancer cell line.
    - In response to the comments from reviewers, the author included additional explanations or clarifications on the methods and approaches used in the case study
- Discussion on the case study's uncertainties has been expanded to the extent possible and an uncertainty table for the case study as a whole has been added. For example, additional details on the areas of uncertainty in the toxicogenomics data are now included.
- It was requested at the meeting to highlight which data and methods are the most important and which are supportive information, especially under chapter 6.2 (Conclusion on the usefulness of NAMs to support read across) as well as Table 4. (*In vivo* data matrix) This would better help to support the understanding of the weight of evidence for the category definition.
- The uncertainty analysis was performed for the NAMs used in this case study, including oestrogen receptor activity, ToxCast, ADME, toxicogenomics and PBBK.

The main uncertainties identified for the case study were as follows:

- The selected *in vivo* study ranked Klimish score 3 (non-guideline, no dose-response) would typically result in significant uncertainty. However, the author adopted this study as conservative PoD for risk assessment of BP and PP.

- There are uncertainties in the aggregate exposure assessment, which were reduced by creating a 'tiered' occurrence value, and rounding the result up. The decision was made to make conservative assumptions, such as rounding up, in the exposure assessment so as not to underestimate the consumer exposure. This approach results in uncertainty in accuracy of the exposures, which realistically would result in lower occurrence/less exposure, but assures that the risk assessment is protective. In the event that the Margin of Exposure (MoE) was not large enough, the exposure estimate could also be refined.

The main comments on the use of the case study in other member countries' regulatory context are as follow:

- ***Australia:*** Yes - we have used consideration of metabolic transformation of an applied chemical in assessments to develop MoE calculations of identified toxicants. However, the question whether the common metabolite PHBA is indeed non-toxic needs to be addressed.
- ***Canada (Health Canada):*** Aspects of this case study could be used in a weight of evidence approach or as supporting information.
- ***Canada (Environment and Climate Change Canada):*** Yes, some of the information (including NAMs) and approaches described for the paraben category can be applied to risk assessment.
- ***Japan:*** Similarity hypothesis is well justified with NAM data. That would be acceptable in our regulatory context. However, conservative safety assessment with the toxicity study data of Klimisch score 3 needs to be taken into account with more reliable test data.
- ***United States (Office of Research and Development /US EPA):*** Concern that the *in vivo* data is insufficient. However, if the butylparaben no-observed-adverse-effect-level (NOAEL) was well established, then the extrapolation made was probably fine for regulatory purposes.
- ***United States (Office of Chemical Safety and Pollution Prevention/US EPA):*** No Depending on EPA regulatory program and underlying statute(s), the information in this case study would also be subject to further weight-of the evidence evaluation with other available evidence/data with the possibility of additional data gathering activities.

Based on the experience reviewing this case study, the following areas were identified for potential guidance development:

- Clearer definition of NAM
- Guidance on how to report physiology-based pharmacokinetic (PBPK) modelling
- Guidance on how to report critical information vs supporting information
- Guidance on how to evaluate ToxCast data.

3.2. Case Study 2: Case Study on the use of Integrated Approaches for Testing and Assessment for Systemic Toxicity Arising from Cosmetic Exposure to Caffeine

The purpose of this case study is to assess the potential risk from consumer exposure to caffeine by using a read-across approach which utilised *in vivo* data from close structural

analogues while assuming that no *in vivo* repeated dose toxicity data were available for caffeine. The exposure assessment covered aggregate exposure from dermal (cosmetics) and oral (food/drink) exposure. On the basis of an analysis of the pivotal mode of action (MOA) of caffeine and other methyl xanthines, theophylline appeared to be the most potent methyl xanthine. In order to increase the confidence in the read-across approach, *in silico*/*in vitro* biokinetic and toxicodynamic data were generated and a physiologically-based biokinetic (PBBK) model was established to enable a robust estimate of the internal concentration of caffeine after both dermal and oral exposure. Based on the toxicity data for theophylline, a no-observed-adverse-effect-level (NOAEL) modelled plasma concentration was derived which was compared with the modelled plasma concentrations resulting from human dermal and oral exposure to achieve a Margin of Internal Exposure (MoIE) value for the safety assessment.

Please refer to ENV/JM/MONO(2020)17 Series on Testing & Assessment No. 321 for the case study to put the following points into context.

The strongest aspects of the case study were identified as follows:

- The strongest aspect of the case study has been the value added by NAMs (*in silico* tools, *in vitro* ADME, PBBK modelling, ToxCast, transcriptional profiling, read across) in opening new perspectives to adapt the traditional safety assessment approach using a conservative aggregate consumer exposure assessment.
- The PBBK modelling used to estimate the internal concentration of caffeine was a strength of the case study.
- The level of conservatism considered in this assessment was strong. For example, the consideration of an aggregate exposure scenario from cosmetics represented a strength in the report as this would represent a conservative, worst-case scenario. As another example, the analogue/metabolite with the lowest critical effect level (i.e., theophylline) was selected. This analogue selection was substantiated by the discussion on the relative potency factors for caffeine and its metabolites. This information provided additional confidence in the analogue selection.
- The strongest aspect of this case study is the description of the read-across analysis. The data suggested that compounds with high structural similarity show similar mode of action. On the other hand, it could be discussed how experts judge what is "similar"; for example, which *in silico* tools or similarity calculation.
- Profiling the potential endocrine activity by using multiple tools: the QSAR Toolbox, VEGA, and Endocrine Disruptome.

The main points discussed for revising the case study were as follows:

- Additional information is delivered regarding the PBBK model including model development, parameterisation methods, assumptions made, validation methods, and sensitivity analysis. Additionally, more explanations were requested regarding the difference with the PBBK modelling published by JRC.
- The transcriptomics profiling of chemicals (in well studies cell lines) was a screening method to identify potential MOA for unknown chemicals. For the sake of streamlining the case study the text has been revised accordingly, the section on transcriptional profiling approach has been removed.

- The author clarified why positive assays in ToxCast were not used by providing the explanation that many tests are only borderline active, often only the highest concentration is above baseline, or it is stated that the result is potentially confounded by overfitting.
- The author provided an explanation as to why theophylline proved to be the most appropriate analogue. However, it was requested at the meeting to add more explanation for selection of the closest analogue.
- In response to the reviewers' comments, the author added summary tables (annex 3) regarding details on the systemic and reproductive toxicity data of caffeine as well as of the analogues theophylline, theobromine and paraxanthine.
- In order to help validate the case study, it was requested to present what conclusion the use of *in vivo* data would have led to, as a comparison.

The main uncertainties identified for the case study were as follows:

- It was indicated that the one of most sensitive areas of uncertainty was the Margin of Internal Exposure (MoIE) calculation and the inputs used to derive the MoIE, including realistic exposure scenarios and the worst-case aggregate exposure (oral, dermal) to caffeine through the use of PBBK modelling. A relatively small variation in these inputs may alter the outcome of the risk assessment. In addition, other reviewers pointed out that there were no or limited explanations on the PBBK modelling such as evaluation, validation and sensitivity analyses. The author revised the case study by providing discussion and references to studies regarding the impact of inter-individual differences on the result of PBBK modelling in general in order to reduce the uncertainty.
- Only two *in vivo* studies are available for the structural analogues selected, increasing uncertainty. The availability of additional *in vivo* data for structural analogues would reduce the uncertainty. In addition, using data from minor metabolites (theophylline accounts for 4% and theobromine for 10%) was identified as one of the dominant and most relevant areas of uncertainty. The author attempted to reduce this area of uncertainty by adding the point of departure for the caffeine safety assessment on the basis of the *in vivo* theophylline data.
- It was indicated that discussion of uncertainty should include uncertainty related to the MOA since this area was a key component of the case study, and one of the areas of uncertainty the authors attempted to address with NAMs.

The main comments on the use of the case study in other member countries' regulatory contexts are as follows:

- ***Australia:*** Certain approaches described in this case study are useful in our regulatory context. This includes the use of structural similarity and metabolites and weight of evidence approach in risk assessment and the use of *in silico* tools for identifying analogues and relevant information.
- ***Canada (Health Canada):*** No. The approach found no risk for theophylline, which is a substance that is prohibited for use in cosmetics in Canada (on Canada's Cosmetic Ingredient Hotlist). Further, a comparison of the MoIE and MOE values showed that the proposed approach as less conservative than traditional approach. Therefore, it is felt that further refinement of this approach would be needed prior to its application in the regulatory context.

- ***Germany:*** Might be useful for screening purposes.
- ***Netherlands:*** The concept of using the comparison of human maximum plasma concentration (C_{max}) and NOAEL C_{max} from animal studies is certainly appealing. However, for Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) and industrial chemicals the reliability of PBBK modelling would have to be high. As toxicokinetics is not a REACH endpoint, and industrial chemicals are not measured routinely in human plasma, modelling would be difficult to validate.

Based on the experience reviewing this case study, the following areas were identified for potential guidance development:

- Guidance on Margin of Internal Exposure (MoIE)
- Guidance on how to evaluate ToxCast data.
- Guidance on application of modelling and the interpretation of the results of PBBK modelling.
- Guidance on PBBK modelling/simulation incorporating *in vitro* and *in vivo* data for chemical structures

3.3. Case Study 3: Case Study on the Use of Integrated Approaches for Testing and Assessment for 90-Day Rat Oral Repeated-Dose Toxicity of Chlorobenzene-Related Chemicals

The IATA presented in this case study illustrated data gap filling for 90-day rat oral repeated-dose toxicity for chlorobenzene-related chemicals based on a grouping approach. Chlorinated benzenes are metabolised in the liver to an epoxide form via oxidation by cytochrome P450. The epoxide is further metabolised to quinone and glutathione (GSH) conjugates, which induce hepatotoxicity and nephrotoxicity. Twelve chlorobenzene-related chemicals were identified as the category members. The data from *in vivo* studies, metabolites and physical chemical properties were gathered. A negative correlation was observed between the no-observed-adverse-effect-level (NOAEL) and logKow based on toxicological investigation. Read-across with consideration of the logKow values was performed to derive NOAELs for 3 chemicals.

Please refer to ENV/JM/MONO(2020)18 Series on Testing & Assessment No. 322 for the case study to put the following points into context.

The strongest aspects of the case study were identified as follows:

- The use of *in silico* and *in vivo* data to show that members within the group have similar metabolic pathways.
- The read-across of NOAEL values which decrease with increasing logKow.
- The mechanistic hypothesis used as the basis of the read-across and the supporting evidence.

The main points discussed for revising the case study were as follows:

- In terms of Member 1 of the category, from the available data it was concluded that there was weight loss of the spleen and no NOAEL could be determined. The NOAEL (60 mg/kg/day) derived from liver effects in member 3 was used as a PoD

based on read-across. A reviewer commented it was not necessary to use read-across for Member 1 since originally, the NOAEL for liver effects for this compound was 60 mg/kg/day. In response to this comment, author concluded not to determine a read-across NOAEL for Member 1.

- In terms of Member 2, a reviewer suggested to use kidney effect for the NOAEL instead of liver effects in member 3 as the accumulation of hyaline droplets was not seen in females but in males only and this compound did not reveal the liver effects at all according to *in vivo* study. Therefore, the author derived the NOAEL of the kidney effects from Member 5.
- It was requested to use the correct terminologies related to AOP, molecular initiating event (MIE), MOA and etc. by making reference to Annex I of the Guidance document on Developing and Assessing AOP (ENV/JM/MONO(2013)6). In addition, it was pointed out that Figure 1 is describing the metabolic pathway of chlorinated benzene, and cannot be considered a putative AOP. An AOP should be described according to the AOP Users' Handbook (ENV/JM/MONO(2016)12).
- Sub-categorisation on Group 1 was suggested based on metabolites (e.g. forming quinone or epoxide) because the different metabolites may induce different target effects or different toxicological effect level. The QSAR Toolbox can support this type of subcategorisation.
- Justification for group 2 for p-dialkoxy chlorobenzenes is not robust because there is only one source chemical and one category member (Member 15). This group likely follows a different MOA based on the results of the *in vitro* toxicogenomics testing. Since there is no additional evidence to support the hypothesis of Group 2, Group 2 should be removed from the case study.
- More details on the *in vivo* data used in the case study in Annex 1 was required so that the reviewers could check the reliability of the *in vivo* data.

The main uncertainties identified for the case study were as follows:

- The uncertainty of read-across for Group 2 members is very high due to the inconsistencies of the toxicogenomics data, the low number of category members and only one member with *in vivo* data for repeated dose toxicity. As Member 15 was shown to be an outlier with respect to the chosen MOA/AOP pathway, more substances should have been included in Group 2 to strengthen the conclusions, or more research should have been done to support why this substance differs from others in the grouping. In response to the comments, the author has deleted Group 2 from the case study.
- Discussion on the correlation between logPow and the NO(A)EL for Group 1 chemicals (variously substituted chlorobenzenes) was uncertain. Multiple metabolites with different mode of actions could be involved in their toxicity to liver and/or kidney. Representing such complex toxicity by simple logPow may not be sufficient. To represent it, additional and stronger scientific evidences (e.g. major toxic metabolite(s), their amount, distribution, clearance, etc.) need to be discussed, otherwise the uncertainty remains very high. In addition, the other largest source of uncertainty for Group 1 is the issue of the applicability domain of the semi-quantitative QSAR approach. For target chemical 12 with the highest logKow, the NOAEL was defined as >0.34 mg/kg (the lowest reported NOAEL in

the chemical group). This suggested that the defined chemical group/category is insufficient to provide a prediction for all target chemicals (i.e. target chemical 12).

The main comments on the use of the case study in other member countries' regulatory contexts are as follows:

- ***Australia:*** Some of the approaches described in this study would be useful in our regulatory context. However, the application of these approaches as such may be limited by the resources required and the urgency to complete the assessment sooner. The other concepts described in this case study can be explored.
- ***Canada (Health Canada):*** The authors have used existing *in vivo* studies to fill data gaps for substances that do not have these types of studies. The authors use the logKow as their measure of similarity between substances. Health Canada would use additional data to support the identification of analogues). At the time of review, it was not all always clear how all of the data presented in the case study fed into the final conclusion.
- While this approach would not be sufficient to assess safety on its own, aspects of this case study could serve as one of the data inputs in our regulatory context.
- ***Italy:*** Yes, especially concerning Group 1. The case study on Group 2 could be used as an element of a weight of evidence approach.
- ***Japan:*** No, discussion of metabolic activation pathways for both groups of chemicals could be insufficient. It is clear that number of metabolites having different AOPs are involved in the resulting toxicity in combination. Some metabolites are also expected to be distributed in different organs, which makes the prediction of toxicity rather complicated.
- ***Netherlands:*** The description of the science is sound and for this reason the case study might be useful for REACH registration. However, for REACH, the reporting of the results and justification as presented in the case study would not be sufficient to support read-across. The description of the applied methodology is not available at a sufficient level. Robust study summaries and reliability (Klimisch scores) are not provided.
- ***United States (Office of Research and Development /US EPA):*** Based on the regulatory needs of the particular program partners, the results and conclusions of the case study would be insufficient. A derivation of toxicity values for the target chemicals is required, which involves a more comprehensive analysis of the toxicity profile (beyond a 90-day rat study) of the target and analogues and a quantitative characterisation of uncertainty.
- ***ECHA:*** By putting all chlorobenzenes together (with differences in effects), and because there are not sufficient and consistent high-tier bridging information between the subgroups of the sources and the targets, the hypothesis is related to considerable uncertainty. It is agreed there is uncertainty with the toxicokinetics, not only because for Group 2 experimental toxicokinetic data are missing but also between the chlorobenzenes there seems not to be sufficiently clear. The interpretation of the alternative data is controversial. There is not a sufficient level of detail in some reported data, e.g. liver and kidney effects cannot be distinguished from some of the provided summaries. Also there are terms like semi-QSAR from one chemical that is used and such terminology is unclear and possibly covers

unjustified read-across from one substance. The 1,4-dichlorobenzene has harmonised classification for carcinogenicity (Carc. 2), and hexachlorobenzene has harmonised Carc. 1B classification. These aspects probably have to be checked before determining the PoD for general toxicity, because the PoDs might be underestimated in this study. One of the target substance is a pesticide, and we suggest that authors seek alignment with pesticides regulations. Some data might be useful in screening assessments or for classification.

Based on the experience reviewing this case study, the following areas were identified for potential guidance development:

- Selection of primary target toxicity endpoint for read-across
- Category justification based on common or similar toxicity mechanism
- Guidance on the metabolic similarity assessment

3.4. Case Study 4: Case Study on the Use of Integrated Approaches for Testing and Assessment to Inform Read-across of p-Alkylphenols: Repeated-Dose Toxicity

This case study was developed to illustrate how to use *in vitro* methods for biological similarity assessment in the category approach for data gap filling that may be used in a regulatory context in member countries. The purpose of the IATA presented in this case study is to fill a data gap for sub-chronic rat oral repeated dose toxicity for p-alkylphenols based on grouping approach combined with use of MOA-based *in vitro* methods. It is well known that p-alkylphenols induce hepatotoxicity by formation of reactive metabolite quinone methide (QM) and its binding to nucleophiles such as glutathione (GSH). The category member were selected based on the alkyl chain substituted at the position 2, 4 and 6. Various *in silico* tools were used to simulate structural alerts, absorption, and metabolism. *In vitro* testing was performed to investigate the amount of dansylated glutathione (dGSH) adducts formed, and how this affected the cytotoxicity *in vitro*. Although *in silico* alerts for the rate of intestinal absorption were found to be similar among all category members, *in silico* alerts for hepatotoxicity were found to be different between the QM forming category members and the no-QM forming category members. However, the amount of dansylated (dGSH) adducts and cytotoxicity decreased according to the elongation or structural hindrance at the 4-position. This data suggests that QM-induced hepatotoxicity would be prevented by structural hindered substituents at the 4-position. Considering the chemical reactivity of substituents, the lowest NOAEL values were applied to target members which had more complex structure than source members.

Please refer to ENV/JM/MONO(2020)19 Series on Testing & Assessment No. 323 for the case study to put the following points into context.

The strongest aspects of the case study were identified as follows:

- The authors demonstrated with appropriate references that the quinone methides produced at position 4 form the basis of hepatotoxicity. The different *in silico*, *in vitro*, and *in vivo* results relate back to this mechanism. The use of metabolic profiling and *in silico* tools for metabolite analysis demonstrate the utility and practical application of new approach methodologies (NAMs) in strengthening the weight of evidence for the selection of source analogues and the overall read-across justification. Furthermore, targeted assays to interrogate the production of QMs and subsequent cell viability significantly strengthen the arguments made, and round out the analysis.

- The authors divided the alkylphenols into sub-categories based on structural similarity (position 2 and 6 alkyl groups) and derived an appropriate worst-case NOAEL for each category.
- The author provided clear explanations of the *in vitro* and *in silico* methodologies used in the case study. For example, the table regarding models used for prediction and justification in this read across is very clear. It would be really helpful if this kind of table could be made for all sections of the case study in which data/tools are applied/used.
- "For the category approach to read-across, category members must be grouped not only according to structural but also chemical, biological, and toxicological similarities" – this is a great premise to follow.

The main points discussed for revising the case study were as follows:

- In response to the comments, the author revised the case studies as follows:
  - Added a workflow figure to make the read-across approach clear.
  - Provided more details on the category definition.
  - Added a figure describing the putative AOP on hepatotoxicity induced by p-alkylphenols to provide an overview of hepatotoxic mechanism of p-alkylphenols.
- Regrouping the category members was requested at the meeting. Justification of the category definition for category members having a long chain or a branched chain at position 4 as well as category members for group 3 is not enough because these chemicals are not metabolised to quinone methide or do not form GSH conjugate. It was suggested to exclude these chemicals from the category members.

The main uncertainties identified for the case study were as follows:

- Making the foundation of this category sub-grouping based on substitution at position 4 and not at 2, 6 was proposed. According to the results of the dGSH trapping test, compounds with straight and short hydrocarbon chain at position 4 seem to be oxidised most easily and therefore seem to show the highest/most probable hepatotoxicity. However, the hepatotoxicity of 2,4,6-Tri-tert-butylphenol cannot be explained by the proposed mechanism of action since QM cannot be formed due to the quaternary carbon atom at position 4.
- The largest uncertainty was the lack of *in vivo* and *in vitro* data for target compounds.
  - Regarding *in vivo* data, *in vivo* reference values are missing for the longer alkyl chain substituents compound in Group 1. Groups 2 and 3 contain a limited number of category members. Moreover the groups comprise both QM and not QM forming members.
  - Regarding *in vitro* data, only 7 chemicals were tested and the reasoning was never specified. In addition, No. 2 (2,4-dimethyl-6-tert-butylphenol shows reduced cell viability. The overall results do seem to support the methyl group at position 4 to be the most reactive, and therefore, choosing NOAELs based on these source compounds can be considered as protective. Hence, reduction of this uncertainty may not lead to a different conclusion as NOAEL is already conservative. However, some

justification for the selection of target compounds would be beneficial. According to the available *in vivo* data, however, the NOAELs cannot be correlated with the substituent at position 4.

- A key uncertainty is in the domain of applicability of the (Q) SARs, particularly for metabolite predictions (e.g. are p-alkylphenols well represented in the training/validation sets).

The main comments on the use of the case study in other member countries' regulatory contexts are as follows:

- ***Australia:*** Yes, the results reflect a sufficiently detailed but conservative approach and the justifications for the approach appear sufficiently robust for regulatory decision-making.
- ***Canada (Health Canada and Environment and Climate Change Canada):*** Yes, aspects of this case study could be used in the Canadian regulatory context. The reasoning of the read-across is sound and a precautionary approach is taken to mitigate uncertainty in the absence of *in vivo* data on target compounds.
- ***Germany:*** No, the worst case scenario is rather uncertain since the *in vivo* results do not correlate with *in vitro* results.
- ***Italy:*** Yes, but the robustness of the approach needs to be improved.
- ***Netherlands:*** We proposed to adjust the hypothesis; therefore, it is difficult to say if this case study could be used for REACH. However, the quality of reporting, choice of the methodologies and general presentation of the case are very good, and could be described as almost reaching REACH level.
- ***United States (Office of Research and Development /US EPA):*** Based on the regulatory needs of the particular program partners, the results and conclusions of the case study would be insufficient. A derivation of toxicity values for the target chemicals is required, which involves a more comprehensive analysis of the toxicity profile (beyond a 90-day rat study) of the target and analogues and a quantitative characterisation of uncertainty
- ***United States (Office of Chemical Safety and Pollution Prevention/US EPA):*** No. This is not a validated method. In addition, although the IATA presents a plausible hypothesis, potential competing hypotheses have not been clearly addressed.
- ***ECHA:*** There might be scientific reasons to consider, for example the fact that no other toxicity but hepatotoxicity is investigated. The provided experimental data originates from 28 and reproductive/developmental screening studies, so the systemic toxicity could be underestimated due to lack of higher tier studies. Some data might be useful for screening assessments.

Based on the experience reviewing this case study, the following areas were identified for potential guidance development:

- Guidance on how to deal with compounds that do not fit into hypothetical groupings.
- The dGSH trapping test seems an interesting option for checking GSH reactivity *in vitro* and for finding reactive metabolites.

- The assessment of metabolic similarity.

3.5. Case Study 5: Prediction of a 90 day Repeated Dose Toxicity Study (OECD 408) for 2-Ethylbutyric Acid Using a Read-Across Approach to Other Branched Carboxylic Acids

This case study characterised the hazard of 2-ethylbutyric acid (2-EBA) after 90 days oral exposure by using a read-across approach, with assuming that 2-EBA has to be registered under REACH. Nine aliphatic carboxylic acids with different branched aliphatic side chains are regarded as most similar to the target compound. Two compounds, 2-ethylhexanoic acid (2-EHA) and valproic acid (VPA), have *in vivo* animal studies with repeated oral exposure. The studies show that the two chemicals induce hepatotoxicity. The read-across hypothesis is therefore, that 2-EBA is a liver toxicant with special concern for steatosis. In addition to the nine structural analogues, pivalic acid (PVA) is tested as negative control compound. PVA did not induce any liver toxicity in a subacute study up to the highest tested dose. A negative compound was needed as a negative control for the NAM data.

In this read-across assessment, an AOP network is available for liver steatosis, the primary toxicological effect. From this AOP network, molecular initiation events (MIEs) and one key event (KE) were tested to see if the grouped compounds might induce this adverse outcome. In addition, the perturbation of general biological pathways and cellular functions were tested together with cytotoxicity to identify potential major differences between the grouped compounds. The data from *in vitro* testing showed that with increasing chain length, the number of activated MIEs related to steatosis increased and that lipid accumulation was mainly observed for long chain analogues, whereas short-chain analogues remained inactive. On other hand, no difference was observed within the grouped analogues with regard to cytotoxicity (in liver and kidney cells). The target compound 2-EBA activated only one MIE, PPAR-α. It can therefore not be excluded that a pathway towards lipid accumulation is activated by 2-EBA. However, in the *in vitro* assay regarding the lipid accumulation, 2-EBA was inactive. In the results of the *in vitro* assay, 2-EBA was less toxic then the two liver toxic analogues with *in vivo* animal data, 2-VPA and 2-EHA, which indicated that 2-EBA will not induce liver steatosis up to the highest tested *in vitro* dose. Kidney models did not show any difference in cytotoxicity, with all compounds being of general low cytotoxicity. Also endpoints measuring the perturbation of general biological processes like mitochondrial membrane potential or cytotoxicity support the trend of higher activity with longer side chain length.

Human PBPK models were established for all read-across compounds based on physiochemical properties and *in vitro* data. Human *in vivo* pharmacokinetic data for VPA was identified and verified for good predictive performance based on observed plasma concentration data in humans. Based on this proof of concept, *in vitro* to *in vivo* extrapolation (IVIVE)-PBPK models were used for *in vitro* to *in vivo* extrapolations for all analogues.

The 10th percentile of the most sensitive *in vitro* endpoint of 2-EBA was used to derive an oral equivalent dose. Quantitative IVIVE (QIVIVE) resulted in an oral equivalent dose of 730 to 948.6 mg/kg bw/d for rats, which can be used to fill the data gap of a subchronic toxicity study. Furthermore, QIVIVE was used to determine directly a corresponding human oral equivalent dose, which is 138 mg/kg bw/d. Below this threshold, a risk for humans to develop liver steatosis or general liver/kidney toxicity is not expected.

Please refer to ENV/JM/MONO(2020)20 Series on Testing & Assessment No. 324 for the case study to put the following points into context.

The strongest aspects of the case study were identified as follows:

- Hypothesis building for the toxicity of a target chemical was systematic and scientifically reasonable.
- Selection criteria for category members were strict and systematic (e.g. one branched position at 2, side chain length from C1 to C5, etc., Table1). Hence, chemical variations were minimal and led to consistent chemical and biological trends. A negative control was included to support the accuracy of NAM data.
- The well-defined exposure related effects (i.e. liver hypertrophy, steatosis and cytotoxicity) and the known AOP for steatosis allowed for MOA determination with high confidence.
- The attempt of the authors to gather and to integrate all possible information (physico-chemical properties and other molecular descriptors; Toxicokinetic data - ADME properties by NAMs, Toxikodynamic data - Hazard characterisation by NAMs). For example, the study systematically evaluated MIE profiles of the target compound using *in vitro* NAM data based on the AOP network on liver steatosis (lipid accumulation), and assessed the human toxic threshold based on the human oral equivalent dose (hOED) derived from the PBPK model established in the study.
- The case study shows that activation of an MIE might not be sufficient to determine if the entire pathway will be triggered. A strength of this case study is that the authors conducted targeted assays to follow-up on the MIE that they observed to be activated (i.e., PPAR-α and lipid accumulation assays).
- The establishment of rat and human PBPK models, following the WHO PBPK guidance is a strength. The rat model was used and used to calculate plasma and target organ concentrations. The human model was verified against *in vivo* data for VPA and was used for IVIVE for all analogues.
- Assessment of the uncertainty for each component supports the read-across with high confidence.

The main points discussed for revising the case study were as follows:

- The rational for the selected testing of one key event and several MIEs was added in more detail. Also, it was requested to describe in an appropriate section (e.g. summary, introduction or conclusion) which subset of tests for key events, would be sufficient to evaluate the repeated dose toxicity in this case study.
- Highlighting the testing corresponding to the MIEs/KEs in the AOP in Figure 6 was requested for improved communication.
- More description on AOP uncertainty was requested as these AOPs have not been endorsed.
- A step by step flow diagram of the overall approach to read-across was introduced to explain the workflow.

The main uncertainties identified for the case study were as follows:

- The highest uncertainty pertained to the number of analogues with a lack of suitable *in vivo* endpoint data.
- Metabolites are assumed not to be toxic and are not considered in the PBPK model. This assumption could lead to an underestimation of toxicity.
- The liver toxicity endpoint selected (i.e. liver hypertrophy) from the animal study often has a large sensitivity difference between rodents and humans. However, the author addressed this uncertainty by the use of human *in vitro* data as support for reading across the *in vivo* data.
- The AOP-based network has not been endorsed. Therefore, there are some remaining uncertainties about the individual MIEs and KEs, as well on the completeness of the AOP network.
- The uncertainty of the read-across justification for kidney toxicity is higher compared to that for liver steatosis. This uncertainty can be reduced by adding transcriptome data.
- Uncertainties regarding the *in vitro* methods used were identified. The author attempted to reduce the uncertainties by providing a stronger justification for the *in vitro* methods used for the MIE/KE's measured and on the biological plausibility and essentiality of the measured KE.

The main comments on the use of the case study in other member countries' regulatory contexts are as follows:

- ***Australia:*** Yes the approaches used in this case study could be used.
- ***Canada (Health Canada):*** Yes these approaches could be used for the assessment of Existing Substances. However, all endpoints would need to be considered for a risk assessment, not just systemic toxicity following a 90-day exposure. Therefore, this approach could be used together with additional data streams.

  Canadian NSN Regulations currently permit the use of such alternative data if it can be proven to be equally or better suited to measure the endpoint under investigation. However, it is uncommon to have a substance set as structurally similar as in this case study.
- ***Italy:*** The results could be used for screening purposes. Once the AOPs for hepatic steatosis will be fully endorsed by the OECD, the results of this case study could be very useful in our regulatory context.
- ***Japan:*** Yes, for screening level assessment. The way of selecting category members (how to build criteria) in the study will be useful for a read-across in general. It could reduce chemical and biological variations and make their trends relatively consistent. However, in order to accept the final results, it would be necessary to build a team of experts of PBPK, *in vitro* toxicology, toxicity mechanisms and *in vivo* toxicology and experts who can interpret them in an integrated manner.
- ***Netherlands:*** This case study cannot be considered for submission for REACH, since it focuses only on liver toxicity. We have reviewed the case study as a case study for this specific endpoint only and in that sense, this case study serves as a proof-of-concept and may provide some interesting recommendations for reporting future read-across cases.

What is missing in the case study is support for the hypothesis that liver steatosis is the critical, and only relevant, endpoint for repeated dose toxicity. In the discussion of the (limited) *in vivo* data kidney toxicity is mentioned. The case study therefore cannot be seen as sufficient to waive repeated dose testing, but might offer enough information to support that the liver toxicity (as a part of repeated dose testing) is likely to be critical for 2-EBA.

Based on the experience reviewing this case study, the following areas were identified for potential guidance development:

- Use of positive and negative reference compounds
- Targeted testing focusing on biological read-across setting
- NAM testing mapped to either endorsed or non-endorsed AOPs and documentation required for using non-endorsed AOPs
- How to report a new model, such as Dempster Schaefer Theory
- How to report PBBK modelling
- How to consider effects on organs other than the target organ in the context of the case study

3.6. Case Study 6: Read-across Based Filling of Developmental and Reproductive Toxicity Data Gap for Methyl Hexanoic Acid

This case study was developed to fill the data gap of developmental and reproductive toxicity (DART) for 2-methylhexanoic acid (MHA). Structural analogues with DART data were identified in order to explore the possibility to read across information of these source chemicals to MHA. Eight structurally related aliphatic carboxylic acids were selected that have *in vivo* developmental toxicity data. Four analogues proved to be clear developmental toxicants, while others were identified as not being toxic to development. Since structural similarity alone doesn't allow for the derivation of a conclusion on the developmental toxicity of MHA, MHA and all the selected source chemicals were tested in a battery of *in vitro* tests with clear relevance to developmental toxicity. These results were combined with toxicokinetic models to calculate effective cellular concentrations and associated *in vivo* exposure doses. With these new approach methodologies (NAM), the case study was exploring whether they could correctly predict the *in vivo* developmental toxic properties of these aliphatic carboxylic acids, and thus could be used to predict the *in vivo* developmental toxicity of MHA itself. This data would also allow to further explore the relationship between structure and developmental toxicity within this series of aliphatic carboxylic acids. For that reason 2-methylpentanoic acid (MPA) has also been tested, despite the absence of *in vivo* data. In addition, the potential to inhibit histone deacetylase activity in Zebrafish Embryo Test (ZET), Embryonic Stem cell Test (mEST), and iPSC-based neurodevelopmental (UKN1) model has been investigated, as this enzyme is postulated to be the molecular initiating target leading to neural tube defects observed with these analogues.

The NAM results showed that the analogues with *in vivo* developmental and/or reproductive toxicity data were correctly predicted. The NAM results suggest that MHA may not be fully negative for developmental toxicity, while MPA is predicted as negative.

Please refer to ENV/JM/MONO(2020)21 Series on Testing & Assessment No. 325 for the case study to put the following points into context.

The strongest aspects of the case study were identified as follows:

- Hypothesis was well outlined.
- Very clear selection of source and control compounds, representing both positive and negative outcomes of *in vivo* testing.
- Development and use a battery of complementary *in vitro* and *in silico* methods for addressing the data gap for developmental toxicity
- General concordance between *in vitro* and *in vivo* (mouse exencephaly model) results.
- Use of the Dempster-Shafer Theory (DST) approach to provide a weight-of-evidence (WoE) estimate from different *in vitro* assays. Overall, this method could help dealing with uncertainties in chemical hazard assessment, and also support the combination of multiple sources of evidence to obtain a weight of evidence conclusion.

The main points discussed for revising the case study were as follows:

- It was indicated that the scope of this case study is on developmental toxicity, not reproductive toxicity. Therefore, the author refocused the text more on developmental toxicity. However, since there is still a lot of text regarding reproductive toxicity in the report in the revised version, it was requested to update wording in the report, particularly on the Chemical Activated LUciferase gene eXpression (CALUX) assays.
- Additional justification was provided to explain the difference of *in vivo* response or potency observed/predicted across the category, including the discussion on the relationship between structure and developmental toxicity of the aliphatic carboxylic acids chosen for the read-across.
- Since there is still uncertainty as to the exact AOPs that is operational, linking *in vitro* assays to MIE and /or KEs of these AOPs is not quite justified yet. While the author added two postulated AOPs submitted to AOP-Wiki, one for craniofacial development (#274), and one for neural tube defects (#275), both having histone deacetylase (HDAC) inhibition as the MIE, other parameters /AOPs might also be involved. However, there was a further request to include discussion on AOPs, especially an explanation on why the case study did not directly use the putative AOPs. There was some confusion on the AOP context because the information was used initially as a way to choose the assays for the case study.

The main uncertainties identified for the case study were as follows:

- The category is small as there are only four source chemicals and only one of these is negative. Therefore, read across by interpolation is less likely to be robust. As a trend for potency is being established in this case, a larger number of source chemicals and more than just one negative source chemical may be required to justify the assignment of a low uncertainty to this aspect. The author addressed the uncertainty by including negative controls in this case study, i.e. structurally related (yet no category member) with *in vivo* data, such as pentenoic acid (PA), dimethyl pentanoic acid (DMPA).

- For the four source chemicals, the non-standard assay of mouse exencephaly-induction is the only *in vivo* assay available to score for response and potency. As only two of the four source chemicals (both positive) have data from OECD guideline or equivalent studies of developmental toxicity, this reduces confidence in the predictive power of the approach for the category. The author added discussion on the uncertainty related to this category restriction to 2-branched aliphatic carboxylic acid structures.
- There is a degree of uncertainty concerning the enantiomeric composition of the test substances used in the mouse exencephaly model. The remark was added regarding asymmetric carbon atoms for some aliphatic carboxylic acids.
- Limitations in the *in vitro* assays as described: limited metabolic capacity (mEST, UKN1, CALUX reporters), specificity and sensitivity (ZET reporter, UKN1). For example, in initial version, how limited metabolic capacities regarding mEST and ZET would impact the read-across assessment is not discussed. In response to this comment, the author added this indication to the uncertainty section.
- In some of the *in vitro* tests (ZET jitter/tremor, ZET scoliosis/lordosis, mEST) results for MHA (predicted negative) are not significantly different to those for EHA (positive). From this, it is rather difficult to draw a final conclusion for MHA.
- To further strengthen the read-across it would be useful to discuss why the postulated AOP pathway was considered most relevant, as this might not be the only AOP that could lead to (neuro)developmental toxicity. It would also be useful to present the postulated adverse outcome pathway more upfront, including the MIE, KEs and AO and to link the different methods / assays to these events, also in a graphical presentation. However, there still is quite some uncertainty as to the exact AOP that is operational. Therefore, linking *in vitro* assays to MIE and /or KEs of these AOPs is not quite justified yet.

The main comments on the use of the case study in other member countries' regulatory contexts are as follows:

- ***Australia:*** The approach described in this study could be used for screening and prioritising of chemicals for potential development toxicity in our regulatory context although it would be resource intensive.
- ***Canada (Health Canada):*** The results were found to be sound, well-arrived, and supported by multiple means. As such, the results could be used as part of the weight of evidence in the Existing Substances regulatory context.

  Yes, the New Substances Assessment and Control Bureau of Health Canada could use these results in the current regulatory context (New Substances Notification Regulations). The information could be used as a qualitative line of evidence (i.e., not expected to cause certain developmental effects/malformations, uncertainty remains if lesser effects would be seen) in relation to the target compound.
- ***Netherlands:*** This case study aims to fill the data gap for MHA for developmental toxicity (OECD Test No. 414: Prenatal Developmental Toxicity Study) for REACH, using structurally similar chemicals that have this developmental properties information.

  The *in vitro* and *in vivo* data indicate that MHA is not a very potent developmental toxicant and unlike VPA will most likely do not cause neurodevelopmental toxicity

due to inhibition of HDAC. However, based on the aspects mentioned in the scientific review results we are not convinced that for MHA any developmental effects can be excluded and could still be a weak developmental toxicant. Therefore, from a regulatory perspective it might therefore still be questioned whether an *in vivo* study is still not warranted.

Based on the experience reviewing this case study, the following areas were identified for potential guidance development:

- How to report PBBK modelling
- Criteria to choose the relevant and necessary information needed to support read-across from all available data
- Guidance on how to use data from NAMs to support the justification for the read-across.

3.7. Case Study 7: Identification and Characterisation of Parkinsonian Hazard Liability of Deguelin by an AOP-based Testing and Read Across Approach

The aim of this case study was to assess the application of an AOP approach in a read across safety assessment of structurally related mitochondrial complex I inhibitors, deguelin and rotenone. The AOP relied on in the case study was published by the OECD as Series on Adverse Outcome Pathways No. 7 (OECD, 2018f). Rotenone was selected as the source chemical since rotenone is known as a chemical which induces Parkinson phenotypes. Whether deguelin has such a parkinsonian hazard liability in humans is currently unclear and therefore deguelin was the target substance for this case study.

Based on the AOP, structural modelling approaches and multiple human-based *in vitro* test systems were established and applied to assess effects by probing the molecular initiating event and various key events. Both biokinetic evaluation of cellular exposure to rotenone and deguelin as well as PBPK modelling have been used to evaluate the relevance of the observed effects *in vitro* towards a likely *in vivo* exposure situation. Results from *in silico* docking, *in vivo* metabolism profiles and biokinetic behaviour indicated that deguelin has a similar MOA as rotenone. However, deguelin is less potent.

It was concluded that the integration of specific technologies and test systems that are mapped to test the MIE and KEs of a specific AOP, might find broader application in a read across safety assessment of structurally related substances.

Please refer to ENV/JM/MONO(2020)22 Series on Testing & Assessment No. 326 for the case study to put the following points into context.

The strongest aspects of the case study were identified as follows:

- The use of a well outlined scientific hypothesis, combined use of *in silico* and *in vitro* NAM and the application of an endorsed and published AOP-based approach to conduct a read-across hazard assessment of structurally related mitochondrial complex I inhibitors.
- Using these assays, the authors provided compelling reasoning for the use of rotenone data (with respect to Parkinsonian effects) for read-across to deguelin as a conservative approach. This strategy provided a strong rationale that future *in vivo* testing in conventional animal toxicity studies would not be warranted.

- The authors provided excellent summaries of the complex, AOP-based bioassays within the report, which provided sufficient context to scientists of all levels of expertise. In addition, the following sections were comprehensive, thorough, and relevant:
    - Endpoint(s) for which the read-across is performed
    - Kinetics: Absorption, distribution, metabolism and excretion
    - Chemical/biological interaction
    - Responses found in alternative assays (e.g., experimental (NAM) data, *in silico*).
    - Data and methods
- Very interesting concept of read-across based on the modelled interactions of target and source chemicals´ 3D molecules with mitochondrial complex I and supported by NAMs and *in vivo* data
- The case study highlighted that rotenone and deguelin are highly similar in structure, metabolism, toxicokinetics, phys-chem properties and biological effects.

The main points discussed for revising the case study were as follows:

- It was suggested to include other Complex I inhibitors to probe if other structurally similar compounds would have the same MOA. The author responded that other complex I inhibitors that are structurally dissimilar from rotenone have been tested. However, the information was not included in the case study due to their structural dissimilarity.

The main uncertainties identified for the case study were as follows:

- The case study utilised a large number of approaches including *in vitro*, *in silico* model, PBPK modelling, and also an AOP-based testing strategy. While it is likely that these approaches would have contributed to the uncertainty, it is noted that the case study has addressed all the uncertainty elements appropriately and the overall level of confidence in the read-across approach has also been reported as being of low uncertainty.
- As noted by the authors, only one source compound was used, and therefore it was not possible to define inclusion and exclusion criteria for when structural similar compounds would have the same mode of action. A difference in potency was also noted between the target and the source compounds, thus introducing uncertainty. But there was consistency of effects, and a high degree of similarity was also observed between the two chemicals. In this way, the confidence in the read-across prediction was increased.
- Reliability and validity of the *in vitro* bioassays. Uncertainty can be reduced by the use of validated assays but this requirement would undermine the utility of novel methods. If other positive chemicals as well as negative chemical had been used for each *in vitro* method, the results of the *in vitro* approach would be more reliable. However, the reduction of this uncertainty is not anticipated to lead to a different conclusion of the case study.
- Consideration of data quality and the applicability domain of *in silico* models were considered to be uncertainties in the case study.

- There was a concern regarding the influence of the choice of solvent on determination or prediction of the relative strength of deguelin vs rotenone. The author added an explanation about the limited influence by dimethyl sulfoxide (DMSO) on the result of the *in vitro* testing in the uncertainty table.
- Lack of KE5 testing. KE5 represents a key event in the AOP and having no data or testing on this endpoint introduces uncertainty in the assessment. The author added an explanation as to why KE5 was not tested and a statement that the lack of this measurement would not hamper the overall conclusions.
- There are uncertainties regarding the hypothesis that rotenone induced Parkinson disease (PD) induces exclusively an inhibition of mitochondrial complex I. It has been identified that microtubule depolymerisation and the accumulation of cytosolic dopamine and reactive oxygen species as alternative mechanisms underlying rotenone-induced dopamine neuron death. Rotenone induced PD is likely a combination of disrupting microtubule dynamics and inhibiting complex I. In order to address the uncertainty, the author explained that literature data indicates that disrupting microtubule dynamics take place at higher concentration than inhibiting complex I. Hence it was assumed that these are off-targets effects that rather take place at higher concentration and therefore bear little relevance for the current case study results and conclusions

The main comments on the use of the case study in other member countries' regulatory contexts are as follows:

- ***Australia:*** Some of the approaches described in this study may be useful in our regulatory context. However, we are unable to comment on the application of the approaches in the pesticide area as pesticides are regulated under a different scheme.
- ***Canada (Health Canada):*** Yes, in the presence of the appropriate AOP, this approach could be used. However, the ability to generate this type of data in-house would be limited. Further, in order to perform a risk assessment, a full analysis of all endpoints would be needed to determine the critical effects.

  This approach could be used if there was a question of a substance inhibiting complex I of the mitochondrial respiratory chain. However, there is less confidence in the potency estimation and it does not provide a numerical point of departure. It is also highly specific to one class of neurotoxic compounds, so a negative result would not necessarily indicate negative neurotoxicity.
- ***Denmark:*** The results could be used as a part of weight of evidence
- ***Netherlands:*** In the current form, this case study would be very difficult to be used for REACH screening purposes. All test results were to prove the hypothesis but not a read-across approach to fill in data gap or to conduct an assessment.

Based on the experience reviewing this case study, the following areas were identified for potential guidance development:

- Guidance on reporting docking/modelling
- How many key events should be tested.
- How many assays would be needed for testing of each key event

- Use of endorsed or non-endorsed AOPs to inform/guide NAM testing in read-across and the value of testing structural analogues that act through the same pathway
- Use of biological based read-across.

3.8. Case Study 8: Mitochondrial Complex-III-Mediated Neurotoxicity of Azoxystrobin - Read-Across to Other Strobilurins

The scientific hypothesis of the case study was to confirm whether the absence of a neurotoxic potential (as detected with a TG424 study) mediated by inhibition of Complex III of the mitochondria can be predicted by toxicodynamic and toxicokinetic NAM data. It was known that the strobilurins bind to the quinol oxidation site of cytochrome b of complex III (CIII) of the mitochondria, which may mediate some signals of potential neurotoxicity. Azoxystrobin was selected as the target chemical. The formation of the category is based on the hypothesis that the compounds share the similarity on chemical structure, mode of action, toxicophore, neurotoxic potential and toxicokinetics. The 4 other strobins showing no signs of neurotoxicity were chosen as source compounds. Furthermore, Antimycin A, a well-established CIII inhibitor with neurotoxic effects, served as a reference compound for this mode of action.

The overall structural similarity of the compounds is low, however, they have the same pesticidal mode of action and toxophore. The hypothesis is supported by mechanistic data, anchored to a putative AOP (based on the recently OECD adopted AOP on CI inhibition leading to parkinsonian disorder (OECD, 2018f)), and kinetic PBTK data. Based on the putative AOP, inhibition of CIII complexes, effects on membrane potential and effects on glycolysis and cell viability were measured. The generated effect data showed that there is no evidence for a stronger neurotoxic potential of azoxystrobin mediated by a complex III inhibitory mode of action as compared to the source compounds. The kinetic data and simulations confirmed comparable kinetics and that the exposure of the brain to the strobilurins is limited being approx. twice the plasma concentration.

Since the source compounds do not show neurotoxicity *in vivo*, it is concluded that also the target compound azoxystrobin is not a neurotoxicant.

Please refer to ENV/JM/MONO(2020)23 Series on Testing & Assessment No. 327 for the case study to put the following points into context.

The strongest aspects of the case study were identified as follows:

- The *in vivo* ADME data showing that the strobilurins have low presence in the brain.
- The detailed explanation of the PBPK model
- This case study is a good example describing the opportunities for waiving neurotoxicity and to elaborate what information is needed to do so.
- This is a focused IATA specific to strobilurin fungicides which is anchored to a putative AOP.
- The layout of this case study and the underlying evidence put forth to support the hypothesis are presented in a readable and easy to follow manner.

- Category members share a common toxicophore, and the same mode of action (mitochondrial complex III inhibition), therefore the confidence in the read-across was increased;
- Consideration of a positive reference compound was useful for supporting the AOP, and for comparing the results of target and source compounds.
- Useful docking methods to predict the MIE were used.

The main points discussed for revising the case study were as follows:

- A figure describing the putative AOP and a more comprehensive description of the AOP were added in the purpose section in order to increase the understanding of complex III inhibition mediated neurotoxicity. In addition, more explanation was provided under section "Data Gap filling and Justification", regarding the AOP-based testing strategy, especially the relationship between *in silico*/*in vitro* methods and MIE/KEs. This hypothesis was strengthened by adding the available data on Antimycin A, a well-established CIII inhibitor with neurotoxic effects.
- It was clarified why brain concentration was chosen as the main end point for the putative neurotoxicity. It was assumed that the neurotoxic effects would arise from exposure to local concentrations in the brain.
- Two tables summarising *in vivo* ADME and neurotoxicity data have been inserted in order to facilitate the comparison between the source chemicals.
- The title of the case study should be narrowed down because the case study does not capture all aspects of neurotoxicity and applies a putative AOP. The focus should be on Complex III mediated neurotoxicity. In addition, "waiving" should be removed from the title as well as the main text.

The main uncertainties identified for the case study were as follows:

- Although the target chemical, azoxystrobin, is structurally not extremely similar to the source compounds, all available functional and ADME data show that the group of strobilurins (despite structural diversity) behave in a similar way. Moreover, the biological similarity, i.e. the mode of action, is the same, as measured by *in vitro* assays.
- There is uncertainty concerning the role of other AOPs (i.e. a mode of action not linked to CIII inhibition or mitochondrial impairment). The used functional assays that determine neuronal integrity (KE3 assays), are not only sensitive towards mitochondrial toxicants, but are capable of detecting a wide range of modes of actions. Therefore, the probability of azoxystrobin being neurotoxic by another mechanism is low, but cannot be entirely excluded.
- The AOP applied in the case study was putative, not reviewed and endorsed. Although the assays applied for KE3 (viability assays and neurodegeneration/neurite outgrowth assays) are closely related to the human AO as they directly assess neuronal integrity, the reviewers indicated that there is uncertainty concerning the putative AOP since MIE and AO are different from the endorsed AOP.
- There is only information on molecular docking of the positive control.
- Statistical analysis of the *in vitro* data is lacking.

The main comments on the use of the case study in other member countries' regulatory contexts are as follows:

- ***Canada (Health Canada):*** Aspects of this case study could be used to form a weight of evidence in a human health risk assessment of Existing Substances.

  In the Canadian New Substances Notification Regulations, data using NAMs may be provided to meet a regulatory data requirement if it can be proven to be equally or better suited to measure the endpoint under investigation. These data would not be sufficient to eliminate the need for an OECD 424 study if neurotoxicity was suspected.

- ***Netherlands:*** It would be very difficult to use this methodology for REACH to waive TG424 study. Neurotoxicity testing is not directly addressed within REACH, however when alerts are observed based on organ specific toxicity studies then neurotoxicity assessment has to be performed.

- ***United States:*** No. This is not a validated method.

Based on the experience reviewing this case study, the following areas were identified for potential guidance development:

- Guidance on reporting of docking/modelling
- How much evidence is needed to show low toxicity
- How to perform *ab initio* testing?
- How many key events should be tested
- Level of test method description, performance and validation
- Guidance on when to search for metabolism information
- Guidance on when data for PBPK modelling is needed, especially for the low toxicity chemicals
- Conduct of biologically based read-across

4. LEARNINGS AND LESSONS

4.1. Summary of the Case Studies Reviewed in All the Review Cycles

This chapter summarises learnings and lessons stemming from the case studies of the project including the fifteen case studies from the past four review cycles. Table 2 shows a summary of the twenty-three case studies reviewed up to date.

The assessment approaches illustrated by the case studies are classified into four types: data-gap filling by read-across based on grouping of chemicals (17 case studies), grouping of chemicals for cumulative risk assessment, not for read-across (Case Study 2016-2), safety assessment workflow (Case Study 2016-5 and 2019-1), screening of chemicals (Case Study 2017-1, 2017-2 and 2018-2). Although Case Study 2017-1 focused on a prioritisation scheme for potential endocrine active chemicals, it illustrated two systematic methodologies for identifying analogues with cheminformatics and data analysis tools. In addition, Case Study 2016-5 and 2019-1 include the read-across approach in the concept of a safety assessment workflow and Case Study 2019-1 focused on the read-across approach for exploring the endocrine activity of the target chemical. Case study 2017-3 illustrated the approach on read-across for nano-TiO_2 considering nano-specific properties, such as crystal type, surface coating and size. Case Study 2018-2 included the elements of the defined approach for identifying oestrogen receptor active chemicals.

The target endpoints of the case studies were: repeated dose toxicity (11 case studies), neurotoxicity (3 case studies), reproductive toxicity (2 case studies), oestrogenicity (2 case studies), mutagenicity (1 case study), bioaccumulation (1 case study), genotoxicity (1 case study), ecotoxicity (1 case study) and developmental toxicity (1 case study).

Every case study addresses some challenging topics related to IATA, including use of MOA/AOP (14 case studies), capturing and communicating uncertainty (21 case studies), use of new approach methodologies (18 case studies) and low/no toxicity prediction (11 case studies).

Identified areas for further developing guidance from the 23 case studies are summarised in section 4.2.

Table 2. Summary of the Case Studies Reviewed in the Past Five Review Cycles

Year-No. (Lead)	Assessment Approach	Endpoint	IATA Topics				Reference
			AOP[1]	UR[2]	NAM[3]	L/N[4]	
2019-1 (BIAC)	Safety assessment workflow	Reproductive toxicity	X	X	X	X	OECD, 2020a
2019-2 (BIAC)	Read-across	Repeated dose toxicity	X	X	X		OECD, 2020b
2019-3 (BIAC)	Read-across	Repeated dose toxicity	X	X			OECD, 2020c
2019-4 (BIAC)	Read-across	Repeated dose toxicity	X	X	X		OECD, 2020d
2019-5 (BIAC)	Read-across	Repeated dose toxicity	X	X	X	X	OECD, 2020e
2019-6 (BIAC)	Read-across	Developmental toxicity	X	X	X	X	OECD, 2020f
2019-7 (BIAC)	Read-across	Neurotoxicity	X	X	X		OECD, 2020g
2019-8 (BIAC)	Read-across	Neurotoxicity	X	X	X	X	OECD, 2020h
2018-1 (Japan)	Read-across	Reproductive toxicity	X	X			OECD, 2019b
2018-2 (US)	Prioritisation and screening	Oestrogenicity	X	X	X	X	OECD, 2019c
2017-1 (Canada/US)	Prioritisation and hazard characterisation	Oestrogenicity	X	X	X	X	OECD, 2018b
2017-2 (Canada)	Prioritisation of chemicals	Ecotoxicity	X	X	X	X	OECD, 2018c
2017-3 (JRC)	Read-across	Genotoxicity for nano-TiO_2		X	X		OECD, 2018d
2017-4 (ICAPO)	Read-across	Repeated dose toxicity		X	X	X	OECD, 2018e
2016-1 (Japan)	Read-across	Repeated dose toxicity		X	X		OECD, 2017b
2016-2 (US)	Grouping for cumulative risk assessment	Neurotoxicity	X		X		OECD, 2017c
2016-3 (ICAPO)	Read-across	Repeated dose toxicity		X	X	X	OECD, 2017d
2016-4 (ICAPO)	Read-across	Repeated dose toxicity		X	X	X	OECD, 2017e
2016-5 (JRC/BIAC)	Safety assessment workflow	Repeated dose toxicity	X		X		OECD, 2017f
2015-1 (Canada/US)	Read-across	Mutagenicity	X	X			OECD, 2016b
2015-2 (Canada)	Read-across	Repeated dose toxicity		X	X		OECD, 2016c
2015-3 (Japan)	Read-across	Repeated dose toxicity	X	X			OECD, 2016d
2015-4 (Japan)	Read-across	Bioaccumulation		X		X	OECD, 2016e

*1: AOP: Use of mode of action/adverse outcome pathways
*2: UR: Uncertainty reporting
*3: NAM: Use of new approach methodologies
*4: L/N: Low/no toxicity prediction

4.2. Update of the Identified Areas for Further Developing Guidance

In the past four review cycles, the following 9 areas for further developing guidance were identified as priority areas from the 15 case studies (OECD, 2016a, 2017a 2018a and 2109a). Note that although these are areas in which potential guidance could be developed, the intent is not to address all of these aspects within future OECD guidance documents, but rather note a potential need that has been identified.

1. Describing scope and context for read-across
2. Building hypotheses based on MOA/AOP
3. Definition of analogues/category boundaries
4. Justification of data gap filling
5. Incorporation of new approach methodologies
6. Decision logic for low/no toxicity predictions
7. Uncertainty analysis and reporting
8. Integrated conclusion
9. Reporting templates for IATA based on a building block approach

The following areas for further developing guidance were identified based on the review experience of the 8 case studies in the fifth review cycle:

1. ***Reporting new approach methodologies and the data generated by NAMs*:** All the case studies in the fifth review cycle except Case Study 2019-3 apply various types of new approach methodologies (NAMs). These methodologies and the data they provide are important to justify data gap filling and similarity assessment within the category members. However, since the NAMs are not familiar to all project members, it was agreed that guidance documents on how to describe the methodologies and to report the data from NAMs are needed. Based on the approaches applied in the case studies in the fifth review cycle, the following areas for further developing guidance were identified:
    - Guidance for evaluating ToxCast data
    - Guidance on reporting of docking/modelling approaches
    - Definition of NAM (will be addressed in IATA Concepts Document); to be published in the OECD series on testing and assessment
    - Description of PBPK modelling (WHO document for *in vivo* PBPK description; on-going OECD project for *in vitro* based model descriptions)
    - Use of gene expression data in IATA (note on-going OECD project on templates for transcriptomics to capture studies for regulatory use)
    - Establishing a list of chemicals (comprising data rich chemicals with various MOAs) to be used as standards for NAM validation
2. ***Using an AOP for building a hypothesis and testing strategy*:** While Case Study 2019-7 is based on an AOP endorsed by OECD for building the hypothesis and testing strategy, Case Study 2019-5 and 2019-8 applied non-endorsed/putative AOPs. It was concluded that all AOPs including the non-endorsed AOPs are useful for building hypothesis and testing strategies. However, a certain level of

uncertainty associated with the non-endorsed AOPs should be considered in the overall uncertainty analysis since their level of review is not clear. In order to help case study submitters with the use of non-endorsed AOP within an IATA, guidance on using non-endorsed AOPs regarding documentation of their inherited uncertainty was identified as an area for future work. Harmonisation of the terminology used at the AOPs was also a point for discussion and further improvements.

Regarding an AOP based testing strategies, there was discussion on the extent of coverage of key events (KEs) within an AOP or especially within an AOP network that would be required. Although the authors considered that the *in vitro* assays covered the majority of KEs, some reviewers indicated that there was no full coverage of the KEs of a specific AOP as some KEs were not measured due to lack of associated and suitable and *in vitro* assays. In addition, it was discussed that even when *in vitro* assays are available, it is necessary to better explain why the selected assays in an AOP based testing strategy are were fit for purpose. Therefore, guidance on the extent of coverage of KEs with an AOP or AOP network would be helpful to better understand the appropriate testing battery. In this topic, the following two areas were identified for further developing guidance.

- Tips on using non-endorsed AOPs regarding documentation/uncertainty/terminology
- Coverage of KEs in an AOP based testing strategies

3. ***Justification of negative predictions*****:** Case Study 2019-1, 2019-5, 2019-6 and 2019-8 demonstrated negative predictions. Generally, the justification for a no effect prediction was made based on read-across from the source chemicals, which have no or less effects on the target endpoints. Moreover, Case Study 2019-5 and 2019-6 used negative controls for *in vitro* assays to increase the justification for negative predictions. Negative predictions are important for decision making in the regulatory context. Therefore, this area was identified for further developing guidance.

4. ***Important information required for justification of approaches and conclusions*****:** Most of the case studies included a wealth of data and information. Some of reviewers requested the authors to highlight the main (crucial) data that was used to justify the conclusions in order to increase understanding of the key lines of evidence relied upon in the case studies.

   In some case studies (Case Study 2019-1, 2019-5, 2019-6, 2019-7 and 2019-8), it was assumed that the metabolites would not be toxic or they were not directly considered. Reviewers questioned this assumption. Therefore, guidance on when information on metabolites should be included is needed.

   Case Study 2019-7 and 2019-8 demonstrated read-across based on biological similarity. The biological read-across was performed by illustrating that the *in vitro* assays results were similar for both the target chemical and the source chemicals. In Case Study 2019-8, this was the case even with lower structural similarity but a similar toxicophore. Biological read-across would be useful in the risk assessment under the regulatory context. Therefore, guidance on what is needed to support biologically based read-across would be important in the future.

In this topic, the following three areas were identified for further developing guidance.

- How to report key data vs supporting data
- When should information on metabolites be included?
- What is needed to address biological read-across

Moreover, at the 5th meeting of the IATA Case Studies Project, the project members discussed prioritisation of the areas for developing guidance. The following areas were identified as high priority:

- Revising the OECD grouping guidance including an uncertainty section, grouping based on metabolites, guidance on biological read-across etc. Likely additional past issues identified for read-across guidance could also be included in an update of the guidance.
- Justification of negative predictions.

A summary of updated areas for further developing guidance incorporating the above issues is shown in Table 3 (the underlined and bold areas were identified in the 2019 review cycle):

Table 3. Summary List of the Areas for the Development of Further Guidance

Areas for the development of further guidance	Related case study	High Priority
1. Describing scope and context for read-across	2015-1, 2015-2, 2015-3, 2015-4, 2016-1, 2016-3, 2016-4, 2017-3, 2017-4, 2018-1, 2019-2, 2019-3, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	High (included in revising the OECD grouping guidance)
a. Rationale for the selected endpoint	2015-1, 2015-2, 2015-3, 2015-4, 2016-1, 2016-3, 2016-4, 2017-3, 2017-4, 2018-1, 2019-2, 2019-3, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	High (included in revising the OECD grouping guidance)
b. Considerations for justifying focus of an IATA (e.g. choosing 'major' effect vs 'minor' effect); providing explanation why a certain effect is considered the most relevant (toxicological response observed at a lower dose), while others are minor (occurring at a higher dose)	2015-2, 2015-3, 2016-1, 2016-3, 2016-4, 2017-4, 2019-2, 2019-3	High (included in revising the OECD grouping guidance)
2. Building hypotheses based on MOA/AOP	2015-1, 2015-3, 2016-2, 2016-5, 2017-1, 2017-2, 2018-1, 2018-2, 2019-3, 2019-4, 2019-5, 2019-7, 2019-8	
a. Hypothesis for category formation that includes the use of omics data	2016-1	High (included in revising the OECD grouping guidance)
3. Definition of analogue/category boundaries	2015-1, 2015-2, 2015-3, 2015-4, 2016-1, 2016-2, 2016-3, 2016-4, 2017-1, 2017-3, 2017-4, 2018-1, 2019-1, 2019-2, 2019-3, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	High (included in revising the OECD grouping guidance)
a. Defining boundaries based on- phys/chem properties, toxicokinetics, toxicodynamics, bioavailability and metabolism, or , nanomaterials-specific parameters	2015-1, 2015-3, 2015-4, 2016-1, 2016-2, 2016-3, 2016-4, 2017-1, 2017-3, 2017-4, 2018-1, 2019-3, 2019-4, 2019-7, 2019-8	High (included in revising the OECD grouping guidance)
4. Justification of data gap filling	All case studies	
a. Reporting of QSAR prediction results	2015-1, 2015-4, 2016-1, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-4, 2018-1, 2019-1, 2019-2, 2019-4, 2019-5, 2019-6	
b. How much to report on reliability	2015-1, 2015-2, 2015-3, 2015-4, 2016-1, 2016-3, 2016-4, 2017-3, 2017-4, 2018-1,	

		2018-2, 2019-1, 2019-2, 2019-3, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	
	c. Use of New Approach Methodology (NAM) data, TTC approach and PBPK models (e.g. How to integrate NAM data – for example via linking to mechanistic relevance (interpretation))	2015-2, 2016-1, 2016-2, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-3, 2017-4, 2018-2, 2019-1, 2019-2, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	
	d. Guidance for describing New Approach Methodology data in the context of IATA case studies	2015-2, 2016-1, 2016-2, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-3, 2017-4, 2018-2, 2019-1, 2019-2, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	
	e. **Decision logic for low/no toxicity predictions**	2015-4, 2016-3, 2016-4, 2017-1, 2017-2, 2017-4, 2018-2, 2019-1, 2019-5, 2019-6, 2019-8	High
	f. Guidance on when *in vitro* data could be further generated to support read-across	2015-2, 2016-1, 2016-3, 2016-4, 2017-4, 2018-2, 2019-1, 2019-2, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	
	g. Guidance for use and reporting of results of HTS and HTTK assays	2015-2, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-4, 2018-2, 2019-1, 2019-2, 2019-5, 2019-6	
	h. **What is needed to address biological read-across**	2019-1, 2019-2, 2019-5, 2019-6, 2019-7, 2019-8	
5.	Uncertainty Analysis	2015-1, 2015-2, 2015-3, 2015-4, 2016-1, 2016-3, 2016-4, 2017-1, 2017-2, 2017-3, 2017-4, 2018-1, 2018-2, 2019-1, 2019-2, 2019-3, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	
	a. Exposure route, including route to route extrapolation	2015-4, 2016-5, 2017-1, 2017-4, 2019-1, 2019-2	
	b. Use of data from different test conditions for read-across for a target endpoint	2015-1, 2015-2, 2015-3, 2016-1, 2016-3, 2016-4, 2017-3, 2017-4	High (included in revising the OECD grouping guidance)
	c. How uncertainties impact on overall conclusion	2015-1, 2015-2, 2015-3, 2015-4, 2016-1, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-3, 2017-4, 2018-1, 2018-2, 2019-1, 2019-2, 2019-3, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	
	d. Guidance for evaluating the reliability/robustness of data including TK/TD data ● Similarity of metabolic pathways ● Whether differences in the structure of target chemicals would have any significant impact on the metabolic pathway ● **When should information on metabolites be included?**	2015-1, 2015-2, 2015-3, 2016-1, 2016-3, 2016-4, 2017-4, 2018-1, 2019-1, 2019-2, 2019-3, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	
	e. Reporting of uncertainty of read-across (e.g. Ranking of uncertainty vs descriptive analysis/ quantitative vs qualitative analysis)	2015-1, 2015-2, 2015-3, 2015-4, 2016-1, 2016-3, 2016-4, 2017-3, 2017-4, 2018-1, 2019-2, 2019-3, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	High (included in revising the OECD grouping guidance)
	f. Consider approaches in: AOP handbook (OECD, 2016g) and scientific papers (Wu et al., 2010; Blackburn & Stewart, 2014; Schultz et al., 2015)	2015-1, 2015-2, 2016-3, 2016-4, 2017-4	
	g. Uncertainty framework (Overall uncertainty in the assessment resulting from the combined uncertainties of the different IATA components and data types)	2015-1, 2015-2, 2015-3, 2015-4, 2016-1, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-3, 2017-4, 2018-1, 2018-2, 2019-1, 2019-2, 2019-3, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8	
	h. **Tips on using non-endorsed AOPs regarding documentation/uncertainty/terminology**	2015-1, 2016-2, 2016-5, 2017-1, 2018-1, 2018-2, 2019-3, 2019-4, 2019-6, 2019-8	
6.	Integrated Conclusion	All case studies	
	a. Combining approaches/methodologies for predicting bioaccumulation	2015-4	

	b. Integrating QSAR predictions, including when to use consensus modelling or not	2015-1, 2015-4, 2016-1, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-4, 2018-1, 2019-1, 2019-2, 2019-4, 2019-5, 2019-6
	c. Guidance on deriving integrated conclusions from the different components of the IATA, including harmonised uncertainty assessment	2015-1, 2015-2, 2015-3, 2015-4, 2016-1, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-3, 2017-4, 2018-1, 2018-2, 2019-1, 2019-2, 2019-3, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8
	d. How to define acceptable uncertainty	2015-1, 2015-2, 2015-3, 2015-4, 2016-1, 2016-3, 2016-4, 2017-1, 2017-2, 2017-3, 2017-4, 2018-1, 2018-2
7.	Others	
	a. Relevance of change in pH to prediction of degradation products (e.g. in the environment)	2015-4
	b. UVCBs, multi-constituents coverage (composition coverage, methodology and other)	2015-2, 2017-2
	c. Level of detail needed in case studies according to the defined purpose	All case studies
	d. How to include data on/predictors for metabolism when building IATAs according to the defined purpose	2015-2, 2015-3, 2017-4, 2018-1, 2019-3, 2019-4
	e. How to describe the rationale for justification of the BMD and point of departure used	2016-2
	f. Reporting template for IATA based on a building blocks approach	2016-2, 2016-5, 2017-1, 2017-2, 2018-2, 2019-1
	g. Guidance on developing prioritisation scheme based on IATA	2017-1, 2017-2
	h. Guidance on use or reporting new approach methods (chem-informatics tools, HTS, HTTK assays)	2015-2, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-3, 2017-4, 2018-2, 2019-1, 2019-2, 2019-5, 2019-6, 2019-7, 2019-8
	i. Guidance on how to develop integrated testing strategies (ITS) and data interpretation procedures (DIP)	2018-2
	j. Guidance on how to combine *in vitro* and computational information into an integrated report, including applicability domain	2018-2
	k. **Guidance for evaluating ToxCast data**	2015-1, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-4, 2018-2, 2019-1, 2019-2
	l. **Guidance on reporting of docking/modelling approaches**	2019-7, 2019-8
	m. **Coverage of KEs in AOP based testing strategy**	2018-2, 2019-5, 2019-7, 2019-8
	n. **How to report key vs supporting data**	All case studies
	o. **Establishing a list of chemicals (comprising data rich chemicals with various MOAs) to be used as standards for NAM validation**	2015-2, 2016-1, 2016-2, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-3, 2017-4, 2018-2, 2019-1, 2019-2, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8
8.	Areas related to ongoing OECD projects	
	a. **Definition of NAM (will be addressed in IATA Concepts Document)**	2015-2, 2016-1, 2016-2, 2016-3, 2016-4, 2016-5, 2017-1, 2017-2, 2017-3, 2017-4, 2018-2, 2019-1, 2019-2, 2019-4, 2019-5, 2019-6, 2019-7, 2019-8
	b. **Description of PBPK modelling (WHO document for *in vivo* PBPK description; on-going OECD project for *in vitro* based model descriptions)**	2016-5, 2019-1, 2019-2, 2019-5, 2019-6, 2019-7, 2019-8
	c. **Use of gene expression data in IATA (note on-going OECD project on transcriptomics templates to capture studies for regulatory use)**	2016-1, 2019-1
9.	Areas related to other working party	
	a. Nanomaterials (NMs) *1 • Standardised guidelines for characterisation and testing of NMs, • NM-specific and relevant protocols • Increase data availability/quality	2017-3
	b. Guidance on the interpretation of NM-related data *1	2017-3

c. Guidance for reporting from exposure simulation models (e.g. environmental concentrations) *2	2017-2

**1: The area is related to the Working Party on Manufactured Nanomaterials*
**2: The area is related to the Working Party on Exposure Assessment*

4.3. Considerations from the case studies in the fifth review cycle

This section describes the learnings gained through the review experience of the eight case studies in the fifth review cycle in 2019 which all addressed read-across approaches.

Case Study 2019-1 and 2019-2 include the Margin of Internal Exposure Concept for calculation of the ratio of a measure of internal exposure, such as blood concentration or target-tissue dose, to a corresponding toxicological point of departure.

Case Study 2019-5, 2019-7 and 2019-8 applied AOP based *in vitro* testing batteries. Throughout the review, questions were raised: how many Key Events (KEs) should be covered and how many tests should be enough to cover the relevant KEs. In addition, there was discussion on use of endorsed AOP versus non-endorsed AOP, including having a structural analogue of stressor of an AOP as a strength of the case study. Case Study 2019-5 applied AOP networks for liver steatosis.

The advantages identified in these approaches should be taken into account for the development of future IATA. As a result, the following four topics are discussed in this section:

1. Margin of Internal Exposure concept
2. Coverage of KEs in an AOP based testing strategy
3. Use of endorsed versus non-endorsed AOPs including having a structural analogue of a stressor of an AOP as the as a strength
4. Advantage and coverage of AOP networks

4.3.1. Margin of Internal Exposure concept

Case Studies 2019-1 and 2019-2 applied a physiologically-based biokinetic (PBBK) model to estimate blood concentrations following exposures to the target or source chemicals in experimental animals and humans. The estimated blood concentration allows to derive a Margin of Internal Exposure (MoIE) that is calculated as the ratio of a measure of internal exposure, such as blood concentration or target-tissue dose to a corresponding toxicological point of departure. This ability to rely on a measure of internal exposure reduces the uncertainty in the risk assessment by incorporating chemical-specific information on the absorption, distribution, metabolism and excretion (ADME) of the chemical in both the experimental animal and the human. Calculation of internal exposures with a PBBK model can be used to replace the default uncertainty factor of 4 for interspecies differences in TK (IPCS 2005, WHO 2010). The US EPA follows this practice in determining Reference Concentrations and Reference Doses (US EPA 2011). Thus, a MoIE of 25 was considered equivalent to the default external dose MoE of 100 in the case studies.

In the two case studies, the MoIE was calculated by the below formula:

$$\text{MoIE} = \text{Cmax NOAEL animal study of source chemical (SC)} / [(\text{Cmax human of target chemical (TC)}) \times (\text{Relative Potency of TC /SC})]$$

In Case Study 2019-1, the value of C_{max} NOAEL animal study of source chemical, butylparaben (BP), was calculated by simulating the exposure scenario in the rat toxicity study (Fisher *et al*., 1999). The value of C_{max} human of target chemical, propylparaben (PP), from the consumer exposure estimates was calculated by human PBBK simulation of internal plasma concentration. Relative Potency PP/BP was calculated based on the AC10[3] median values in the ToxCast oestrogen receptor activity assays.

In Case Study 2019-2, the value of C_{max} NOAEL animal study of source chemical, theophylline, was observed in the rabbit developmental toxicity (Shibata et al, 2000). The value of C_{max} human of target chemical, Caffeine, was calculated by human PBBK simulation of internal plasma concentration following estimated oral and whole body dermal exposure. Relative Potency Caffeine/Theophylline was calculated based on the inhibitory constant (Ki) for the A1-adenosine receptor in rat brain.

PBBK models play a key role in estimation of C_{max}. In order to develop PBBK model, the ADME parameters are important, such as absorption rate, liver parameters and intrinsic clearance. In Case Study 2019-1, the rat subcutaneous injection dosing route has high uncertainty in the PBBK model because there are no rat subcutaneous kinetic data to address this uncertainty.

Table 4 shows a summary of the two case studies applied MoIE.

Table 4. A summary of case studies applying MoIE

Case Study	Target Chemical	Source Chemical	C_{max} NOAEL animal study source chemical	C_{max} human of target chemical	Relative Potency	Reference
2019-1	Propylparaben	Butylparaben	Simulation of the exposure scenario in the rat toxicity study	Human PBBK simulation of internal plasma concentration following whole body exposure in lotion	AC10 median values in the ToxCast oestrogen receptor activity assays	OECD, 2020a
2019-2	Caffeine	Theophylline	Rabbit developmental toxicity	Human PBBK simulation of internal plasma concentration following estimated oral and whole body dermal exposure	inhibitory constant (Ki) for the A1-adenosine receptor in rat brain	OECD, 2020b

4.3.2. Coverage of KEs in AOP based strategy

Several case studies in the fifth review cycle (Case Study 2019-5, 2019-7, 2019-8) applied in vitro testing batteries based on AOPs for the hazard characterisation. In addition, Case Study 2018-2 applied an in vitro test battery based on a putative AOP. Table 5 provides the summary of the coverage of MIE and KEs in AOP based testing strategy in the case studies.

Case Study 2018-2 (OECD, 2019c) applied 16 *in vitro* assays covering the MIE and 4 KEs of an AOP for the oestrogen receptor agonist pathway. All the MIE/KEs, except KE4, (cell proliferation) were tested by using multiple *in vitro* assays. The ER pathway model was analysed by using *in vitro* reference chemicals and *in vivo* reference chemicals in order to determine whether a subset of assays could achieve equivalent performance to the full 16 *in vitro* HTS assays. The results of the analysis showed that as few as 4 assays are sufficient by comparing the results from various combinations of 1 to 15 assays to results from the 16-assay ER pathway model. This flexibility was archived by the subset of assays for different points on the ER pathway and incorporation of different technologies.

[3] Concentrations Associated with 10% of maximum activity

A steatosis AOP network built based on 6 AOPs that are currently under development in the AOP-Wiki was used to select MIEs/KEs to be measured in Case Study 2019-5. Since it is not necessary, also potentially not technically feasible, to measure all the MIEs and KEs of an AOP network, a set of early events (MIEs) together with lipid accumulation (late KE close to the adverse outcome) in the cell were selected. Activation of selected MIEs was detected by testing one of the source compounds, valproic acid, which is known to induce liver steatosis. In this case study, besides MIEs, the perturbation of general cellular function (lipid accumulation) was tested together with cytotoxicity and revealed potential major differences between the grouped compounds.

From the results of the MIEs, it was observed that with increasing chain length the number of activated MIEs related to steatosis increased. The KE measurement showed that the two compounds with *in vivo* data on liver steatosis induced lipid accumulation, whereas the *in vivo* negative compound was inactive up to the highest dose. In terms of the target compound 2-Ethylbutyric acid (2-EBA), while it cannot be excluded that a pathway towards lipid accumulation is activated by 2-EBA since 2-EBA activated one of the tested MIEs, 2-EBA was inactive for lipid accumulation up to the highest dose.

The case study showed that activation of an MIE might not be sufficient to determine if the entire pathway will be triggered. A strength of this case study is conducting the assays for the target KE lipid accumulation, including positive and negative compounds, to follow-up on the MIE that they observed to be activated for the target compound. However, it was indicated that there are uncertainties on the coverage of the KEs. The uncertainties could be reduced by additional *in vitro* testing for other KEs or if a stronger justification on the essentiality of the measured MIE or KE is provided. The uncertainties identified in this testing strategy could be further reduced in case of MIE linkage to the KE (lipid accumulation) by a non-adjacent KER with strong biological plausibility or by the identification of other essential KEs and empirical evidence described within an endorsed AOP.

Case Study 2019-7 included an *in silico* assay for testing the MIE and a number of *in vitro* assays for testing four out of five KEs of an endorsed AOP. For Parkinson motor deficits induced by Complex I inhibition, no data was presented for KE5 since there are no appropriate assays available to measure this KE (neuroinflammation) *in vitro*. The author assumed that the lacking of this KE testing will not hamper the overall conclusions as neuroinflammation (KE5) is triggered early in the neurodegenerative process (KE4) and then exacerbates it significantly, meaning that the impact of KE5 to the pathway can be indirectly captured in KE4.

Case Study 2019-8 included an *in silico* assay for testing the MIE and a certain number of *in vitro* assays for testing all the KEs of a putative AOP. (Mitochondrial complex III inhibition leads to neuronal toxicity). The assays for the MIE, KE1 and KE2 were used to investigate whether there might be a molecular cause for neurotoxic mode of action of the tested chemicals. The assays applied for KE3 were closely related to the human AO as they directly assess neuronal integrity and can be potentially affected by multiple upstream MIEs and KEs.

Table 5. Case studies using AOP based testing strategy

Case Study	AOP information	Coverage of MIE/KEs						
		MIE	KE1	KE2	KE3	KE4	KE5	Other KEs
2018-2	Endocrine Receptor (ER) Pathway	Receptor binding	Receptor dimerisation/ Protein stabilisation	Cofactor recruitment	Transcription activation	Cell proliferation	-	-
	(putative AOP)	(3 *in vitro* assays)	(10 *in vitro* assays for KE1 and KE2)		(2 *in vitro* assays)	(1 *in vitro* assay)		
2019-5	Steatosis AOP network	6 MIEs - Nrf2, PXR; PPAR-α and γ, AhR, LXR.	Main KE - Lipid accumulation	-		-	-	Cytotoxicity and 4 cellular mechanisms (GSH depletion, phospholipidosis, mitochondrial superoxide, MMP)
	(AOPs under development; AOP-wiki 34*1, 36*2, 57*3, 58*4, 60*5 and 61*6)	(CALUX and GFP)	(HepG2, HepaRG, PHH)					(CALUX, HepG2, HepaRG, PHH, RPTEC cell)
2019-7	Inhibition of the mitochondrial complex I of nigro-striatal neurons leading to parkinsonian motor deficits	Interaction with complex I	Inhibition of complex I	Mitochondrial dysfunction	Impaired proteostasis	Degeneration of DA neurons nigrostriatal pathway	Neuroinflammation	-
	(Published No.7) (OECD, 2018f)	(Receptor docking and Similarity)	(Seahorse assay)	(Mitochondrial membrane potential)	(Protease activity and CHOP-GFP)	(Viability and Neuronal health)	No available	
2019-8	Mitochondrial complex III inhibition leads to neuronal toxicity	Binding of inhibitor to cytochrome bc1 complex	Mitochondrial complex III inhibition	Mitochondrial dysfunction	Neuronal degeneration	-	-	-
	(putative AOP)	(Receptor docking and Similarity)	(Seahorse assay)	(Mitochondrial membrane potential)	(Viability and Neuronal health)			

*1: https://aopwiki.org/aops/34
*2: https://aopwiki.org/aops/36
*3: https://aopwiki.org/aops/57
*4: https://aopwiki.org/aops/58
*5: https://aopwiki.org/aops/60
*6: https://aopwiki.org/aops/61

4.3.3. Use of endorsed/non-endorsed AOPs including having a the structural analogue of the stressor of an AOP as a strength

The considerations document in the fourth review cycle (OECD, 2019a) described how to use information from an AOP under development (status on AOP-wiki). Generally, the read-across case studies applied MOA/AOP to formulate a category by identifying members which were assumed to induce a toxicological effect via the same MOA/AOP.

All the case studies in the fifth review cycle demonstrated how to use an MOA/AOP in read-across. Similar to past read-across case studies, Case Study 2019-3 and 2019-4 used MOA information for identifying category members. However, in Case Study 2019-1 and 2019-2, the authors used the common MOA between the target chemicals and the source chemicals and to support the calculation of the relative potency in order to estimate MoIE as described in section 4.3.1. Case Study 2019-5, 2019-7 and 2019-8 demonstrated to use of AOPs as anchors for the test strategy. While Case Study 2019-7 applied an endorsed AOP to establish the testing strategy, Case Study 2019-5 and 2019-8 relied on non-endorsed AOPs. It was indicated that there were uncertainties associated with the maturity and level of development and acceptance of putative AOPs since these AOPs have not been reviewed and endorsed by the OECD.

Case Study 2019-7 applied an endorsed AOP on inhibition of the mitochondrial complex I of nigro-striatal neurons leading to parkinsonian motor deficits (OECD, 2018f). The target chemical was deguelin that has a high similarity with rotenone, a stressor used to provide empirical evidence for this AOP. While only rotenone could be used as analogue for the deguelin, the case study demonstrated that deguelin would cause nigrostriatal neuronal system degeneration since the results of the *in silico* and *in vitro* assays showed similar biological activities between deguelin and rotenone.

Case Study 2019-5 compiled 6 non-endorsed AOP on liver steatosis and built an AOP network. The author conducted testing using *in vitro* assays covering the 6 MIEs and 1 late KE, (lipid accumulation) for the target chemical and the source chemicals as well as a negative control. As described in section 4.3.2, these assays showed the correct prediction for all the three compounds with *in vivo* data, either those that induce or do not induce liver steatosis. Even though there was uncertainty regarding the AOP network formed by putative AOPs, the assay results for the positive and negative compounds could reduce this area of uncertainty.

Case Study 2019-8 applied the putative AOP on mitochondrial complex III inhibition leading to neuronal toxicity, which was built based on the endorsed AOP on inhibition of the mitochondrial complex I of nigro-striatal neurons leading to parkinsonian motor deficits (OECD, 2018f). The target chemical was azoxystrobin and the source compounds were selected based on the same toxicophore, the (E)-β-methoxyacrylate group. In addition, Antimycin A was used as a positive compound since the source chemicals showed lower neurotoxicity. *In vitro* assays suggested that strobilurins would be less neurotoxic compared with the source chemicals. Similar to Case Study 2019-5, while there was uncertainty on the putative AOP, the author tried to reduce the area of uncertainty by using a positive control chemical.

Table 6. Case Studies Using Information from an MOA/AOP

Case Study	Target Chemical	Stressor	AOP information (Status AOP-wiki)	How to use AOP
2015-1 (Canada)	Azo direct dyes releasing DMOB via metabolic activity	DMOB	DNA mutation associated with DNA binding (No relevant AOP on the AOP-wiki)	Hypothesis for the category
2015-3 (Japan)	Esters of single allyl alcohol and saturated aliphatic carboxylic acid	1. Allyl Alcohol 2. Carbon tetrachloride 3. Retinol 4. Dimethyl nitrosamine 5. Thioacetamide	Protein Alkylation Leading to Liver Fibrosis (Published as AOP No.2 (OECD,2016f) (No 38 at AOP-wiki)*1)	Hypothesis for the category
2016-2 (US)	Organophosphate pesticides.	Organophosphate pesticides (No information at AOP-Wiki)	Acetylcholinesterase inhibition leading to acute mortality (Under development as No. 16 at AOP-wiki)*2	Hypothesis for the category
2016-5 (JRC/BIAC)	piperonyl butoxide	No information at AOP-Wiki 1. Allyl Alcohol 2. Carbon tetrachloride 3. Retinol 4. Dimethyl nitrosamine 5. Thioacetamide	1. LXR activation leading to hepatic steatosis (Under development as No. 34 at AOP-wiki)*3 2. Protein Alkylation Leading to Liver Fibrosis (Published as AOP No.2 (OECD,2016f) (No 38 at AOP-wiki)*1)	Targeted testing based on the broad screening results from *in silico* or *in vitro* HTS assays
2017-1 (Canada/US)	• 4-Tert-Butylphenol • 2,4-Di-TertButylphenol • Octabenzone	Non-hindered phenols	Binding of oestrogen leading to reproductive or developmental effects (No relevant AOP on the AOP-wiki)	Building data gathering strategy
2018-1 (Japan)	Ethylene glycol methyl ether (EGME)-related category members	- - 1. Methoxyacetic acid 2. Butyrate 3. Trichostatin A 4. Valproate	1. Testicular toxicity caused by inhibition of lactate production in Sertoli cells and the transport of lactate to spermatocytes (No relevant AOP on the AOP-wiki) 2. Sarcosine dehydrogenase inhibition leading to testicular toxicity (No relevant AOP on the AOP-wiki) 3. Histone deacetylase inhibition leading to testicular toxicity (Under review as No. 212 at AOP-wiki)*4	Hypothesis for the category
2018-2 (US)	Environmental and commercial chemicals	-	Endocrine Receptor (ER) mediating reproductive, developmental and other health effects. (Nonspecific AOP)	Building testing strategy
2019-1 (BIAC)	Propylparaben	17Beta-estradiol	Endocrine Receptor (ER) mediating reproductive effects. (Nonspecific AOP)	Calculating relative potency scaling factors for oestrogen receptor

2019-2 (BIAC)	Caffeine	-	Adenosine receptor inhibition leading to prenatal developmental toxicity (Nonspecific AOP)	Calculating relative potency scaling factors for oestrogen receptor
2019-3 (BIAC)	Chlorinated benzenes	Chlorinated benzenes	Chlorinated benzenes induced hepatotoxicity (No relevant AOP on the AOP-wiki)	Hypothesis for the category
2019-4 (BIAC)	2,4,6-tri-alkylphenols	Quinone methide	Quinone methide induced hepatotoxicity (No relevant AOP on the AOP-wiki)	Hypothesis for the category
2019-5 (BIAC)	2-Ethylbutyric acid	No information at AOP-Wiki	1. LXR activation leading to hepatic steatosis (Under development as No. 34 at AOP-wiki)*3	Compiling AOP on liver steatosis to select the MIEs/KE for testing
		No information at AOP-Wiki	2. Peroxisomal Fatty Acid Beta-Oxidation Inhibition Leading to Steatosis (Under development as No. 36 at AOP-wiki)*5	
		No information at AOP-Wiki	3. AhR activation leading to hepatic steatosis (Under development as No. 57 at AOP-wiki)*6	
		No information at AOP-Wiki	4. NR1I3 (CAR) suppression leading to hepatic steatosis (Under development as No. 58 at AOP-wiki)*7	
		No information at AOP-Wiki	5. NR1I2 (Pregnane X Receptor, PXR) activation leading to hepatic steatosis (Under development as No. 60 at AOP-wiki)*8	
		No information at AOP-Wiki	6. NFE2L2/FXR activation leading to hepatic steatosis (Under development as No. 61 at AOP-wiki)*9	
2019-6 (BIAC)	2-Methylhexanoic acid	1. Valproic acid 2. Butyrate 3. Trichostatin A 4. Suberoylanilide hydroxamic acid 5. MS-275 6. Apicidin	Histone deacetylase inhibition leads to impeded craniofacial development (Under development as No. 274 at AOP-wiki)*10	Investigation to verify the correlation between HDAC inhibition and exencephaly induction
		No information at AOP-Wiki	Histone deacetylase inhibition leads to neural tube defects (Under development as No. 275 at AOP-wiki)*11	
2019-7 (BIAC)	deguelin	1. MPP+ 2. Rotenone	Inhibition of the mitochondrial complex I of nigro-striatal neurons leading to parkinsonian motor deficits (Published as AOP No.7(OECD,2018f) (No 3 at AOP-wiki)*12)	Building testing strategy
2019-8 (BIAC)	azoxystrobin	-	Mitochondrial complex III inhibition leads to neuronal toxicity (No relevant AOP on the AOP-wiki)	Building testing strategy

*1: https://aopwiki.org/aops/38 *2: https://aopwiki.org/aops/16 *3: https://aopwiki.org/aops/34 *4: https://aopwiki.org/aops/212 *5: https://aopwiki.org/aops/36
*6: https://aopwiki.org/aops/57 *7: https://aopwiki.org/aops/58 *8: https://aopwiki.org/aops/60 *9: https://aopwiki.org/aops/61 *10: https://aopwiki.org/aops/274

*11: https://aopwiki.org/aops/275 *12: https://aopwiki.org/aops/3

4.3.4. Advantage of AOP networks

Connecting multiple AOPs through AOP network formation can provide some benefits for considering various MIEs/KEs leading to common AO for evaluating complex endpoints. Case Study 2019-5 demonstrated how an AOP network is used within an IATA framework. In Case Study 2019-5, the liver steatosis AOP network was built based on several AOPs. While there was uncertainty in the AOPs since these AOPs are under development, there were advantages of the use of a network. The network helped to guide the selection of *in vitro* assays addressing different MIEs and common KEs leading to transparency about which assay was used for which purpose. As well, the two high throughput assays, the CALUX and Green Fluorescent Protein (GFP) reporter assays, measured six of the MIEs from the AOP network leading to steatosis, which allowed evaluating systematically MIE profiles of the target and source compounds. Therefore, the MOA induced by the target compounds was determined with high confidence due to the network approach.

5. USEFUL TOOLS FOR IATA

This chapter highlights useful tools for IATA, which were presented and demonstrated at webinars. The webinar for demonstration of IATA tools was agreed at the fourth meeting of the IATA Case Studies Project in 2018 in order to share additional information on IATA tools within the project team and to promote the use of these tools for developing IATAs. So far, two webinars were held and the following four tools were introduced and demonstrated:

1. Hazard Evaluation Support System Integrated Platform (HESS) [Japan]
2. Computational model for Estrogen Receptor (ER) pathway [the United States]
3. Consexpo [the Netherlands]
4. European Union System for the Evaluation of Substances (EUSES) [the Netherlands]

Table 7. IATA tools demonstrated at the webinar shows more detailed information on the above tools.

Table 7. IATA tools demonstrated at the webinar

Tool	Description
Hazard Evaluation Support System Integrated Platform (HESS)*1 (Japan)	HESS allows chemicals to be categorised on the basis of structural, physicochemical and mechanistic similarities and helps predict the repeated dose toxicity of untested chemicals by means of the category approach.
Computation model for Estrogen Receptor (ER) pathway*2 (the United States)	The model used 16 *in vitro* assays which are a subset of a larger collection of assays used in the US EPA ToxCast program to identify and quantify the ER agonist activity of a chemical. The application of the model is for screening of environmental chemicals based on their ER agonist activity and for determining whether further evaluation of endocrine-related activity in higher tier *in vivo* tests (e.g., female pubertal assay, two generation reproductive toxicity study) is needed.
Consexpo*3 (the Netherlands)	A computer program that enables the estimation and assessment of exposure to substances from consumer products such as paint, cleaning agents and personal care products;
European Union System for the Evaluation of Substances (EUSES)*4 (the Netherlands)	Model for the evaluation of the risks of substances to man and the environment. The main outputs of EUSES are local and regional risk characterisation ratios (RCRs) for several environmental compartments: air, surface water, sediment, soil, biota.

*1: https://www.nite.go.jp/en/chem/qsar/hess-e.html
*2:https://www.epa.gov/endocrine-disruption/use-high-throughput-assays-and-computational-tools-endocrine-disruptor
*3: https://www.rivm.nl/en/consexpo
*4: https://ec.europa.eu/jrc/en/scientific-tool/european-union-system-evaluation-substances

6. CONCLUSION

Eight case studies were reviewed in the fifth review cycle of the project in 2019. While Case Study 2019-1 applied a safety assessment workflow, the hazard of the target chemical was characterised based on a read-across approach. The other 7 case studies also applied a read-across approach. However, each of the read-across approaches was diverse. Case Study 2019-1 and 2019-2 explored a common mode of action between the target and source chemicals based on the gathered information and then read-across was performed by calculating MoIE based on the *in vivo* data, results of PBBK modelling and the relative potency. In Case Study 2019-3 and 2019-4, the categories were built based on the chemicals having a common structural scaffold, which was expected to induce the same toxicological effect Case Study 2019-5, 2019-6, 2019-7 and 2019-8 established *in vitro* testing strategies. Especially, 2019-5, 2019-7 and 2019-8 used AOPs for building the testing strategies. These case studies demonstrated that category approaches based only on structural similarity alone are not sufficient and that the applied *in vitro* testing strategies could predict *in vivo* toxicity correctly. The various lessons and learnings gained from the review experience of the case studies can be considered for improvement of the grouping guidance (OECD, 2014a) in the future.

The eight case studies in the fifth review cycle utilised various type of data from NAMs, some employed AOP based testing strategies and some focused on negative prediction. The five areas for further developing guidance on the use or reporting of NAMs data were identified. The project team agreed that better understanding how to evaluate ToxCast data and reporting of docking/modelling approaches, which were identified as new areas, would benefit from further guidance. Key findings were gained from the case studies that applied AOP based testing strategies such as elements surrounding documentation and uncertainty when using non-endorsed AOPs, and the necessary level of coverage of KEs in AOP based strategy. Based also on discussions regarding presentation of crucial data and information, three areas for further guidance development in this area were identified. In addition, the project team identified revision of the grouping guidance and justification of negative prediction as high priorities.

Four consideration topics from the fifth review cycle are discussed in this document. Margins of Internal Exposure can be considered in the risk assessment if the data on blood concentrations are available for the target and source chemicals. Coverage of the key events in AOP is important for developing an AOP based testing strategy. While covering all the MIE/KEs is ideal, it would be technically difficult. In this case, a justification on coverage of MIE/KEs should be provided. Also, it was indicated that there is uncertainty on the use of non-endorsed AOPs. Use of a stressor of the AOP as a structural analogue could reduce this type of the uncertainty. Finally, the advantage of the use of AOP networks was discussed in this document.

In summary, useful lessons and learnings were gained through the review of the eight read-across case studies. These lessons and learnings provide key considerations for the use of the NAMs in the context of the IATAs as well as improvements of read-across approaches.

7. REFERENCES

Blackburn, K. and S.B. Stuard (2014) A Framework to Facilitate Consistent Characterization of Read Across Uncertainty. *Regulatory Toxicology and Pharmacology.* Vol. 68, No. 3, pp 353-62.

ECHA (2017). Read-Across Assessment Framework (RAAF), ECHA-17-R-01-EN, European Chemicals Agency, Helsinki. https://doi.org/10.2823/619212

Fisher, J. S., K. J. Turner, D. Brown, and R. M. Sharpe (1999). Effect of Neonatal Exposure to Estrogenic Compounds on Development of the Excurrent Ducts of the Rat Testis through Puberty to Adulthood, *Environ Health Perspect*, Vol. 107, No. 5, pp. 397-405. https://doi.org/10.1289/ehp.99107397.

IPCS (2005). Chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration–response assessment. International Programme on Chemical Safety Workshop on Incorporating Uncertainty and Variability into Risk Assessment (2000: Berlin, Germany), WHO, Geneva, Switzerland. https://apps.who.int/iris/handle/10665/43294

OECD (2014a). Guidance on Grouping of Chemicals, Second Edition, Series on Testing & Assessment No. 194. ENV/JM/MONO(2014)4, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2014b). Weight of Evidence Assessment for the Skin Sensitisation Potential of 4-Isopropylaniline (Cumidine, CAS 99-88-7), Series on Testing & Assessment No. 199. ENV/JM/MONO(2014)5, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2014c). Guidance Document for Describing Non-Guideline *In vitro* Test Methods, Series on Testing and Assessment No. 211, ENV/JM/MONO(2014)35, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2016a). Report on Considerations from Case Studies on Integrated Approaches for Testing and Assessment (IATA), First Review Cycle (2015), Case Studies on Grouping Methods as a Part of IATA, Series on Testing & Assessment No. 250. ENV/JM/MONO(2016)48, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2016b). Case Study on the Use of Integrated Approaches for Testing and Assessment for *In vitro* Mutagenicity of 3,3' Dimethoxybenzidine (DMOB) Based Direct Dyes, Series on Testing & Assessment No. 251. ENV/JM/MONO(2016)49, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2016c). Case Study on the Use of Integrated Approaches for Testing and Assessment for Repeat Dose Toxicity of Substituted Diphenylamines (SDPA), Series on Testing & Assessment No. 252. ENV/JM/MONO(2016)50, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2016d). Case Study on the Use of an Integrated Approach to Testing and Assessment for Hepatotoxicity of Allyl Esters, Series on Testing & Assessment No. 253. ENV/JM/MONO(2016)51, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2016e). Case Study on the Use of an Integrated Approach to Testing and Assessment of the Bioaccumulation Potential of Degradation Products of 4,4'-Bis(Chloromethyl)-1,1'- Biphenyl, Series on Testing & Assessment No. 254. ENV/JM/MONO(2016)52, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2016f). Adverse Outcome Pathway on Protein Alkylation Leading to Liver Fibrosis, Series on Adverse Outcome Pathways No. 2. OECD, Paris. https://dx.doi.org/10.1787/5jlsvwl6g7r5-en

OECD (2016g). Users' Handbook supplement to the Guidance Document for developing and accessing Adverse Outcome Pathways, Series on Testing & Assessment No. 233, OECD Series on Adverse Outcome Pathways No. 1, ENV/JM/MONO(2016)12, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2016h). Guidance Document on the Reporting of Defined Approaches to Be Used within Integrated Approaches to Testing and Assessment No. 255, ENV/JM/MONO(2016)28, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2016i). Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to Be Used within Integrated Approaches to Testing and Assessment (IATA) for Skin Sensitisation, Series on Testing and Assessment No. 256, ENV/JM/MONO(2016)29, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2017a). Report on Considerations from Case Studies on Integrated Approaches for Testing and Assessment (IATA), Second Review Cycle (2016), Series on Testing & Assessment No. 270. ENV/JM/MONO(2017)22, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2017b). Case Study on the Use of an Integrated Approach to Testing and Assessment for Repeated-Dose Toxicity of Phenolic Benzotriazoles, Series on Testing & Assessment No. 271. ENV/JM/MONO(2017)23, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2017c). Case Study on the Use of Integrated Approaches for Testing and Assessment for Pesticide Cumulative Risk Assessment & Assessment of Lifestage Susceptibility, Series on Testing & Assessment No. 272. ENV/JM/MONO(2017)24, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2017d). Case Study on the Use of Integrated Approaches for Testing and Assessment of 90-Day Rat Oral Repeated-Dose Toxicity for Selected n-Alkanols: Read-Across, Series on Testing & Assessment No.273. ENV/JM/MONO(2017)25, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2017e). Case Study on the Use of Integrated Approaches for Testing and Assessment of 90-Day Rat Oral Repeated-Dose Toxicity for Selected 2-Alkyl-1-alkanols: Read-Across, Series on Testing & Assessment No. 274. ENV/JM/MONO(2017)26, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2017f). Chemical Safety Assessment Workflow Based on Exposure Considerations and Non-Animal Methods, Series on Testing & Assessment No. 275. ENV/JM/MONO(2017)27, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2018a). Report on Considerations from Case Studies on Integrated Approaches for Testing and Assessment (IATA), Third Review Cycle (2017), Series on Testing & Assessment No. 289.

ENV/JM/MONO(2018)25, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2018b). Case Study on the Use of Integrated Approaches for Testing and Assessment (IATA) for Estrogenicity of Substituted Phenols, Series on Testing & Assessment No. 290. ENV/JM/MONO(2018)26, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2018c). Prioritisation of chemicals using the Integrated Approaches for Testing and Assessment (IATA)-based Ecological Risk Classification, Series on Testing & Assessment No. 291. ENV/JM/MONO(2018)27, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2018d). Case Study on Grouping and Read-across for Nanomaterials Genotoxicity of Nano-TiO_2, Series on Testing & Assessment No. 292. ENV/JM/MONO(2018)28, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2018e). A Case Study on the Use of Integrated Approaches for Testing and Assessment for Sub-Chronic Repeated-Dose Toxicity of Simple Aryl Alcohol Alkyl Carboxylic Esters: Read-Across, Series on Testing & Assessment No. 293. ENV/JM/MONO(2018)29, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2018f). Adverse Outcome Pathway on Inhibition of the mitochondrial complex I of nigro-striatal neurons leading to parkinsonian motor deficits, Series on Adverse Outcome Pathways No. 7., OECD, Paris. https://doi.org/10.1787/2415170X

OECD (2019a). Report on Considerations from Case Studies on Integrated Approaches for Testing and Assessment (IATA), Fourth Review Cycle (2018), Series on Testing & Assessment No. 307. ENV/JM/MONO(2019)26, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2019b). Case Study on the Use of an Integrated Approach to Testing and Assessment for Testicular Toxicity of Ethylene Glycol Methyl Ether (EGME)-Related Chemicals, Series on Testing & Assessment No. 308. ENV/JM/MONO(2019)27, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2019c). Case Study on the Use of Integrated Approaches for Testing and Assessment for Estrogen Receptor Active Chemicals, Series on Testing & Assessment No. 309. ENV/JM/MONO(2019)28, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2020a). Case Study on use of an Integrated Approach to Testing and Assessment (IATA) and New Approach Methods to Inform a Theoretical Read-Across for Dermal Exposure to Propylparaben from Cosmetics, Series on Testing & Assessment No. 320. ENV/JM/MONO(2020)16, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2020b). Case Study on the use of Integrated Approaches for Testing and Assessment for Systemic Toxicity Arising from Cosmetic Exposure to Caffeine, Series on Testing & Assessment No. 321. ENV/JM/MONO(2020)17, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2020c). Case Study on the Use of Integrated Approaches for Testing and Assessment for 90-Day Rat Oral Repeated-Dose Toxicity of Chlorobenzene-Related Chemicals, Series on Testing & Assessment No. 322. ENV/JM/MONO(2020)18, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2020d). Case Study on the Use of Integrated Approaches for Testing and Assessment to Inform Read-Across of p-Alkylphenols: Repeated-Dose Toxicity, Series on Testing & Assessment No. 323. ENV/JM/MONO(2020)19, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2020e). Case Study on the Use of Integrated Approaches to Testing and Assessment for Prediction of a 90 day Repeated Dose Toxicity Study (OECD 408) for 2-Ethylbutyric acid Using a Read-Across Approach to Other Branched Carboxylic acids, Series on Testing & Assessment No. 324. ENV/JM/MONO(2020)20, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2020f). Case Study on the Use of Integrated Approaches to Testing and Assessment for Read-Across Based Filling of Developmental Toxicity Data Gap for Methyl Hexanoic acid, Series on Testing & Assessment No. 325. ENV/JM/MONO(2020)21, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2020g). Case Study on the Use of Integrated Approaches to Testing and Assessment for Identification and Characterisation of Parkinsonian Hazard Liability of Deguelin by an AOP-based Testing and Read Across Approach., Series on Testing & Assessment No. 326. ENV/JM/MONO(2020)22, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

OECD (2020h). Case Study on the Use of Integrated Approaches to Testing and Assessment for Mitochondrial Complex-III-Mediated Neurotoxicity of Azoxystrobin - Read-Across to Other Strobilurins, Series on Testing & Assessment No. 327. ENV/JM/MONO(2020)23, OECD, Paris. Available at http://www.oecd.org/chemicalsafety/testing/series-testing-assessment-publications-number.htm.

Schultz T.W., Richarz A.-N. and Cronin M. T.D. (2019) Assessing uncertainty in read-across: Questions to evaluate toxicity predictions based on knowledge gained from case studies. Computational Toxicology 9 (2019) 1-11. https://doi.org/10.1016/j.comtox.2018.10.003

Schultz TW, Amcoff P, Berggren E, et al. (2015). A strategy for structuring and reporting a read-across prediction of toxicity. *Regul Toxicol Pharmacol.* Vol. 72, No. 3, pp. 586 - 601. doi:10.1016/j.yrtph.2015.05.016

Shibata M, Wachi M, Kawaguchi M, Kojima J, Onodera K (2000). Teratogenic and fetal toxicity following intravenous theophylline administration in pregnant rabbits is related to maternal drug plasma levels. Methods Find Exp Clin Pharmacol 22: 101–107

US EPA (2011). Recommended use of body weight3/4 as the default method in derivation of the oral reference dose. Risk Assessment Forum. EPA/100/R11/0001. United States Environmental Protection Agency, Washington D.C.

WHO (2010). Characterization and Application of Physiologically Based Pharmacokinetic Models in Risk Assessment. IPCS Harmonization Project Document No. 9, World Health Organization, Geneva. https://apps.who.int/iris/handle/10665/44495

Wu, S., K. Blackburn, J. Amburgey, J. Jaworska and T. Federle (2010) A Framework for Using Structural, Reactivity, Metabolic and Physicochemical Similarity to Evaluate the Suitability of Analogs for SAR-based toxicological assessments. *Regulatory Toxicology and Pharmacology*. Vol. 56, No 1, pp 67-81.

Annex 1. Template for IATA Case Studies on Chemical Grouping (Read-across)

Title: Case Study on the use of Integrated Approaches for Testing and Assessment for "Target Endpoint(s)" of "Target Chemical(s)/Category"

(N.B. The following template should not be viewed as a strict structure, but rather identifies the information that should be included in this type of case study. Depending on the specific case study additional information/(sub)section(s) may be required or particular subsections may not apply. The order of the (sub)sections of the template can be changed and two or more (sub)sections of the template can be merged, as necessary. The titles of a (sub)section can be changed as necessary. The template will be revised based on experience with use. A case study based on the template is expected to be assessed as stand-alone, thus needs to contain all necessary information and appropriate links for a detailed assessment.)

Abstract / Synopsis / Executive summary

(This section should provide a brief overview of the case study, including the objectives, concepts, methodologies, outcomes and conclusion in about 300 words.)

Introduction

(This should include a very short summary of the background/problem formulation, purpose, endpoints covered and description of the target chemical(s)/category)

Table of Contents

A. Purpose

a. Purpose of use

Specify the purpose of use of the IATA (e.g. regulatory context: hazard identification, hazard characterisation, risk assessment, screening etc.). If in a regulatory context, provide a short description of any (e.g. legal) requirements for the IATA approach to be accepted.

b. Target chemical(s)/category definition [See 3.2.3.1 "Chemical identity and composition" of the grouping guidance (OECD, 2014a)]

- For analogue approach, provide the chemical descriptor common identifiers (including CAS number, name and composition including impurities) and chemical structure(s) of the target substance(s).
- For category approach, provide a summary of the common features of the category members; describe the boundaries; allowed variations (e.g. in chemical structures); composition including impurities; and if known, any limitations in the information.

c. Endpoint(s)

- Identify the endpoint(s) for which the analogue/category approach is applied. Endpoint-specific considerations/approaches may be needed if more than one endpoint is addressed by the read-across.

d. Exposure information (if needed)

- Provide the considered exposure for the grouping/read-across, such as route(s) of administration covered by the experimental model (e.g. oral), the population of interest (.e.g. human, ecological), and as relevant, any route to route or *in vivo/in vitro* extrapolations that were applied to inform the grouping/read-across

Tip

- The description of the purpose of use is important for considering the acceptable uncertainty of the case study, which could be linked to the uncertainty assessment. For example, if the conclusion derived by case study is renewable in a framework such as tiered-approach, this needs to be clearly stated (see case studies OECD, 2016b and 2016c).
- As the goal of the OECD IATA Case Studies project is to discuss case studies which would lead to regulatory application a description of the regulatory relevance is important to contextualise the case and discuss the further development of guidance and how use IATA for regulatory purpose.
- It is recommended to specify the analogues and justification for data gap filling, used for each addressed endpoint, in order to identify for what endpoints is the analogue/category being applied.

Tip for nanomaterials

- The parameters to be considered for grouping and read-across of nanoforms and their relevance for human health and environmental endpoints are for example surface chemistry, size, shape and surface area, along with physical/chemical properties. (See *"1.2 Target chemicals"* of the case study 2017-3 (OECD, 2018d))
- For the complete list of parameters and more information on grouping of nanomaterials please, see "ECHA (2017), Guidance on information requirements and chemical safety assessment Appendix R.6-1 for nanomaterials applicable to the Guidance on QSARs and Grouping of Chemicals", Appendix I, Table 2: Key physicochemical parameters to be considered for grouping and read-across of nanoforms and their relevance for human health and environmental endpoints.

B. Hypothesis for the analogue approach/category [See 2.4 "The mechanistic basis of using analogues or chemical categories" and 3.2.1 "Hypothesis and evidence based approaches" of the grouping guidance (OECD, 2014a)]

- For an analogue approach, describe the characteristics that a substance must have to be suitable as a source substance, including a description of the composition of the source substance (e.g. level of purity). Provide the hypothesis for why read-across can be performed between the source and target chemicals [See 4.2.2 "Step 1: Identification of potential analogues" of the grouping guidance (OECD, 2014a)].
- For a category approach, provide the hypothesis for why the category was formed including the relational features of the category. Provide the hypothesis for why read-across can be performed within the category [See 5.2.2 "Step 1: Develop

category hypothesis and definition and identify category members" of the grouping guidance (OECD, 2014a)].

- These hypotheses can be argued by a number of elements as follows [See 3.2.3 "Elements for a read-across justification" of the grouping guidance (OECD, 2014a)]. Chemical identity and composition, including level of purity
  - Chemical identity and composition, including level of purity [See 3.2.3.1 "*Chemical identity and composition* of the grouping guidance (OECD, 2014a)]
  - Physical-chemical properties and other molecular description [See 3.2.3.2 "*Physical-chemical properties* of the grouping guidance (OECD, 2014a)]
  - Kinetics: Absorption, distribution, metabolism and excretion [See 3.2.3.3 "*Absorption, distribution, metabolism and excretion* of the grouping guidance (OECD, 2014a)]
  - Mode/Mechanism of action or adverse outcome pathways (MOA/AOP) [See 3.2.3.4 "*Mode/ mechanisms of action or adverse outcome pathways (MOA/AOP)* of the grouping guidance (OECD, 2014a)]
  - Chemical / biological interaction [See 3.2.3.5 "*Chemical / biological interaction* of the grouping guidance (OECD, 2014a)]
  - Toxicological and epidemiological information, along with information from new approach methodologies (NAMs) [See 3.2.3.6 "*Responses found in in vitro methods* of the grouping guidance (OECD, 2014a)]
  - Information obtained from other endpoints/species/routes
  - Information on fate in the environment (hydrolysis, biodegradation)
  - The route and duration of expected exposure

Ideally, all elements relevant for the assessment should be addressed. In addition, it is recommended to describe how the (combination of) elements support the hypothesis (see for more detail OECD, 2014a).

- Especially, hypothesis of mechanism(s) (AOP/MOA) for how the target chemical induces target endpoint toxicity need to be described in this section. Hypothesis of structural boundaries and limitations for the approach should also be clearly described, including possible impact of structural dissimilarities. The graphical representation of the AOP would be helpful for the reader and key references (See "Graphical Representation of the AOP" at section 1- AOP Description (OECD, 2016g)). The tools in the AOP-KB should be referred to as appropriate (e.g. AOP wiki, Effectopedia etc.). Identifying the relevant AOP from AOP wiki is required. Please provide the AOP number, status on AOP-wiki and the link. For AOPs that are not documented, consider the "Users' Handbook supplement to the Guidance Document for developing and accessing Adverse Outcome Pathways" (OECD, 2016g) - although an entire AOP description is not the purpose here.
- Describe how a data gap is intended to be filled for the purpose of read-across (the prediction model used - worst case scenario, regression etc.). Here it could also be justified as to why read-across is sufficient, and why further testing is not needed.

Tip

- Hypothesis needs to be described as a testable format.
- For the hypothesis that metabolite induces target effect, the effects induced by other metabolites other than the toxicant need to be considered (see "("*2.2 Elements for a read-across hypothesis* of the case study 2016-3" (OECD, 2016d)).

Tip for nanomaterials

- Provide an explanation which parameters are critical for the analogue approach/category hypothesis.
- Hypothesis could be argued using for example the following physicochemical and chemical properties (list is not exhaustive) (see for example "2.2 Characterisation of the analogue nanoforms" of 2017-3 (OECD 2018d)):
  - Chemical composition
  - Surface chemistry (including coating chemicals and the coating ratio)
  - Impurity
  - Size (including primary particle diameter)
  - Shape (including surface chemistry)
  - Surface area
  - Solubility
  - Hydrophobicity
  - Zeta potential
  - Dispersibility
  - Dustiness
  - Physical hazard
  - Biological (re)activity
  - Photoreactivity
- For the complete list of parameters and more information on grouping of nanomaterials please, see "ECHA (2017), Guidance on information requirements and chemical safety assessment Appendix R.6-1 for nanomaterials applicable to the Guidance on QSARs and Grouping of Chemicals", Appendix I, Table 2: Key physicochemical parameters to be considered for grouping and read-across of nanoforms and their relevance for human health and environmental endpoints.

C. Source chemicals/Category members [See 2.3 "Selecting analogues/Creating chemical categories and setting boundaries", 4.2.2 "Step 1: Identification of potential analogues" and 5.2.2 "Step 1: Develop category hypothesis and definition and identify category members" of the grouping guidance (OECD, 2014a)]

a. Identification and selection of source chemicals/category members

- Provide the selection criteria, based on the hypothesis described in section B, that were used to identify the source chemicals/category members.
- Provide the rationale for selection of analogue(s)/category members with respect to the defined purpose and endpoint.

- Provided consideration of selection bias in the choice of source chemicals when using the analogue or category approach (e.g. data quality and completeness, support for hypothesis etc.).
- Describe the methods used to identify the source chemicals/category members (e.g. inventories and tools used should be provided). Listing search criteria to establish initial pool of candidate analogues is helpful.

b. List of source chemicals/ category members

- Provide the common chemical identifiers (including CAS number, name and composition including impurities) and chemical structure(s) of the source chemicals/category members. (See 3.2.3.1.3 "Examples of categories and structural relationships" of the grouping guidance (OECD, 2014a); example of the chemical identifiers for UVCBs)

Tip

- Not only structural similarity but also impacts of structural differences to the target effect need to be considered when selecting analogues. A clear description of boundaries is also important.

D. Justification of data gap filling

a. Data gathering [See 4.2.3 "Step 2: Data gathering for the analogues" and 5.2.3 "Step 2: Gather data for each category member" of the grouping guidance (OECD, 2014a)]

- Provide the methods used for gathering the data for target and source chemicals/category members (e.g. selection criteria of the data, data source).

b. Data and methods [See 4.2.4 "Step 3: Evaluation of available data for adequacy", 4.2.5 "Step 4: Construct a matrix of data availability" (analogue approach); 5.2.4 "Step 3: Evaluate available data for adequacy"; 5.2.5 "Step 4: Construct a matrix of data availability" (category approach) of the grouping guidance (OECD, 2014a)]. Provide a matrix of data (see data matrix template) with the following:

- If mass unit such as mg/kg-bw is used in the data, it should also be expressed in molar units such as mmol/kg-bw.
- Provide detailed data as necessary (in case that the detailed data are used for the justification of the hypothesis). The appropriate degree of detail of the data should be considered in the context of the purpose of case study. Examples of reports of detailed data can be found in past IATA case studies[4].
- If data from non-guideline test methods are included, provide descriptions of the methods or links to sources that summarise the methods. The appropriate degree of detail of the description should be considered in the context of the purpose of the case study. A template for the description is available in the OECD guidance document No. 211 (OECD, 2014c) Examples of description using the template can found in JRC EURL ECVAM Database service on Alternative Methods to animal

[4] http://www.oecd.org/chemicalsafety/risk-assessment/iata-integrated-approaches-to-testing-and-assessment.htm#casestudies

experimentation (DB-ALM)[5] and U.S. EPA Toxicity ForeCaster (ToxCast™) Data[6].

- If QSAR data are included, provide the name, version, owner of the models used for deriving QSAR estimation data QSAR models if not already described elsewhere should be reported using the QSAR Model Reporting Format (QMRF), and individual predictions, if applicable, should be reported using the QSAR Prediction Reporting Format (QPRF). A QMRF inventory is maintained by JRC that can be utilised as a resource of QMRFs and its reference number can be referred to JRC QSAR Model databases[7].
- If data derived from defined approaches of IATA are included, provide the descriptions of the defined approaches. A template for the description and case study examples are available in OECD guidance documents (OECD, 2016h; 2016i). In this section, please describe the individual information sources used and data interpretation procedure applied (See "6. Description of the individual information sources used (see Annex II)" and "7. Data interpretation procedure applied" of the OECD guidance (OECD, 2016h)).
- Provide all available suitable information regarding the defined purpose, including the data from *in silico*, *in vitro* and *in vivo*. If possible, the cells of the data matrix should also indicate the available key study results.

c. Justification [See 2.5 "Robustness of a chemical category and of an analogue approach", 2.6 "The interdependence between categories and (Q)SARs.", 4.2.6 "Step 5: Assess the adequacy of the analogue approach and fill the data gap" and 5.2.6 "Step 5: Perform a preliminary evaluation of the category and fill data gaps" of the grouping guidance (OECD, 2014a)]

- Based on the data matrix, summarise how these data support the hypothesis described in section B.
- Identify similarities/trends in the experimental data of the endpoint(s) for the chemicals in the data matrix and verify their concordance with hypothesis described in section 2.
- Identify which elements drive the toxicity/endpoint.
- For category approach, describe the set of inclusion and/or exclusion rules that identify the boundaries within which reliable estimations can be made for category members. A broader consideration including mechanistic information, profiling computational methods, screening with non-standard *in vitro* tests should be given. Clearly indicate the boundaries of the category and for which substances the category does not hold i.e. substances outside scope of predictions e.g. by endpoint [See 5.2.4 "Step 3: Evaluate available data for adequacy" of the grouping guidance (OECD, 2014a): example of outlier].
  The applicability domain of each estimation method including QSAR and alternative methods should be discussed based on the consistency between the estimation data and the experimental data of the source chemical(s)/category members.

[5] JRC, EURL ECVAM Database service on Alternative Methods to animal experimentation (DB-ALM). https://ecvam-dbalm.jrc.ec.europa.eu/

[6] U.S. EPA, Toxicity ForeCaster (ToxCast™) Data https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data

[7] JRC, QSAR Model Database. https://qsardb.jrc.ec.europa.eu/qmrf/

Tip

- Reliability of each QSAR prediction result needs to be described in terms of the applicability domain of QSARs. For example, it can be discussed by the coverage of the fragments in the training sets (See the case study 2015-4 (OECD, 2016e)).
- It is recommended that every approach be described separately, e.g. if read-across, QSAR and *in vitro* tests are used, every one of these approaches would need to be described separately before combining in IATA.
- Please explain how satisfying comprehensiveness/coverage of the data gathering is achieved.
- For transparency, the data reporting is an important aspect. For example, if estimation relies on qualitative/semi-quantitative explanations, it is important to explain how these support quantitative estimations where needed for that purpose. Further, to demonstrate coherence of findings and similarity/trend/strength of effects sufficient reporting of the experimental data is needed (e.g. type, degree and dose levels). If data reveal inconsistencies or similar studies show different concerns this would also benefit from explanation.
- Please, try to ensure maximal use of existing experimental information before considering QSAR predictions.
- Alert-based system work best for predicting an alert and not lack of it, unless there are structure-specific definitions for lack of activity

Tip for nanomaterials (See "5. JUSTIFICATION OF DATA GAP FILLING" of the case study 2017-3 (OECD, 2018d))

- Describe methods used for measuring the endpoints
- It is recommended to describe which methodologies for measurements of the relevant parameters are applied, and to describe what are differences between the methodologies are, if applicable.
- Identify which parameters are relevant to which endpoints, if possible.
- For the complete list of parameters and more information on grouping of nanomaterials, please see ECHA (2017) "Guidance on information requirements and chemical safety assessment Appendix R.6-1 for nanomaterials applicable to the Guidance on QSARs and Grouping of Chemicals", Appendix I, Table 2: Key physicochemical parameters to be considered for grouping and read-across of nanoforms and their relevance for human health and environmental endpoints.

E. Strategy for and integrated conclusion of data gap filling

a. Uncertainty

- Discuss the uncertainty of each factor for the read- across. For the given purpose, it seems that the consideration of uncertainty may start from the choice of hypothesis (like in Appendix 1). Another consideration includes severity of effect, if it is present. (e.g. Does the number of targets matter? Could all targets meet all sources? How read-across could be addressed (e.g. subgrouping)?)

- Aspects can include uncertainty and confidence associated with the data (e.g. applicability domain, type and quality) and assumptions used to develop the similarity rationale of the analogues/category members and uncertainty and confidence associated with the underlying data used for read across from the source chemicals.
- The following is an example of reporting uncertainty (Please modify as appropriate and also it is recommended to describe what is not addressed.): Examples of modified templates, which were used for past case studies, are shown in Appendix 1, 2,and 3. Appendix 4 lists a series of questions to guide through the assessment of uncertainties. and 4 Also, refer to the case studies published in the past[8].

Factor	Uncertainty (low, medium, high)	Comment
Hypothesis used for the read across		
Structural similarity		
Similarity of physico-chemical properties		
Similarity of toxicokinetics data		
Similarity of other supportive data (e.g. data related to key event)		
Number of analogues used for the read across		
Quality of the endpoint data used for the read across		
Similarity of the endpoint data (among source chemicals)		
Concordance and weight of evidence of all data used for justifying the hypothesis		
Overall uncertainty of the read-across		

Tip

- When using ranks to indicate uncertainties (e.g. low, medium, high), definitions should be provided.

[8] http://www.oecd.org/chemicalsafety/risk-assessment/iata-integrated-approaches-to-testing-and-assessment.htm#casestudies

Tip for nanomaterials

- In addition to the above mentioned aspects, the following should be considered in the characterisation of uncertainties related to the analogue/category approach for nanomaterials (See "7. UNCERTAINTY ASSESSMENT" of the case study 2017-3 (OECD, 2018d)):
  - Complexity of nanostructures: similarity, category boundaries and members
  - Identity characterisation of the nanomaterials
  - Variability of the measurements, test system relevance for nanomaterials and possible nanospecific artefacts in assays

For more information on grouping of nanomaterials please, see "ECHA (2017), Guidance on information requirements and chemical safety assessment Appendix R.6-1 for nanomaterials applicable to the Guidance on QSARs and Grouping of Chemicals"

b. Integrated conclusion

- Provide the strategy used to fill the data gap and integrated conclusion of data gap filling, including description how the data gap is actually filled (e.g. average, most sensitive, similarity weighted, qualitative). In case of category approach, indicate proposed conclusion/value for each data gap. If prediction models were used, please describe the satisfaction with parameters related to the prediction.
- Give discussion of remaining uncertainties and how they might be addressed.
- Finally, provide a short conclusion wrapping up the outcome of the evaluation with linking to the given purpose.

References

Annex

Appendix 1: Example of Reporting Template of Uncertainty (1)

The template was prepared based on the following frameworks and was used for the case studies 1&2 in 2015 of the project (OECD, 2016b; 2016c).

- Wu, S., K. Blackburn, J. Amburgey, J. Jaworska and T. Federle (2010) A Framework for Using Structural, Reactivity, Metabolic and Physicochemical Similarity to Evaluate the Suitability of Analogs for SAR-based toxicological assessments. Regulatory Toxicology and Pharmacology. Vol. 56, Issue 1, pp 67-81.
- Blackburn, K. and S.B. Stuard (2014) A Framework to Facilitate Consistent Characterization of Read Across Uncertainty. Regulatory Toxicology and Pharmacology. Vol. 68, Issue 3, pp 353-62.

An overview of the template is shown below. Please refer to the original papers and the case studies above for details.

Part 1: Analogue suitability rating for read-across [a]

Evaluation Criteria [b]	Question [c]	Uncertainty [d]
Structure and reactivity	Do the target & analogue have similar structural features& chemical reactivity?	
Metabolism	Do the target & analogue have similar metabolic pathways?	
Physicochemical Properties	Do the target & analogue have similar phys-chem properties?	
…….		
Overall "suitability rating" [e]		

a This table is based on the decision tree of the framework by Wu et al. (2010)
b Criteria used for evaluating the suitability of analogues.
c Question and answer used for evaluating the criteria.
d Description of the uncertainties in the answer to the question.
e Rank (Suitable, Suitable with interpretation, Not suitable, Suitable with preconditions) derived from the decision tree.

Part 2: Uncertainty associated with the prediction of hazard using read across [e]

Analogue Data Set Characteristics [f]	Comment [g]
Number of analogues contributing data	
Robustness of analogue data set	
Concordance of effect(s)	
…….	
Overall uncertainty of read across prediction [h]	

e This table is based on the framework by Blackburn and Stuard (2014).
f Analogue data set characteristics used for evaluating overall uncertainty of read across prediction.
g Description of the evaluation results of the analogue data set characteristics obtained by answering the questionnaire of the framework.
h Rank of overall uncertainty of read across prediction derived from the evaluation results of analogue data set characteristics (Low, Low to Moderate, Moderate, High) with the description of the reason.

Appendix 2: Example of Reporting Template of Uncertainty (2)

The template was developed in the following framework and was used for the case studies 3&4 in 2016 of the project (OECD 2017d; 2017e) as well as in case study 4 in 2017 (OECD 2018e).

- Schultz, T.W., P. Amcoff, E. Berggren, F. Gautier, M. Klaric, D.J. Knight, C. Mahony, M. Schwarz, A. White and M.T.D. Cronin (2015), A Strategy for Structuring and Reporting a Read-across Prediction of Toxicity. Vol. 72, Issue 3, pp 586-601.

An overview of the template is shown below. Please refer to the original paper and the case studies above for details.

Part 1: Parameters and associated uncertainty used to justify category membership

Justification Parameter [a]	Data Uncertainty [b]	Strength of Evidence [c]	Comment [d]
Structural Similarity	Table Cell (Alt+E)		
Phys/Chem Properties			
Metabolic Similarity			
Mechanistic Similarity			
Trends in Effects			
…….			
Overall uncertainty in similarity of category members			

a Similarity parameter used for justifying the category.
b Rank of uncertainty (low, medium, high) associated with underlying data used for analysis
c Rank of consistency (low, medium, high) within the data
d Description of the reason for the assignment of the ranks of the uncertainty and strength of evidence
e Rank of overall uncertainty (low, medium, high) and description of the reason

Part 2: Uncertainty associated with the prediction of hazard and dose-response using read-across

Factor [e]	Uncertainty [f]	Comment [g]
Number of analogues contributing data		
Robustness of analogue data set		
Concordance of effects		
Concordance of potency		
Severity of critical effect		
…….		
Overall uncertainty of read-across (low, medium, high)		

e Uncertainty factor associated with the prediction of hazard and dose-response using read-across.
f Rank of uncertainty (low, medium, high)
g Description of the reason for the assignment of the ranks of the uncertainty
h Rank of overall uncertainty (low, medium, high) and description of the reason

Appendix 3. Examples for reporting uncertainty from the ECHA Read-across Assessment Framework (RAAF)[9] (3)

Examples of assessment elements (AEs) for an analogue approach, for all RAAF read-across scenarios and detailed description of the AEs see (ECHA, 2017).
Assessment Elements for Scenario 1 (read-across based on hypothesis for (bio)transformation to common products)

AE A.1 Common AE: Identity and characterisation of source substance
AE A.2 Common AE: Link of structural similarity and differences with the proposed prediction
AE A.3 Common AE: Reliability and adequacy of the source study
AE 1.1 Scenario-specific AE: Formation of common (identical) compound(s)
AE 1.2 Scenario-specific AE: The biological targets for the common compound(s)
AE 1.3 Scenario-specific AE: Exposure of the biological target(s) to the common compound(s)
AE 1.4 Scenario-specific AE: The impact of parent compounds
AE 1.5 Scenario-specific AE: Formation and impact of non-common compounds
AE A.4 Common AE: Bias that influences the prediction

Assessment Elements for Scenario 2 (read-across based on hypothesis that different compounds have the same type of effects)

AE A.1 Common AE: Identity and characterisation of source substance
AE A.2 Common AE: Link of structural similarity and differences with the proposed prediction
AE A.3 Common AE: Reliability and adequacy of the source study
AE 2.1 Scenario-specific AE: Compounds the test organism is exposed to
AE 2.2 Scenario-specific AE: Common underlying mechanism, qualitative aspects
AE 2.3 Scenario-specific AE: Common underlying mechanism, quantitative aspects
AE 2.4 Scenario-specific AE: Exposure to other compounds than to those linked to the prediction
AE 2.5 Scenario-specific AE: Occurrence of other effects than covered by the hypothesis and justification
AE A.4 Common AE: Bias that influences the prediction

[9] https://echa.europa.eu/support/registration/how-to-avoid-unnecessary-testing-on-animals/grouping-of-substances-and-read-across

Appendix 4. Examples for reporting uncertainty (4)

30 questions relating to 12 types of uncertainty were identified to be addressed in assessing uncertainties of a read-across in the following study:

- Terry W. Schultz, Andrea-Nicole Richarz, Mark T.D. Cronin (2019) Assessing uncertainty in read-across: Questions to evaluate toxicity predictions based on knowledge gained from case studies. Computational Toxicology, Vol. 9, pp. 1-11

Uncertainty in Read-Across	Uncertainty in Read-Across
The context of, and relevance to, the regulatory use of the read-across prediction as defined by appropriate problem formulation	• Is the regulatory purpose of the read-across prediction clearly defined? • Is the acceptable level or degree of uncertainty for the stated purpose defined? • Is the stated acceptable level or degree of uncertainty appropriate for the stated regulatory purpose?
Type of category/group including the definition of the applicability domain	• Is the read-across approach (e.g., analogue or category) clearly reported? • Are the target and source chemicals clearly identified? • Is the applicability domain of the analogue or category defined? • Do target and source chemicals fit within the defined applicability domain?
The premise or hypothesis of the read-across.	• Is the hypothesis on which the read-across is based clearly stated and presented in sufficient detail to be assessed?
Mechanistic plausibility including completeness of the understanding of the MoA or AOP	• How clearly does the hypothesis state the chemical and biological mechanisms underpinning the toxic effect being read across? • Is there sufficient experimental information provided to support the proposed chemical and toxicological mechanisms? • How extensively does the experimental information provided support the mechanistic plausibility and / or the AOP or MoA on which the read-across is based?
Similarity in chemistry	• Are the chemical structures (i.e., 2D structure, isomers, SMILES and molecular formula) reported for the derivatives used in the read-across? • Are the dissimilarities in chemical structure reported and are they toxicologically relevant? • Are the relevant molecular and physico-chemical properties (e.g., for molecular size, hydrophobicity, solubility, volatility, degradation etc.) reported for the derivatives used in the read-across? • Are the dissimilarities in molecular and physico-chemical properties reported and are they toxicologically (or pharmacokinetically) relevant?
Toxicodynamic similarity	• Is there sufficient and consistent toxicodynamic information provided to establish similarity in the hazard of the derivatives used in the read-across?
Toxicokinetic similarity	• Is there sufficient ADME information provided to establish toxicokinetic similarity for the derivatives used in the read-across? • Are any dissimilarities in ADME properties (and, as appropriate, metabolism / degradation) toxicologically relevant?
The quality of the apical endpoint data used to fill the data gap	• Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the data read across reported clearly? • Has the quality of the data to be read across been assessed and are they sufficient to meet the purpose of the exercise i.e., complete and of sufficient quality?
The consistency in the effects and severity of the apical *in vivo* hazard and their concordance with regards to the intermediate and apical effects and potency data	• Is the qualitative expression of the data reported and is it consistent among the source chemicals? • Is the potency of the hazard reported and is it consistent among the source chemicals? • What are the temporal relationships between relevant endpoints? • What are the dose–response relationships between relevant endpoints?
Strength or robustness of the supporting datasets	• How extensively are the relevant or key events either empirically measured and/or modelled by appropriate *in silico*, *in chemico* and *in vitro* data? • Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the supporting methods adequately reported?
The Weight-of-Evidence (WoE) supporting the prediction	• Is there consistency in the supportive information (e.g., structural alerts) between analogues or within the category? • How many and how large are the dissimilarities in the supporting information (i.e., data gaps)?
Documentation and written evidence provided	• Is the read-across prediction adequately documented? • Does the evidence support the hypothesis that the uncertainty is acceptable for the stated purpose (as per Question 1)?

Data matrix for analogue approach

Data matix, IATA for "indication of title of case study"

Chemical ID									
		Source1	Target	Source2	Source3	Source4	Source5	Outlier1	Outlier2
CAS									
Name									
Structure									
Summary of data gap filling									
		Source1	Target	Source2	Source3	Source4	Source5	Outlier1	Outlier2
Target endpoint1	Experimental result	value, unit, test method (eg. test guide line)		value, unit, test method (eg. test guide line)		value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)
	Integrated conclusion (eg. read-across)		derived result						
Target endpoint2	Experimental result	value, unit, test method (eg. test guide line)		value, unit, test method (eg. test guide line)		value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)
	Integrated conclusion (eg. read-across)		derived result						
Molecular profiling related to the analogue approach hypothesis									
Parent chemical	Profiler 1 (name, version)								
	Expert system 1 (name, version)								
Metabolite*	Profiler 1 (name, version)								
	Expert system 1 (name, version)								
Physical-chemical data									
Melting point									
Boiling point									
Density									
logPow (measured value)									
logPow (calculated value)									
…									
Kinetics**									
Absorption									
Distribution									
Metabolism									
Excretion									
Supporting data related to the target endpoint(s)									
		Source1	Target	Source2	Source3	Source4	Source5	Outlier1	Outlier2
In vivo	Toxicogenomics								
	…								
In vitro	Alternative method A								
	…								
In chemico	…								
In silico	QSAR1 (Target endpoint1)								
	QSAR2 (Target endpoint1)								
	QSAR3 (Target endpoint2)								
	QSAR4 (*In vitro* endpoint)								
	…								
Other data	Battery approach								
	Defind approach of IATA								
	…								

* More relevant metabolite such as toxicant

**General outline of relative comparative kinetics

Data matrix for category approach

Data matix, IATA for "indication of title of case study"

Chemical ID									
		Member 1	Member 2	Member 3	Member 4	Member 5	Member 6	Member 7	Member 8
CAS									
Name									
Structure									
Summary of data gap filling									
		Member 1	Member 2	Member 3	Member 4	Member 5	Member 6	Member 7	Member 8
Target endpoint1	Experimental result	value, unit, test method (eg. test guide line)		value, unit, test method (eg. test guide line)		value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)
	Integrated conclusion (eg. read-across)		derived result		derived result				
Target endpoint2	Experimental result	value, unit, test method (eg. test guide line)		value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)		value, unit, test method (eg. test guide line)	value, unit, test method (eg. test guide line)
	Integrated conclusion (eg. read-across)		derived result				derived result		
Molecular profiling related to the category hypothesis									
Parent chemical	Profiler 1 (name, version)								
	Expert system 1 (name, version)								
Metabolite*	Profiler 1 (name, version)								
	Expert system 1 (name, version)								
Physical-chemical data									
Melting point									
Boiling point									
Density									
logPow (measured value)									
logPow (calculated value)									
…									
Kinetics**									
Absorption									
Distribution									
Metabolism									
Excretion									
Supporting data related to the target endpoint(s)									
		Member 1	Member 2	Member 3	Member 4	Member 5	Member 6	Member 7	Member 8
In vivo	Toxicogenomics	result	result	result	result	result	result	result	result
	…								
In vitro	Alternative method A		result	result	result				
	…								
In chemico	…								
In silico	QSAR1 (Target endpoint1)	result	result	result	result	result	result	result	result
	QSAR2 (Target endpoint1)	result	result	result	result	result	result	result	result
	QSAR3 (Target endpoint2)	result	result	result	result	result	result	result	result
	QSAR4 (*In vitro* endpoint)	result	result	result	result	result	result	result	result
	…								
Other data	Battery approach	result	result	result	result	result	result	result	result
	Defind approach of IATA								
	…								

* More relevant metabolite such as toxicant

**General outline of relative comparative kinetics

Annex 2. General Template for IATA case Studies - Building Blocks

Title: Case Study on the use of Integrated Approaches for Testing and Assessment for "Target Endpoint(s)" of "Target Chemical(s)"

(N.B. The following template should not be viewed as a strict structure, but rather identifies the information that should be included in this type of case study. Depending on the specific case study additional information/(sub)section(s) may be required or particular subsections may not apply. The order of the (sub)sections of the template can be changed and two or more (sub)sections of the template can be merged, as necessary. The titles of a (sub)section can be changed as necessary. The template will be revised based on experience with use).

Abstract / Synopsis / Executive summary

(This section should provide a brief overview of the case study, including the objectives, concepts, methodologies, outcomes and conclusion in about 300 words.)

Introduction

(This should include a summary of the background/problem formulation, purpose, endpoints covered and description of the target chemical(s)/category, assessment approach)

Table of Contents

A. Purpose

a. Purpose of use

Indicate the regulatory relevance (i.e. intended application) of the IATA. This may be: a)screening for priority setting in view of further evaluation; b) hazard identification/characterisation; c) risk assessment; d) other (please specify). If more than one purpose is possible, please specify the purpose as d) other.

If in a regulatory context, provide a short description of any (e.g. legal) requirements for the IATA approach to be accepted.

b. Target chemical(s)

Provide the chemical descriptor common identifiers (including CAS number, name and composition including impurities [See 3.2.3.1 "*Chemical identity and composition* of the grouping guidance (OECD, 2014a)]) and chemical structure(s) of the target substance(s). In some case studies, target chemicals may be entire chemical classes or the IATA illustrated may be generic. Or if there are no specific target chemicals, example chemicals can be used to illustrate the IATA (SEE "1. PURPOSE" or "3. RESULTS OF ERC PRIORITISATION" of the case study 2017-2 (OECD, 2018c) and "b. Target Chemical(s)" at the section "A. Purpose" of the case study 2018-1).

c. Endpoint(s)

Identify the endpoint(s) for which the IATA is applied.

d. Exposure information (if needed)

Provide the considered exposure, such as route of exposure (dermal, oral and inhalation), type of exposure (consumer, occupational and environment), for example, if the case study addresses prioritisation or chemical assessment work flows. The inclusion of this section and its level of detail/quantification will depend on the case study.

If relevant, please describe extrapolation from *in vitro* into *in vivo*.

B. Hypothesis for performing IATA

- Provide the hypothesis for performing IATA for the identified purpose
- Describe how the IATA will be performed for the specific purpose.

C. Approaches used (Potential Blocks for Inclusion)

Describe which approaches are applied for assessing the chemicals under the provided hypothesis:

- **AOP/MOA**: Description of potential mechanism(s) for the target chemicals to induce target endpoint toxicity. In particular, the graphical representation of the AOP would be helpful for the reader and key references (See "Graphical Representation of the AOP" at section "1- AOP Description" of "User's Handbook supplement to the Guidance Document for developing and assessing Adverse Outcome Pathways" (OECD, 2016g)). The tools in the AOP-KB should be referred to as appropriate (e.g. AOP wiki[10], Effectopedia[11] etc.).
  Identifying the relevant AOP from AOP wiki is required. Please provide the AOP number, status on AOP-wiki and the link. For AOPs that are not documented, consider the "Section 1-AOP Description" of "Users' Handbook supplement to the Guidance Document for developing and accessing Adverse Outcome Pathways" (OECD, 2016g) - although an entire AOP description is not the purpose here.
- **Defined Approach**: If a defined approach is included, please refer to the ANNEX I: TEMPLATE FOR REPORTING DEFINED APPROACHES TO TESTING AND ASSESSMENT BASED ON MULTIPLE INFORMATION SOURCES" of "Guidance Document on the Reporting of Defined Approaches to be used within Integrated Approaches to Testing and Assessment" (OECD, 2016h). Please copy into this section the "*5. Rationale underlying the construction of the defined approach*" from the above mentioned template (OECD, 2016h), completed with proper explanations.
- **Workflow**: If an IATA workflow is included, provide a schematic and explanation of the elements of the workflow including input, decision and exit points. If prioritisation is the goal of IATA workflow, provide an explanation of how to classify the hazard and exposure profiling and potential risk classification. Please refer to the section "CHEMICAL SAFETY ASSESSMENT WORKFLOW" of the case study 2016-5 (OECD, 2017f), "3.3 IATA Workflow" of the case study 2017-1 (OECD, 2018b) and the section "2. PRIORITISATION OF CHEMICALS USING AN IATA-BASED ERC APPROACH" of the case study 2017-2 (OECD, 2018c).

[10] AOP Wiki. https://aopwiki.org/

[11] Effectopedia. https://www.effectopedia.org/

- **Read-across**: If a read-across is included, use elements of the template for IATA case studies on Read-Across or the grouping guidance (OECD, 2014a). Please refer to "*4. Identification of analogues, suitability assessment and existing data*" of the case study 2016-5 (OECD, 2017f) and "4.1. Analogue chemicals" of the case study 2017-1 (OECD, 2018b)

D. Data/Information gathering

In this section, please describe the test methods or data sources used for gathering data for target chemicals

a. Data/Information

- Provide the methods used for gathering the data for target chemical(s) (e.g. selection criteria of the data, data source).
- Provide the data gathered using appropriate reporting format. The levels details for reporting the data should be considered depending on the purpose of the IATA.
- If data from non-guideline test methods are included, provide descriptions of the methods or links to sources that summarise the methods. The appropriate degree of detail of the description should be considered in the context of the purpose of the case study. A template for the description is available in a OECD guidance document (OECD, 2014c). Examples of description using the template can be found in JRC EURL ECVAM Database service on Alternative Methods to animal experimentation (DB-ALM)[12] and U.S. EPA Toxicity ForeCaster (ToxCast™) Data[13].
- If QSAR data are included, provide the name, version, owner of the models used for deriving QSAR estimation data. QSAR models if not already described elsewhere should be reported using the QSAR Model Reporting Format (QMRF), and individual predictions, if applicable, should be reported using the QSAR Prediction Reporting Format (QPRF).). A QMRF inventory is maintained by JRC that can be utilised as a resource of QMRFs and its reference number can be referred to JRC QSAR Model databases[14].
- If the exposure elements are included, provide the methods used for the data generation (e.g. data source, exposure models/tools.) Please refer to "2. Identification of the use scenario of the case study 2016-5 (OECD, 2017f)" and "*Exposure profiling*" of the case study 2017-2 (OECD, 2018c)
- If a defined approach is included, please refer to the template of "Guidance Document on the Reporting of Defined Approaches to be used within Integrated Approaches to Testing and Assessment" (OECD, 2016h). In this section, please describe the individual information sources used and data interpretation procedure applied (See *"6. Description of the individual information sources used (see Annex II)"* and *"7. Data interpretation procedure applied"* of the OECD guidance (OECD, 2016h).

b. Analogue chemicals.

[12] JRC, EURL ECVAM Database service on Alternative Methods to animal experimentation (DB-ALM). https://ecvam-dbalm.jrc.ec.europa.eu/.

[13] U.S. EPA, Toxicity ForeCaster (ToxCast™) Data https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data

[14] JRC, QSAR Model Database. https://qsardb.jrc.ec.europa.eu/qmrf/

- If the data of analogue chemicals were used for the IATA, provide the selection criteria that were used to identify the analogue chemicals. This can be based on the hypothesis described in section B.
- Provide rational for selection of analogue(s) with respect to the defined purpose and endpoint.
- Consider selection bias selecting analogue chemicals in relation to employment of the IATA (e.g. data completeness, support for hypothesis etc.).
- Describe the methods used to identify the analogue chemicals (e.g. inventories and tools used should be provided). Listing search criteria to establish initial pool of candidate analogues is helpful.
- Provide the common chemical identifiers (including CAS number, name and composition including impurities) and chemical structure(s) of the analogue chemicals.

E. Application of IATA

a. Summary of data

- Provide a summary of data in a suitable format for the purpose of IATA.
- Reliability of data should be discussed.
- The applicability domain of each estimation method including QSAR and alternative methods should be discussed
- Provide analysis of the available information for suitability regarding the defined purpose. If possible, the available key study results should be indicated.

b. Application of IATA

- Describe how to apply IATA based on the hypothesis and the data gathered.
- Describe the result of IATA.
- Refine the hypothesis used, if necessary.

c. Uncertainty

- Discuss the uncertainty of each element of the IATA.
- Aspects can include uncertainty and confidence associated with the data and assumptions.
- The magnitude and impact of the sources of uncertainty should be considered and to the extent possible, how the individual sources of uncertainty affect the overall uncertainty in the final outcome of the IATA. OECD guidance documents on defined approaches of IATA *("11. Consideration of uncertainties associated with the application of the defined approach"* of OECD, 2016h; "*11. Consideration of uncertainties associated with the application of the defined approach"* of CASE STUDY I-XII of OECD, 2016i) might be helpful for considering uncertainties related to non-guideline test methods as well as the approaches outlined in the template for IATA case studies on Read-Across

d. Strategy and integrated conclusion

- Describe the strategy used to develop the integrated conclusion.
- Discuss how/if to further address the uncertainties.
- Finally, provide a short conclusion wrapping up the outcome of the evaluation.

References

(*See OECD style guide third edition, p.56 "Bibliographical referencing: Sources and citations"*)

Annex