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

12-Sep-2016

English - Or. English

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

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

JT03400412

ENV/JM/MONO(2016)48
Unclassified
English - Or. English

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 250

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

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 2016

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 35 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 Organizations.

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

FOREWORD

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 (Q)SAR, chemical categories and Adverse Outcome Pathways (AOPs) 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 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 four case studies, listed below, submitted to the 2015 review cycle of the IATA Case Studies project. The topics discussed in this document include the strongest aspects and uncertainties of each case study, and the document identifies areas for developing further guidance on IATA.

1. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR IN VITRO MUTAGENICITY OF 3,3' DIMETHOXYBENZIDINE (DMOB) BASED DIRECT DYES, ENV/JM/MONO(2016)49, Series on Testing & Assessment No. 251.

2. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR REPEAT DOSE TOXICITY OF SUBSTITUTED DIPHENYLAMINES (SDPA), ENV/JM/MONO(2016)50, Series on Testing & Assessment No. 252.

3. CASE STUDY ON THE USE OF AN INTEGRATED APPROACH TO TESTING AND ASSESSMENT FOR HEPATOTOXICITY OF ALLYL ESTERS, ENV/JM/MONO(2016)51, Series on Testing & Assessment No. 253.

4. CASE STUDY ON THE USE OF AN INTEGRATED APPROACH FOR TESTING AND ASSESSMENT OF THE BIOACCUMULATION POTENTIAL OF DEGRADATION PRODUCTS OF 4,4'-BIS (CHLOROMETHYL)-1,1'-BIPHENYL, ENV/JM/MONO(2016)52, Series on Testing & Assessment No. 254.

This document has been prepared by a project team of the Task Force on Hazard Assessment and was endorsed at the 9th Task Force on Hazard Assessment meeting in June 2016.

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

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 (Q)SAR, chemical categories and Adverse Outcome Pathways (AOPs) as a part of Integrated Approaches for Testing and Assessment (IATA). For example, concurrently to this project, a guidance document for reporting of defined approaches and individual information sources, based on skin sensitization case studies, is ongoing. 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.

In 2014, the Task Force for Hazard Assessment (TFHA) proposed an IATA Case Studies Project as a one of the high priority projects of the revised Cooperative Chemicals Assessment Programme (CoCAP)[1] to increase experience with the use of IATA by developing case studies. The proposed project was endorsed at the 52nd Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology (2-4 November 2014).

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. It is anticipated that 3-4 case studies will be reviewed each year.

The project team was established in February 2015 consisting of representatives nominated by the TFHA from Australia, Canada, Denmark, Japan, Netherlands, Sweden, United States, European Union/European Commission (EU/EC), European Union/Joint Research Centre (EU/JRC), European Union/European Chemicals Agency (EU/ECHA), Business and Industry Advisory Committee to the OECD (BIAC) and International Council for Animal Protection in OECD Programmes (ICAPO). In addition, Germany, Italy and United Kingdom participated in the review meeting of the first set of case studies. In 2015, the four case studies shown in Table 1 were reviewed. All of these case studies focus on the application of grouping methods as replacements of animal testing to reach a conclusion on a hazard endpoint. The studies were developed based on actual cases of the regulatory use of IATA in the lead counties. The final case studies are published [ENV/JM/MONO(2016)49-52, Series on Testing and Assessment non 251-254]. 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 jurisdictions. In addition, these cases studies should not be interpreted as official regulatory decisions made by the authoring member countries. This document reports the learnings and lessons obtained from the review experiences of the case studies.

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

Table 1. List of Case Studies Reviewed in 2015

No.	Title	Lead Country	Purpose of Use
1	In Vitro Mutagenicity of 3,3' Dimethoxybenzidine (DMOB) Based Direct Dyes	Canada United States	Hazard characterization for a screening level risk assessment under Canada's Chemicals Management Plan
2	Repeat Dose Toxicity of Substituted Diphenylamines (SDPA)	Canada	Hazard identification for data poor chemicals. Elements of the case study may be used to support screening level risk assessment under Canada's Chemicals Management Plan
3	Hepatotoxicity of Allyl Ester Category	Japan	Hazard identification for a risk assessment under Japan's Chemical Substances Control Law
4	Bioaccumulation Potential of Biodegradation Products of 4,4'-Bis (chloromethyl)-1,1'-biphenyl	Japan	Assessment of bioaccumulation of new chemical substances under Japan's Chemical Substances Control Law

2. PROCESS FOR REVIEWING THE CASE STUDIES

The template (Annex) used for the case studies was developed by the OECD secretariat in collaboration with the lead countries of the case studies based on the reporting format in the OECD Guidance on Grouping of Chemicals (OECD 2014a) and a case study document on IATA (OECD 2014b). 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)

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 would strengthen the case study?

6. 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?

7. What areas of the case study were particularly challenging to review?

8. Would you use the results of such a case study in your regulatory context? If no, why not (legislative/policy/scientific reasons)?

9. Does the template work well?

10. 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 in the first meeting of the IATA Case Studies Project (19-20 November 2015) in order to finalize the case studies and summarise the learnings and lessons obtained from the review experience. The case studies were revised based on the comments at the meeting.

3. SUMMARY OF REVIEW RESULTS

3.1. Case Study 1: In Vitro Mutagenicity of 3,3' Dimethoxybenzidine (DMOB) Based Direct Dyes [Canada & United States]

This case study was developed based on a hazard characterization for a screening level risk assessment conducted under Canada's Chemicals Management Plan (CMP)[2]. In addition, experiences of the US EPA's Action Plans[3] for the target category were incorporated. A category for in vitro mutagenicity consisting of 13 members of 3,3' dimethoxybenzidine (DMOB) based direct dyes was formed. The category members have a common metabolite causing the target effect. Using read-across based on three category members with experimental data, 10 other category members were evaluated as positive for the target effect. QSAR prediction results were used for supporting the read-across.

[2] Canada, Chemicals Management Plan. http://www.chemicalsubstanceschimiques.gc.ca/plan/index-eng.php

[3] US EPA, Action Plans. http://www.epa.gov/assessing-and-managing-chemicals-under-tsca/current-chemical-risk-reduction-activities

Please refer to ENV/JM/MONO(2016)49, Series on Testing & Assessment No. 251 for the case study to put the following points into context.

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

- Clear presentation of multiple lines of evidences (structural, physico-chemical, in vivo, in vitro, in silico) used in the justification.
- The steps used to build a category are accounted for.
- Convincing evidence is presented in the justification for the use of read-across:
  - The inclusion of data from the common metabolite is beneficial
  - Structural boundaries are described
  - Source substances cover the target substances
- A discussion on the exclusions from the category is included
- Subgrouping is based on chemical structures and well justified

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

- The data for the target effect of the category clearly showed a trend supporting the mechanism hypothesised. It was pointed out that the data also imply the existence of another mechanism and the authors included a possible explanation for this in the revised case study.
- A consensus approach was used for deriving a conclusion from the prediction results of four QSAR models. Some reviewers pointed out that it was not very clear how to integrate the prediction result. The authors removed the consensus approach and provided descriptions in regards to which models support the read across as an alternative approach.
- Uncertainty analysis was reported using qualitative labels (Low/Medium/High) for data uncertainty and the strength of evidence around the justification elements. Some reviewers pointed out that how to assign the qualitative labels was not clear. The authors revised their approach to report the uncertainty using descriptive language for each criterion.

The following areas of uncertainty were identified for the case study:

- The applicability domain of the in silico models used for predicting in vitro mutagenicity and metabolites.
- Other areas of genotoxicity (e.g. mutagenicity in mammalian cells and an in vitro test for chromosomal aberrations) are not included.

It was considered that such a case study could be used for some regulatory purposes in member countries (e.g. under REACH, if its legislative requirements are satisfied. In addition, compliance of the case study with the recently published ECHA Read-Across Assessment Framework (RAAF) (ECHA, 2015) would need to be assessed first to gain confidence in the possible use of the case study for REACH purposes).

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

- Further incorporating Adverse Outcome Pathway (AOP) concepts into grouping approaches
  - For example, other genotoxicity endpoints and carcinogenicity data could potentially be incorporated into the IATA by organizing information around an AOP.
- Uncertainty communication and/or framework (e.g. Reporting format for uncertainty)
- Defining applicability domains
  - Boundaries for physicochemical properties (currently the functionalities that must be present to increase solubility are qualitatively mentioned – but offer no strict cut-off for logarithm of the distribution coefficient (LogD), water solubility (WS), etc.)
- Level of detail required in study summaries to support the purpose of the case studies
- How to improve incorporation of QSAR results into integrated conclusions
  - Discussing reliability of each model result and how to combine predictions, and other data types, for integrated conclusions (e.g. consensus vs. weighting)

In addition, it was suggested that the uncertainty analysis of the case study could be used as a model for future cases to gain experience before developing guidance.

3.2. Case Study 2: Repeat Dose Toxicity of Substituted Diphenylamines (SDPA) [Canada]

This case study was developed based on information gathered for hazard characterisation for a screening level risk assessment conducted under Canada's CMP. The approach outlined in the case study was used to support further discussion and guidance development under the OECD Case Study Project. Accordingly, the final approach applied for the planned Screening Assessment Report currently under development by Health Canada and Environment and Climate Change Canada may differ. The document is not intended to provide complete characterization of health effects. Also, it does not provide information regarding exposure of the general population of Canada to SDPAs. These elements, along with risk characterizations, will be presented in the subsequent draft screening assessments and related documents developed under Canada's Chemicals Management Plan.

A category for oral repeat dose toxicity consisting of 14 members of substituted diphenylamines (SDPA) including UVCBs was formed. The category was divided into 4 subcategories. Since there are limited data with respect to mode of action (MOA) or AOP interpretations for the effects observed across the SDPAs, the hypothesis of the subcategory was mainly built based on structural similarity including considerations of the types of functional groups present. Similarities in physicochemical properties, oral bioavailability and observed toxicological effects were used for justifying the subcategories. The effect levels for six category members without test data were predicted by read–across within each subcategory.

Please refer to ENV/JM/MONO(2016)50, Series on Testing & Assessment No. 252 for the case study to put the following points into context.

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

- Provides an example of application of read-across for UVCB chemicals
- The hypothesis is well justified by using multiple types of information
- Multiple substances have high quality study data (all structural features are covered in tested substances)
- Provides an example of the analysis of high-throughput screening data acknowledging the need for this source of data for an increased number of group members
- Clear presentation and use of templates (data matrix and reporting template)
- Clear stepwise approach following defined criteria
- Detailed uncertainty analysis was included
- Subgrouping is based on chemical structure and well described

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

- In the first draft of the case study, the most conservative effect levels were selected by examining major effects across the group, to fill data gaps for category members without data. Most reviewers suggested the use of subgroups to better account for observed or potential differences in chemical structure, physicochemical properties, bioavailability and systemic effects. The authors explained that the purpose of the case study in its original form was to illustrate the application of conservative read across, fit for purpose in a tiered approach to screening level risk assessments where refinements would be conducted ,if required, based on the derived Margin of Exposure (MOE). To support further discussion and guidance development within the OECD IATA Case Studies project, the authors have further sub-grouped the substances and demonstrated the use of a closest neighbour approach for read across.
- With regard to the description of structural similarity of the category members, some reviewers suggested to include a discussion about the influence of the structural differences among the category members on the expected toxicological profile (e.g. alkyl versus phenyl substitution). The authors have formed subgroups as noted above to better account for structural differences.
- Some reviewers pointed out the descriptions of similarity in toxicological effects of the category members were not sufficiently detailed because it was not described how observed findings support the similarity in toxicological effects of the category members although effects on the liver clearly dominate. Especially, it was suggested that the differences in response of various organs should be discussed. The authors conducted a more in-depth analysis and sub-grouped the substances, including more detail on the other effects observed.
- In the first draft of the case study, uncertainty was reported using qualitative labels (Low/Medium/High) for data uncertainty and strength of evidence around the justification elements. Some reviewers pointed out that it was unclear how the qualitative labels were assigned. The authors revised their approach to report the uncertainty using descriptive language for each justification criterion rather than qualitative labels.

The following areas of uncertainty were identified for the case study:

- Potential impacts of the structural differences of subcategory members on toxicity.
- Level of similarity in metabolism, physicochemical properties and toxicokinetics parameters, for which not much empirical data is available. In addition, the applicability domains of the in silico models used for predicting these endpoints are not clear.
- The difficulty in identifying representative structures covering all UVCB components.

It was considered that such a case study could be used for some regulatory purposes in member countries. It cannot be immediately used for some regulatory purposes under specific legislation, such as under REACH for replacement of a standard experimental test result, or data gap filling for Annex XI requirements (registration), which has the highest requirements for confidence in the read-across prediction in terms of possible uses of read-across in REACH activities (please see REACH Read Across Assessment Framework (ECHA, 2015) for more details).

However, in addition to the Canadian regulatory use, as stated in the case study, the submitted category with corresponding data and read across has also been used by EPA's Office of Pollution Prevention and Toxics (OPPT) program to develop a screening-level hazard characterization (US EPA 2009).

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

- Guidance on how to address UVCBs – e.g. incorporation of reaction schemes and selection of representative structures. Possible use of UVCB G-Graph.
- Uncertainty communication and/or framework (e.g. Reporting format for uncertainty)
- Guidance on the level of detail required in these case studies when describing QSAR models.
- Incorporation of AOP information into a hypothesis and justification and accounting for lack of AOP or mechanistic information in the uncertainty assessment.
- Incorporation of different types of 'omics approaches and in vitro data, including high throughput data, to build better justifications

3.3. Case Study 3: Hepatotoxicity of Allyl Ester Category [Japan]

This case study was prepared to be used for hazard identification in chemical risk assessment under Japan's Chemical Substances Control Law (CSCL)[4]. A category consisting of 19 allyl esters was formed for the target endpoint of repeated-dose hepatotoxicity, based on a hypothesis of an adverse outcome pathway, in which the hepatotoxicant is acrolein, a common metabolite of allyl esters. The category is subcategorized into 2 subcategories: allyl esters with linear alkyl chain and allyl esters with branched alkyl chain. The effect levels for 16 category members without test data were predicted by read–across.

Please refer to monograph ENV/JM/MONO(2016)51, Series on Testing & Assessment No. 253 for the case study to put the following points into context.

[4] Japan, Chemical Substances Control Law. http://www.meti.go.jp/policy/chemical_management/english/cscl/

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

- The elements for a read-across hypothesis are well described based on metabolism and MOA/AOP and there is a clear understanding of the metabolic consequences.

- The rationale for read-across is well justified. There is convincing evidence and overall data at different levels in order to conclude that category members will show similar effects following repeat dose toxicity with respect to the endpoints investigated.

- High confidence that can be attributed to subcategory 1, which is a highly similar structural group with quantitative metabolic hydrolysis pathway information on some members. The hepatotoxicity of these allyl esters correlates well with the rate of hydrolysis to allyl alcohol.

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

- Most reviewers pointed out that the structural boundaries of the subcategory were not clearly defined (e.g. chain length of carboxylic acid moiety of allyl ester). The authors clarified the boundaries for the subcategory of allyl esters with linear alkyl chain although it could not be clarified for the subcategory of allyl esters with branched alkyl chain due to structural variation.

- It was pointed out that there is a lack of substantiation of the hepatotoxic effects of acrolein being the critical effect. The authors clarified this point in the revised case study.

- Reviewers noted that the discussion regarding the hypothesis - that metabolites other than the toxicant do not induce other toxic effects - was insufficient. The authors enhanced the discussion on this point including more references.

- Reviewers requested further description of the way data gathering was conducted (e.g. which databases were used, how to select the data used). The authors enhanced this description.

The following areas of uncertainty were identified for the case study:

- The rate determining factor for formation of acrolein and what effect the ADME might have on toxicity.

- Lack of knowledge on the range of ester hydrolysis rate and no clear match with structural complexity.

- Lack of understanding of the mechanism of bile duct hyperplasia observed in the repeated dose toxicity test of some category members.

- The hypothesis that metabolites other than the toxicant do not induce other toxic effects.

- Human relevance.

It was considered that such a case study could be used for some regulatory purposes in member countries, such as screening assessment, prioritisation, and classification, although the results of such a case study cannot be used for registration purposes under REACH, because it would not likely fulfil the information requirements for registration, similar to Case Study 2. In the near future, the results of such a case study could be considered to be used under Japanese Chemical Substances Control Law.

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

- Building a hypothesis for read-across based on an AOP
- Description of the structural boundaries of a category
- Use of supportive data for justifying a category based on an AOP
- Uncertainty communication and/or framework (e.g. Reporting format for uncertainty)

3.4. Case Study 4: Bioaccumulation Potential of Biodegradation Products of 4,4'-Bis (chloromethyl)-1,1'-biphenyl [Japan]

This case study was developed based on an assessment of 4,4'-bis (chloromethyl)-1,1'-biphenyl under Japan's CSCL. The assessment was conducted according to the bioaccumulative analogy rule that is used for the judgment of new chemicals under the CSCL (Japan, 2013). The purpose of the case study is to evaluate the bioaccumulation potential of biodegradation products of 4,4'-bis (chloromethyl)-1,1'-biphenyl, which has bioconcentration test data. The biodegradation products (target chemicals) include 3 compounds with known structure and 1 compound with unknown structure. The parent chemical and 2 analogues with bioconcentration test data were used as source chemicals for the read-across. QSAR prediction results and HPLC data of the biodegradation test were used for justifying the read-across. The bioaccumulation potential of all target chemicals were qualitatively evaluated as “Low - Not highly bioaccumulative" using CSCL criteria.

Please refer to ENV/JM/MONO(2016)52, Series on Testing & Assessment No. 254 for the case study to put the following points into context.

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

- The methodology used follows a clear stepwise approach and the criteria for identifying analogues and developing justifications are clear.
- Demonstration of an effective way to use the analytical results (HPLC data) of a biodegradation study (OECD 301C) on a source substance to estimate bioconcentration factors (BCF) of a parent substance and its metabolites by comparing their relative hydrophobicity (log Kow).
- Comparison of experimental and modelled values: Good correlation between the experimental BCF values and the QSAR predicted values for structurally similar source chemicals increases the reliability of the prediction result of the target chemicals.
- The taking into account of all information, providing sound justification and combining them to reach a conclusion.

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

- It was commented that the criteria of Japan’s bioaccumulative analogy rule could lead to differences in assessment outcomes due to variability in expert judgment. The authors explained that the criteria are relatively new (announced in 2013) and that there is room for improvement. At this moment, the rule depends heavily on expert judgment, but by implementing this rule, Japan is gaining experience which would help to update the rule to be more robust in the future.

- It was suggested that the following hypotheses was not adequately justified: that the structural differences between target and source chemicals do not result in a significant effect on bioconcentration potential. However, in order to support the hypothesis, the authors provided another example of a group of chemicals with BCF data showing that such structural differences did not affect bioconcentration.

- The (Q)SAR prediction BCF values of one of the source chemicals are much higher than the experimental value. It was commented that the reason for this needs to be discussed to reduce the uncertainty of the (Q)SAR prediction BCF values of the target chemicals. The authors highlighted that the source chemical has a fragment that is not covered by the chemicals in the training sets of the (Q)SAR models and explained that this could be the reason for the higher predictions. In this regard, it was also suggested to state whether each target chemical is within the applicability domain of the QSAR models.

- The bioaccumulation potential of one target chemical with unknown structure was assessed by read-across based on the similarity of hydrophobicity (logKow measured by HPLC). It was commented that the minimum requirement of the analogy rule, to know the structure of the molecule, is not fulfilled. In response to the comment, the authors provided possible chemical structures of the target chemicals by predicting the metabolites of its parent chemical by using the "Microbial metabolic simulator" of the OECD QSAR Toolbox; and provided low bioconcentration prediction result for all the estimated metabolites in order to reduce the uncertainty.

The following areas of uncertainty were identified for the case study:

- Uncertainty related to the potential differences between bioaccumulation and bioconcentration that would result depending on the route of exposure. It was suggested that comparison of this method with alternative methods in assessing bioaccumulation/bioconcentration for data poor substances would strengthen the case study.

- There is uncertainty related to the potential differences in degradation products measured under the biodegradation test conditions (OECD 301C) and existing in the aquatic environmental compartments (e.g. caused by the differences in pH).

It was considered that such a case study could be used for some purposes in member countries. Canada commented that currently Environment and Climate Change Canada uses other means in evaluating bioaccumulation potential, using another well-developed approach that considers multiple uptake/elimination pathways. However, there is potentially an opportunity to use both approaches in supportive ways.

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

- Identification of degradation products and validation of the endpoint of interest to address the bioaccumulation potential (e.g. BCF, BAF…) by considering the route of exposure

- How to derive an integrated conclusion to be used for multiple purposes (e.g. both a conclusion under the Japanese framework and a more general conclusion for other purposes)

- Uncertainty communication and/or framework (e.g. Reporting format for uncertainty)How to discuss the inconsistency between QSAR prediction values and experimental values and report QSAR predictions

- Guidance on determining similarity of structural features when considering bioconcentration and the application of read-across (e.g. identifying subgroups contributing to increase/decrease of bioconcentration).When to gather/produce more experimental data where the uncertainty is high.

4. LEARNINGS AND LESSONS

4.1. Methods Illustrated by the Case Studies

Different strong aspects were identified for each case study. All the case studies have either been applied in a regulatory context or have the potential to be used for different purposes in the regulatory environment. The case studies illustrate pragmatic use of grouping methods while addressing some challenging topics such as the use of AOPs and complexities in addressing UVCBs. Table 2 shows where topics in the OECD Guidance on Grouping of Chemicals (OECD 2014a) have been illustrated by the case studies. The guidance document usually describes concepts for each topic and the case studies allow for the comparison and contrast of approaches used in various contexts. These examples help strengthen the use of read-across based on the guidance in one's specific situation.

Table 2. Examples of the Topics in the OECD Guidance on Grouping of Chemicals (OECD 2014a) Illustrated by the Case Studies

Topics in the Grouping Guidance	Case Study
2.3.2. Category and subcategory membership and applicability domain 2. Subcategories	2, 3
2.4. The mechanistic basis of using analogues or chemical categories	1, 3
3.4. Computational methods based on external models	1, 2, 4
6.2. Metabolic or degradation pathways and toxicokinetics	1, 2, 3, 4
6.6.1 General guidance on developing categories for organic UVCBs	2
7.1. Reporting Format for analogue approach	4
7.2. Reporting format for chemical categories	1,2,3
10. Table 14. Specific aspects of endpoint read-across justifications (Genotoxicity, Repeated dose toxicity, Bioaccumulation)	1, 2, 3, 4

4.2. Areas Identified for Further Guidance Development

Through the review of each case study, six areas for further developing guidance were identified:

1. Describing scope and context for read-across

The review experience highlighted the importance of the need for a very clear description of the application, scope and framework for which the case study is used. This provides context for the reader in regards to what aspects will be addressed by the case study, what level of detail in reporting is required and what level of uncertainty might be accepted. Further guidance for authors on this point could be developed.

2. Building hypotheses based on MOA/AOP

Building hypotheses based on mechanistic information was identified as a specific topic in all case studies that could benefit from the development of further guidance. More elaborated hypotheses would strengthen the similarities (and potential differences) with respect to target endpoints of the category members in each case study. In addition, uncertainties regarding human or environmental relevance identified in the case studies could be clarified for each of the mechanistic key events. Strengthening the mechanistic basis of the case studies can extend the use of the case studies.

OECD member counties have already identified the usefulness of AOPs in forming chemical categories (OECD 2011). OECD has published guidance for developing AOPs (OECD 2013) and a number of AOPs are under development[5]. It is expected that AOPs can be applied to support grouping methods, however there is a need to continue to demonstrate how AOP information can be incorporated in IATA. The descriptions for building hypotheses based on AOPs in the OECD Guidance on Grouping of Chemicals (OECD 2014a) are limited and conceptual (Sections 2.4.2, 3.2.3.4 and 5.2.4.1.1). These are necessary to be elaborated in the future.

3. Definition of analogues/category boundaries

Reviewers suggested that all the case studies should have more detailed description on the definition of the structural boundaries and physicochemical properties of the analogues. How to describe clear category boundaries is common issue for all endpoints.

Especially, it was identified that most case studies lacked a discussion on the structural differences in the chemical structures of analogues whereas their structural similarities were well discussed. Table 3 contains examples of analogues whose structural differences were discussed in the review process. There are several useful tools such as OECD QSAR Toolbox[6] to identify substructures leading to a variation in toxicological effect. However, acceptable structural differences for analogues are typically defined by expert judgement. The considerations taken when applying expert judgement should be documented.

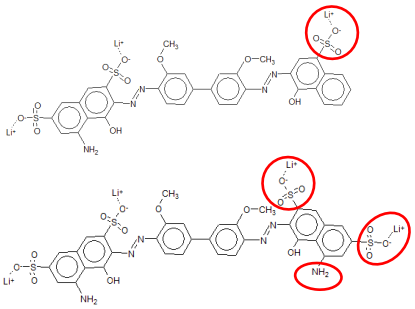
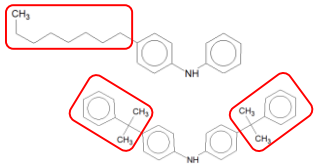
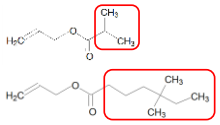
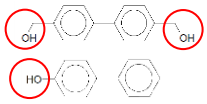
Guidance on "similarity" of structural features is also an important topic since the selection of source chemicals is one of the most important issues in data gap filling by grouping and directly affects the conclusion.

This is a high priority area for further developing guidance. In the OECD Guidance on Grouping of Chemicals (OECD 2014a), there is general guidance for definition of analogues/category boundaries (sections 2.3, 4.2.2 and 5.2.2). These need to be elaborated (e.g. by endpoint) in the future.

[5] OECD, Adverse Outcome Pathways, Molecular Screening and Toxicogenomics. http://www.oecd.org/chemicalsafety/testing/adverse-outcome-pathways-molecular-screening-and-toxicogenomics.htm

[6] OECD QSAR Toolbox. http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm

Table 3. Examples of Structural Differences of Analogues Discussed in the Case Studies

Case Study /End Point	Example of Analogues Compared (Highlighted Structural Differences)	Example of Reviewer's Comment
Case Study 1 /In vitro mutagenicity		It is stated that solubility is an important factor impacting the target endpoint. However, there is no indication of whether the extent of transformation of category members into DMOB is influenced by the number of sulfonate salt substituents in the azo substance and subsequent solubility.
Case Study 2 /Repeat Dose Toxicity		Can we consider alkyl chain and phenyl substituted derivatives to be similar? (Subgroups were formed to better account for these differences.)
Case Study 3 /Repeat Dose Toxicity		There is uncertainty with respect to the range of the ester hydrolysis rate between analogues and no clear match with structural complexity.
Case Study 4 /Bioaccumulation		Can we consider these structural differences to be minor with respect to the target endpoint?

4. Justification of data Gap filling

Conventional chemical categories (e.g. in CoCAM) have been developed mainly based on *in vivo* test data. Since one of the most important purposes of incorporating IATA concepts into grouping methods is to strengthen or extend chemical groups by using all available data, such as QSAR data and 'omics data, further development of guidance for justifying chemical groups based on many different kinds of data is necessary. From the review results of the four case studies the following specific issues were identified in this area.

- How to describe the similarity/trend of the observed effect of the target endpoint, when different information is available for substances. This issue of integrating contradictory results was especially highlighted by the Case Studies 2 and 3 when examining more complex repeated dose toxicity endpoints.

- The extent of data related to the target endpoint to be used in the data-gap filling justification. For example, reviewers suggested including the following additional data for each case study.

  – Case Study 1: Other mutagenicity test data

  – Case Study 2: Other human health endpoint data

- Case Study 3: Genotoxicity and developmental toxicity of metabolites
- Case Study 4: Bioconcentration data for other species than carp

- How to incorporate novel types of in vitro data: Extensive initiatives are ongoing in this field such as ToxCast[7] and SEURAT-1[8]. Although high throughput screening (HTS) data were used for only one chemical in this review cycle (Case Study 2), the reviewers and the authors strongly expect that when further *in vitro* data is available it can be more effectively used to strengthen complex endpoint categories in the future. Further case studies are needed to investigate the integration of HTS data.

- How to report QSAR prediction results: The importance of documenting if the modelled chemicals are within the applicability domain of the QSAR models used was emphasized. One way to do this is to show the coverage of the fragments used in the training set of the QSAR models. Inconsistency of QSAR predictions can be explained by doing so (see Case Study 4). There are tools for assessing if the predicted substance is within/outside the applicability domain of the model used. One of the most frequent sources of uncertainty identified by reviewers was the reliability of QSAR results.

- How to integrate data derived from different methods or models (e.g. when integrating various QSAR conclusions, experimental data).

The general guidance on data gap filling in the OECD Guidance on Grouping of Chemicals (OECD 2014a) (sections 2.5, 2.6, 4.2.6 and 5.2.6) need to be elaborated by incorporating IATA concepts based on the case studies in the future.

5. Uncertainty Analysis and Reporting

It is not the intent of the case studies to provide scientifically ideal examples of IATA, but to demonstrate examples of pragmatic uses of IATA in a regulatory context that are scientifically sound in terms of a fit for purpose approach. Therefore each case study contains different uncertainties as the data or the resources used for the case studies vary under each regulatory context. Whereas there is no guidance on documenting uncertainty in the OECD Guidance on Grouping of Chemicals (OECD 2014a), a section of uncertainty analysis was prepared in the template used for the case studies.

The following types of dominant uncertainty were identified by the authors or the reviewers.

1. Level of confidence in the hypothesis for read-across including:
   - knowledge of mechanistic basis
   - correlation of structural features to metabolism and adverse effects observed
   - amount of data available for supporting the hypothesis
   - reliability of data for supporting the hypothesis, especially QSAR predictions

[7] US EPA, Toxicity ForeCaster (ToxCast™) Data.
http://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data

[8] EU SEURAT-1. http://www.seurat-1.eu/

- level of similarity in observed effects for supporting the hypothesis

2. Relationship between hypothesis and human or environmental relevance

Some uncertainties identified in the original forms of the case studies were reduced by different methods such as incorporating sub categories, providing more detailed discussions or including supportive QSAR predictions. This demonstrates that consideration of uncertainty is helpful to improve the case studies. On the other hand, uncertainties remain and in this case, the uncertainty analysis helped the reviewers to consider the impacts of uncertainty and the acceptable degree of uncertainty with respect to the purpose of use of the case studies. The importance of uncertainty description and communication is recognized and identified as a high priority area for gaining further experience in the IATA context and then further developing guidance.

There were discussions on possible reporting formats of uncertainty in the context of grouping approaches in Case Studies 1 and 2. These range from more descriptive approaches to ranking approaches (Schultz et. al, 2015, Wu *et. al*, 2010 and Blackburn and Stuard, 2014). Most reviewers thought the descriptive analysis is more useful because it was viewed to be more transparent and objective. On the other hand, the ranking analysis was helpful due to its solid clear grades (levels). Therefore, it was suggested that the reporting formats used in the case studies should be explored in future case studies as the appropriate format depends on the purpose of case study and both formats are mutually complementary. It is suggested to develop guidance for reporting of uncertainty, in order to harmonise this aspect in future case studies.

The guidance documents of ECHA (ECHA, 2012) related to uncertainty and the case studies of SEURAT-1 can also be referred to in future activities on uncertainty analysis or explored as additional resources/examples on which to develop further considerations to inform recommendations and/or guidance development. The activities should aim at the inclusion of uncertainty analysis in the OECD grouping guidance.

6. Integrated Conclusion

The case studies were developed based on use in certain regulatory contexts of the lead countries. However, due to their specificities, the reviewer countries at times identified possible challenges in applying the results in their regulatory contexts. Thus, it was recommended that if the purpose of the case study is very specific, general conclusion for other purposes could be separately described. In addition, it would be helpful to develop guidance on how the methodology could be combined with other approaches in order to apply it in different regulatory frameworks.

5. CONCLUSION

The lessons learned from the cooperative review of four IATA case studies on grouping methods have increased experience in the application of these approaches. Case studies based on actual use in the lead countries provided concrete examples of how to use the grouping methods in a regulatory context. This experience provided insight into the importance of considering the difference between pragmatic approaches for a specified purpose which may be used in each regulatory context and perfect read-across. Understanding of the background of the regulatory framework and purpose of the case study helped the

reviewers to explore the issues regarding the practical use of the methods. Comparison between case studies with different purposes and target endpoints helped to identify common challenges with grouping methods, which were shared between the member countries.

The experience gained and shared through these case studies demonstrates the value of working collaboratively through case studies as a promising way for expanding the use of alternative methods in the member countries.

However, it has also been recognized that more case studies are needed for developing general guidance. Therefore, there is a need to build upon aspects identified review of this first set of case studies by incorporating further review experiences of future case studies. It is envisioned that future rounds of case studies will include grouping and read across approaches but also address a wider diversity of case studies examining additional IATA approaches and the use of novel approaches.

REFERENCES

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.

ECHA (2012), Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.8: Characterisation of dose [concentration]-response for human health. http://echa.europa.eu/guidance-documents/guidance-on-information-requirements-and-chemical-safety-assessment

ECHA (2015), Read-Across Assessment Framework (RAAF). http://echa.europa.eu/documents/10162/13628/raaf_en.pdf

Japan (2013), Concerning the assessment of bioaccumulation of new chemical substances by analog approach, etc. (Announcement). http://www.meti.go.jp/policy/chemical_management/english/files/laws/bioaccumulation_analog_approach.pdf

OECD (2011), Report of the Workshop on Using Mechanistic Information in Forming Chemical Categories. No. 138, Series on Testing & Assessment. ENV/JM/MONO(2011)8, OECD, Paris.

OECD (2013), Guidance Document on Developing and Assessing Adverse Outcome Pathways. No. 184, Series on Testing & Assessment. ENV/JM/MONO(2013)6, OECD, Paris.

OECD (2014a), Guidance on Grouping of Chemicals, Second Edition, No. 194, Series on Testing & Assessment. ENV/JM/MONO(2014)4, OECD, Paris.

OECD (2014b), Weight of Evidence Assessment for the Skin Sensitisation Potential of 4-Isopropylaniline (Cumidine, CAS 99-88-7), No. 199, Series on Testing & Assessment. ENV/JM/MONO(2014)5, OECD, Paris.

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 (2014), A Strategy for Structuring and Reporting a Read-across Prediction of Toxicity. Regulatory Toxicology and Pharmacology, Vol. 72, Issue 3, pp 586-601.

US Environmental Protection Agency (2009), Hazard Characterization Document: Screening Level Hazard Characterization Substituted Diphenylamines Category. Washington (DC): US EPA, Office of Pollution Prevention and Toxics.

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.

ANNEX: TEMPLATE USED FOR THE 2015 CASE STUDIES

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. The template will be revised based on experience with use and depending on the specific case study additional information may be required or particular sections may not apply).

Foreword

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

Table of Contents

1. Purpose

1.1. Purpose of use

- Specify the purpose of use of the IATA (eg. regulatory context, hazard identification, hazard characterization, risk assessment, screening etc.).

1.2. Target chemical(s)/category definition

- For analogue approach, provide the chemical descriptor common identifiers (including CAS number and name) 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 (eg. in chemical structure); and if known, any restrictions.

1.3. Endpoint(s)

- Identify the endpoint(s) for which the analogue/category approach is applied.

2. Hypothesis for the analogue approach/category

- For an analogue approach, describe the characteristics a substance must have to be suitable as a source substance. Provide the hypothesis for why read-across can be performed between the source and target chemical.
- 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.
- These hypotheses can be argued by the number of elements as follows (See Chapter 3 of the grouping guidance document).
    - Chemical identity and composition
    - Physical-chemical properties and other molecular description

- Kinetics: Absorption, distribution, metabolism and excretion
- Mode/Mechanism of action or adverse outcome pathways (MOA/AOP)
- Chemical / biological interaction
- Responses found in alternative assays
- Information obtained from other endpoints/species/routes
- The route and duration of expected exposure

Ideally, all elements should be addressed.

3. Source chemicals/Category members

3.1. Identification and selection of source chemicals/category members

- Provide the selection criteria, based on the hypothesis described in section 2, that were used to identify the source chemicals/category members.
- 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.

3.2. List of source chemicals/ category members

- Provide the common chemical identifiers (including CAS number and name) and chemical structure(s) of the source chemicals/category members.

4. Justification of data gap filling

4.1. Data gathering

- Provide the methods used for gathering the data for target and source chemicals/category members (eg. selection criteria of the data, data source).
- Provide the name, version and owner of the models used for deriving QSAR estimation data (Provide QMRF to the annex or reference to QMRF inventory maintained by the JRC).

4.2. Data matrix

- Provide a matrix of data (see data matrix template).
- Provide detailed data in an annex, as necessary.

4.3. Justification

- Based on the data matrix, summarize how these data support the hypothesis described in section 2.
- 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 range of values within which reliable estimations can be made for category members. Clearly indicate the borders of the category and for which substances the category does not hold.
- 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. Strategy for and integrated conclusion of data gap filling

5.1. Integrated conclusion

- Provide the strategy used to fill the data gap and integrated conclusion of data gap filling.

5.2. Uncertainty

- Discuss the uncertainty of the integrated conclusion.
- Aspects can include uncertainty and confidence associated with the data 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.

References

Annex

Data matrix for analogue approach

Data matrix, IATA

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 (GLP)	result		result		result	result	result	result
	Experimental result (non-GLP)				result	result			
	Integrated conclusion (eg. read-across)		result						
Target endpoint2	Experimental result (GLP)	result		result	result	result		result	result
	Experimental result (non-GLP)	result					result		
	Integrated conclusion (eg. read-across)		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	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
	…								

* More relevant metabolite such as toxicant

**General outline of relative comparative kinetics

Data matrix for category approach

Data matrix, IATA

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 (GLP)	result		result		result	result	result	result
	Experimental result (non-GLP)				result	result			
	Integrated conclusion (eg. read-across)		result		result				
Target endpoint2	Experimental result (GLP)	result		result	result	result		result	result
	Experimental result (non-GLP)	result					result		
	Integrated conclusion (eg. read-across)		result				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
	…								

* More relevant metabolite such as toxicant

**General outline of relative comparative kinetics