Background document

This is a thoughtstarter document based on a report developed as part of an ECETOC Task Force (TF) on Categories, read-across and (Q)SAR approaches. It is provided solely for the purposes of this Workshop to help frame and facilitate the breakout group scientific discussions. This document may not be copied, reproduced, or distributed further. The content provided in this thoughtstarter has been a work product of the individual experts who contributed to the Task Force and as such forms an Industry contribution to the development of further guidance for read-across approaches. The full report is expected to be published by ECETOC later in October. Please see www.ecetoc.org for more information.

Summary

Considerable practical experience has been gained in applying non-testing approaches for regulatory purposes, most recently driven in particular by the demands of the REACH legislation (EC, 2006). Currently read-across strategies that have been applied by companies for REACH have tended to rely on the 'analogue approach' as opposed to a 'category approach', in addition to (Q)SAR approaches.

(Q)SAR approaches have been extensively relied upon to address data gaps for physicochemical properties such as log K_{ow}, environmental fate parameters including biodegradation, hydrolysis, bioaccumulation potential and ecotoxicity endpoints: acute aquatic toxicity in the standard species (fish, *daphnia* and algae). For the aforementioned properties, such (Q)SAR approaches have typically been used as direct replacements to experimental testing. There are many expert systems available to facilitate these assessments including the OECD (Q)SAR Toolbox or the EPA's Episuite etc. For mammalian endpoints, (Q)SARs have been applied less frequently with exception to endpoints such as Ames mutagenicity and, to a lesser extent, skin sensitisation where the mechanisms are reasonably well understood and where the underlying data is more readily available.

Read-across approaches have been considered to help address data gaps for longer-term studies such as the 90-day repeated-dose study or reproductive/developmental effects. For such complex endpoints, (Q)SARs have been applied, but their role has been as "supporting information" to highlight potential chemical modes of toxicity or to offer insights to help demonstrate similarity in effect.

Information on likely transformation products and the rate of formation of these products derived from experimental studies are strongly recommended to substantiate the overall

read-across justification. Whilst toxicokinetic information is not a requirement under REACH *per se*, such information is viewed as key to help rationalise read-across approaches, particularly for endpoints like reproductive/developmental effects where current (Q)SAR approaches are still in early development.

Absence of toxicity is a particular challenge to justify. Despite provisions in REACH calling for the use of read-across and (Q)SAR for both the absence and presence of toxicity (see Annex XI in EC, 2006), the justification of 'absence' is not to be underestimated. Toxicokinetic information or physiologically based pharmacokinetic (PBPK) modelling is considered valuable to provide supporting information.

Read-across approaches for longer-term effects should ideally be structured to present an overall 'weight of evidence' (WoE) argument (SCENIHR, 2012). A justification needs to rely on several lines of corroborating evidence whether it be consistent metabolic profiles, similarity in effects at shorter exposures, (Q)SAR estimates or other supporting analogues with experimental data that are not necessarily part of the category/analogue approach. In the latter case, tools such as Toxmatch, Leadscope or the OECD (Q)SAR Toolbox can prove helpful to identify related analogues (Patlewicz et al, 2011).

Justifications should ideally be structured using a template such as the category/analogue reporting formats (CRF/ARF) which are outlined in both the OECD and REACH guidance documents. These templates are effective means of structuring the justifications on an endpoint per endpoint basis as well as presenting an overall data matrix for the analogue or category member under evaluation. A justification will be strengthened by the presentation of an explicit data matrix and one which demonstrates a consistent hazard profile for the members of the category or analogues under consideration. Presenting data for the source analogues for the endpoint that is proposed to be read-across alone may not be sufficient.

By default, read-across is considered to be associated with additional uncertainty due to the fact that information on a target substance is being inferred from that available on a source substance(s). Whilst assessment factors can be a route by which uncertainty is addressed, these should be used on a case-by-case basis and driven by the confidence associated with the underlying similarity hypothesis as well as the quality of the study data forming part of the supporting WoE information.

In the future, less 'classical' toxicity data will be anticipated for each individual analogue member, and rather more 'omics' will be available (van Ravenzwaay et al, 2012). Thus, there may well be a commensurate shift towards deriving larger categories as contrasted

with analogue approaches. This should facilitate analysis of trends, although the data gap-filling approaches will likely be contingent on the application of non-standard, alternative toxicity testing data including that from high throughput/high content technologies. Whilst this will be a challenge in interpretation, it does present a cost-efficient means of generating data in a relatively short timeframe.

Guidance and experience will continually evolve as Tox21 activities progress. EPA's Toxcast is one such example (Judson et al, 2010) and the OECD's adverse outcome pathway (AOP) work programme is another (OECD, 2011). Both will have an impact on the formation of read-across justifications. AOPs are foreseen as having the potential to provide the conceptual framework for how to utilise alternative data in the appropriate biological context as well as the chemical anchor for molecular initiating events (MIEs). Datasets such as those generated in Toxcast and related programmes may ultimately formulate one of the practical strategies in quantifying AOPs. OECD's grouping guidance, which is currently under revision, discusses AOPs as a means towards developing new categories and read-across that are more mechanistically based (OECD, 2011). AOPs will also be implemented in some fashion in the OECD (Q)SAR Toolbox to extend the scope of its functionality. It is anticipated that regulatory agencies may start to consider these approaches as part of their evaluations. Indeed, the US EPA have alluded to a shift in the development of their chemical categories from those that are purely based on structure and physicochemical properties to ones that rely on the concepts of AOP information to inform their development and evaluation (Seed, 2012).

This thoughtstarter presents some highlights from the ECETOC TF report that is currently being finalised for publication.

Table of Contents

Motivations for Practical Guiding Principles

The OECD HPV manual for chemical categories formed the starting point for the development of the REACH guidance (ECHA, 2008) as well as an updated OECD guidance (OECD, 2007) for chemical categories. The aim was to develop more practical guidance for developing, justifying and documenting chemical categories. Whilst the guidance was a step change in terms of cementing (Q)SAR principles, it does not provide details for evaluating the suitability of an analogue, the extent of detail that is required to document an approach or any specific examples to guide those needing to develop category approaches (Wu et al, 2010).

The ECETOC TF report that was prepared aims to address some of these issues with practical insights derived from examples companies have worked on. This thoughtstarter presents highlights from each of the main chapters.

Considerations for the use of Non-Testing Approaches

Regulatory programmes, such as REACH (EC, 2006), mandate that vertebrate animal testing should be conducted only as a last resort. However, it is not always possible or appropriate to utilise non-testing approaches in lieu of animal testing. Therefore, when deciding whether or not to use non-testing approaches, the first consideration should always be whether such approaches are scientifically plausible. If the answer is 'no' then more traditional approaches encompassing a combination of *in vitro* and *in vivo* toxicity testing should be considered. If it is determined that non-testing approaches can be exploited then a number of other factors should be taken into account when constructing the justification.

The benefits include:

- Reduce Animal testing – legal and 'reputational' obligation to avoid 'unnecessary' testing in animals; avoiding animal testing due to legal restrictions (e.g. Cosmetics Directive)
- Time – non-testing approaches and read-across take less time to implement – particularly if using these approaches to characterise endpoints such as repeated dose toxicity and reproductive/developmental toxicity
- Money – toxicity tests are expensive. By utilising non-testing and read-across approaches multiple substances can be 'hazard characterised' for far less money than running full toxicological studies on every substance

It is important to note that for regulatory programmes such as the EU REACH regulation, registrants must have legitimate access to refer to the data they require to support their registration. As such, where an analogue or category has been used, the registrant would need to purchase access to all the data necessary to support their registration dossier. Doing so may involve significant up-front costs, however this cost is typically lower than that associated with performing a new set of studies.

The number of data gaps that need to be filled may also drive the type of non-testing approach - category approach or (Q)SARs. Obviously the latter will depend on the availability of actual valid (Q)SAR models that can be applied and the level of granularity they provide in terms of their predictions. Some (Q)SARs may only be capable of hazard identification whereas others might provide a quantitative measure of toxicity to permit a risk assessment.

At the same time there are a number of risks and implications associated with utilising non-testing approaches, acknowledging these may shape the strategy taken to justify their use.

Two of the most obvious risks and their implications are that the read-across approach is rejected, or that in using read-across or non-testing approaches, the hazards of the substance are mis-characterised – i.e. by either being too conservative (over-classifying) or not being sufficiently conservative. The rejection of a read-across approach is clearly a major concern since the outcome could lead to significant additional testing required in order to 'fill the identified gaps' following rejection of read-across. Alternatively, if the use of this approach mis-characterises the hazards of the substance by not being sufficiently conservative, then workers, consumers or the environment *may* be placed at increased risk unknowingly. This is likely the primary reason why non-testing approaches are scrutinised to the extent they are. Additionally if a read-across approach is rejected then the upfront time and money invested in developing the read-across justification and purchasing access to data has essentially been wasted.

These risks drive the need to make sure that when non-testing approaches are utilised, they are well supported with sound and robust justifications. Thus, it is vital that sufficient time, effort and expertise are invested in order to ensure that where these approaches are employed, they are done so in a scientifically rigorous and supportable fashion. This should increase the likelihood of regulatory acceptance as well as reduce the possibility that hazards are mis-classified (for better or for worse) or that resources are wasted.

Systematic Workflow

There are a number of different steps that are involved in deriving a category/analogue approach. The first step is the identification of an analogue(s) from which to form a category/analogue approach and the overarching considerations that could be taken to evaluate the relevance of those analogues. The next step evaluates what the underlying rationale(s) could be for forming the category/analogue approach, the scope of the category/analogue – whether it should be restricted to certain endpoints; mammalian or environmental effects and how a read-across might be substantiated for different endpoints in turn. Other aspects to the workflow include considerations around classification and labelling (C&L) of the category members as well as their impurities.

Analogue identification

Analogue identification is a critical first step when undertaking a grouping approach. For situations where the starting point is a single chemical under evaluation, identification of analogues and evaluating their relevance is fundamental. Depending on the end application, the strategy of analogue identification can vary considerably.

To date, many examples under REACH are thought to be analogue approaches highlighting a reluctance to broaden the scope to more than a handful of substances to manage the information and data that needs to be submitted in support of a substance and ensure that the relevant study summaries and access to the data usage is undertaken in an appropriate manner. Hence in the case of regulatory applications, it may be that *apriori* the substance(s) to be registered and the analogues to be used in support of the registration are already defined. In the cases where there is no presumption/restriction of what analogues to use, a myriad of tools and techniques may be relied upon to assist and facilitate the identification of analogues.

The most common analogue identification approaches still rely on structural similarity despite the fact that this is known to be only one criterion in identifying and evaluating analogues for their suitability for read-across.

Rationale for grouping: Analogue evaluation

Once potential analogues have been identified, a next step is to determine their suitability for the specific purpose in mind. Current regulatory guidances provide rationales that can

be used as the underlying hypotheses to support a category or analogue approach. Those highlighted in the REACH technical guidance are:

- common functional group(s) (e.g. aldehyde, epoxide, ester, specific metal ion)
- an incremental and constant change across the category (e.g. a chain-length category), often observed in physicochemical properties, e.g. boiling point range
- common constituents or chemical classes, similar carbon range numbers. This is frequently the case with complex substances often known as substances of Unknown or Variable composition, Complex reaction products or Biological material (UVCB substances)
- the likelihood of common precursors and/or breakdown products, via physical or biological processes, which result in structurally similar chemicals (e.g. the metabolic pathway approach of examining related chemicals such as acid/ester/salt)

Whilst a category/analogue hypothesis may principally make reference to one of these "similarity" rationales, in practice endpoint justifications and supporting information will be multifaceted. Arguably multiple justifications should serve to increase the overall confidence in the category/analogue approach.

General considerations that need to be factored into any category/analogue justification and the tools and approaches that can facilitate this step are discussed in more detail in the main report together with suggestions of factors that can help increase confidence in a category approach. Special attention is drawn to particular considerations for a 'metabolic category' as well as metals/metal compounds.

List of endpoints covered

The read-across within a category is dependent on an underlying rationale to justify the similarity between members for a specific endpoint of concern i.e. an endpoint specific category. As such the similarity context may differ depending on what the MOA and molecular initiating event is for the particular toxicity. As such some members of a category could be excluded for a given endpoint or in contrast not all endpoints might be relevant for all category members. This should be discussed in the overall justification presented.

Endpoint by endpoint justification

Whilst a rationale needs to be proposed to underpin a given category, a critical component in demonstrating that the category is robust and scientifically credible lies in providing a comprehensive justification on an endpoint by endpoint basis. The decision on which and how many endpoints to fill in data gaps by non-testing methods depends on the ultimate purpose of forming the category. For regulatory purposes, e.g. REACH where a larger range of endpoints is required to be addressed, a larger base set of test data (e.g. physicochemical, basic (eco)toxicity and environmental fate) would be needed for each chemical before one can confidently decide to group chemicals into a single category. There is also higher confidence in the robustness of the category when the number of category members is not too large. Before filling in data gaps for more complex endpoints e.g. reprotoxicity, development toxicity, one may require the existence of similar trends in test data for lower endpoints, e.g. genotoxicity, skin sensitisation.

There are a number of considerations that may come into play for the evaluation of analogues for each key endpoint within the category. (Q)SAR information in terms of what structural/physicochemical information is pertinent for a given endpoint can provide a means to demonstrate commonality in chemical mechanisms or modes of action. Suggestions for what endpoint justifications might be helpful are discussed in more detail in the main ECETOC TF report.

Recommendations

There are many endpoints to which read-across can be applied and these range in complexity and sophistication from simple physicochemical and acute/local effects to repeat dose/systemic, reproductive toxicity etc., this range means that there is naturally a range in the complexity of approaches that need to be developed.

Data on toxicokinetics could be a key piece of evidence to support read-across justifications. Toxicokinetic data can help demonstrate that substances in a category are metabolised in a similar manner, have similar absorption and excretion kinetics and similar distributions and therefore would be expected to have similar toxicological properties. These data can also aid in defining trends across a category for example, confirming that as molecular weight increases, bioavailability decreases, thus reducing the potential for systemic toxicity. Having this information therefore has the potential to reduce the uncertainty in assessing read-across and reduce the need for intensive toxicological studies on every substance.

Research Needs/Activities

Historically, read-across has been limited to the analogue approach and to small chemical categories of structurally similar substances. With Tox21 coming into effect, new approaches and emerging technologies will shape the sorts of toxicity testing that is carried out in the future and thus a question exists on how this type of information will become integrated to inform read-across approaches. The report proposes some practical approaches.

References

Commission of the European Communities. (2006). Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Off J Eur Union, L 396/1 of 30.12.2006.

ECHA REACH TGD. Technical Guidance Documents for the implementation of REACH. Guidance on information requirements and chemical safety assessment. European Chemicals Agency, Helsinki, Finland. (http://echa.europa.eu/web/guest/guidance-documents/guidance-on-information-requirements-and-chemical-safety-assessment)

Judson R, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, Reif D, Rotroff D, Shah IA, Richard AM, Dix DJ. (2010). In Vitro Screening of Environmental Chemicals for Targeted Testing Prioritization: The ToxCast Project. Environ. Health Perspect. 118(4): 485-492.

OECD (2007). Series on Testing and Assessment Number 80: Guidance on Grouping of Chemicals. ENV/JM/MONO(2007)28. Organisation for Economic Co-operation and Development. Paris, France.

OECD (2011). OECD Series on Testing and Assessment Number 138. Report of the Workshop on Using Mechanistic Information in Forming Chemical Categories.

ENV/JM/MONO(2011)8. Organisation for Economic Co-operation and Development. Paris, France.

OECD (2012) Series on Testing and Assessment No.168: The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins Part 1: Scientific Evidence. ENV/JM/MONO(2012)10/PART1. Organisation for Economic Co-operation and Development. Paris, France.

Patlewicz G, Chen MW, Bellin CA. (2011). Non-testing approaches under REACH - help or hindrance? Perspectives from a practitioner within industry. SAR QSAR Environ Res 22(1-2): 67-88.

Scientific Committee on Emerging and Newly Identified Health Risks SCENIHR. (2012). Memorandum on the use of the scientific literature for human health risk assessment purposes – weighing of evidence and expression of uncertainty. 17th plenary of 19 March 2012

Seed J. (2012) Presentation made at OECD Expert Group on Molecular Screening and Toxicogenomics meeting. 7-8th June, 2012, Paris, France.

van Ravenzwaay B, Herold M, Kamp H, Kapp MD, Fabian E, Looser R, Krennrich G, Mellert W, Prokoudine A, Strauss V, Walk T, Wiemer J. (2012). Metabolomics: A tool for early detection of toxicological effects and an opportunity for biology based grouping of chemicals-From QSAR to QBAR. Mutat Res. 15 746(2): 144-150.

Wu S, Blackburn K, Amburgey J, Jaworska J, Federle T. (2010). A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. Regul. Toxicol. Pharmacol. 56: 67-81.