Assessing Uncertainty in Read-Across: Questions to Evaluate Toxicity Predictions Based on Knowledge Gained from Case Studies

Terry W. Schultz[1], Andrea-Nicole Richarz[2], and Mark T.D. Cronin[3*]

[1]The University of Tennessee, College of Veterinary Medicine, 2407 River Drive, Knoxville, TN 37996-4543 USA

[2]Directorate for Health, Consumers and Reference Materials, Joint Research Centre, European Commission, Ispra, Italy;

[3]Liverpool John Moores University, School of Pharmacy and Biomolecular Sciences, Byrom Street, L3 3AF Liverpool, England

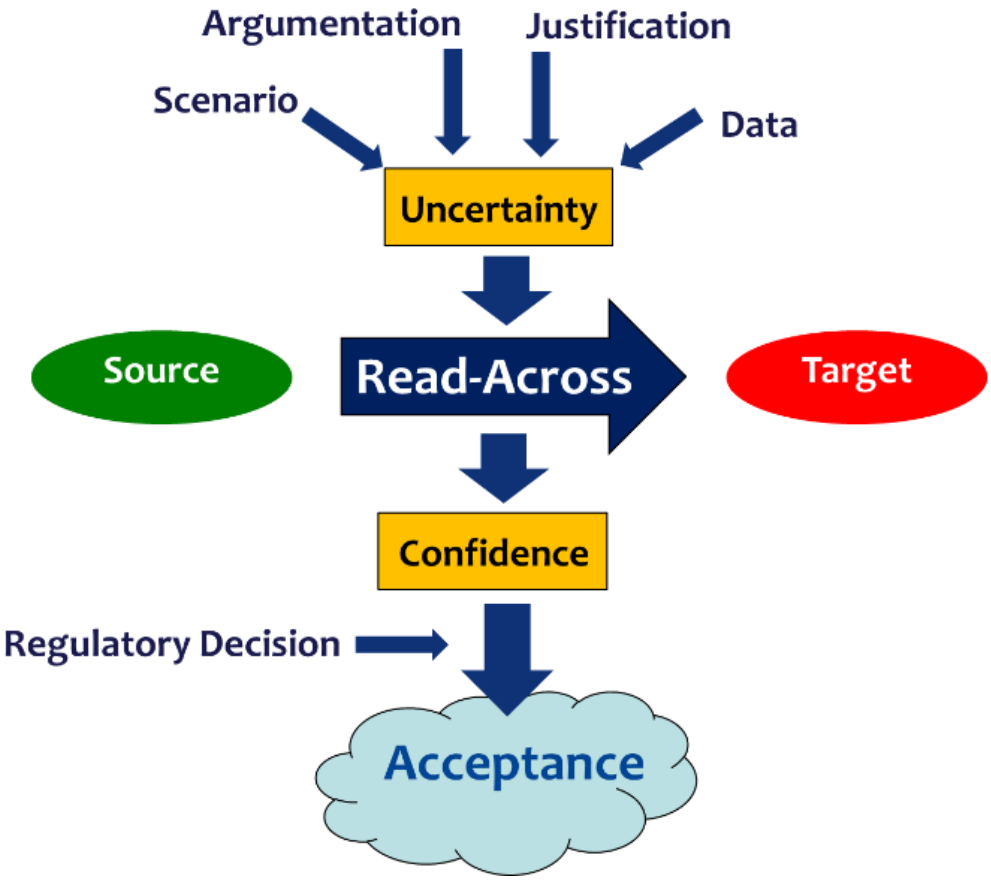
* Corresponding author.

E-mail address: m.t.cronin@ljmu.ac.uk

Short title: Assessing Uncertainty in Read-Across

Disclaimer: The views expressed in this document are those of the authors and do not necessarily represent the official views of the European Commission.

Graphical Abstract



Abstract

Read-across as an alternative assessment method for chemical toxicity has growing interest in both the regulatory and industrial communities. The pivotal means of acquiring acceptance of a read-across prediction is identifying and assessing uncertainties associated with it. This study has identified and summarised in a structured way the variety of uncertainties that potentially impact acceptance of a read-across argument. The main sources of uncertainty were established and divided into four main categories: i) the regulatory use of the prediction, ii) the data for the apical endpoint being assessed, iii) the read-across argumentation, and iv) the similarity justification. Specifically, the context of, and relevance to, the regulatory use of a read-across will dictate the acceptable level of uncertainties. The apical endpoint (or other) data must be of sufficient quality and relevance for data gap filling. Read-Across argumentation uncertainties include: 1) mechanistic plausibility (i.e., the knowledge of the chemical and biological mechanisms leading to toxicity), 2) completeness of the supporting evidence, 3) robustness of the supporting data, and 4) Weight-of-Evidence. In addition, similarity arguments for chemistry, physico-chemical properties, toxicokinetics and toxicodynamics are linked to these read-across argumentation issues. To further progress in this area, a series of questions are proposed with the goal of addressing each type of uncertainty.

Keywords: read-across, toxicity, sources of uncertainty, types of uncertainty, guiding questions

Highlights:

- Six read-across case studies were reviewed to establish overarching uncertainties.
- Twelve types of uncertainties identified for read-across for toxicity prediction.
- Questions were formulated to assist in assessing uncertainties in read-across.
- Comparison with existing schemes for read-across uncertainty is given.

Introduction

Toxicologically data-poor substances dominate the universe of industrial chemicals. The paucity of data and information has motivated the use of chemical grouping and read-across as a data gap filling technique [1]. However, whilst it is a commonly applied technique, among stakeholders using it there is agreement that for better regulatory acceptance of read-across more experience is needed on how confidence in the prediction could be enhanced i.e., how to lower uncertainties [2-5]. In order to achieve this goal, for regulatory application at least, the Organisation for Economic Co-operation and Development (OECD) has developed technical guidance describing a workflow to identify analogues, or form categories, of similar chemical structure to allow for read-across predictions to be made to fill toxicological data gaps [1]. In a related activity, in 2015 the OECD launched the Integrated Approach to Testing and Assessment (IATA) Case Studies Project under its revised Cooperative Chemicals Assessment Programme. The aim of this project was to increase experience of the use of IATA by developing case studies which, in turn, could help inform the better use of read-across and hence development of guidance. The case studies developed were intended to be examples of read-across predictions that would be fit for regulatory use. Individual case studies submitted from member countries and other organisations were reviewed within the IATA Case Studies Project. The review process concentrated on the: 1) the strongest aspects of case study, 2) the areas of uncertainty within the case study, 3) the identification of, and agreement for, areas of further development of guidance, and 4) the possibility of the use of the case study in various regulatory contexts.

The IATA Case Studies Project enabled the OECD to take a more reflective and dynamic analysis of the application of read-across for toxicological data gap filling than would otherwise have been possible [1, 4, 6]. Specifically, the analysis focused on establishing consistency in the information required to demonstrate: 1) how to document the justification for a read-across, 2) how to perform read-across for more complex endpoints (e.g., repeated dose toxicity), 3) how to develop and gain support for quantitative read-across which may be applicable for hazard characterisation, 4) how to take mechanistic considerations into account in grouping chemicals, 5) how to derive weight-of-evidence (WoE) conclusions based on the results from alternative methods (and New Approach Methodologies (NAMs)), and 6) how to assess the uncertainty associated with the read-across.

Additional to the on-going activities described above there have been other recent efforts to address the issues raised by the OECD within academia, government and industries. These efforts began the process

of the refinement and improvement of how read-across is documented and concentrated on clarifying the seminal issues and overcoming the challenges posed by the OECD (see Patlewicz et al. [7, 8] and Schultz and Cronin [9]). Several different frameworks for undertaking read-across are available, as described in both regulatory technical guidance and peer-reviewed scientific literature [6, 10-16]. Such guidance includes two multi-step procedures, one for the analogue- and one for the category-approach that includes several key features. Patlewicz and co-workers [7, 8] note that the OECD framework is an endpoint specific, bottom-up approach that is a generalisation of the ECHA approach. These frameworks have recently been summarised and a more generic framework proposed [8].

A consistent issue among these frameworks is that of defining and, where possible, quantifying the different elements and overall uncertainty within a read-across hypothesis, justification and prediction. Specifically, given the challenges of gaining regulatory acceptance of read-across-based predictions, there have been efforts to explore how read-across justifications can be evaluated critically so that uncertainties can be explicitly identified and practical strategies can be brought to bear on reducing those uncertainties [6, 10-15]. The criteria covered by five key publications [6, 10, 12, 13, 15] are summarised in Tables S1-S5 in the Supplementary Information. High confidence (i.e., low concerns about potential error in the prediction) is linked to a read-across argument where there is information and data showing why the prediction is valid. To some extent at least, the development by the European Chemicals Agency (ECHA) of the Read-Across Assessment Framework (RAAF) was a further attempt to facilitate regulatory acceptance. The RAAF provides a means to assess (quantitatively) the documentation and evidence of a number of read-across scenarios, however it does not provide a comprehensive review of the uncertainties associated with read-across [15].

The purpose of this investigation was to review the issues of uncertainty as related to read-across for a complex health endpoint with the specific objective of using this knowledge to assist in the development of guidance on assessing uncertainty in read-across for supporting regulatory decisions, e.g. (quantitative) risk assessment, classification and labelling, screening and prioritisation etc. The overall aim was to report all sources and types of uncertainty previously identified in case studies that may affect the overall uncertainty. Further, for each type of uncertainty a series of questions was proposed which were intended to practically assist both the developer and user in assessing the uncertainty surrounding a read-across prediction.

Selection and Analysis of Read-Across Case Studies

Selection of Case Studies

The intention of selecting read-across case studies was not to be all inclusive but rather to examine studies where the hypothesis and justification for grouping has a similar level of complexity. The regulatory context of each case study was developing a point of departure suitable for quantitative risk assessment, although the uncertainties were developed in terms of being usable for other regulatory contexts. Among the criteria used in selecting the case studies reviewed in this investigation was that the OECD "Template for IATA Case Studies on Read-Across" [5, 6] or a highly similar template for documenting the read-across prediction was available and had been utilised to guide the analysis. The case studies were selected such that appropriate information was available from the (OECD or other) template relating to:

1) the purpose of the read-across,
2) the hypothesis for the analogue approach/category,
3) details of the source chemicals/category members,
4) justification of data gap filling, and
5) the strategy for, and integrated conclusion of, data gap filling.

The last factor was particularly important with regard to the assessment of uncertainties.

Four other criteria were also included in the evaluation process: that the data gap was for a complex toxicological endpoint, in this case repeated-dose toxicity; the case studies considered were comparable in terms of their purpose and how they were undertaken; the category approach was applied; and that different read-across hypotheses were considered (e.g., based on different modes of action and/or different toxicokinetics) to cover a wide range of possible scenarios of chemical similarity for read-across.

Case Studies Considered

Following the selection process, six read-across case studies were identified that met the selection criteria – it is recognised that this was not an exhaustive list and there is a bias to those arising from the OECD IATA programme [14, 17-19] and the authors' own work [20, 21]. However, these case studies are representative of the current state of the art and focus on chemical group-based read-across with an array of chemicals with data gaps (i.e., target chemicals). Inevitably, the case studies differ in the number of chemicals with data (i.e., the source substances), which varies from one to several as well as (as defined

by the selection criteria) the read-across hypothesis. The six read-across case studies (CSs) considered were all for repeated dose toxicity and are summarised in Table 1.

Table 1. Summary of the six read-across case studies for repeated dose toxicity evaluated as part of the identification, evaluation and assessment of sources and types of uncertainties.

Case Study (CS) Number	Compounds Considered	Read-Across Hypothesis and The Proposed Mechanism of Action Exhibited by the Compounds	Reference
CS1	n-Alkanols	Basal cytotoxicity / non-polar narcosis Mode-of-Action (MoA) with no systemic toxicity and high No Observed Effects Levels (NOEL) linked to Phase 1 oxidative metabolism to CO_2	OECD [17]
CS2	2-Alkyl-1-alkanols	Basal cytotoxicity / non-polar narcosis MoA with no systemic toxicity and moderate NOEL values linked to Phase 2 glucuronidate metabolism	OECD [18]
CS3	Aryl alcohol alkyl carboxylic esters	Basal cytotoxicity / non-polar narcosis MoA with no systemic toxicity and moderate NOEL linked to ester hydrolysis	OECD [19]
CS4	Short-chain mono-alkylphenols	Polar narcosis MoA with no systemic toxicity and moderate NOEL values linked to Phase 2 glucuronidate metabolism	Mellor et al. [21]
CS5	Allyl esters	An electrophilic MoA with systemic toxicity (i.e., liver fibrosis) and low NOEL values. The justification was that ester hydrolysis leads to a common metabolite, allyl alcohol, which was metabolised to the same definitive toxicant - acrolein	OECD [14]
CS6	β-Olefinic alcohols	An electrophilic MoA with systemic toxicity (liver fibrosis) and low NOEL values linked to oxidative metabolism of the parent alcohol leading to a different definitive electrophilic	Przybylak et al. [20]

		toxicant, exhibiting the same reactive mechanism - Michael Addition	

Analysis of the Selected Read-Across Case Studies Focussing on Uncertainty

The uncertainty of each read-across case study was analysed with regard to the fifth topic of the OECD Template [6], namely the consideration of "the strategy for and integrated conclusion of data gap filling". This section is of two interrelated parts which firstly discusses the uncertainty associated with each part of the read-across and secondly makes an integrated conclusion which draws together all aspects of the read-across to assess its validity or potential fitness of purpose. Specifically for this analysis, the integrated conclusion was the focus of the analysis of the determination of uncertainties.

Results and Discussion

This study has assessed six read-across case studies for the prediction of repeated dose effects. The read-across exercises have all been recorded on standardised templates which has allowed for an assessment of their relative uncertainties. It is important to note that this exercise was not intended as a retrospective exercise to evaluate, or otherwise, the quality or robustness of the read-across case studies, but to identify and compile a comprehensive list of uncertainties associated with read-across as a means to stimulate the creation of guidance.

An examination of the six case studies revealed that all six studies addressed, to some extent at least, the overarching question - is the uncertainty acceptable to use the read-across prediction(s) to fill the data gap? All six case studies related to reading across a NOEL value; this is considered to be one of the most challenging applications of read-across and that uncertainties identified in this analysis could be applied broadly to other endpoints and uses. The answer to the question of whether the uncertainties are acceptable is on one hand part of the risk management decision or dictated by the relevant chemical legislation or required purpose. On the other, practical, hand, to address this question, different sources of uncertainty were identified and a variety of types of uncertainty described. Four main sources of uncertainty were identified and are listed in Table 2.

Table 2. Summary of the main overarching sources of uncertainty identified in the read-across case studies.

Main Sources of Uncertainty in Read-Across
I. Uncertainty related to the regulatory use (i.e., the impact the regulatory scenario will place on acceptable levels of uncertainty.
II. Uncertainty related to the data for the endpoint under consideration for the source compound(s).
III. Uncertainty related to the argumentation of the read-across including, but not limited to, data quality. This can be sub-divided into the following types:
a) The approach taken to read-across, i.e. the type of similarity and grouping applied
b) The mechanistic plausibility of the read-across argument
c) The completeness of the argument provided to support the read-across
d) The robustness of the argument from the hypothesis to execution of the read-across
e) The Weight-of-Evidence presented
IV. Uncertainty related to the justification of similarity between the target and source compounds. This can be sub-divided into the following types:
a) Definition and demonstration of similarity in chemical structure between the target and source compounds
b) Definition and demonstration of similarity in physico-chemical properties between the target and source compounds
c) Definition and demonstration of similarity in toxicodynamics including species and assay differences between the target and source compounds
d) Definition and demonstration of similarity in toxicokinetics, focussed on Absorption, Distribution, Metabolism and Excretion (ADME) properties between the target and source compounds

Based on our review, we recognised the sources of uncertainty are associated with types (or elements) of uncertainty (Table 3). While not all the case studies identified or addressed all the types of uncertainty, and in some case studies they were even combined, Table 3 reports all the uncertainties found one or more times in our review. The sources and their interrelationships with types of uncertainty are summarised schematically in Figure 1.

INSERT FIGURE 1 HERE

Table 3. Definition of types of uncertainties identified in the read-across case studies.

Number	Individual Uncertainties in Read-Across	Description and Relevance	Source of Uncertainty – Numbered as per Table 2
1	The context of, and relevance to, the regulatory use of the read-across prediction as defined by appropriate problem formulation.	Regulatory use is usually defined at the start of the read-across but may also be applied at the end. The use defines the level of uncertainty that may be acceptable. Acceptable uncertainty is lower for risk assessment than for prioritisation and screening, respectively. Especially for risk assessment, the acceptable level of uncertainty may not be reached for all members of a category. The read-across must be relevant to the endpoint and question to be addressed.	I
2	Type of category / group including the definition of the applicability domain	The number of target and source chemicals (i.e., the size of the applicability domain), impacts on uncertainty [13, 15]. Uncertainty increases, and is more difficult to assess, in the following order of scenarios: 1) Many source substances to one target chemical, 2) One source substance to one target chemical, 3) Many source substances to many target chemicals, 4) One source substance to many target chemicals. Uncertainty in a read-across prediction is defined by the extent of the interpretability and defensibility of the applicability domain which is described in a transparent manner. Whilst interpretability often decreases in going from one to many target chemicals, defensibility is typically increased. The applicability domain would normally be described in terms of the structural features and properties of the category.	IIIa, IVa

3	The premise or hypothesis of the read-across	A well-stated premise or hypothesis, described in a testable format is essential to assessing uncertainty in read-across. The hypothesis of the read-across is typically linked to other uncertainties. Uncertainty associated with the hypothesis of the read-across often (but not always) centred on the identification of the correct MoA (e.g., confirmation of the molecular initiating event and/or key events along an adverse outcome pathway). The hypothesis typically includes statements about uncertainties 4-11 in this Table. Overall, the premise or hypothesis must be relevant to the endpoint and use scenario.	IIIa, IIId
4	Mechanistic plausibility including completeness of the understanding of the MoA or AOP	Mechanistic plausibility is a key element of defining and justifying the read-across hypothesis (Uncertainty 3). Assessment of the uncertainty of the mechanistic plausibility is based largely on the knowledge of the chemical and biological mechanism(s) resulting in the toxicity. Mechanistic assessments take account of the strength, consistency, and selectivity of the experimental evidence association with sets of data/information typically taken from a MoA. Strength is related to the number of intermediate or key events tested and the number of analogues tested within an assay or key event of an AOP – as well as how complete the AOP is. Consistency is related to data uniformity within a key event/mechanistic test and data regularity between different tests. Selectivity is related to the ability to discriminate between known positive (i.e., category members) and known negative (i.e., substances known to be outside the domain of the read-across).	IIIb
5	Similarity in chemistry	The uncertainty in chemical similarity/dissimilarity is a key element of assessing the robustness of a read-across prediction. The underlying philosophy of read-across is that substances which are similar in chemical structure will have similar chemical properties and thereby, similar toxicological properties [13] (it is recognised that biological similarity may also be used as a means to justify a read-across). Thus, demonstrating similarities in chemistry is essential and becomes the starting point in any similarity argument unless there is clear evidence that similarity can be defined in terms of biology (e.g., similarity in an omics profile). Assessing similarity in chemistry is typically undertaken by examining structural information and relevant physico-chemical properties. This implies that, where practically possible, the chemical identity of target and sources chemicals are correct and stated explicitly (noting potential issues with mixtures and chemical substances of Unknown or Variable Composition, Complex Reaction	IVa, IV4b

		Products and Biological Materials (UVCB Substances)). The number and type of factors relating to chemical similarity must be relevant to the endpoint under consideration and will themselves affect the quality of the read-across [22, 23]. In using the analogue approach, i.e. a one-to-one read-across, establishing similarity in chemistry may be sufficient to establish the similarity justification to meet acceptable uncertainty. It is also often important to show that dissimilarities in chemistry are not toxicologically relevant.	
6	Toxicodynamic similarity	Similarity in toxicodynamics is inter-related to several uncertainties in this table (notably Uncertainty 4). However, its separate assessment with regard to uncertainty may include information from studies with different species (e.g., rat, mouse or rabbit), different exposure durations (e.g., 45 to 54-day, 90-day and 2-year) and different exposure schemes (e.g., gavage, feed or drink, as well as dermal).	IVc
7	Toxicokinetic similarity	Similarity in toxicokinetics, especially ADME properties, is seen as crucial to assessing uncertainty. Metabolism is often seen as the most contentious aspect of the toxicokinetic similarity justification. Specifically, when the hypothesis proposes that a metabolite induces the apical endpoint, the metabolic pathways and effects induced by metabolites need to be considered. In addition, similarity or a predictable trend in the rate of transformation has also to be considered. There are seldom experimental ADME data for all the target materials. Experimental metabolism data for one or more of the source substances are, however, critical to gaining acceptable uncertainty. Furthermore, it is critical to gaining acceptable uncertainty to have highly similar metabolic simulations of the target material(s) and source analogue(s).	IVd
8	The quality of the apical endpoint data used to fill the data gap	The quality of the *in vivo* apical endpoint data read across provides one of the fundamental source of uncertainty in a read-across. Data with low uncertainty are sought, which may include aspects of the reliability of the data and relevance to the endpoint and context, as well as for human health / environmental effects. For example, in the context of the case studies considered in this investigation and repeated-dose toxicity in particular, low uncertainty can be considered to be associated with (amongst other possibilities) high quality *in vivo* data derived from GLP studies following an OECD test guideline (e.g., TG 408, 90-day Subchronic Oral Toxicity). For a	II

		GLP study the manner (e.g., reliability accuracy, precision, repeatability and reproducibility) by which the apical *in vivo* data are generated should be well documented. For data from a non-standard protocol or non-guideline studies, or performed according to outdated protocols, which may be used to support a read-across, expert judgment is required. A decision must be made as to whether the data can be used reliably without restrictions, the data can be used but with restrictions (e.g., Weight of Evidence) or the data are not usable (e.g., flawed experimental design, lack of understanding what the results mean, etc.). The decision on non-standardised data will also determine if the read-across is sufficient to meet the regulatory purpose.	
9	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	If there is more than one source substance, the consistency or inconsistency in the outcomes and severity of the apical *in vivo* hazard affects uncertainty. Low uncertainty is associated with consistent phenotypic expression and consistent potency of the hazard among the source substances. Conversely, higher uncertainty is associated with inconsistent phenotypic expression and/or varied potency of the hazard among the source substances. Again, for the example of repeated-dose toxicity, it is not only crucial to have similar NOEL values but also having similar LOEL effects based on the same effects. A distinction is not made between adverse and non-adverse effects. For other endpoints, communality in effects would be sought. Concordance with regards to the intermediate and apical effects and potency data has been described by Blackburn and Stuard [12], Schultz et al. [13] and OECD [6], (see Tables S2, S3 and S5 in the Supplementary Information, respectively). A complete data matrix and significant dose-response relationships are required to assess concordance and determine uncertainty.	II, IIIc
10	Strength or robustness of the supporting data sets	The type, quality and robustness of the supporting data sets provide an important means of reducing uncertainty. Supportive data are typically relevant *in silico*, *in chemico* and *in vitro* data and may also include other (relevant and acceptable) non-standard *in vivo* data. When data consistency is observed across the category, uncertainty is reduced.	IIIc, IIId

11	The Weight-of-Evidence (WoE) supporting the prediction	The WoE supporting the read-across prediction is garnered from all the information and data used in the read-across. The most important means of reducing uncertainty associated with the WoE was to supplement the read-across with relevant data from additional, possibly targeted, studies using alternative methods. Having consistency across the applicability domain for information from alternative methods is important to reduce uncertainty associated with the overall WoE.	IIIc, IIId, IIIe
12	Documentation and written evidence provided	Uncertainty is reduced when the evidence is presented clearly and unambiguously including cogent descriptions, arguments and justifications. The evidence should relate the uncertainty in the read-across to that deemed to be acceptable and demonstrate that the level of uncertainty is acceptable for the stated purpose.	Applicable to all sources and IIIa, IIIb, IIIc in particular

In general, examination of the six case studies that formed the basis of this investigation showed that confidence in a read-across prediction can be enhanced by providing mechanistic transparency and using appropriate experimental data from structural analogues. In addition, confidence is enhanced by using appropriate toxicokinetic properties in the form of absorption, distribution, metabolism and excretion (ADME) information, as well as, relevant *in vitro, in chemico* data, structure-activity relationships (SARs), and high throughput screening (HTS) information, especially as a means of substantiating chemical and biological similarity and increased WoE [9].

Specific analysis of a series of 30 questions that are proposed (in Table 4) to address the 12 uncertainties listed in Table 3 was performed. Each case study was evaluated according to the questions in Table 4 based on the information provided in the source document with regard to the assessment of uncertainties. The findings for each case study are presented in the Supplementary Information Table S6-S11 and are summarised in Table 4. Table 4 confirms that increased uncertainty was often associated with the plausibility of the mechanism of action, toxicokinetics as well as the quality and relevance of the data. These are well established areas of concern (cf [9]) and indicate where future effort is required. Analysis of a read-across prediction using the questions in Table 4 proved to be a rapid and efficient means of determining and analysing the uncertainty(ies). The questions are practical and pragmatic and cover all aspects required by e.g. the ECHA RAAF and other templates.

The questions in Table 4 and overall analyses find commonalities with the main types of uncertainty and characteristics of the previously published schemes [6, 10, 12-15] summarised in Tables S1-S5 (of Supplementary Information) respectively. It is clear that the previous schemes take very different approaches to what types of uncertainty and characteristics may be addressed and how to do this, although there are overlaps in coverage of areas of uncertainties. The six schemes have been mapped, as completely as possible, onto the questions developed in this study and the results of this mapping are shown in Table 5. This shows the ECHA RAAF to be the most comprehensive, although it is noted that this assessment is from the questions derived from all possible RAAF scenarios (as in Table S4). The OECD Template for addressing uncertainties (see Table S5) is a brief set of questions but is intended to support the overall documentation process. The uncertainties identified and questions developed in this study (Table 4) have unified the assessment of uncertainty and provide a comprehensive and usable generic means to assess all relevant aspects of a read-across prediction. As such, responses to the questions in Table 4 provide a comprehensive means of assessing a read-across prediction that can be applied by a risk assessor or toxicologist that is favourable to ECHA's RAAF whilst covering the same information. The

questions, and associated flowchart for implementation (Figure 2), allow for a decision on the suitability of a read-across for a particular purpose to be made by ensuring that assessor places the decision to be made at an appropriate level of uncertainty.

INSERT FIGURE 2 HERE

Whilst not an overall goal of the investigation, analysis of the case studies using the questions in Table 4 allowed for a rating, or grading, of the responses. Specifically in the case studies assessed in this investigation, qualitative grading schemes (e.g., low, medium and high) were utilised as proposed by Schultz et al [15] following review of the practice at that time. Assignment of such a grading is inevitably subjective and the allocation of uncertainty (in this case low, moderate or high) can be undertaken with more or fewer classifications, or even a numerical scoring scheme, according to need, context and requirement. However, in this assessment the gradings were interpreted to provide regulatory relevance, e.g., an overall uncertainty of "low" has been stated [12, 13] to indicate that the overall the read-across prediction is equivalent to the information that would be provided from a standardised *in vivo* test (i.e., for these examples this implies OECD TG 408, Subchronic Oral Toxicity: 90-Day Study). This is in accordance with the requirements of legislation such as REACH in the EU. Likewise, assigning an overall uncertainty of "medium" means the read-across prediction is likely to be similar to doing a standardised *in vivo* test. An overall uncertainty of "high" means it is not possible to assess the uncertainty in relation to doing a standardised *in vivo* test.

The assessment of uncertainties using the questions in Table 4 demonstrates the importance of understanding the context of the read-across and how the problem was formulated. The case studies analysed in this investigation required a high level of confidence, associated with low uncertainty, to be acceptable. For other regulatory contexts and decisions, a different level of uncertainty will be acceptable. Further work in this particular area is needed to determine the level of uncertainty, and how that can be determined, for acceptance of read-across different types of regulatory decisions and scenarios. In addition, as well as being applicable to the range of regulatory decisions, the questions are equally applicable to endpoints and effects other than repeated-dose toxicity. It is also important to consider that whilst in this investigation the uncertainties have been anchored on *in vivo* data, this need not be the case and relevant *in vitro*, molecular biology (omics) or other data may be appropriate. Thus, the questions in Table 4 are broadly applicable and adaptable to a whole variety of scenarios.

Table 4. Questions to address in assessing uncertainties of a read-across and summary of responses from the six case studies.

Number	Uncertainty in Read-Across	Questions that Need to be Addressed Regarding Uncertainty	Summary of the Level and Types of Uncertainties Identified in the Case Studies Following Analysis of the Questions (see Supplementary Information Tables S6-S11)
1	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?	Low uncertainty, all case studies related to risk assessment
2	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?	Low uncertainty, whilst varying between case studies, chemicals were clearly identified. Applicability domains were well defined (in terms of chemistry and properties); target and sources compounds fell within the stated domains.
3	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?	Low uncertainty, all case studies had clearly stated hypotheses.

4	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?	Low – medium uncertainty. All case studies were based around a mechanistic hypothesis. However levels of experimental evidence, and hence uncertainty varied. Only CS1 and CS3 had low uncertainty, other case studies had medium uncertainty due to the lack of supporting evidence.
5	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?	Low uncertainty. All case studies reported target and source compounds accurately with extensive information on structure and properties. Dissimilarity in structures was well described and the differences accounted for.
6	Toxicodynamic similarity.	• Is there sufficient and consistent toxicodynamic information provided to establish similarity in the hazard of the derivatives used in the read-across?	Low uncertainty. Within the context of mechanistic plausibility, all case studies reported toxicodynamic properties sufficiently well to establish similarity in hazard.

7	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?	Low to medium uncertainty. CS2, CS3 and CS5 all had medium levels of uncertainty due to the lack of toxicokinetic / ADME data to demonstrate similarity between the target and source compounds.
8	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?	Low uncertainty. All cases studies were associated with high quality data.
9	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?	Low to high uncertainty. There was generally low uncertainty with regard to effects and potency data. However, no case studies discussed the dose-response relationships. CS2 and CS5 had medium uncertainty due to differences in length of the tests. CS4 and CS6 showed differences in potency, effects and hazard and were assigned medium potency as a result.
10	Strength or robustness of the supporting data sets.	• How extensively are the relevant or key events either empirically measured and/or modelled by appropriate *in silico*, *in chemico* and *in vitro* data?	Low uncertainty. All case studies were well supported by other data, including those from New Approach Methodologies (NAMs).

		• Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the supporting methods adequately reported?	
11	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)?	Low uncertainty. All case studies showed a strong weight of evidence, combining multiple lines of evidence that supported the read-across hypothesis.
12	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)?	Low uncertainty. All case studies were thoroughly documented in terms of argumentation and the presentation of data.

Table 5. Mapping of information relevant to identified uncertainties to existing read-across templates and schemes.

Number	Uncertainty in Read-Across	Questions that Need to be Addressed Regarding Uncertainty	Wu et al [10] Summarised in Table S1	Blackburn and Stuard [12] Summarised in Table S2	Schultz et al [13] Summarised in Table S3	ECHA RAAF [15] Assessment Elements Summarised in Table S4	(Draft) OECD Template for Reporting Uncertainty [6, 14] Summarised in Table S5
1	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?	No	No	Yes	Yes	No
		Is the acceptable level or degree of uncertainty for the stated purpose defined?	No	No	Yes	Yes	No
		Is the stated acceptable level or degree of uncertainty appropriate for the stated regulatory purpose?	No	No	Yes	Yes	No
2	Type of category / group including the definition of the applicability domain.	Is the read-across approach (e.g., analogue or category) clearly reported?	No	Yes	Yes	Yes	Yes
		Are the target and source chemicals clearly identified?	Yes	Yes	Yes	Yes	Yes
		Is the applicability domain of the analogue or category defined?	No	No	Yes	Yes	No
		Do target and source chemicals fit within the defined applicability domain?	No	No	Yes	Yes	No

3	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?	No	No	Yes	Yes	Yes
4	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?	No	No	Yes	Yes	Yes
		• Is there sufficient experimental information provided to support the proposed chemical and toxicological mechanisms?	No	No	Yes	Yes	Yes
		• How extensively does the experimental information provided support the mechanistic plausibility and / or the AOP or MoA on which the read-across is based?	No	No	Yes	Yes	Yes
5	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?	Yes	Yes	Yes	Yes	Yes
		• Are the dissimilarities in chemical structure reported and are they toxicologically relevant?	Yes	Yes	Yes	Yes	Yes
		• Are the relevant molecular and physico-chemical properties (e.g., for molecular size, hydrophobicity, solubility,	Yes	Yes	Yes	Yes	Yes

		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?	Yes	Yes	Yes	Yes	Yes
6	Toxicodynamic similarity.	• Is there sufficient and consistent toxicodynamic information provided to establish similarity in the hazard of the derivatives used in the read-across?	No	No	Yes	Yes	No
7	Toxicokinetic similarity.	• Is there sufficient ADME information provided to establish toxicokinetic similarity for the derivatives used in the read-across?	No	No	Yes	Yes	Yes
		• Are any dissimilarities in ADME properties (and, as appropriate, metabolism / degradation) toxicologically relevant?	No	No	Yes	Yes	Yes
8	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?	No	Yes	Yes	Yes	Yes
		• 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?	No	Yes	Yes	Yes	Yes

9	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?	No	Yes	Yes	Yes	Yes
		• Is the potency of the hazard reported and is it consistent among the source chemicals?	No	Yes	Yes	Yes	Yes
		• What are the temporal relationships between relevant endpoints?	No	Yes	Yes	Yes	Yes
		• What are the dose-response relationships between relevant endpoints?	No	Yes	Yes	Yes	Yes
10	Strength or robustness of the supporting data sets.	• How extensively are the relevant or key events either empirically measured and/or modelled by appropriate *in silico*, *in chemico* and *in vitro* data?	No	No	Yes	Yes	Yes
		• Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the supporting methods adequately reported?	No	No	Yes	Yes	Yes
11	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?	No	Yes	Yes	Yes	Yes
		• How many and how large are the dissimilarities in the supporting information (i.e., data gaps)?	No	Yes	Yes	Yes	Yes

12	Documentation and written evidence provided	• Is the read-across prediction adequately documented?	No	No	No	No	No
		• Does the evidence support the hypothesis that the uncertainty is acceptable for the stated purpose (as per Question 1)?	No	No	Yes	Yes	No

Conclusions

Read-across has become an essential tool to fill toxicity data gaps, in particular supporting a number of regulatory applications including, but not limited to, (quantitative) risk assessment, classification and labelling as well as screening and prioritisation. The acceptance of a read-across prediction for regulatory purposes is a complex and occasionally ambiguous process. However, it is acknowledged that for a read-across to be accepted for a specific regulatory purpose uncertainties must be identified and be appropriate. As such, uncertainties for read-across must be established and ideally quantified. To this end, a number of schemes have been published ranging from assessment of individual aspects of describing the read-across [10, 12] through to more comprehensive assessments with a particular focus on enabling regulatory acceptance [6, 13, 15]. This study attempted to harmonise all the current knowledge of uncertainties by focussing on the uncertainties that could be identified in case studies for one of the most challenging toxicological endpoints (repeated dose toxicity).

Based on a review of existing knowledge and how it was applied to the information presented in six case studies, a variety of uncertainties were identified that potentially impact on the regulatory acceptance of read-across predictions. Four main sources of uncertainty were identified including that associated with the regulatory use of the read-across prediction; the quality and relevance of the apical endpoint data being read across; the read-across argumentation; and the similarity justification. A total of twelve types of uncertainty were defined that cross over and describe these four sources fully. In order to enable use of this knowledge and to facilitate determination and evaluation of uncertainties, a series of 30 questions was formulated that guide the read-across developer and assessor through the assessment of all relevant uncertainties. The six case studies were assessed in terms of the questions developed demonstrating that areas such as mechanistic plausibility, the quality of the data being read across and inclusion of information on toxicokinetics were important uncertainties to reduce. The questions cover and extend all the uncertainties documented in existing schemes and place them into a rapidly applied generic framework. This findings of this study, whilst focussed on read-across for repeated dose toxicity, have the possibility to be applied to other endpoints, effectsand read-across scenarios.

Acknowledgment

This manuscript benefitted considerably from the input, advice and wisdom of Dr Grace Patlewicz (United States Environmental Protection Agency) and the anonymous reviewers.

References

[1] Organisation for Economic Cooperation and Development (OECD), Guidance on Grouping of Chemicals, Second Edition, No. 194, Series on Testing & Assessment. ENV/JM/MONO(2014)4, OECD, Paris, 2014.

[2] G. Patlewicz, N. Ball, P. Boogaard, R.A. Becker, B. Hubesch, Building scientific confidence in the development and evaluation of read-across, Regul. Toxicol. Pharmacol. 72 (2015) 117-133.

[3] N. Ball, M.T.D. Cronin, J. Shen, K. Blackburn, E.D. Booth, M. Bouhifd, E. Donley, L. Egnash, C. Hastings, D.R. Juberg, A. Kleensang, N. Kleinstreuer, E.D. Kroese, A.C. Lee, T. Luechtefeld, A. Maertens, S. Marty, J.M. Naciff, J. Palmer, D. Pamies, M. Penman, A.-N. Richarz, D.P., Russo, S.B. Stuard, G. Patlewicz, B. van Ravenzwaay, S. Wu, H. Zhu, T. Hartung, Toward Good Read-Across Practice (GRAP) guidance. ALTEX 33 (2016) 149-166.

[4] Organisation for Economic Cooperation and Development (OECD), 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, No. 250, Series on Testing & Assessment. ENV/JM/MONO(2016)48, OECD, Paris. 2016.

[5] Organisation for Economic Cooperation and Development (OECD), Report on Considerations from Case Studies on Integrated Approaches for Testing and Assessment (IATA), Second Review Cycle (2016), Case Studies on Grouping Methods as a Part of IATA, No. 270, Series on Testing & Assessment. ENV/JM/MONO(2017)22, OECD, Paris, 2017.

[6] Organisation for Economic Cooperation and Development (OECD), Report on Considerations from Case Studies on Integrated Approaches for Testing and Assessment (IATA), Third Review Cycle (2017), Case Studies on Grouping Methods as a Part of IATA, No. XXX, Series on Testing & Assessment. ENV/JM/MONO(2018)XX, OECD, Paris. In press, 2018.

[7] G. Patlewicz, G. Helman, P. Pradeep, I. Shah, Navigating through the minefield of read-across tools. A review of in silico tools for grouping, Comput. Toxicol. 3 (2017) 1-18.

[8] G. Patlewicz M.T.D. Cronin G. Helman J.C. Lambert, L.E. Lizarraga I. Shah, Navigating through the minefield of read-across frameworks: A commentary perspective. Comput. Toxicol. 6 (2018) 39-54.

[9] T.W. Schultz, M.T.D. Cronin, Lessons learned from read-across case studies for repeated-dose toxicity, Regul. Toxicol. Pharmacol. 88 (2017) 185-191.

[10] S. Wu, K. Blackburn, J. Amburgey, J. Jaworska, T. Federle, 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 (2010) 67-81.

[11] N.C. Wang, Q. Jay Zhao, S.C. Wesselkamper, J.C. Lambert, D. Petersen, J.K. Hess-Wilson, Application of computational toxicological approaches in human health risk assessment. I. A tiered surrogate approach, Regul. Toxicol. Pharmacol. 63 (2012) 10-19.

[12] K. Blackburn, S.B. Stuard, A framework to facilitate consistent characterization of read across uncertainty, Regul. Toxicol. Pharmacol. 68 (2014) 353-362.

[13] T.W. Schultz, P. Amcoff, E. Berggren, F. Gautier, M, Klaric, D.J. Knight, C. Mahony, M. Schwarz, A. White, M.T.D. Cronin, A strategy for structuring and reporting a read-across prediction of toxicity, Regul. Toxicol. Pharmacol. 72 (2015) 586-601.

[14] Organisation for Economic Cooperation and Development (OECD), Case Study on the Use of an Integrated Approach to Testing and Assessment for Hepatotoxicity of Allyl Esters, No. 253, Series on Testing & Assessment. ENV/JM/MONO(2016)51, OECD, Paris, 2016.

[15] European Chemicals Agency (ECHA) Read-Across Assessment Framework (RAAF). European Chemicals Agency, Helsinki, 2017.

[16] European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC), Technical Report 116 Category Approaches, Read-Across, (Q)SAR. ECETOC, Brussels, 2012.

[17] Organisation for Economic Cooperation and Development (OECD), Organisation for Economic Cooperation and Development (OECD), 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, No. 273, Series on Testing & Assessment. ENV/JM/MONO(2017)25, OECD, Paris, 2017.

[18] Organisation for Economic Cooperation and Development (OECD), 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, No. 274, Series on Testing & Assessment. ENV/JM/MONO(2017)26, OECD, Paris, 2017.

[19] Organisation for Economic Cooperation and Development (OECD), 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, No. XXX, Series on Testing & Assessment. ENV/JM/MONO(2018)XX, OECD, Paris. In press, 2018.

[20] K.R. Przybylak, T.W. Schultz, A.-N. Richarz, C.L. Mellor, S.E. Escher, M.T.D Cronin, Read-across of 90-day rat oral repeated-dose toxicity: A case study for selected β-olefinic alcohols. Comput. Toxicol. 1 (2017) 22-32.

[21] C.L. Mellor, T.W. Schultz, K.R. Przybylak, A.-N. Richarz, M.T.D Cronin, Read-across for rat oral gavage repeated-dose toxicity for short-chain mono-alkylphenols: A case study. Comput. Toxicol 2 (2017) 1-11.

[22] P. Pradeep, K. Mansouri, G. Patlewicz, R. Judson, A systematic evaluation of analogs and automated read-across prediction of estrogenicity: A case study using hindered phenols. Comput. Toxicol 4 (2017) 22-30.

[23] G. Helman, I. Shah, G. Patlewicz, Extending the Generalised Read-Across approach (GenRA): A systematic analysis of the impact of physicochemical property information on read-across performance, Comput. Toxicol 11(2018) in press.

Figure Titles.

Figure 1. Schematic depiction of the interrelationships between the sources and types on uncertainty in a read-across. Solid boxes with Roman numbers (I – IV) represent the main sources as identified in Table 2. Arabic numbers (1 – 12) represent the types of uncertainties as described in Table 3.

Figure 2. Workflow for applying the questions related to uncertainty in Table 4. The numbers in parenthesis relate to the questions in Table 4. It is intended that Question 1 should inform Question 12, i.e. the context and problem formulation will inform the level of uncertainty acceptable for the purpose of the read-across and whether that is achieved.

Figure 1.

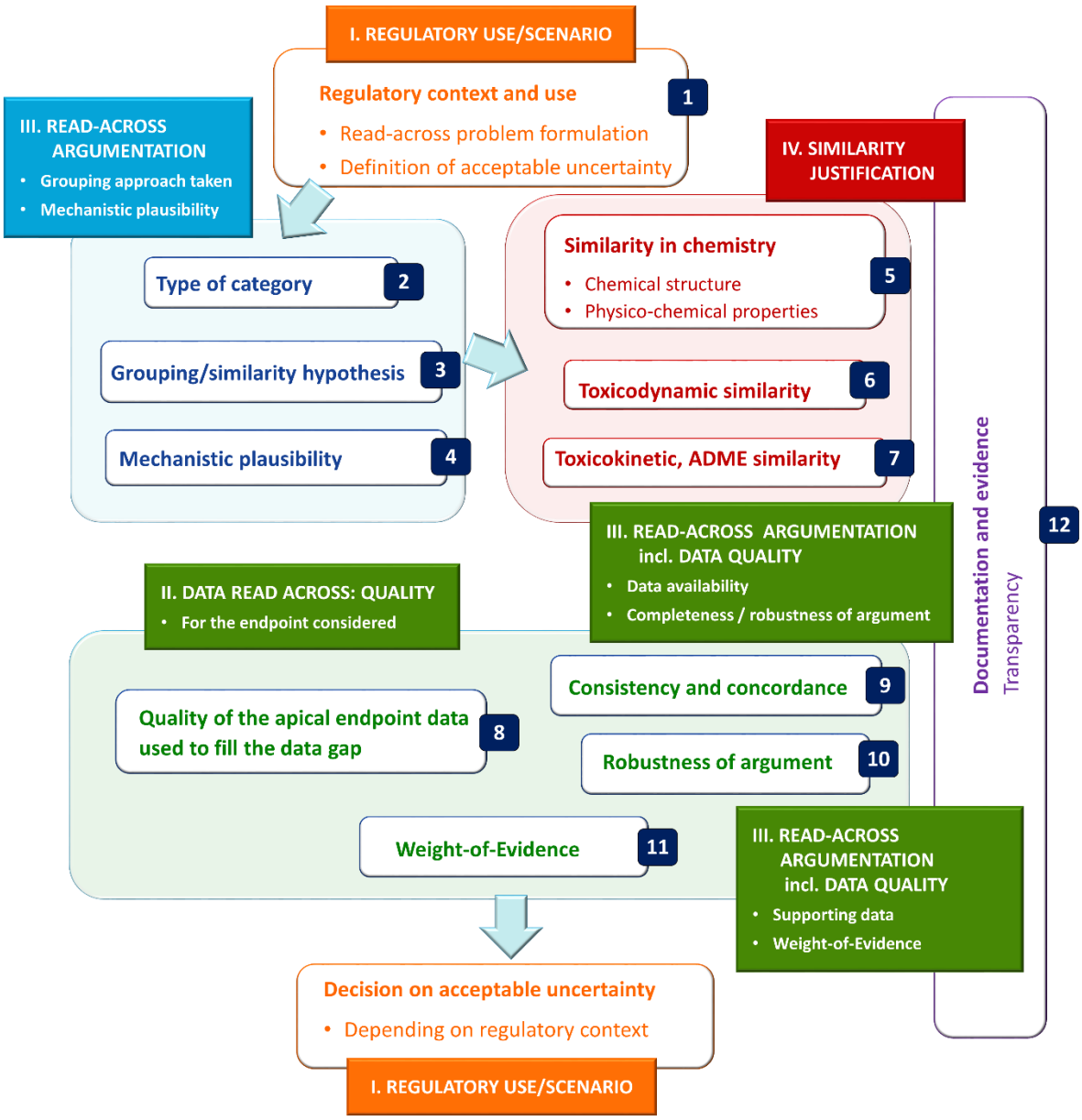
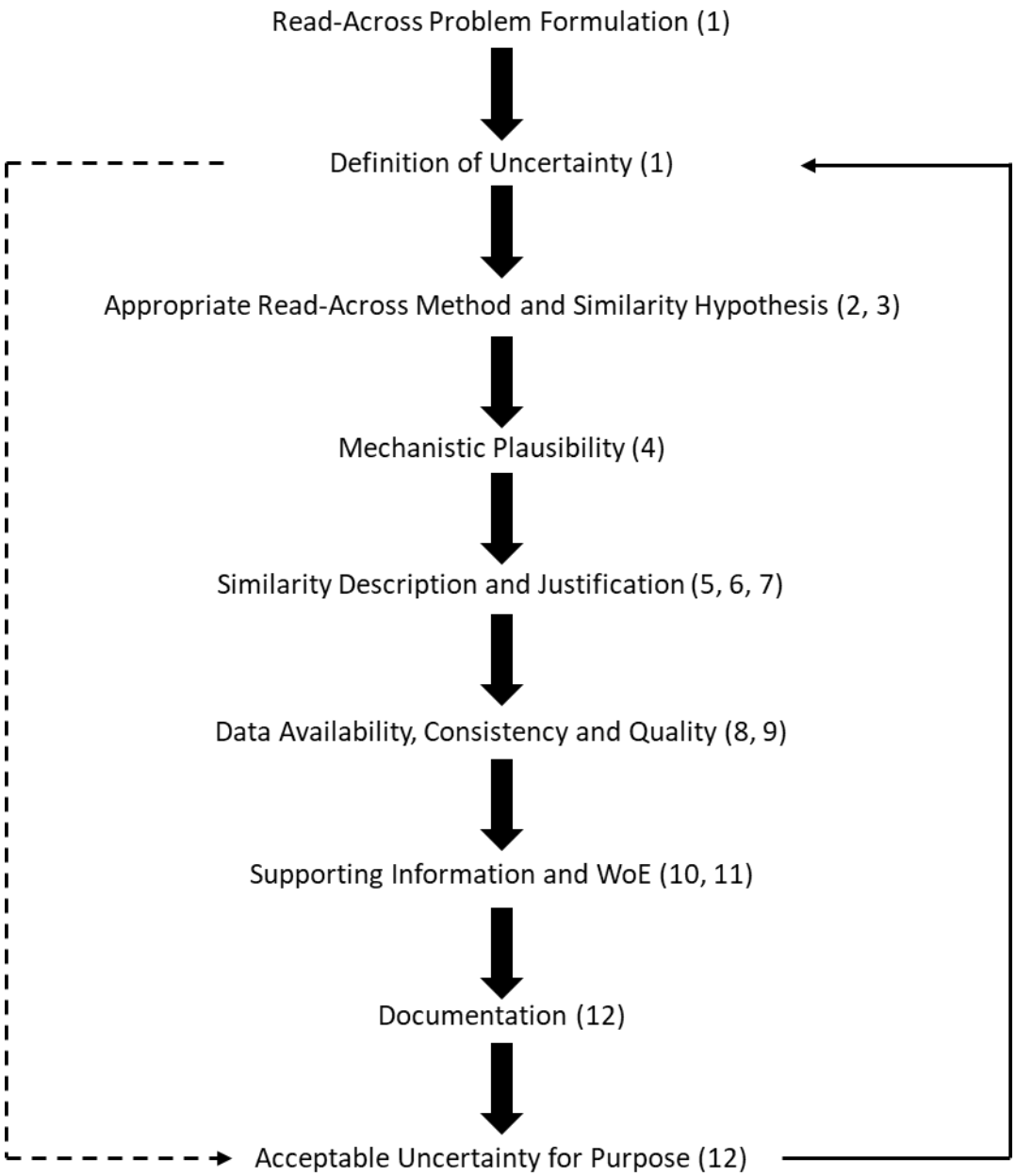

I. REGULATORY USE/SCENARIO
Regulatory context and use
1
• Read-across problem formulation
• Definition of acceptable uncertainty
III. READ-ACROSS ARGUMENTATION
• Grouping approach taken
• Mechanistic plausibility
IV. SIMILARITY JUSTIFICATION
Type of category
2
Grouping/similarity hypothesis
3
Mechanistic plausibility
4
Similarity in chemistry
5
• Chemical structure
• Physico-chemical properties
Toxicodynamic similarity
6
Toxicokinetic, ADME similarity
7
III. READ-ACROSS ARGUMENTATION incl. DATA QUALITY
• Data availability
• Completeness / robustness of argument
II. DATA READ ACROSS: QUALITY
• For the endpoint considered
Quality of the apical endpoint data used to fill the data gap
8
Consistency and concordance
9
Robustness of argument
10
Weight-of-Evidence
11
III. READ-ACROSS ARGUMENTATION incl. DATA QUALITY
• Supporting data
• Weight-of-Evidence
Documentation and evidence
Transparency
12
Decision on acceptable uncertainty
• Depending on regulatory context
I. REGULATORY USE/SCENARIO

Figure 2.


Read-Across Problem Formulation (1)
Definition of Uncertainty (1)
Appropriate Read-Across Method and Similarity Hypothesis (2, 3)
Mechanistic Plausibility (4)
Similarity Description and Justification (5, 6, 7)
Data Availability, Consistency and Quality (8, 9)
Supporting Information and WoE (10, 11)
Documentation (12)
Acceptable Uncertainty for Purpose (12)

Supplementary Information

Assessing Uncertainty in Read-Across: Questions to Evaluate Toxicity Predictions Based on Knowledge Gained from Case Studies

Terry W. Schultz[1], Andrea-Nicole Richarz[2], and Mark T.D. Cronin[3*]

[1]The University of Tennessee, College of Veterinary Medicine, 2407 River Drive, Knoxville, TN 37996-4543 USA

[2]Directorate for Health, Consumers and Reference Materials, Joint Research Centre, European Commission, Ispra, Italy;

[3]Liverpool John Moores University, School of Pharmacy and Biomolecular Sciences, Byrom Street, L3 3AF Liverpool, England

Table S1. Criteria related to uncertainty of a read-across based on the discussion, decision tree and templates described by Wu et al. [10]*

Evaluation Criteria	Features Evaluated
Structure and reactivity	Commonality of structural alerts
	Commonality of key functional groups
	Commonality in position of double bonds
	Effects of additional functional groups
Physicochemical properties	Commonality of properties including molecular weight, distribution coefficient (log D) (combining the ionisation constant (pKa) and partition coefficient (log P)) and aqueous solubility
Metabolism	Commonality of metabolic pathways
	Potential for the analogue to metabolise to the target or a highly related compound as well as that for the target to metabolise to the analogue
	Potential for the metabolism of the analogue and the target to converge on a common stable metabolite or reactive metabolite with the same mode of action
	Potential for the metabolism of the analogue and the target to diverge on different bioactivation pathway to alter the toxicological profile of the analogue or target.

***** S. Wu, K. Blackburn, J. Amburgey, J. Jaworska, T. Federle, 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 (2010) 67-81.

Table S2. Criteria related to uncertainty of a read-across based on the templates described by Blackburn and Stuard [12]*

Analogue Data Set Characteristics [f]
Number of analogues contributing data
Robustness of analogue data set
Concordance of effect(s)
Concordance of potency
Severity of critical effects

* K. Blackburn, S.B. Stuard, A framework to facilitate consistent characterization of read across uncertainty, Regul. Toxicol. Pharmacol. 68 (2014) 353-362

Table S3. Criteria related to uncertainty of a read-across based on the templates described by Schultz et al [13]*

Part 1: Data Uncertainty and Weight-of-Evidence Associated with the Fundamentals of Chemical, Transformation/Toxicokinetic and Toxicological Similarity.

Similarity Parameter
Substance identification, structure and chemical classifications
Physio-chemical and molecular properties
Substituents, functional groups and extended structural fragments
Transformation/toxicokinetics and metabolic similarity
Potential metabolic products
Toxicophores /mechanistic lerts
Mechanistic plausibility and AOP-related events
Other relevant, *in vivo*, *in vitro* and *ex vivo* endpoints

Part 2: Template for Assessing Uncertainty Associated with Mechanistic Relevance and Completeness of the Read-Across.

Uncertainty[e]
The problem and premise of the read-across
Number of analogues in the source set
Quality of the *in vivo* apical endpoint data read across
Severity of the apical *in vivo* hazard
Robustness of analogue data set
Concordance with regard to the intermediate and apical effects and potency data
Weight of Evidence

* T.W. Schultz, P. Amcoff, E. Berggren, F. Gautier, M, Klaric, D.J. Knight, C. Mahony, M. Schwarz, A. White, M.T.D. Cronin, A strategy for structuring and reporting a read-across prediction of toxicity, Regul. Toxicol. Pharmacol. 72 (2015) 586-601.

Table S4. Criteria related to the confidence (which are analogous to the uncertainties in this investigation) that may be assigned to a read-across based on ECHA's RAAF Assessment Elements [15]*

Assessment Element (AE) Number(s)	Assessment Element
Scientific Assessment of Human Health Effects	
Analogue Approach, Scenarios 1-2	
AE A.1	Identity and characterisation of the source substance
AE A.2	Link of structural similarities and differences with the proposed prediction
AE A.3	Reliability and adequacy of the source study
AE A.4	Bias that influences the prediction
AE 1.1	Formation of common (identical) compound(s)
AE 1.2	The biological targets for the common compound(s)
AE 1.3	Exposure of the biological target(s) to the common compound(s)
AE 1.4	The impact of parent compounds
AE 1.5	Formation and impact of non-common compounds
AE 2.1	Compounds the test organism is exposed to
AE 2.2	Common underlying mechanism, qualitative aspects
AE 2.3	Common underlying mechanism, quantitative aspects
AE 2.4	Exposure to other compounds than to those linked to the prediction
AE 2.5	Occurrence of other effects than covered by the hypothesis and justification
Category Approach, Scenarios 3-6	
AE C.1	Substance characterisation
AE C.2	Structural similarity and differences within the category
AE C.3	Link of structural similarities and structural differences with the proposed regular pattern
AE C.4	Consistency of effects in the data matrix
AE C.5	Reliability and adequacy of the source study(ies)
AE C.6	Bias that influences the prediction
AE 3.1 AE 5.1	Formation of common (identical) compound(s)
AE 3.2 AE 5.2	The biological target(s) for the common compound(s)
AE 3.3 AE 5.3	Exposure of the biological target(s) to the common compound(s)
AE 3.4 AE 5.4	The impact of parent compounds
AE 3.5 AE 5.5	Formation and impact of non-common compounds
AE 4.1 AE 6.1	Compounds the test organism is exposed to
AE 4.2 AE 6.2	Common underlying mechanism, qualitative aspects
AE 4.3 AE 6.3	Common underlying mechanism, quantitative aspects
AE 4.4 AE 6.4	Exposure to other compounds than those linked to the prediction
AE 4.5 AE 6.5	Occurrence of other effects than covered by the hypothesis and

	justification
Scientific Assessment of Environmental Fate and Effects	
Analogue Approach, Scenarios 1-2	
AE A.1	Characterisation of source and target substances
AE A.2	Link of structural similarities and structural differences with the proposed prediction (presence of hypothesis)
AE A.3	Impact of impurities on the prediction
AE A.4	Consistency of properties in the data matrix
AE A.5	Reliability and adequacy of the source data AE
AE A.6	Bias that influences the prediction
AE 1.1	Formation of common (identical) and non-common compounds
AE1.2	Degradation of non-common compounds
AE 1.3	Bioaccumulation potential of non-common compounds
AE 1.4	Impact of non-common compound
AE 2.1	Degradation
AE 2.2	Bioaccumulation potential
AE 2.3	Common underlying mechanism, qualitative aspects
AE 2.4	Common underlying mechanism, quantitative aspect
Category Approach, Scenarios 3-6	
AE C.1	Characterisation of source and target substances
AE C.2	Structural similarity and dissimilarity within the category (category description)
AE C.3	Link of structural similarities and structural differences with the proposed regular pattern (presence of hypothesis)
AE C.4	Impact of impurities on the prediction
AE C.5	Consistency of properties in the data matrix
AE C.6	Reliability and adequacy of the source data
AE C.7	Bias that influences the prediction
AE 3.1 AE 5.1	Formation of common (identical) and non-common compound(s)
AE 3.2 AE 5.2	Degradation of non-common compounds
AE 3.3 AE 5.3	Bioaccumulation potential of non-common compounds
AE 3.4 AE 5.4	Impact of non-common compounds
AE 4.1 AE 6.1	Degradation
AE 4.2 AE 6.2	Bioaccumulation potential
AE 4.3 AE 6.3	Common underlying mechanism, qualitative aspects
AE 4.4 AE 6.4	Common underlying mechanism, quantitative aspects

* European Chemicals Agency (ECHA) Read-Across Assessment Framework (RAAF). European Chemicals Agency, Helskinki, 2017.

Table S5. Criteria related to uncertainty of a read-across extracted from the OECD [6, 14]* template for of reporting uncertainty.

Uncertainty
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

* Organisation for Economic Cooperation and Development (OECD), Case Study on the Use of an Integrated Approach to Testing and Assessment for Hepatotoxicity of Allyl Esters, No. 253, Series on Testing & Assessment. ENV/JM/MONO(2016)51, OECD, Paris, 2016.

Organisation for Economic Cooperation and Development (OECD), Report on Considerations from Case Studies on Integrated Approaches for Testing and Assessment (IATA), Third Review Cycle (2017), Case Studies on Grouping Methods as a Part of IATA, No. XXX, Series on Testing & Assessment. ENV/JM/MONO(2018)XX, OECD, Paris. In press, 2018.

Table S6. Uncertainties relating to Case Study 1 for n-alkanols (OECD [17]*) – for illustrative purposes only.

Number	Uncertainty in Read-Across	Questions that Need to be Addressed Regarding Uncertainty	Brief Response	Associated Uncertainty	Source of Evidence or Supporting Information
1	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?	Yes	Low	Section 1 (page 13)
		• Is the acceptable level or degree of uncertainty for the stated purpose defined?	Yes	Low	Section 1.1 (page 13)
		• Is the stated acceptable level or degree of uncertainty appropriate for the stated regulatory purpose?	Yes	Low	Section 1.1 (page 13)
2	Type of category / group including the definition of the applicability domain.	• Is the read-across approach (e.g., analogue or category) clearly reported?	Yes	Low	Section 1.2 (page 13); Section 2 (page 14)
		• Are the target and source chemicals clearly identified?	Yes	Low	Table 3 (page 13)
		• Is the applicability domain of the analogue or category defined?	Yes	Low	Section 2.2 (page 15)
		• Do target and source chemicals fit within the defined applicability domain?	Yes	Low	Section 2.2 (page 15)
3	The premise or hypothesis of the RA.	• Is the hypothesis on which the RA is based clearly stated and presented in sufficient detail to be assessed?	Yes	Low	Section 2 (page 14)
4	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?	Very clearly stated	Low	Section 3 (pages 16-17); Table 5 (pages 21-22)
		• Is there sufficient experimental information provided to support the proposed chemical and toxicological mechanisms?	Yes	Low	Section 3 (pages 16-17); Table 5 (pages 21-22)
		• How extensively does the experimental information provided support the mechanistic plausibility and / or the AOP or MoA on which the RA is based?	Very extensively stated	Low	Section 3 (pages 16-17); Table 5 (pages 21-22)

5	Similarity in chemistry.	• Are the chemical structures (i.e., 2D structure, isomers, SMILES and molecular formula) reported for the derivatives used in the RA?	Yes	Low	Table 3 (page 13); Table 4 (pages 19-20)
		• Are the dissimilarities in chemical structure reported and are they toxicologically relevant?	Yes, relevant only to toxicokinetics	Low	Table 4 (pages 19-20)
		• 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 RA?	Yes	Low	Table 4 (pages 19-20)
		• Are the dissimilarities in molecular and physico-chemical properties reported and are they toxicologically (or pharmacokinetically) relevant?	Yes, relevant only to toxicokinetics	Low	Table 4 (pages 19-20)
6	Toxicodynamic similarity.	• Is there sufficient and consistent toxicodynamic information provided to establish similarity in the hazard of the derivatives used in the RA?	Yes	Low	Section 3 (pages 15-18); Table 5 (pages 21-22)
7	Toxicokinetic similarity.	• Is there sufficient ADME information provided to establish toxicokinetic similarity for the derivatives used in the RA?	Yes	Low	Section 3 (pages 15-18); Table 4 (pages 20-21)
		• Are any dissimilarities in ADME properties (and, as appropriate, metabolism / degradation) toxicologically relevant?	No – or as stated	Low	Section 3 (pages 15-18); Table 4 (pages 20-21)
8	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?	Yes	Low	Introduction (pages 8-12); Table 5 (pages 21-22)
		• 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?	Yes	Low	Introduction (pages 8-12); Table 5 (pages 21-22)

9	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?	Yes	Low	Introduction (pages 8-12); Table 5 (pages 21-22)
		• Is the potency of the hazard reported and is it consistent among the source chemicals?	Yes	Low	Introduction (pages 8-12); Table 5 (pages 21-22)
		• What are the temporal relationships between relevant endpoints?	All 90 day	Low	Introduction (pages 8-12); Table 5 (pages 21-22)
		• What are the dose-response relationships between relevant endpoints?	Not reported	High	
10	Strength or robustness of the supporting data sets.	• How extensively are the relevant or key events either empirically measured and/or modelled by appropriate *in silico*, *in chemico* and *in vitro* data?	Well reported	Low	Introduction (pages 8-12); Table 5 (pages 21-22)
		• Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the supporting methods adequately reported?	Yes	Low	Introduction (pages 8-12); Table 5 (pages 21-22)
11	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?	Yes	Low	Introduction (pages 8-12); Table 5 (pages 21-22)
		• How many and how large are the dissimilarities in the supporting information (i.e., data gaps)?	Small level of dissimilarity	Low	Introduction (pages 8-12); Table 5 (pages 21-22)
12	Documentation and written evidence provided	• Is the RA prediction adequately documented?	Yes	Low	Complete document
		• Does the evidence support the hypothesis that the uncertainty is acceptable for the stated purpose (as per Question 1)?	Yes	Low	Section 5 (pages 22-23)

* Organisation for Economic Cooperation and Development (OECD), Organisation for Economic Cooperation and Development (OECD), 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, No. 273, Series on Testing & Assessment. ENV/JM/MONO(2017)25, OECD, Paris, 2017.

Table S7. Uncertainties relating to Case Study 2 for 2-alkyl-1-alkanols (OECD [18]*) – for illustrative purposes only.

Number	Uncertainty in Read-Across	Questions that Need to be Addressed Regarding Uncertainty	Brief Response	Associated Uncertainty	Source of Evidence or Supporting Information
1	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?	Yes	Low	Section 1 (page 12)
		• Is the acceptable level or degree of uncertainty for the stated purpose defined?	Yes	Low	Section 1.1 (page 12)
		• Is the stated acceptable level or degree of uncertainty appropriate for the stated regulatory purpose?	Yes	Low	Section 1.1 (page 12)
2	Type of category / group including the definition of the applicability domain.	• Is the read-across approach (e.g., analogue or category) clearly reported?	Yes	Low	Section 1.2 (page 12); Section 2 (pages 13-14)
		• Are the target and source chemicals clearly identified?	Yes	Low	Table 3 (page 12)
		• Is the applicability domain of the analogue or category defined?	Yes	Low	Section 2.2 (page 15)
		• Do target and source chemicals fit within the defined applicability domain?	Yes	Low	Section 2.2 (page 15)
3	The premise or hypothesis of the RA.	• Is the hypothesis on which the RA is based clearly stated and presented in sufficient detail to be assessed?	Yes	Low	Section 2 (page 13-14)
4	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?	Very clearly stated	Medium	Section 3 (pages 15-21); Table 5 (pages 25-26)
		• Is there sufficient experimental information provided to support the proposed chemical and toxicological mechanisms?	Yes	Medium	Section 3 (pages 15-21); Table 5 (pages 25-26)

		• How extensively does the experimental information provided support the mechanistic plausibility and / or the AOP or MoA on which the RA is based?	Very extensively stated	Medium	Section 3 (pages 15-21); Table 5 (pages 25-26)
5	Similarity in chemistry.	• Are the chemical structures (i.e., 2D structure, isomers, SMILES and molecular formula) reported for the derivatives used in the RA?	Yes	Low	Table 3 (page 12); Table 4 (pages 23-24)
		• Are the dissimilarities in chemical structure reported and are they toxicologically relevant?	Yes, relevant only to toxicokinetics	Low	Table 4 (pages 23-24)
		• 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 RA?	Yes	Low	Table 4 (pages 23-24)
		• Are the dissimilarities in molecular and physico-chemical properties reported and are they toxicologically (or pharmacokinetically) relevant?	Yes, relevant only to toxicokinetics	Low	Table 4 (pages 23-24)
6	Toxicodynamic similarity.	• Is there sufficient and consistent toxicodynamic information provided to establish similarity in the hazard of the derivatives used in the RA?	Yes	Low	Section 3 (pages 15-21); Table 5 (pages 25-26)
7	Toxicokinetic similarity.	• Is there sufficient ADME information provided to establish toxicokinetic similarity for the derivatives used in the RA?	Yes	Medium	Section 3 (pages 15-18); Table 4 (pages 23-24)
		• Are any dissimilarities in ADME properties (and, as appropriate, metabolism / degradation) toxicologically relevant?	No – or as stated	Medium	Section 3 (pages 15-18); Table 4 (pages 23-24)
8	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?	Yes	Low	Introduction (pages 9-11); Table 5 (pages 25-26)
		• 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?	Yes	Low	Introduction (pages 9-11); Table 5 (pages 25-26)

9	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?	Yes	Low	Introduction (pages 9-11); Table 5 (pages 25-26)
		• Is the potency of the hazard reported and is it consistent among the source chemicals?	Yes	Low	Introduction (pages 9-11); Table 5 (pages 25-26)
		• What are the temporal relationships between relevant endpoints?	All 90 day	Low to Medium	Introduction (pages 9-11); Table 5 (pages 25-26)
		• What are the dose-response relationships between relevant endpoints?	Not reported	High	
10	Strength or robustness of the supporting data sets.	• How extensively are the relevant or key events either empirically measured and/or modelled by appropriate *in silico*, *in chemico* and *in vitro* data?	Well reported	Low	Introduction (pages 9-11); Table 5 (pages 25-26)
		• Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the supporting methods adequately reported?	Yes	Low	Introduction (pages 9-11); Table 5 (pages 25-26)
11	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?	Yes	Low	Introduction (pages 9-11); Table 5 (pages 25-26)
		• How many and how large are the dissimilarities in the supporting information (i.e., data gaps)?	Small level of dissimilarity	Low	Introduction (pages 9-11); Table 5 (pages 25-26)
12	Documentation and written evidence provided	• Is the RA prediction adequately documented?	Yes	Low	Complete document
		• Does the evidence support the hypothesis that the uncertainty is acceptable for the stated purpose (as per Question 1)?	Yes	Low	Section 5 (page 26)

*Organisation for Economic Cooperation and Development (OECD), 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, No. 274, Series on Testing & Assessment. ENV/JM/MONO(2017)26, OECD, Paris, 2017.

Table S8. Uncertainties relating to Case Study 3 for aryl alcohol alkyl carboxylic esters (OECD [19]*) – for illustrative purposes only.

Number	Uncertainty in Read-Across	Questions that Need to be Addressed Regarding Uncertainty	Brief Response	Associated Uncertainty	Source of Evidence or Supporting Information
1	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?	Yes	Low	**NB Examples / page numbers not finalised as document is under review*
		• Is the acceptable level or degree of uncertainty for the stated purpose defined?	Yes	Low	
		• Is the stated acceptable level or degree of uncertainty appropriate for the stated regulatory purpose?			
2	Type of category / group including the definition of the applicability domain.	• Is the read-across approach (e.g., analogue or category) clearly reported?	Yes	Low	
		• Are the target and source chemicals clearly identified?	Yes	Low	
		• Is the applicability domain of the analogue or category defined?			
		• Do target and source chemicals fit within the defined applicability domain?			
3	The premise or hypothesis of the RA.	• Is the hypothesis on which the RA is based clearly stated and presented in sufficient detail to be assessed?	Yes	Low	
4	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?	Very clearly stated	Low	
		• Is there sufficient experimental information provided to support the proposed chemical and toxicological mechanisms?	Yes	Low	

		• How extensively does the experimental information provided support the mechanistic plausibility and / or the AOP or MoA on which the RA is based?	Very extensively stated	Low	
5	Similarity in chemistry.	• Are the chemical structures (i.e., 2D structure, isomers, SMILES and molecular formula) reported for the derivatives used in the RA?	Yes	Low	
		• Are the dissimilarities in chemical structure reported and are they toxicologically relevant?	Yes, relevant only to toxicokinetics	Low	
		• 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 RA?	Yes	Low	
		• Are the dissimilarities in molecular and physico-chemical properties reported and are they toxicologically (or pharmacokinetically) relevant?	Yes, relevant only to toxicokinetics	Low	
6	Toxicodynamic similarity.	• Is there sufficient and consistent toxicodynamic information provided to establish similarity in the hazard of the derivatives used in the RA?	Yes	Low	
7	Toxicokinetic similarity.	• Is there sufficient ADME information provided to establish toxicokinetic similarity for the derivatives used in the RA?	Yes	Low to Medium	
		• Are any dissimilarities in ADME properties (and, as appropriate, metabolism / degradation) toxicologically relevant?	No – or as stated	Low to Medium	
8	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?	Yes	Low	
		• 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?	Yes	Low	

9	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?	Yes	Low	
		• Is the potency of the hazard reported and is it consistent among the source chemicals?	Yes	Low	
		• What are the temporal relationships between relevant endpoints?	All 90 day	Low to Medium	
		• What are the dose-response relationships between relevant endpoints?	Not reported	High	
10	Strength or robustness of the supporting data sets.	• How extensively are the relevant or key events either empirically measured and/or modelled by appropriate *in silico*, *in chemico* and *in vitro* data?	Well reported	Low	
		• Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the supporting methods adequately reported?	Yes	Low	
11	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?	Yes	Low	
		• How many and how large are the dissimilarities in the supporting information (i.e., data gaps)?	Small level of dissimilarity	Low	
12	Documentation and written evidence provided	• Is the RA prediction adequately documented?	Yes	Low	
		• Does the evidence support the hypothesis that the uncertainty is acceptable for the stated purpose (as per Question 1)?			

* Organisation for Economic Cooperation and Development (OECD), 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, No. XXX, Series on Testing & Assessment. ENV/JM/MONO(2018)XX, OECD, Paris. In press, 2018.

Table S9. Uncertainties relating to Case Study 4 for short-chain mono-alkylphenols (Mellor et al, [21]*) – for illustrative purposes only.

Number	Uncertainty in Read-Across	Questions that Need to be Addressed Regarding Uncertainty	Brief Response	Associated Uncertainty	Source of Evidence or Supporting Information
1	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?	Yes	Low	Introduction (pages 1-2)
		Is the acceptable level or degree of uncertainty for the stated purpose defined?	No	High	
		• Is the stated acceptable level or degree of uncertainty appropriate for the stated regulatory purpose?	No	High	
2	Type of category / group including the definition of the applicability domain.	• Is the read-across approach (e.g., analogue or category) clearly reported?	Yes	Low	Introduction (pages 1-2) and Hypothesis (page 4)
		• Are the target and source chemicals clearly identified?	Yes	Low	Table 1 (page 4)
		• Is the applicability domain of the analogue or category defined?	Yes	Low	Results (page 4)
		• Do target and source chemicals fit within the defined applicability domain?	Yes	Low	Results (page 4)
3	The premise or hypothesis of the RA.	• Is the hypothesis on which the RA is based clearly stated and presented in sufficient detail to be assessed?	Yes	Low	Introduction (pages 1-2) and Hypothesis (page 4)
4	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?	Very clearly stated	Low-to-Medium	Results (pages 4-9); Table 4 (pages 8)
		• Is there sufficient experimental information provided to support the proposed chemical and toxicological mechanisms?	Yes	Low-to-Medium	Results (pages 4-9); Table 4 (pages 8)

		• How extensively does the experimental information provided support the mechanistic plausibility and / or the AOP or MoA on which the RA is based?	Very extensively stated	Low-to-Medium	Results (pages 4-9); Table 4 (pages 8)
5	Similarity in chemistry.	• Are the chemical structures (i.e., 2D structure, isomers, SMILES and molecular formula) reported for the derivatives used in the RA?	Yes	Low	Supplementary Information: Table 1
		• Are the dissimilarities in chemical structure reported and are they toxicologically relevant?	Yes, relevant only to toxicokinetics	Low	Results (page 4)
		• 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 RA?	Yes	Low	; Supplementary Information Table 2
		• Are the dissimilarities in molecular and physico-chemical properties reported and are they toxicologically (or pharmacokinetically) relevant?	Yes, relevant only to toxicokinetics	Low	Results (pages 4-9); Table 4 (pages 8)
6	Toxicodynamic similarity.	• Is there sufficient and consistent toxicodynamic information provided to establish similarity in the hazard of the derivatives used in the RA?	Yes	Low	Results (pages 4-9); Table 4 (pages 8)
7	Toxicokinetic similarity.	• Is there sufficient ADME information provided to establish toxicokinetic similarity for the derivatives used in the RA?	Yes	Low	Results (pages 4-9); Table 4 (pages 8)
		• Are any dissimilarities in ADME properties (and, as appropriate, metabolism / degradation) toxicologically relevant?	No – or as stated	Low	Results (pages 4-9); Table 4 (pages 8)
8	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?	Yes	Low	Results (pages 4-9); Table 5 (pages 9)
		• 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?	Yes	Low	Results (pages 4-9); Table 5 (pages 9)

9	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?	Yes	Low	Results (pages 4-9); Table 5 (pages 9)
		• Is the potency of the hazard reported and is it consistent among the source chemicals?	Yes	Medium	Results (pages 4-9); Table 5 (pages 9)
		• What are the temporal relationships between relevant endpoints?	All 90 day	Low	Results (pages 4-9); Table 5 (pages 9)
		• What are the dose-response relationships between relevant endpoints?	Not reported	High	
10	Strength or robustness of the supporting data sets.	• How extensively are the relevant or key events either empirically measured and/or modelled by appropriate *in silico*, *in chemico* and *in vitro* data?	Well reported	Low	Results (pages 4-9); Table 5 (pages 9); Supplementary Information
		• Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the supporting methods adequately reported?	Yes	Low	Results (pages 4-9); Table 5 (pages 9)
11	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?	Yes	Low	Results (pages 4-9); Table 5 (pages 9)
		• How many and how large are the dissimilarities in the supporting information (i.e., data gaps)?	Small level of dissimilarity	Low	Results (pages 4-9); Table 5 (pages 9); Supplementary Information
12	Documentation and written evidence provided	• Is the RA prediction adequately documented?	Yes	Low	Complete document
		• Does the evidence support the hypothesis that the uncertainty is acceptable for the stated purpose (as per Question 1)?	No	High	

* C.L. Mellor, T.W. Schultz, K.R. Przybylak, A.-N. Richarz, M.T.D Cronin, Read-across for rat oral gavage repeated-dose toxicity for short-chain mono-alkylphenols: A case study. Comput. Toxicol 2 (2017) 1-11.

Table S10. Uncertainties relating to Case Study 5 for allyl esters (OECD [14]*) – for illustrative purposes only.

Number	Uncertainty in Read-Across	Questions that Need to be Addressed Regarding Uncertainty	Brief Response	Associated Uncertainty	Source of Evidence or Supporting Information
1	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?	Yes	Low	Section 1 (page 9)
		• Is the acceptable level or degree of uncertainty for the stated purpose defined?	No	High	
		• Is the stated acceptable level or degree of uncertainty appropriate for the stated regulatory purpose?	No	High	
2	Type of category / group including the definition of the applicability domain.	• Is the read-across approach (e.g., analogue or category) clearly reported?	Yes	Low	Section 2 (pages 9-12)
		• Are the target and source chemicals clearly identified?	Yes	Low	Table 1 (page 12-13)
		• Is the applicability domain of the analogue or category defined?	Yes	Low	Section 4.3 (page 15)
		• Do target and source chemicals fit within the defined applicability domain?	Yes	Low	Section 4.3 (page 15)
3	The premise or hypothesis of the RA.	• Is the hypothesis on which the RA is based clearly stated and presented in sufficient detail to be assessed?	Yes	Low	Section 2 (pages 9-12)
4	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?	Very clearly stated	Medium	Section 4 (pages 13-15); Table 3 (pages 19-20)
		• Is there sufficient experimental information provided to support the proposed chemical and toxicological mechanisms?	Yes	Medium	Section 4 (pages 13-15); Table 3 (pages 19-20)
		• How extensively does the experimental information provided support the mechanistic plausibility and / or the AOP or MoA on which the RA is based?	Very extensively stated	Medium	Section 4 (pages 13-15); Table 3 (pages 19-20)

5	Similarity in chemistry.	• Are the chemical structures (i.e., 2D structure, isomers, SMILES and molecular formula) reported for the derivatives used in the RA?	Yes	Low	Table 1 (page 12-13)
		• Are the dissimilarities in chemical structure reported and are they toxicologically relevant?	Yes, relevant to toxicokinetics and mechanisms dependent on sub-category	Low	Table 1 (page 12-13); Section 4 (pages 13-15); Section 5 (pages 15 -16)
		• 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 RA?	Yes	Low	Table 1 (page 12-13); Section 4 (pages 13-15); Section 5 (pages 15 -16)
		• Are the dissimilarities in molecular and physico-chemical properties reported and are they toxicologically (or pharmacokinetically) relevant?	Yes, relevant only to toxicokinetics	Low	Table 1 (page 12-13); Section 4 (pages 13-15); Section 5 (pages 15 -16)
6	Toxicodynamic similarity.	• Is there sufficient and consistent toxicodynamic information provided to establish similarity in the hazard of the derivatives used in the RA?	Yes	Low	Table 1 (page 12-13); Section 4 (pages 13-15); Section 5 (pages 15 -16)
7	Toxicokinetic similarity.	• Is there sufficient ADME information provided to establish toxicokinetic similarity for the derivatives used in the RA?	Yes	Medium	Table 1 (page 12-13); Section 4 (pages 13-15); Section 5 (pages 15 -16)
		• Are any dissimilarities in ADME properties (and, as appropriate, metabolism / degradation) toxicologically relevant?	No – or as stated	Medium	Table 1 (page 12-13); Section 4 (pages 13-15); Section 5 (pages 15 -16)
8		• Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the data read across reported clearly?	Yes	Low	Section 2 (pages 9-12); Annex 1

	The quality of the apical endpoint data used to fill the data gap	• 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?	Yes	Low	Section 2 (pages 9-12); Annex 1
9	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?	Yes	Low	Section 4.3 (pages 14-15))
		• Is the potency of the hazard reported and is it consistent among the source chemicals?	Yes	Low	Section 4.3 (pages 14-15))
		• What are the temporal relationships between relevant endpoints?	All 90 day	Low to Medium	Section 4.3 (pages 14-15))
		• What are the dose-response relationships between relevant endpoints?	Not reported	High	
10	Strength or robustness of the supporting data sets.	• How extensively are the relevant or key events either empirically measured and/or modelled by appropriate *in silico*, *in chemico* and *in vitro* data?	Well reported	Low	Table 1 (page 12-13); Section 4 (pages 13-15); Section 5 (pages 15 -16)
		• Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the supporting methods adequately reported?	Yes	Low	Table 1 (page 12-13); Section 4 (pages 13-15); Section 5 (pages 15 -16)
11	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?	Yes	Low	Section 5 (pages 15 -16)
		• How many and how large are the dissimilarities in the supporting information (i.e., data gaps)?	Small level of dissimilarity	Low	Section 5 (pages 15 -16)
12	Documentation and written evidence provided	• Is the RA prediction adequately documented?	Yes	Low	Complete document
		• Does the evidence support the hypothesis that the uncertainty is acceptable for the stated purpose (as per Question 1)?	No	High	

* Organisation for Economic Cooperation and Development (OECD), Case Study on the Use of an Integrated Approach to Testing and Assessment for Hepatotoxicity of Allyl Esters, No. 253, Series on Testing & Assessment. ENV/JM/MONO(2016)51, OECD, Paris, 2016.

Table S11. Uncertainties relating to Case Study 6 for β-olefinic alcohols (Przybylak et al, [20]*) – for illustrative purposes only.

Number	Uncertainty in Read-Across	Questions that Need to be Addressed Regarding Uncertainty	Brief Response	Associated Uncertainty	Source of Evidence or Supporting Information
1	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?	Yes	Low	Introduction (pages 22-23)
		• Is the acceptable level or degree of uncertainty for the stated purpose defined?	Yes	Low	Introduction (page 23)
		• Is the stated acceptable level or degree of uncertainty appropriate for the stated regulatory purpose?	Yes	Low	Introduction (page 23)
2	Type of category / group including the definition of the applicability domain.	• Is the read-across approach (e.g., analogue or category) clearly reported?	Yes	Low	Introduction (pages 22-23) and Hypothesis (pages 23-24)
		• Are the target and source chemicals clearly identified?	Yes	Low	Table 1 (page 24)
		• Is the applicability domain of the analogue or category defined?	Yes	Low	Results (page 25)
		• Do target and source chemicals fit within the defined applicability domain?	Yes	Low	Results (page 25)
3	The premise or hypothesis of the RA.	• Is the hypothesis on which the RA is based clearly stated and presented in sufficient detail to be assessed?	Yes	Low	Introduction (pages 22-23) and Hypothesis (page 23-24)
4	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?	Very clearly stated	Medium	Results (pages 24-31); Table 4 (page 29)
		• Is there sufficient experimental information provided to support the proposed chemical and toxicological mechanisms?	Yes	Medium	Results (pages 24-31); Table 4 (page 29)

		• How extensively does the experimental information provided support the mechanistic plausibility and / or the AOP or MoA on which the RA is based?	Very extensively stated	Medium	Results (pages 24-31); Table 4 (page 29)
5	Similarity in chemistry.	• Are the chemical structures (i.e., 2D structure, isomers, SMILES and molecular formula) reported for the derivatives used in the RA?	Yes	Low	Table 1 (page 24); Supplementary Information: Table 1
		• Are the dissimilarities in chemical structure reported and are they toxicologically relevant?	Yes, relevant only to toxicokinetics	Low	Results (pages 24-31)
		• 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 RA?	Yes	Low	Supplementary Information Table 2
		• Are the dissimilarities in molecular and physico-chemical properties reported and are they toxicologically (or pharmacokinetically) relevant?	Yes, relevant only to toxicokinetics	Low	Results (pages 24-31); Table 4 (page 29)
6	Toxicodynamic similarity.	• Is there sufficient and consistent toxicodynamic information provided to establish similarity in the hazard of the derivatives used in the RA?	Yes	Low	Results (pages 24-31); Table 4 (page 29)
7	Toxicokinetic similarity.	• Is there sufficient ADME information provided to establish toxicokinetic similarity for the derivatives used in the RA?	Yes	Low	Results (pages 24-31); Table 4 (page 29)
		• Are any dissimilarities in ADME properties (and, as appropriate, metabolism / degradation) toxicologically relevant?	No – or as stated	Low	Results (pages 24-31); Table 4 (page 29)
8	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?	Yes	Low	Results (pages 24-31); Table 5 (page 30)
		• 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?	Yes	Low	Results (pages 24-31); Table 5 (page 30)

9	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?	Yes	Low	Results (pages 24-31); Table 5 (page 30)
		• Is the potency of the hazard reported and is it consistent among the source chemicals?	Yes	Medium	Results (pages 24-31); Table 5 (page 30)
		• What are the temporal relationships between relevant endpoints?	All 90 day	Low	Results (pages 24-31); Table 5 (page 30)
		• What are the dose-response relationships between relevant endpoints?	Not reported	High	
10	Strength or robustness of the supporting data sets.	• How extensively are the relevant or key events either empirically measured and/or modelled by appropriate *in silico*, *in chemico* and *in vitro* data?	Well reported	Low	Results (pages 24-31); Table 5 (page 30) Supplementary Information
		• Is the performance (e.g., reliability, accuracy, precision, repeatability and reproducibility) of the supporting methods adequately reported?	Yes	Low	Results (pages 24-31); Table 5 (page 30)
11	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?	Yes	Low	Results (pages 24-31); Table 5 (page 30)
		• How many and how large are the dissimilarities in the supporting information (i.e., data gaps)?	Small level of dissimilarity	Low	Results (pages 24-31); Table 5 (page 30) Supplementary Information
12	Documentation and written evidence provided	• Is the RA prediction adequately documented?	Yes	Low	Complete document
		• Does the evidence support the hypothesis that the uncertainty is acceptable for the stated purpose (as per Question 1)?	Yes	Low	Conclusions (page 25)

* K.R. Przybylak, T.W. Schultz, A.-N. Richarz, C.L. Mellor, S.E. Escher, M.T.D Cronin, Read-across of 90-day rat oral repeated-dose toxicity: A case study for selected β-olefinic alcohols. Comput. Toxicol. 1 (2017) 22-32.