1

A Strategy for Structuring and Reporting
a Read-Across Prediction of Toxicity

T.W. Schultz,[a] P. Amcoff,[b] E. Berggren,[c] F. Gautier,[d] M. Klaric,[b] D. J. Knight,[e] C. Mahony,[f] M. Schwarz,[g] A. White[h] and M.T.D. Cronin[i*]

[a]The University of Tennessee, College of Veterinary Medicine, Knoxville, TN, 37996-4500, USA. e-mail: tschultz @utk.edu;

[b]Cosmetics Europe, Avenue Herrmann Debroux 40, 1160 Brussels, Belgium. e-mails:pamcoff@cosmeticseurope.eu & mklaric@cosmeticseurope.eu;

[c]Joint Research Centre, European Commission, 21021 Ispra, Italy. e-mail: elisabet.berggren@ec.europa.eu;

[d]L'Oréal, 25-29 Quai Aulagnier, 92600 Asnières-sur-Seine, France. e-mail: fgautier@rd.loreal.com;

[e]ECHA, Annankatu 18, P.O. Box 400 Helsinki, Finland. e-mail Derek.Knight@echa.europa.eu;

[f]Procter & Gamble, Technical Centres Ltd, Whitehall Lane, Egham, Surrey, England. e-mail: mahony.c@pg.com;

[g]Eberhard Karls University of Tübingen, Institute of Experimental and Clinical Pharmacology and Toxicology, Department of Toxicology, Wilhelmstr. 56, 72074 Tübingen, Germany. e-mail: Michael.schwarz@uni-tuebingen.de;

[h]Unilever PLC, Colworth Science Park, Sharnbrook, Bedfordshire MK44 1LQ, England. e-mail: Andrew.white@unilever.com.

[i]School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, England. e-mail: m.t.cronin@ljmu.ac.uk;

26

* Corresponding Author: Mark Cronin: m.t.cronin@ljmu.ac.uk

Disclaimer: The views expressed in this document are those of the authors and do not necessarily represent the official views of Cosmetics Europe or its Members or the European Chemicals Agency.

Word Count for Abstract: 253 words

Word Count for Text: 10,699 words

Word Count for References: 1,379 words

Abstract

Category formation, grouping and read across methods are broadly applicable in toxicological assessments and may be used to fill data gaps for chemical safety assessment and regulatory decisions. In order to facilitate a transparent and systematic approach to aid regulatory acceptance, a strategy to evaluate chemical category membership, to support the use of read-across predictions that may be used to fill data gaps for regulatory decisions is proposed. There are two major aspects of any read-across exercise, namely assessing similarity and uncertainty. While there can be an over-arching rationale for grouping organic substances based on molecular structure and chemical properties, these similarities alone are generally not sufficient to justify a read-across prediction. Further scientific justification is normally required to justify the chemical grouping, typically including considerations of bioavailability, metabolism and biological/ mechanistic plausibility. Sources of uncertainty include a variety of elements which are typically divided into two main issues: the uncertainty associated firstly with the similarity justification and secondly the completeness of the read-across argument. This article focuses on chronic toxicity, whilst acknowledging the approaches are applicable to all endpoints. Templates, developed from work to prepare for the application of new toxicological data to read-across assessment, are presented. These templates act as proposals to assist in assessing similarity in the context of chemistry, toxicokinetics and toxicodynamics as well as to guide the systematic characterisation of uncertainty both in the context of the similarity rationale, the read across data and overall approach and conclusion. Lastly, a workflow for reporting a read-across prediction is suggested.

Highlights

- A strategy to evaluate chemical category membership is presented
- Templates to assess similarity and characterise uncertainty are developed
- A strategy to apply new toxicological data to strengthen read-across predictions
- A workflow for reporting a read-across prediction is described
- Read-across prediction to aid in regulatory decisions

1. Introduction and Problem Formulation

Legislative requirements for registration and safety assessment of chemicals have demonstrated the need for a new way of thinking to obtain toxicological information without resorting to animal testing. The grouping of substances allowing read-across of toxicity is a valuable method to obtain such information and potentially has a number of regulatory applications. The underlying philosophy of read-across is that substances which are similar in chemical structure will have similar properties and thereby, have similar toxicokinetic and toxicodynamic properties. Therefore, experimentally-derived toxicological properties from one substance, often referred to as the source chemical, can be read across to fill the data gap for a second substance, the target chemical, which has a similar chemical structure and for which a toxicology study may be lacking.

Despite the fact that read-across has been used for several years, a number of challenges remain. For instance, when applying read-across to make a prediction of toxicity, a number of questions arise, for which answers may be difficult to arrive at or to document; including:

1) Can a robust group of chemicals (often referred to as a chemical category) be formed to include the target chemical?
2) Is the category formed relevant for the toxicology of the endpoint under assessment?
3) Are there appropriate toxicology studies of high enough quality for the source chemical(s) to allow a meaningful read-across?
4) What is the uncertainty and is it acceptable to use the read across prediction to fill the data gap for a specific regulatory purpose?

To begin to address these questions a flexible strategy for developing and reporting a read-across prediction has been created. The strategy focuses on the two main elements of any read-across estimation, namely assessing (1) the similarity between target(s) and source substance(s) and (2) the uncertainties in the read-across process and ultimate prediction. While the standards for accepting a read-across prediction can vary between regulatory agencies, a good basis is the standard required for filling a REACH registration information requirement (EC, 2006). Conceptually, this means, for example, that in the context of a safety assessment for a complete set of results it should be possible to read-across the findings of a 28-/90-day repeated-dose oral rat toxicity study on the source substance(s) to the target substance(s). As such, the aim of the

98 read-across is to provide a prediction(s) that is (more or less) equivalent to the omitted standard
99 animal study and hence be acceptable for regulatory purposes.

100 The intent of this document is to establish a strategy which may be used to conduct and
101 document read-across predictions for data gap filling. As such, it provides guiding principles for
102 developing read-across predictions for discrete organic compounds. Where possible, emphasis
103 has been placed on undertaking and describing the read-across prediction in the best manner to
104 facilitate regulatory acceptance. This document represents, in part, discussions in and progress
105 made in the European Commission and Cosmetics Europe funded SEURAT-1 Cluster
106 (www.seurat-1.eu). As such, the primary focus of this document is directed towards read-across
107 predictions for chronic toxicity, or improving the possibility to read-across from repeat dose
108 toxicity tests. However, in order to achieve this aim, the document draws upon current expertise
109 and knowledge from other toxicological endpoints and the information, templates and work
110 plans contained herein are generally applicable to all read-across scenarios and endpoints.

111 In order to facilitate regulatory acceptance, a read-across prediction needs to be justified in all
112 aspects. Briefly, the justification of a read-across prediction needs to be robust, reliable and
113 easily explicable. Key principles of similarity need to be clearly documented and, where
114 possible, supported by scientific literature and data. Sources of uncertainty need to be identified
115 and accommodated; these can typically be divided into two main types: 1) the uncertainty
116 associated with the justification of similarity between the source and target structures, and 2) the
117 uncertainty associated with the application of the particular read-across exercise.

118 Whilst no consensus has been reached by stakeholders and users, there is growing agreement that
119 when read-across is applied to make predictions to fulfil information requirements, this must be
120 done on an endpoint-by-endpoint basis, i.e. for the particular toxicology study to be predicted.
121 This approach to apply to endpoints individually is due, even when there is an over-arching
122 category hypothesis, to different applicability domains, different source chemicals and/or
123 different Weights-of-Evidence (WoE) which may apply to making predictions for different
124 endpoints. Obviously, there will be occasions where one or more endpoints will be closely
125 related and knowledge may be transferable, thus allowing read-across arguments to build,
126 partially, on each other.

127 It is generally agreed that the acceptability of a read-across prediction relies on the explanation

128 of the similarity which forms the basis of the read-across, as well as the description of the type

129 and degree of uncertainty associated with the particular read-across. Therefore, it is important to

130 address these two elements in a transparent and consistent manner. The use of templates or work

131 plans facilitate the elucidation of the transparency and consistency in read-across. Existing

132 templates or reporting formats for read-across vary in detail, however, it is generally agreed that

133 they aim to:

1) Describe the rationale for the similarity between the source and target chemical in a
   transparent manner.

2) Document the logic and data leading to the read-across prediction so that, if required, it
   can subsequently be recreated.

3) Describe the uncertainties in the prediction; specifically separating the uncertainties in
   data and definition of similarity from procedural uncertainty.

4) Clarify the roles of any endpoint specific and/or endpoint non-specific factors affecting
   the assessment.

2. Background

Read-across is an alternative method for filling data gaps based on an analogue or chemical

category approach (van Leeuwen et al., 2009). It is the process of assessing a toxic endpoint of

an untested substance (i.e., target chemical) based on the results for the same endpoint for a

tested substance (i.e., source chemical) considered to be "similar" in the context of structure,

properties and/or activities (Dimitrov and Mekenyan, 2010). It is recognised that forming a

chemical category and data gap filling by interpolation within the category, especially for hazard

assessments, is not a new concept (OECD, 2014a). However, greater emphasis has now been

placed on the resultant read-across prediction due to legislative pressure, especially within

Europe, and especially for classification and labelling, and risk assessment. Currently, there is

growing interest in several national Governmental regulatory agencies to establish best practices

for conducting and evaluating read-across within the context of, and to enable, regulatory

decisions. Published exercises and case studies using the OECD QSAR Toolbox (cf. Enoch et

al., 2013) have demonstrated that category-based read-across can be used to establish that a

substance is associated with potentially hazardous properties. However, it is more difficult to

158 show that a substance is not potentially hazardous. In order to address this issue, the more recent

159 literature has identified some of the challenges which need to be taken into account when

160 preparing a read-across justification (cf. Patlewicz et al., 2013a; 2014); specifically, case studies

161 have described the process to create a read-across prediction increasing the likelihood of

162 regulatory acceptance (cf. Ball et al., 2014).

163 Much guidance on grouping of chemicals and read-across is already available (ECETOC, 2012;

164 ECHA, 2009, 2011; OECD, 2007, 2011, 2014a) and the key strategic documents have been

165 summarised in Table 1.4 of Cronin (2013a). This is a fast moving field and the formation of

166 chemical categories, or the grouping of molecules, especially to allow for the filling of data gaps

167 by read-across, has advanced markedly since the start of the 21st Century. Background

168 information on the processes of grouping and read-across has been detailed by Cronin et al.

169 (2013). It is clear that interest in chemical category formation, coupled with read-across for

170 toxicological data gap filling, has grown for a number of reasons (Cronin, 2013a). However, the

171 primary drivers of this expansion are legislation, which has forced the need for non-test methods

172 to assess chemical safety and the willingness of regulatory bodies, although it is cautious, to

173 accept read-across-based submissions in lieu of test results. While there are various advantages

174 and disadvantages to using the category-based read-across approach in toxicology (Patlewicz et

175 al., 2013a, 2013b; Cronin, 2013a), the advantages appear to out-weight the disadvantages. As

176 additional case studies demonstrating the utility and practical application of read-across become

177 available, the advantages will become more prominent and the challenges more readily

178 addressed.

179 All applications of read-across are context dependent and any read-across adaptation (i.e., the

180 formal process by which a prediction is used for regulatory purposes) is likely to be performed

181 with limited sets of experimental data. Thus, successful adaptations of a read-across are

182 contingent not only on the appropriate selection of the characteristics, measures of similarity and

183 assessment of the uncertainties associated with the prediction, but also on the quality and

184 quantity of the information and data used in the exercise.

185 Within the applicability domain of a chemical category, read-across can be performed to fill data

186 gaps with a number of approaches which can be summarised into the following four techniques:

187   1) one-to-one read-across (i.e., one source substance used to make a prediction for a single
188   target chemical),
189   2) many-to-one read-across (i.e., two or more source substances used to make a prediction
190   for a single target chemical),
191   3) one-to-many read-across (i.e., one source substance used to make a prediction for two or
192   more target chemicals), or
193   4) many-to-many read-across (i.e., two or more source substances used to make predictions
194   for two or more target chemicals).

195 Techniques 3 and 4 may be considered as being multiple simultaneous applications of techniques
196 1 and 2, respectively. Given limited data availability, the "one-to-one", or analogue approach, is
197 often the only viable option. Ideally, however, the "many-to-one" or category approach is
198 preferred as it inherently possesses a greater WoE in that each analogue in the category supports
199 the others.

200 With reference to the above applications (one/many-to-one/many), it is recognised that read-
201 across for toxicity prediction can be qualitative or quantitative in design. A qualitative read-
202 across provides a "yes/no" prediction for an effect; quantitative read-across provides quantitative
203 (i.e., potency) values for an endpoint. When conducting a quantitative read-across exercise, the
204 OECD suggests that there are four main approaches to making the prediction (OECD 2014a):

205   1. reading across from the endpoint value of a similar chemical (e.g., the closest source
206   chemical);
207   2. applying a mathematical scale to the trend in available experimental results from two or
208   more chemicals similar to the target chemical (e.g., trend analysis or structure-activity
209   relationships);
210   3. processing the endpoint values from two or more source chemicals (e.g., by averaging, by
211   taking the most representative value), or;
212   4. when sufficient data allow, taking the most conservative value among the source
213   chemicals within the whole category.

214 Establishing similarity on an apical endpoint-specific basis is essential to successful category
215 formation and read-across (ECETOC, 2012). Chemical similarity can be considered in a number
216 of ways (Enoch and Roberts, 2013). Critical to the justification of analogue(s) selection for read-
217 across is the explanation of seminal criteria of chemical similarity on which the selection is

218 based. The definition of these criteria is an on-going issue since chemical similarity may be

219 assessed in many ways and, even when assessed objectively, not all measures of chemical

220 similarity are of equal importance and there is no simple similarity scale. In the extreme, each

221 chemical can be considered as its own category; however this is obviously not practical for

222 predictive purposes. In addition, it is accepted that simple measures of chemical similarity (e.g.,

223 being a member of a simple organic chemical class, having the same carbon skeleton or same

224 function group) are often not practical for making predictions. Thus, as noted by Enoch and

225 Roberts (2013), in order for any read-across prediction to gain acceptance, it is essential to

226 explain the basis for similarity between the target chemical(s) and source chemical(s) in a robust

227 and reliable manner.

228 After a read-across exercise is carried out, an assessment is undertaken of whether the case

229 supporting the read-across is sufficient for the prediction to be acceptable. This acceptance is

230 often stated in the form of confidence or certainty. While the acceptance of read-across

231 predictions is often made according to a standard procedure (e.g., an assessment framework),

232 ultimately the evaluator(s) must be convinced of the scientific credibility of the premise of the

233 read-across and the supporting data provided. Therefore, assuming the rationale for similarity is

234 accepted (i.e., the category is robust and membership is assured), final acceptance of the read

235 across prediction is contingent on reducing uncertainty. While uncertainty is related to the

236 quality and quantity of the read across endpoint data (Cronin, 2013b; Péry et al., 2013;

237 Blackburn and Stuard, 2014), there are a number of other factors that influence uncertainty.

238

239 *2.1 Regulatory Context and International Efforts to Address Read-Across Predictions*

240

241 In order to understand the context of the development of read-across, it is important to consider

242 how it has been developed and shaped as a data gap filling approach with regard to legislative

243 and regulatory pressure. Globally, a multiplicity of regulatory agencies is applying read-across in

244 their decision making processes. While a number of these agencies are currently focusing efforts

245 on how to best standardise the development and evaluation of read-across predictions, the

246 European CHemical Agency (ECHA), especially through the provisions in Registration,

247 Evaluation, Authorisation and Restriction of Chemicals (REACH) is among the better known.

248 Specifically, REACH allows for adaptations to the standard information requirements by means

249 of read-across of a study conducted on a source substance to a target substance (cf. Annex XI in

250 EC, 2006).

251 The standard ECHA advice to registrants on making and documenting a good-quality read-

252 across/category (ECHA, 2013a; 2013b) refers to the importance of making a clear read-across

253 hypothesis and justification. Non-testing approaches to data gap filling have also garnered much

254 attention at the Organisation for Economic Co-operation and Development (OECD) and among

255 its member countries. Specifically, among the OECD member countries, read-across is used as

256 an alternative method for hazard identification and characterisation in risk assessments; read-

257 across is especially useful when based on grouping approaches, because not every chemical

258 needs to be tested (OECD, 2014c).

259 Since the regulatory use of read-across relies on the scientific validity and the robustness of the

260 justification substantiating the prediction for a given endpoint(s), there are a number of issues

261 associated with read-across which may benefit from international discussion on a broader scale.

262 Experiences reported by the OECD members indicate that there is still a lack of agreement on

263 what "chemical similarity" is. Specifically, the OECD has noted the challenge posed by the facts

264 that: 1) a chemical category is defined by a variety of factors, 2) there are no simple similarity

265 scale(s), and 3) similarity can also depend on the endpoint under consideration (OECD, 2014c).

266 Work at OECD has revealed that similarity hypothesis can be based on a variety of aspects, and

267 definitions, of, chemistry. OECD has also concluded that these methods of assessing similarity

268 are not equal in obtaining a robust chemical category for toxicological read-across. Read-across

269 based on mechanistic similarity (e.g., common chemical interaction with a receptor) is generally

270 considered a better similarity hypothesis than an informatics based similarity metric. However,

271 knowledge of the mode or mechanism of action is not always available, especially for the more

272 complex endpoints such as repeated dose toxicity. Moreover, information on transformation

273 products and the rate of formation of these products is likely to be the key factor in accepting

274 read-across predictions. Thus, information derived from experimental studies, as well as

275 toxicokinetic information and ADME information, will contribute to justify the prediction.

276 The current view of OECD (OECD, 2014c) is that more experience is needed on how the

277 confidence in the prediction could be enhanced by providing more mechanistic transparency,

278 using experimental data from structural analogues, using data that are supplemented by

279 toxicokinetic and ADME information, and using data that are supplemented by relevant *in vitro*

280 and *in chemico* endpoints (i.e., incorporation of more information to increase the WoE). More

281 specifically, the OECD has emphasised the following as being crucial to the successful

282 application of read-across: 1) the process of how to document the justification for a read-across,

283 2) consideration of how to perform read-across for more complex endpoints (e.g., repeated dose

284 toxicity), 3) development of approaches and agreement of use of quantitative read-across for

285 hazard characterisation, 4) methods to better take mechanistic considerations into account in

286 grouping chemicals, and 5) approaches to derive WoE conclusions based on results from

287 alternative methods or supplementary information.

288 While the details may vary, it is obvious from all the regulatory requirements and guidance that

289 any general strategy to assess the justification for a read-across prediction must examine whether

290 or not the key principles of similarity are clearly documented and whether the interpretation is

291 supported by scientific justification based on argumentation, literature and data. Development of

292 the similarity rationale, whether for an analogue or a chemical category, must be performed on a

293 case-by-case basis. This case-by-case basis is likely to be influenced by the availability of

294 suitable data to populate the category and be specific to the regulatory endpoint being evaluated

295 (i.e., complex endpoints may intrinsically require greater confidence in the similarity argument

296 and data). Read-across arguments often adopt a multifaceted approach that combines several

297 similarities into a single rationale. This approach, where similarity between the source and target

298 chemicals is demonstrated across multiple parameters, is designed to reduce uncertainty

299 associated with the read-across prediction.

300 Acceptance of a read-across prediction is often couched in the evaluator's sense of confidence

301 or, more accurately, certainty in the prediction. In the end, high confidence (i.e., low concern

302 about potential error in the prediction) is assigned to a read-across when there is strong proof the

303 prediction is valid (i.e., low uncertainty). This confidence is often gained by identifying and

304 addressing the sources of uncertainty.

305 Finally, it is recognised that the OECD is currently conducting further work on the hazard

306 assessment of chemicals. Through the Task Force on Hazard Assessment, the OECD is

307 developing Integrated Approaches to Testing and Assessment (IATA). Included in this effort is

308   the examination of grouping approaches and the exchange of experiences among the member

309   countries on new hazard assessment methodologies. A goal of this work is to achieve a

310   harmonised approach to the implementation of IATA, so as to ensure consistency in how

311   information is used in regulatory decision-making and to foster mutual acceptance of

312   assessments (OECD 2014c). This knowledge and experience will add to the understanding of the

313   process of category formation and use of read-across.

314

315   ### 3.   Defining the Criteria for Category Membership: Establishing Similarity

316   To meet regulatory needs, the read-across hypothesis, or justification for the read-across within a

317   defined chemical category of discrete organic substances, must include a clear definition of the

318   criteria (i.e., chemical similarity) for membership of the category (i.e., a clear definition of the

319   applicability domain). Within the REACH regulation, read-across is founded on the principle o

320   of "structural similarity" combined with a scientific justification. Therefore, within the OECD

321   guidance for read-across, the basis for assessing similarity is typically elaborated with the

322   possibility of other considerations (e.g., bioavailability, toxicokinetics/metabolism) to assess

323   analogue similarity (OECD, 2014a). Moreover, a useful tool that might be employed for

324   demonstrating commonality in toxic behaviour is through an adverse outcome pathway concept;

325   this implies assessing similarity "via molecular initiating events", "key intermediate events" and

326   "other relative *in vitro*" information and data (OECD, 2013; 2014b). Clearly, the basis for

327   establishing the applicability domain of a category will depend both on the endpoint and

328   chemical and means of forming a category e.g. a specifically vs. broadly defined fragment. Thus,

329   the questions "Can a chemical category be formed?" and "Is the category toxicologically

330   relevant?" are often addressed concurrently.

331   Building on six case studies using the information within the OECD QSAR Toolbox (Enoch et

332   al., 2013) and the earlier work of Blackburn et al. (2011) and Wu et al. (2010), it is clear that

333   chemical category membership can be defined by many factors. Table 1 summarises the factors

334   leading to category membership being adequately defined and supported into three elements.

335

336   TABLE 1 HERE

337

338   While there can be a starting premise or over-arching rationale for grouping organic substances

339   based on molecular structure and chemical properties, these similarities alone are generally not

340   sufficient to justify a read-across prediction. Typically, further information is required to justify

341   the chemical grouping on the basis of considerations such as bioavailability, reactivity, and

342   metabolism. Similarity in bioavailability is also crucial to confirm where possible. Read-across

343   should be performed where similar bioavailability can be demonstrated. Currently, without

344   experimental data, it is difficult to obtain realistic estimates of bioavailability *in silico*, however

345   progress is being made in areas such as predicting metabolism and clearance rates which

346   combined could provide usable descriptors. For read-across predictions for the less complex

347   endpoints (e.g., acute aquatic toxicity), adding these toxicokinetic similarities is often enough to

348   justify a read-across prediction. However, for the more complex endpoints (e.g., chronic health

349   toxicities), additional measures of similarity are necessary for read-across prediction to be

350   acceptable.

351   While there is no definitive list of similarities with in a group, eleven similarities which are

352   proposed that to have an impact on forming the chemical category for a read-across prediction,

353   are summarised in Table 2. In order to be both transparent and comprehensive, it is suggested to

354   collect similarity data for as many criteria as possible. Whilst molecular structure similarity is a

355   highly pragmatic approach to identify potential source analogues, it is not on its own sufficient to

356   justify read across, and indeed it may not be the most important element.

357

358   TABLE 2 HERE

359

360   Data for molecular structure and physico-chemical properties to support grouping hypotheses

361   can be easily obtained *in silico* from software such as the OECD QSAR Toolbox. Using two-

362   dimensional molecular structure, structural data can be organised into groups of atoms

363   representing rings (e.g., benzene or naphthalene), linkers (i.e., atoms in a direct path connecting

364   two ring systems), frameworks (i.e., the combination of ring systems and linkers in a molecule),

365 and side chains (i.e., non-ring, non-linker atoms) (Bemis and Murcko, 1996). These molecular

366 scaffolds provide a basis for assessing similarity. Common constituents include substituents

367 (e.g., the 166 well-characterised, common organic moieties described by Hansch and Leo, 1979)

368 and structural fragments (e.g., the 645 fragments used in the US EPA's the Analog Identification

369 Methodology (AIM)). In addition, physico-chemical and molecular property similarities include

370 properties which are linked to key factors that affect toxicity (e.g., volatility, solubility,

371 reactivity, etc.).(http://www.epa.gov/oppt/sf/tools/aim.htm).

372 Five types of similarity (Items 3-7 in Table 2) are typically considered to meet the similarity

373 hypotheses for grouping chemicals for read-across based on common toxicokinetics and/or

374 abiotic transformation; these factors largely focus on metabolism which often has significant

375 uncertainty associated with it due to the potential difficulty in obtaining experimental or *in silico*

376 data. Transformation similarities focus on the likelihood of attaining common or similar

377 precursors and/or breakdown products, via physical or biological processes. This includes key

378 abiotic transformations (e.g., hydrolysis, autooxidation) and toxicokinetics (ADME), the same

379 key metabolic pathway(s) or pathway inhibition, activation to same or similar reactive chemical

380 species and degradation to the same or similar chemical species.

381 For read-across based on common biological/toxicological factors, three types of similarity;

382 toxicophores, mechanistic plausibility and related endpoints, are mostly considered (Table 2), the

383 most important of which is mechanistic plausibility. The AOP construct, an excellent concept for

384 adding mechanistic understanding into the read-across, is one of several means of establishing

385 mechanistic plausibility. In addition, similarity in the biological (preferably *in vivo*) data, such

386 that are available will provided additional evidence for category membership,

387 In the initial phase of developing a read-across, it is advisable to collect information on similarity

388 and data for as many of the criteria listed in Table 2 as possible. However, it is intuitive that the

389 most critical measurements of similarity are endpoint- and scenario-dependent and hence will

390 require expert judgment and application. In amassing information on similarity (for regulatory

391 applications in particular) it is essential to explain the basis for the similarity between the target

392 chemical(s) and the source chemical(s) in sufficient detail to be able to judge fit for purpose.

393 There are a number of potential regulatory purposes for performing, and uses of, a read-across

394 prediction. The regulatory purposes include: 1) Prioritisation and Screening, 2) Hazard

395 Identification (potential), 3) Hazard Characterisation (potency), and 4) Safety Assessment

396 (potential/potency and exposure). Thus, in assessing the similarity associated with grouping, it is

397 important to do so in the context of the decision being considered and the scope of the problem.

398 The "context" and "scope" significantly influence a number of issues including the similarity

399 rationale(s) required to form the category and identify analogues.

400 The regulatory purpose of the read-across often determines the type(s) of similarity required. It is

401 currently accepted (c.f., Cronin et al. (2013), that there are three broad criteria of similarity: 1)

402 chemistry, 2) transformation, and 3) toxicology. In consideration of Prioritisation and Screening,

403 hazard identification and safety assessments greater and more detailed information is required on

404 similarity as described further in Section 4. Therefore, in order to achieve the goal of "fitness for

405 purpose" (i.e., to be both transparent and comprehensive in the justification of a read-across) it is

406 advisable to collect data for as many of these similarity criteria listed in Table 3 as possible. To

407 assist in this process of collecting and assessing information relating to similarity, a template for

408 assessing similarity of analogues and category members for read-across has been proposed and is

409 reported in Appendix A.

410 This proposed template to collect information to establish similarity includes an overall

411 conclusion regarding the rationale for analogue/category similarity (this is provided as a text box

412 in the Template in Appendix A). The conclusion is intended to summarise all relevant scientific

413 information relating to establishing similarity, in order to clearly justify the analogue(s) selected.

414 The overall rationale for similarity is established by assessing the various criteria for similarity.

415 This is achieved by answering the following questions relating to chemical, transformational and

416 toxicological similarity.

417

418 TABLE 3 HERE

419

4. Confidence and Uncertainty

There is general agreement that increased uncertainty has a strong negative impact on a read-across prediction and often negates the use of the read-across method. For that reason, uncertainties need to be identified and appraised (Cronin et al., 2013; Ball et al., 2014; Blackburn and Stuard, 2014; Patlewicz et al., 2014). However, the concept and definition of uncertainty has been described as ambiguous; it tends to incorporate a variety of methodologies with the aim of meeting different goals (Péry, et al., 2013). As a result, a major challenge for the better use of the read-across approach lies in making the concept of uncertainty more understandable and transparent. Currently, determining how much uncertainty is acceptable for a read-across prediction is still largely subjective. It is defined on a case-by-case basis and influenced heavily by the purpose of the prediction, the endpoint assessed, and whether the read-across predicts the presence or absence of toxicity.

To date, the most comprehensive method for gauging uncertainty for read-across, especially for chronic health effects (e.g., repeated dose toxicity), is in the "framework" of Blackburn and Stuard (2014). This is a prescriptive scheme for addressing the various facets of uncertainty as it pertains to read-across. Specifically, it is designed to: 1) increase transparency of the read-across prediction, 2) provide consistency to the exercise, 3) provide a means of examining robustness and consistency among the key facets of similarity, 4) facilitate review and evaluation of the read-across exercise, and 5) help identify where additional data may be helpful, especially in reducing uncertainty. The Blackburn-Stuard framework does not, however, completely remove subjectivity from the process, as expert judgment is still required to categorise uncertainty. In addition, the Blackburn-Stuard framework defines four levels of uncertainty (i.e., low, low to medium, medium and high) and proposes quantitative factors (i.e., 1, 3 and 10, respectively) for addressing the three lesser levels, with the uppermost level of uncertainty being deemed unsuitable for the application of the read-across method. The numerical uncertainty factors serve to build conservatism into the potency prediction and weigh the unknown associated with the prediction. While the framework is new and largely untested, the scheme appears to be good for repeated dose toxicity endpoints where assessment factors can be applied to NOAELs. More quantitative approaches for assessing uncertainty are provided below.

Sources of uncertainty include a variety of elements which are typically divided into two main issues. The first issue is uncertainty associated with similarity justification, and the second is

451 associated with the overall approach and conclusion. With regard to the uncertainty associated
452 with similarity justification, this implies that there are inherent uncertainties associated with the
453 presumption that the results of the *in vivo* study/ies on the source chemical(s) apply (i.e., can be
454 read across) to the target analogue(s). The justification for this presumption is based on two
455 interrelated rationales: 1) that the target and source materials are sufficiently similar to be
456 toxicologically relevant, and 2) that supporting arguments are provided to justify that the
457 differences in chemical structure do not affect the properties relevant to the specific endpoint
458 under consideration.

459 The assessment of uncertainty associated with similarity justification includes consideration of
460 the information supporting the scientific arguments for similarity and data associated with the
461 chemical, toxicokinetic and toxicodynamic similarity resulting in the toxicity being read across.
462 As stated previously, chemical-based toxicological similarity may be established by responding
463 to the questions posed in Table 3 which may be achieved by following the template presented in
464 Appendix A. Uncertainty associated with the answers to the questions in Table 3 is assessed in a
465 uniform manner and a WoE, indicating consistency in quality and quantification of the data for
466 each feature, assigned (Appendix B, Table B.1)

467 Among the uncertainties are those brought about by deficiencies in the underlying knowledge
468 and data associated with assessing the essential areas of similarity. Chemical similarity, in itself,
469 may never be enough to justify fully a read-across prediction. While molecular structure and
470 physico-chemical properties play a role in assessing similarity, depending on the toxicological
471 endpoint under consideration, these factors by themselves may not be enough. For example, for
472 chronic health endpoints, two structurally similar chemicals may have significant differences in
473 toxicity. In these cases, toxicokinetic and/or biological similarity may be more important. When
474 such information is lacking, specific studies may be necessary to confirm the premise of the
475 similarity justification or, as a minimum, reduce the uncertainty in the similarity to an acceptable
476 level for the intended purpose. Such a confirmation of biological similarity may be obtained
477 from the comparison of toxicological profiles derived from, for instance, non-animal tests.
478 However, in such cases, it may be complex and require expert judgement to select the
479 appropriate *in chemico* method, *in vitro* assay or possibly an *in silico* tool to provide the critical
480 information needed to strengthen a similarity rationale.

481   The second issue of uncertainty is associated with the completeness of the read-across argument.

482   The molecular nature (e.g., complexity of molecular structure) of the target chemical(s), the

483   nature and complexity of the apical endpoint to be read across, the premise or hypothesis of the

484   read-across, the purpose of the prediction as well as the quality and robustness of the data all can

485   have an impact on uncertainty, its definition and acceptability for read-across (Table 5).

486   The molecular nature (e.g., complexity of structure) of the target chemical(s) (2nd bullet in Table

487   5) implies that target chemicals with simple molecular structures (e.g., a hydrocarbon scaffold

488   and one functional group) impart less uncertainty than a more complex molecular structure (e.g.,

489   a heteroatom scaffold with multiple structural groups).

490   In terms of chemistry, the more narrowly defined the applicability domain of the grouping, the

491   greater the confidence can be placed in the group membership and hence, the less the

492   uncertainty. For example, low uncertainty is associated with all category members having the

493   same functional groups and appropriately similar key physico-chemical and molecular properties

494   (e.g., aliphatic aldehydes with C2 to C5).

495   Relating to the problem and premise of read-across (1st bullet in Table 5), it is intuitive that

496   reading across from many-to-one provides lower uncertainty than reading across from one-to-

497   one, assuming that the standard of the available *in vivo* data of the source substances, and the

498   trends within them, are comparable. Further uncertainty may be associated with the apical

499   endpoint itself, which is to be read across. For some endpoints, chemical mechanism and/or

500   biological modes-of-action are well-established (e.g., mutagenicity). However, for other

501   endpoints (e.g., repeated dose toxicity), the lack of a mechanistic understanding tends to

502   introduce greater uncertainty into the similarity rationale. Mechanistic uncertainty is best

503   assessed within the context of an AOP. It is recognised that knowledge of an AOP evolves and,

504   as such, AOP development represents a continuum from less-to-more complete with increasing

505   quality, quantification and strength of key events (KEs) and key event relationships (KERs)

506   (Tollefsen et al., 2014). Confidence in using an AOP is typically informed by: 1) support for the

507   biological plausibility of KEs, KERs in relationship to the *in vivo* apical outcome under

508   consideration, 2) support for the essentiality of the MIE and other KEs, and 3) empirical data

509   quantifying the KEs and support for the KERs.

510 As an example, typically, there is more uncertainty with a developmental toxicity endpoint

511 versus a genotoxic endpoint. A chemical which can cause DNA or chromosomal damage is

512 deemed a genotoxin. As such, many *in vitro* and *in vivo* tests for genotoxicity have been

513 developed with a range of endpoints that either detect DNA or protein damage or a genotoxicity-

514 related biological consequence; causal linkage between the interaction of a chemical with

515 biomolecules at the molecular level and subsequent *in vitro* and *in vivo* genotoxic effects are

516 well-established (Petkov et al., 2015). The net result is that there are practical methods of

517 integrating *in silico* and *in vitro* results to reduce uncertainty in predicting genotoxicity outcomes

518 of untested chemicals. In contrast, there are a variety of interactions of a chemical with

519 biomolecules which can subsequently lead to adverse developmental effects (Wu et al., 2013).

520 Many of the interactions that underpin developmental toxicity may not be defined in detail and it

521 may not be possible to obtain data for Key Events in the AOP, even for well defined events.

522 Thus, the read-across of developmental toxicity is implicitly associated with greater uncertainty

523 than for well described and "modelled" endpoints. Linked to this concept is the realisation that

524 there are several sources of uncertainty in supporting biological justification. These sources,

525 which are relevant for all systemic endpoints, include: 1) incomplete knowledge of the biological

526 mechanism(s) resulting in toxicity, 2) relevance and completeness of the supporting evidence in

527 the form of scientific information and/or test data, and 3) problems with the test data (e.g.,

528 variability in results, lack of understanding what the results mean, etc.). Once the weaknesses or

529 data insufficiencies in the justification are documented, new method evidence can be added to

530 address the shortcomings and reduce the uncertainty.

531 The read across endpoint(s) is another focal point of the exercise. The type of endpoint read-

532 across effects uncertainty and as more complex endpoints are addressed, there will be a greater

533 WoE required to justify category membership. Simpler endpoints (e.g., acute toxicity) may be

534 readily addressed with fewer lines of evidence supporting the biological justification; often, a

535 single toxicity profiler or small group of *in vitro* tests are sufficient to establish the chemical

536 category or analogue and support the read-across. In contrast, for more complex endpoints, such

537 as chronic health effects which are traditionally assessed by higher level *in vivo* tests (e.g., 28-

538 day repeated dose testing), establishing the category is more difficult. In the case of complex

539 endpoints, analogues are often identified by WoE, looking at consistency in empirical and/or

540 model data across a number of mechanistically relevant endpoints. For example, read across for

541   skin sensitisation may require a WoE call after gauging uncertainty in skin metabolism or abiotic

542   oxidation, as well as chemical reactivity leading to protein binding and dendritic cell activation.

543   In contrast, reading across for oral *in vivo* mutagenicity may require gauging uncertainty in

544   microbial transformation in the gut, metabolic activation in the liver and chemical reactivity

545   leading to DNA-binding and would probably require a lower overall WoE than for chronic

546   toxicity. The depth and breathe of the information and empirical data for these different activities

547   affect the overall level of uncertainty allowed, while still accepting the prediction via the WoE.

548   The problem and premise of the read-across significantly influence both the similarity rationale

549   required to form an appropriate chemical category and the empirical data of sufficient quality

550   required for the source chemical.

551

552   TABLE 4 HERE

553

554   Thus, taking the scenarios summarised in Table 4, in Scenario 1 toxicokinetics are less critical to

555   establishing similarity and establishing a source chemical as being of high quality than in

556   Scenario 2. In fact, the absence of toxicokinetic data for Scenario 2 may mean the uncertainty is

557   too great as to prevent the use of read-across without further testing. In addition, a read-across

558   prediction of the absence of an adverse effect carries with it a greater perception of uncertainty.

559   In this case it is not possible to demonstrate with absolute certainty that a target chemical does

560   not elicit a particular *in vivo* adverse effect (Scenario 3), however it may be possible to reduce

561   uncertainty by demonstrating the absence of sub-cellular and cellular responses (i.e., negative

562   results from molecular screening and toxicogenomics). In Scenario 4, one of the key questions to

563   be addressed is whether sub-categorisation is required to reduce the uncertainty associated with

564   the applicability domain of the read-across. The purpose of the prediction also impacts the

565   degree of uncertainty that is acceptable.

566   While most previous publications discussing read-across have focused on its application in safety

567   assessment, read-across may be used to fill other needs. As noted earlier, there are four

568   regulatory uses for using read-across predictions that apply three basic types of similarity. The

569   purpose of the prediction may determine the types of similarity required that can be used, and

570   thus influences uncertainty. Prioritisation and screening may be amenable to prediction based

571   only on information from analogue chemistry. Hazard identification may require information on

572   both chemistry and toxicology similarity. However, hazard quantification for risk assessment

573   will normally needs dosing route and transformation similarity to assess exposure and

574   toxicological similarity; in addition there may be an assessment of mechanistic plausibility,

575   perhaps based on an AOP.

576   The uncertainty that is associated with the *in vivo* toxicology study/ies on the source chemical(s)

577   is always case-specific. Assessments should focus on any deficiencies in the quality of the

578   toxicology data to be read across, especially as compared to what is expected from current

579   standard test methods. Questions 3-4 in Table 5 are designed to address uncertainty associated

580   with the *in vivo* data being read across (a number of methods are available to ascertain toxicity

581   data quality, with the reader being referred to (Klimisch et al., 1997; Przybylak et al., 2012;

582   Steinmetz et al., 2014; Yang et al., 2013) for further information). Conversely, the final three

583   questions in Table 5 are designed to address uncertainty associated with *in chemico*, *in vitro* or *in*

584   *silico* data used to strengthen the similarity rationale. Lower uncertainty may also be assigned

585   when empirical and *in silico* measurements of chemical properties are in good agreement.

586   The qualification of transformation impacts uncertainty, especially with respect to metabolism

587   for the category members without empirical data. For example, low uncertainty is associated

588   when all category members have similar ADME properties. Although there is uncertainty

589   associated with predictions from *in silico* tools, the uncertainty is considered lower when

590   empirical studies (*in vivo* and/or *in vitro*) and model predictions indicate similar metabolism. In

591   addition, information on the purity of compounds being considered and read across must be

592   included as this may affect the certainty.

593

594   TABLE 5 HERE

595

596   The uncertainty associated with a read-across prediction is impacted by several additional

597   features, especially those associated with the completeness and application of the read-across

598   procedure; this knowledge is typically summarised in an overall assessment of the WoE. In

22

599 assessing the uncertainties associated with a particular read-across, it is important to put in

600 context both the problem and premise of the read-across. A statement of the problem includes

601 noting the target chemical(s), the apical endpoint to be read across and the purpose of the

602 prediction. Stating the target chemical(s) is critical, as it is one of the focal points of the exercise.

603 Scaling uncertainty is a formidable challenge (Péry, et al., 2013; Blackburn and Stuard, 2014).

604 While there is much agreement on what the essential issues of the read-across are that need to be

605 considered in assessing uncertainty, there is less agreement on what approach to use. At least

606 three approaches could be applied: 1) a sliding scale, which can be tailored to the particulars of

607 the read-across (i.e., problem and premise), 2) a weighted scale, where some issues or their

608 related narrative and/or question(s) used to frame the issue are weighed more than others, and 3)

609 pre-defined divisions, where all issues or their related narrative and/or questions are assigned a

610 value in a parallel fashion. The first two approaches, while interesting academic exercises are

611 likely to be too complex to be practical. Thus, the third approach, the pre-defined divisions

612 approach, is the most likely to be used. Within the latter approach, there is variability in the

613 number of divisions employed. A dichotomous decision scheme (i.e., accept or reject) does not

614 provide any refinement to the assessment; whereas, a multi-divisional scheme will provide the

615 opportunity to add confidence statements into the assessment (e.g., low. medium, high). A five-

616 division scheme (or larger) may offer too much subjectivity in assigning the division. The four-

617 division scheme (i.e., low, low to moderate, moderate and high) described by Blackburn and

618 Stuard (2014), appears to provide a balance between a high number of possible divisions and

619 reduced subjectivity in assigning the final division. The Blackburn and Stuard scheme provides

620 three divisions of uncertainty where the prediction may potentially be usable; with the fourth

621 division indicating high uncertainty such that the read-across method is unfit for data gap filling.

622 The "characteristics by uncertainty" for the low and low-to-medium divisions are much the

623 same, with latter divisions including a WoE evaluation. Initially, read-across case studies are

624 likely to involve extremely-well studied categories and analogues which fit the low uncertainty

625 division of Blackburn and Stuard (2014). However, in the future, the more common read-across

626 predictions, especially for chronic health effects, should include a WoE evaluation.

627 Uncertainty factors are used to build conservatism into assessments and address the unknown

628 associated with a prediction. Converting uncertainty "divisions" (as reported by Blackburn and

629 Stuard, 2014) to numerical uncertainty factors provides another challenge. Excluding the "high"

630 uncertainty division, since reaching this level of uncertainty precludes using read-across to fill a

631 data gap, one is left with assigning three uncertainty factors. There are a variety of numerical

632 scales (e.g., 1-2-3; 1-10-100; 1-3-10; 1-5-10) which may be employed to cover a three-division

633 scheme. A 1-2-3 method provides insufficient differentiation of uncertainty; conversely, a 1-10-

634 100 provides too much differentiation. The 1-3-10 method proposed by Blackburn and Stuard

635 (2014) remains a pragmatic and usable solution and is recommended for use at this time.

636 However, as case studies become available, especially for those where the read-across is less

637 conclusive (i.e., low-moderate or moderate), further evidence may become available to evaluate

638 this proposal more fully, for example to explore the difference in employing a 1-3-10 versus a 1-

639 5-10 quantification method.

640

641 TABLE 6 HERE

642

643 Table 6 summarises the main similarities that need to be considered when assessing a read-across

644 justification, along with how they may be related to specific levels of uncertainty. Table 6 also

645 demonstrates the value of including novel toxicological data to read-across predictions with the

646 aim of decreasing uncertainty. It is likely that uncertainty associated with core structure and

647 functional groups, as well as physicochemical and molecular properties, can be assessed

648 relatively easily. However, because of information gaps, it is likely that uncertainty associated

649 with comparable toxicokinetics and similar mechanistic and toxicological properties, especially

650 for chronic health endpoints, will be more difficult to assess.

651 Consideration of all the evidence (e.g., the uncertainties defined in queries such as Table 5,

652 supporting data and information etc.) provides the basis for the WoE. It is not only the quantity

653 and quality of evidence that affects WoE but also consistency across all aspects of the

654 information/data used to support the similarity rational and prediction. For example, whilst

655 relative uncertainties may be the same, it is intuitive that reading across from many-to-one with

656 consistent phenotypic expressions of toxicity provides a greater WoE than reading across from

657 many-to-one with varied phenotypic expressions of toxicity. This has particular implications in

658 Scenario 4 of Table 4 where multiple mechanisms of action may be present. In terms of

659 chemistry, a greater WoE is assigned when empirical and *in silico* estimates of chemical

660   properties are in good agreement with measured values. In a similar fashion, the WoE is

661   considered higher when empirical studies of metabolism (*in vivo* and/or *in vitro*) and model

662   predictions indicate similar metabolites. Mechanistic plausibility can be more difficult to

663   consider, however consistent empirical data for the target chemical and, where possible, the

664   target chemical and the source chemical(s) for the MIE and/or other KEs strengthens the WoE.

665   Similar arguments can be made for other relevant, *in vivo*, *in vitro* and *ex vivo* endpoints.

666   Concordance across other endpoints (where data exist) is also a relevant consideration. For

667   example, acute oral LD50 data are not part of the mechanistic understanding for oral repeated

668   dose toxicity but having a consistent trend in empirical data among category members may

669   improve the overall WoE.

670   A template has been provided to identify and assess uncertainty in a comprehensive and

671   transparent manner. The template is available in Appendix B and it is recommended for use to

672   assess the uncertainty associated with each similarity parameter used in the read-across and to

673   summarise these finding in a statement of uncertainty. The first aim of the template was to

674   identify the factors of the read-across that contribute to uncertainty in the prediction. These

675   include uncertainty associated with the scientific justification of the similarity that defines the

676   applicability domain of the category or source and target analogue, as well as the uncertainty

677   associated with the read-across. The second aim was to define levels of uncertainty and propose

678   quantitative factors for addressing each level.

679   Table B.1 of the template in Appendix B lists and describes the key issues of chemical,

680   transformation/toxicokinetic and toxicological similarity proposed to assess data uncertainty and

681   WoE (see tables A.1-A.8). The comment column is not intended to be all inclusive but rather

682   give an indication of the type information that may be included. Table B.2 of the template in

683   Appendix B provides the capability to assess the issues raised in Table 5 above. The aim was to

684   assess the non-similarity-based uncertainty associated with the read-across. The first item in

685   Table 5 focuses on the particular read-across problem being addressed. The second to fourth

686   items address the *in vivo* data relevant to the read-across. Items five and six relate to the

687   mechanistically-related *in chemico*, *in vitro* and "new methods" data. Item seven addresses the

688   overall WoE. While a ranking (i.e., low medium or high) is assigned to each item, the comment

689   section is considered to be more significant and hence of greater value. The overall ranking (low,

690   moderate, high) and a summary of the uncertainty associated with the definition of the similarity

691   of analogues or category members, as reported at the end of the relevant tables, is presented in a

692   text box in Appendix B.

693

5.   Workflow for Reporting a Read-Across Prediction

695   Existing workflows for reporting read-across predictions vary in detail, however the general

696   purpose is to: 1) describe the similarity rationale of the read-across in a transparent manner, 2)

697   document the logic and data leading to the read-across prediction so it can be recreated, 3)

698   describe and address the uncertainties, and 4) clarify the roles of any endpoint specific and/or

699   endpoint non-specific factors affecting the assessment.

700   In order to assist with developing a workflow for reporting, the combined process of chemical

701   category formation and toxicological read-across prediction can be sub-divided into distinct and

702   definable activities. Cronin (2013a) identified six such procedures associated with development

703   of a read-across prediction including: 1) the identification of the effect and/or endpoint to be

704   predicted by read-across and the ''target'' chemical(s), 2) the identification the source

705   chemical(s) and other chemicals "similar" to the target, 3) obtaining toxicity data for the

706   category members identified in 1 and 2, 4) definition of the chemical category, 5) making the

707   prediction of toxicity by read-across, and 6) fully documenting the prediction.

708   More recently, the OECD has provided reporting formats for analogue and chemical category

709   approaches (OECD, 2014a). The documentation of read-across predictions, which are largely

710   based on process of using the OECD QSAR Toolbox, includes a number of steps:

711   1) Formulate the problem (i.e., understanding assessment strategy and identify the critical data
712   needs).
713   2) Curate chemical structure of the target compound(s) and other category members.
714   3) Profile the target compound(s) and other category members.
715   4) Develop the similarity rationale for the read-across prediction.
716   5) Establish the category selection criteria and search for potential source analogues or
717   category members.
718   6) Gather data for the category members and construction of data matrix.
719   7) Assessing the adequacy and uncertainty associated with the read-across.
720   8) Applying read-across to fill the data gap.
721   9) Document the analogue/category and read-across prediction.

722   A workflow proposed for reporting a read-across prediction is presented in Appendix C. This

723   builds on the earlier efforts and reflects the essential points described in this paper to address

724   similarity, the data and to justify the validity of the prediction.

725

6. Discussion

727   A significant proportion of REACH registration dossiers include a read-across prediction

728   intended to fill information requirements for higher-tier toxicological studies. In fact, 75% of

729   registration dossiers include read-across or categorisation reasoning (ECHA 2014) by the

730   registrant.

731   Improvements in methodology to perform and report read-across prediction require an

732   understanding of the process, specifically around the concept of similarity with regard to two or

733   more chemicals. Berggren et al. (2015) noted that in considering chemical similarities there are

734   different aspects that must be assessed to make the read-across prediction scientifically justified.

735   These similarities include aspects of chemical stability, the possible formation of toxic

736   metabolites, different active functional groups that might lead to similar or dissimilar behaviours,

737   possible routes of exposure and concentrations at the target tissue, biotransformation (prior to

738   reaching, or at, the target organ), or observable trends with or without a mechanistic explanation.

739   To improve and standardise the development and reporting of a read-across prediction, it is,

740   therefore, useful to identify different scenarios by which a read-across prediction may develop.

741   While this is possible to do in several ways, the toxicokinetic fate of the substance, such as

742   whether the compound itself would be available in the target organ or whether it would be its

743   metabolites or reaction products leading to adverse effect, is a critical factor, especially for

744   chronic health effects (Berggren et al., 2015). In addition, Berggren et al. (2015) noted that the

745   toxicodynamic behaviour of the substance and compared similarities of chemicals based on their

746   assumed mechanism of action, including lack of biological activity, is critical to establish a read-

747   across justification.

748   Category-based read-across adaptations begin with the definition of a chemical category (i.e.,

749   establishment of the category's applicability domain). This definition is assumed to be related to

750   the toxicological property to be read across, which results from a trend observed when the

751   property to be read across is plotted against another property that is known for all members of

752   the category (i.e., an indication of toxicological relevance). Read-across to a target substance is

753   deemed possible when the target substance is an unambiguous member of the category and there

754   are one or more measured property(ies) to be read across for other members of the category.

755   Therefore, a category-approach read-across is based on grouping and may rely on one or more

756   observed trends. Category-approach read-across also covers cases where substances belonging to

757   a well-defined category all show the same type and value for the toxicological property to be

758   read across or do not show an effect at all (i.e., a 'low-toxicity' read-across case).

759   While there is no consensus, there appear to be four most likely scenarios where chemical

760   category formation and subsequent read-across may be used to fill a data gap, especially for

761   repeated dose toxicity. Scenarios for read-across in general are described in Table 4, more

762   specific scenarios for chronic endpoints are given in Table 7.

763

764   TABLE 7 HERE

765

766   It is important to remember that defining the criteria for category membership for a particular

767   scenario of chemical category formation and read-across is only the beginning of the exercise.

768   Improvement in the confidence of a read-across prediction can be made by added value in the

769   form of increased WoE. This added value may come from suggestions of how targeted testing

770   and "new-approach" data, especially when applied using the logic of the Safety Evaluation

771   Ultimately Replacing Animal Testing (SEURAT) conceptual framework (White and Knight,

772   2013), may be used to improve the read-across justification. The increase in justification will be

773   especially true if targeted testing focuses on the weak steps of the read-across argument. In other

774   words, an understanding of how targeted testing may reduce uncertainty is available, for instance

775   as stated in Table 6. The improvement of the robustness of the read-across predictions, when

776   further evidence is added can, in principle, be examined by various means before and after the

777   addition of further evidence.

778   The intention of this manuscript was to report progress in the development of proposed templates

779   and workflows for recording and evaluating traditional *in vivo* toxicology data, as well as

780   alternative methods (e.g., *in chemico, in vitro*) data. Additionally, the intent was to suggest

781   means to standardise the evaluation of similarity and uncertainty so as to enhance the robustness

782   of the read-across prediction and thereby make it more likely to gain regulatory acceptance.

783 Since there are various over-arching scenarios for category formation and read-across, it is
784 critical to not only state the target chemical and its missing endpoint value but also the
785 hypothesis and assumptions on which the read-across is based. A category/analogue hypothesis
786 typically makes references to several similarity rationales which delineate category membership.
787 For example, for a read-across adaptation of Scenario 3 noted in Table 7 it may be possibly to
788 report:

789 • Members of chemical category **A** are indirect-acting toxicants of n_1 to n_2 carbon atoms in
790 size with a molecular scaffolding of **B** and the primary functional group **C**.

791 • Category members elicit a similar chemical mechanism-of-action (e.g., electrophilic
792 reactivity via mechanism **D**), where metabolism via pathway **E** is the primary factor
793 driving the reactivity leading to oral repeated dose toxicity with symptoms/endpoints **F**.

794 • Category members show rapid and complete absorption from the gut, as the parent
795 compound with first past through oxidative metabolism in the liver to the corresponding
796 electrophile with mechanism **D**. Subsequently, the electrophile elicits the *in vitro*
797 outcome **G** at the cellular level leading to the *in vivo* outcome **F**.

798 • Category members have similar volatility, bioavailability and oral uptake.

799 • Reading repeated dose toxic outcome **F** for the source chemical **X** across to the target
800 chemical **Y** is supported by information and data on **A, B, C, D, E** and **G**.

801 Along the same theme, assessments of uncertainty may reveal there are no deficiencies in the
802 quality of the toxicology data to be read across (**F**), especially as compared to what is expected
803 from current standard test methods. However, assessment of uncertainty associated with
804 similarity justification reveals metabolism via pathway **E** to be the weak step of the read-across
805 argument. New methods data after target testing may reduce the uncertainty by strengthening this
806 step in the similarity argument.

807

813

814　**8. References**

815　Ball, N., Bartels, M., Budinsky, R., Klapacz, J., Hays, S., Kirman, C., Patlewicz, G., 2014. The

816　challenge of using read-across within the EU REACH regulatory framework; how much

817　uncertainty is too much? Dipropylene glycol methyl ether acetate, an exemplary case study.

818　Regul. Toxicol. Pharmacol. 68, 212-221.

819　Bemis, G.W., Murcko, M.A., 1996. The properties of known drugs. 1. Molecular frameworks. J.

820　Med. Chem. 39, 2887-2893.

821　Berggren, E., Amcoff, P., Benigni, R., Blackburn, K. Carney, E. Cronin, M., Deluyker, H.,

822　Gautier, F., Judson, R.S., Kass, G.E.N., Keller, D., Knight, D., Lilienblum, W., Mahony, C.,

823　Rusyn, I., Schultz, T., Schwarz, M., Schüürman, G., White, A., Burton, J., Lostia, A., Munn, S.,

824　Worth, A., 2015. Chemical safety assessment using read-across: How can novel testing

825　methods strengthen evidence base for decision-making? Environ. Health Perspect.,*in press*.

826　Blackburn, K., Bjerke, D. Daston, G., Felter, S. Mahony, C., Naciff, J., Robison, S., Wu, S.,

827　2011. Case studies to test: A framework for using structural, reactivity, metabolic and

828　physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological

829　assessments. Regul. Toxicol. Pharmacol. 60, 120-135.

830　Blackburn, K., Stuard, S.B., 2014. A framework to facilitate consistent characterization of read

831　across uncertainty. Reg. Toxicol. Pharmacol. 68, 353-362.

832　Commission of the European Communities (EC), 2006. Regulation (EC) No 1907/2006 of the

833　European Parliament and of the Council of 18 December 2006 concerning the Registration,

834　Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European

835　Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC)

836　No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive

837　76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC.

838　Off J Eur Union, L 396/1 of 30.12.2006.

839　Cronin, M.T.D., 2013a. An introduction to chemical grouping, categories and read-across to

840　predict toxicity, in: Cronin, M.T.D., Madden, J.C., Enoch S.J., Roberts, D.W. (Eds.) Chemical

841　Toxicity Prediction: Category Formation and Read-Across Applications. The Royal Society of

842　Chemistry, Cambridge, pp. 155-167.

843　Cronin, M.T.D., 2013b. Evaluation of categories and read-across for toxicity prediction allowing

844　for regulatory acceptance, in: Cronin, M.T.D., Madden, J.C., Enoch S.J., Roberts, D.W. (Eds.)

845   Chemical Toxicity Prediction: Category Formation and Read-Across Applications. The Royal
846   Society of Chemistry, Cambridge, pp. 155-167.

847   Cronin, M.T.D., Madden, J.C., Enoch S.J., Roberts, D.W., (Eds.), 2013. Chemical Toxicity
848   Prediction: Category Formation and Read-Across Applications, The Royal Society of Chemistry,
849   Cambridge.

850   Dimitrov, S., Mekenyan, O., 2010. An introduction to read-across for the prediction of effects of
851   chemicals, in: Cronin, M.T.D., Madden, J.C. (Eds.) In Silico Toxicology: Principles and
852   Applications. The Royal Society of Chemistry, Cambridge, pp. 372-384.

853   Enoch, S.J., Przybylak, K.R., Cronin, M.T.D., 2013. Category formation case studies, in: Cronin,
854   M.T.D., Madden, J.C., Enoch S.J., Roberts, D.W. (Eds.) Chemical Toxicity Prediction: Category
855   Formation and Read-Across Applications. The Royal Society of Chemistry, Cambridge, pp. 127-
856   154.

857   Enoch, S.J., Roberts, D.W., 2013. Approaches for grouping chemicals into categories, in:
858   Cronin, M.T.D., Madden, J.C., Enoch S.J., Roberts, D.W. (Eds.) Chemical Toxicity Prediction:
859   Category Formation and Read-Across Applications. The Royal Society of Chemistry,
860   Cambridge, pp. 30-43.

861   European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC), 2012. Category
862   Approaches, Read-across, (Q)SAR. Technical Report No 116. ECETOC, Brussels, Belgium.

863   European Chemicals Agency (ECHA), 2009. Practical Guide 6: How to Report Read-Across and
864   Categories, ECHA, Helsinki, ECHA-10-B-11-EN.

865   European Chemicals Agency (ECHA), 2011. The Use of Alternatives to Testing on Animals for
866   the REACH Regulation 2011, ECHA, Helsinki, ECHA-11-R-004.2-EN.

867   European Chemicals Agency (ECHA), 2013a. Grouping of Substances and Read-Across
868   Approach. Part 1: Introductory Note, ECHA, Helsinki, ECHA-13-R-02-EN.

869   European Chemicals Agency (ECHA), 2013b. Read-Across Illustrative Example. Part 2.
870   Example 1 – Analogue Approach: Similarity Based on Breakdown Products, ECHA, Helsinki,
871   ECHA-13-R-03-EN.

872   European Chemicals Agency (ECHA), 2014. The Use of Alternatives to Testing on Animals for
873   the REACH Regulation - Second report under Article 117(3) of the REACH Regulation.
874   Available:http://echa.europa.eu/documents/10162/13639/alternatives_test_animals_2014_en.pdf.

875   Hansch, C., Leo, A. 1979. Substituents Constants for Correlation Analysis in Chemistry and
876   Biology. John Wiley & Sons, New York.

877 Klimisch, H.-J., Andreae, M., Tillmann, U., 1997. A systematic approach for evaluating the
878 quality of experimental toxicological and ecotoxicological data. Regul Toxicol Pharmcol 25, 1-5.

879 Organisation for Economic Co-operation and Development (OECD), 2007. Guidance on
880 Grouping of Chemical. Environment Health and Safety Publications, Series on Testing and
881 Assessment No. 80, Report No. ENV/JM/MONO(2007)28, JT03232745, OECD, Paris.

882 Organisation for Economic Co-operation and Development (OECD), 2011. Report of the
883 Workshop on Using Mechanistic Information in Forming Chemical Categories. Environment
884 Health and Safety Publications, Series on Testing and Assessment No. 138, Report No.
885 ENV/JM/MONO(2011)8, JT03301985, OECD, Paris.

886 Organisation for Economic Co-operation and Development (OECD), 2013. Guidance Document
887 for Developing and Assessing Adverse Outcome Pathways (AOPs). Environment Health and
888 Safety Publications, Series on Testing and Assessment No. 184, Report No.
889 ENV/JM/MONO(2013)6, JT03338300,OECD, Paris.

890 Organisation for Economic Co-operation and Development (OECD), 2014a. Guidance on
891 Grouping of Chemicals: Second Edition. Environment Health and Safety Publications, Series on
892 Testing and Assessment No.194. Report No. ENV/JM/MONO(2014)4, JT03356214, OECD,
893 Paris.

894 Organisation for Economic Co-operation and Development (OECD), 2014b. Users' Handbook
895 Supplement to the Guidance Document for Developing and Assessing AOPs.
896 [ENV/JM/MONO(2013)6], OECD, Paris.

897 Organisation for Economic Co-operation and Development (OECD), 2014c. Draft Outline of
898 Future Cooperative Work on the Hazard Assessment of Chemicals. 52nd Joint Meeting of the
899 Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology.
900 Report No. ENV/JM/MONO(2014)100, OECD, Paris.

901 Patlewicz, G.Y., Ball, N., Booth, E.D., Hulzebos, E., Zvinavshe, E., Hennes. C., 2013a. Use of
902 category approaches, read-across and (Q)SAR: general considerations. Regul. Toxicol.
903 Pharmacol. 67, 1-12.

904 Patlewicz, G.Y., Roberts, D.W., Aptula, A., Blackburn, K., Hubesh, B., 2013b. Workshop: Use
905 of "read-across" for chemical safety assessment under REACH. Regul. Toxicol. Pharmacol. 65,
906 226-228.

907 Patlewicz, G., Ball, N., Becker, R.A., Booth, E.D., Cronin, M.T.D., Kroese, D., Steup, D., van
908 Ravenzwaay, B., Hartung, T. 2014. Read-across approaches – misconceptions, promises and
909 challenges ahead. ALTEX – Altern. An. Exp. 31 (4/14), 387-396.

Péry, A.R.R., Schüürmann, G., Ciffroy, P., Faust, M., Backhaus, T., Aicher, L., Mombelli, E., Tebby, C., Cronin, M.T.D., Tissot, S., Andres, S., Brignon, J.M., Frewer, L., Georgiou, S., Mattas, K., Vergnaud, J.C., Peijnenburg, W., Capri, E., Marchis, A., Wilks, M.F., 2013. Perspectives for integrating human and environmental risk assessment and synergies with socio-economic analysis. Sci. Tot. Environ. 456-457, 307-316.

Petkov, P.I., Patlewicz, G., Schultz, T.W., Honma, M., Todorov, M., Kotov, S., Dimitrov, S.D. Donner, E.M., Mekenyan, O.G., 2015. A feasibility study: Can information collected to classify for mutagenicity be informative in predicting carcinogenicity? Regulatory Toxicology and Pharmacology (in press).

Przybylak, K.R., Madden, J.C., Cronin, M.T.D., Hewitt, M., 2012. Assessing toxicological data quality: basic principles, existing schemes and current limitations. SAR QSAR Environ. Res. 23, 435-459.

Schultz, T.W., 2014. Chapter 2.6. Read-across as a basis for one of the SEURAT-1 proof-of-concepts and an overview of the outcome of the SEURAT-1 read-across workshop, in Gocht, T., Schwarz, M. (Eds.), Toward the Replacement of in vivo Repeated Dose Systematic Toxicity Testing. Volume 4: Implementation of the Research Strategy. COACH Consortium, Paris, pp. 72-80.

Steinmetz, F.P., Enoch, S.J., Madden, J.C., Nelms, M.D., Rodriguez-Sanchez, N., Rowe, P.H., Wen, Y., Cronin, M.T.D., 2014. Methods for assigning confidence to toxicity data with multiple values – Identifying experimental outliers. Sci. Tot. Environ. 482, 358-365.

Tollefsen, K. E., Scholz, S., Cronin, M. T., Edwards, S. W., de Knecht, J., Crofton, K., Garcia-Reyero, N., Hartung, T., Worth, A., Patlewicz, G., 2014. Applying Adverse Outcome Pathways (AOPs) to support Integrated Approaches to Testing and Assessment (IATA). Regul. Toxicol. Pharmacol. 70, 629-640.

van Leeuwen, K., Schultz, T.W., Henry, T., Diderich, R., Veith, G.D., 2009. Using chemical categories to fill data gaps in hazard assessment. SAR QSAR Environ. Res. 20, 207-220.

White, A., Knight, D., 2013. Chapter 4.10.8.3, Indicative Rational Integrated Assessment Strategy for SEURAT-1, in Gocht, T., Schwarz, M. (Eds.), Toward the Replacement of in vivo Repeated Dose Systematic Toxicity Testing. Volume 3: Implementation of the Research Strategy. COACH Consortium, Paris, pp. 326-331.

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

943 Wu, S., Fisher, J., Naciff, J., Laufersweiler, M., Lester, C., Daston, G., Blackburn, K., 2013.
944 Framework for identifying chemicals with structural features associated with the potential to act
945 as developmental or reproductive toxicants. Chem. Res. Toxicol. 26, 1840-1861.

946 Yang, L., Neagu, D., Cronin, M.T.D., Hewitt, M., Enoch, S.J., Madden, J.C., Przybylak, K.,
947 2013. Towards a fuzzy expert system on toxicological data quality assessment. Mol. Inf. 32, 65-
948 78.

949

950

Table 1. Criteria for Category Membership

1) A description of structural and chemical property similarities and differences among the category analogues and how these similarities and differences are linked to the read-across hypothesis.

   a. Supported by a data matrix of key structural and chemical properties.

2) A description of toxicokinetics and/or abiotic transformation similarities and differences among the category analogues and how these are linked to the read-across hypothesis.

   a. Supported by a data matrix of abiotic and biotic modification properties, including a summary of metabolic pathways and metabolites.

3) A description of the similarity and differences in the bioavailability of the chemical analogues and how these are linked to the read-across hypothesis.

4) A description of biological and toxicological similarities and differences among the category analogues and how these are linked to the read-across hypothesis.

   a. Supported by a data matrix of biological and toxicological properties including a summary of toxicological trends within the category.

951

952
953
954

Table 2. Similarities for Establishing a Toxicological Read-Across

1)	Molecular structure similarity including common chemical class and sub-class(es), similar molecular scaffold(s), similar numbers of carbon atoms and common constituents in the form of key substituent(s), structural fragment(s) and extended structural group(s).
2)	Similar physico-chemical and molecular properties, especially those that are linked to key factors that affect bioavailability toxicity (e.g., volatility, solubility, reactivity, etc.).
3)	Similar toxicokinetics.
4)	The same key abiotic transformation process (e.g., hydrolysis, autooxidation).
5)	The same key metabolic pathway(s) or pathway inhibition.
6)	Biotic and abiotic activation to the same or similar reactive chemical species.
7)	Abiotic (e.g. microbial) degradation to the same or similar chemical species.
8)	Similar structural alert, or toxicophore, (i.e., structural fragment(s) and extended structural group(s) experimentally demonstrated to be associated with a specific toxic effect that is causally linked with the *in vivo* endpoint which is read across).
9)	Mechanistic plausibility, especially in the form of a common Adverse Outcome Pathway (AOP) based Molecular Initiating Event (MIE) and /or key intermediate event(s) causally linked to the *in vivo* endpoint which is the basis of the read-across.
10)	Other data (e.g., *in vitro*, *in chemico*, *in silico*) relevant to the *in vivo* endpoint which are the basis of the read-across.
11)	Similarity in *in vivo* toxicological responses within the category

Table 3. Criteria to Establish Similarities for a Toxicological Read-Across

- What are the chemical identifiers and structure of the target substance(s) and the source analogue(s)? (**see Appendix A, Table A.1**)

- Define the similarity in the physico-chemical and molecular properties of the target substance(s) and the source analogue(s). (**see Appendix A, Table A.2**)

- Define the similarity of the key substituents, functional group(s) or extended fragment, generic class of chemicals and sub-class of the target of the target substance(s) and source analogue(s) have? (**see Appendix A, Table A.3**)

- Identify any structural differences between the target substance and source analogue(s)

- Establish how structural differences may affect toxicity (or otherwise) through similarities, for instance, in *in vivo* data

- Define the similarity in abotic transformations and/or toxicokinetics between the target substance and source analogue(s) (**see Appendix A, Table A.4**)

- Define the similarity in potential metabolic products between the target substance and source analogue(s) (**see Appendix A, Table A.5**)

- Define the similarity in toxicophores or structural alerts for causally-linked toxicological endpoints between the target substance and source analogue(s) (**see Appendix A, Table A.6**)

- Identify whether the target substance(s) and source analogue(s) have the same mechanistical plausibility and can be linked mechanistically to the same AOP, MIE or KEs (**see Appendix A, Table A.7**)

- Identify if the target substance(s) and source analogue(s) are linked by other toxicologically relevant data (**see Appendix A, Table A.8**)

980

981

982

983   Table 4. Summary of the Main Types of Read-Across Scenario

1. Chemical similarity of compounds that do not require (or do not undergo) metabolism to exert a potential adverse human health effect (i.e., direct-acting toxicants with a similar mode of toxic action)
2. Chemical similarity involving metabolism and resulting in exposure to the same/similar toxicant (i.e., indirect-acting toxicants with a similar mode of toxic action based on metabolites with the same mechanism of action)
3. Chemical similarity of compounds with low general or no toxicity (i.e., toxicants with no obvious reactive or specific mode of action)
4. Distinguishing chemicals in a structurally similar category with variable toxicities based on Mode of Action hypothesis (i.e., toxicants with high structural similarity but markedly different potency and/or phenotypic profiles)

984   [a]from Schultz (2014)

985

986   Table 5. Proposed Factors Affecting Uncertainty Associated with the Mechanistic Relevance and
987   Completeness of the Read-Across

1)	The problem and premise of the read-across. What is the level of complexity of the read across endpoint? What is the purpose of the exercise? What is the over-arching premise and scenario of the exercise?
2)	Number of source chemicals and their relative applicability domain(s); is it an analogue-or category-based read-across?
3)	Absence/presence of toxicity and relevant mechanisms e.g. whether mechanisms can be defined for non / low toxicity compounds.
4)	Quality of the *in vivo* apical endpoint data read across to include technical issues related to the performance (e.g., reliability accuracy, precision, repeatability and reproducibility of the manner in which apical *in vivo* data are generated). Is the data to be read across sufficient to meet the purpose of the exercise?
5)	Consistency in the severity of the apical *in vivo* hazard. Is the potency of the hazard consistent among the source chemicals?
6)	Robustness of the (*in chemico*, *in vitro* and/or other) data sets. How extensive are the relevant events empirically measured or modelled? What is the performance (e.g. in terms of reliability and reproducibility) of methodology for establishing these data?
7)	Concordance of the *in chemico*, *in vitro* and/or other data with regard to the intermediate and apical effects and potency data. What is the temporal and dose-response relationship between mechanistically-relevant endpoints?
8)	The overall Weight-of-Evidence (WoE) supporting the prediction. How many and how large are the mechanistically-related data gaps?

988

989

990 Table 6. Proposed key similarities relating to toxicological read-across and criteria for assessing

991 uncertainty (adapted from considerations in Blackburn and Stuard, 2014).

	Low uncertainty	Low-to-Moderate uncertainty	Moderate uncertainty	High uncertainty
Core structural similarity i.e., functional groups, extended fragments (especially those associated with chemical reactivity or its modification)	Highly similar	Highly similar	Similar	Differences in core structure and functional groups
Physico-chemical and molecular properties	Highly similar	Similar, having a consistent trend within values	Minor differences in values	Major differences in values
Abiotic transformation and/or toxicokinetics, especially metabolism e.g., leading to a common metabolite	Evidence demonstrating comparability in abiotic transformation and/or toxicokinetics and the same metabolites, similar bioavailability	Evidence demonstrating comparability in abiotic transformation and/or toxicokinetics and the same metabolites, similar bioavailability	No evidence that abiotic transformation and/or toxicokinetics, especially metabolism are dissimilar	Differences in abiotic transformation and/or toxicokinetics, especially in metabolism
Mechanism of action and toxicological properties	Evidence demonstrating comparability in mechanism supported by an AOP	Evidence demonstrating comparability in mechanism, possibly supported by an AOP	No evidence that mechanisms of action are dissimilar	Differences in mechanism of action and/or toxicological properties

992

993

1002

Table 7. The Most Likely Scenarios for a Chronic Toxicity Endpoint Read-Across

1003

1004

1) A 'low-toxicity' or negative read-across prediction; the category members have structural and chemical similarities, toxicokinetics are simple and based on well-documented or easily predicted (from related chemicals) pathways that lead to rapid degradation and/or elimination and/or generation of non-toxic metabolites and there is no obvious chemical reactivity or bioactivity or specific mode-of-action (i.e., members elicit generic effects but only at high concentrations).

2) A 'toxicity' or positive read-across prediction; the category members are direct-acting toxicants (i.e., no transformation or transformation does not drive the toxicity) with similar chemical mechanism-of-action and mode-of-toxic action (i.e., members elicit specific effects at similar internal concentrations or according to an established structural-related trend) leading to the same read across effect.

3) A 'toxicity' or positive read-across prediction; the category members are indirect-acting toxicants (i.e., transformation is the driver of toxicity), where the definitive toxicants has the same chemical mechanism-of-action and elicits the same mode-of-toxic action leading to the same read across effect.

4) A "toxicity' or positive read-across prediction; the category members are structurally and chemically highly similar and initially considered similar in bioactivity. Subsequently, new methods data reveal dissimilarity in bioactivity, often due to the inhibition of a degradative metabolic pathway. Thus, to obtain the appropriate read across endpoint effect (e.g., target organ and disease) requires sub-categorisation.

1005
1006

1007
1008

1009

1010 **Appendix A: Template for Reporting Data for Assessing Similarity of Analogues and**

1011 **Category Members for Read-Across**

1012 In Table A.1, the substance identification information, 2D structure and molecular formula data

1013 for the target substance(s) and proposed source analogue(s) are presented for comparison. The

1014 purpose of this information is to provide, in a transparent manner, a preliminary basis for

1015 assessing similarity.

1016 **Table A.1: Comparison of Substance Identification, Structure and Chemical Classifications**

	Target Substance	Analogue 1	Analogue n
Name			
CAS No:			
SMILES			
2D Structure			
Molecular Formula:			

1017

1018 In Table A.2, selected physico-chemical and molecular property data for the target substance(s)

1019 and proposed source analogue(s) are presented for comparison. The purpose of this information

1020 is to provide, in a transparent manner, the chemical property basis for assessing similarity. These

1021 data may assist in defining the boundaries of the applicability domain of the category, especially

1022 in regards to *in vivo* (bioavilability) and *in vitro* (solubility) toxicity.

1023 **Table A.2: Comparison of Physico-Chemical and Molecular Properties[1]**

	Target Substance	Analogue 1	Analogue n
Name			
Molecular Weight:			
Log Kow			
Vapor Pressure			
Density			
Melting Point			
Water Solubility			

	Target Substance	Analogue 1	Analogue n
Boiling Point			
pKa			

1024 [1]Value typically derived from EPISuite v4.0; [2]value for OECD QSAR Toolbox v3.3

1025

1026 In Table A.3, substituents, functional groups and extended structural fragments as well as

1027 chemical class data for the target substance(s) and proposed source analogue(s) are presented for

1028 comparison. The purpose of this information is to provide, in a transparent manner, the chemical

1029 structure sub-fragments and chemical class data for assessing similarity. These data may assist in

1030 defining the boundaries of the applicability domain of the category.

1031 **Table A.3: Comparison of Substituents, Functional Groups, and Extended Structural**
1032 **Fragments**

	Target Substance	Analogue 1	Analogue 2
Name			
Key Substituent(s)			
Functional Group(s)			
Extended Fragment(s)			
Chemical Class:			
Chemical Sub-Class:			
Chemical Sub-Class:			

1033

1034 In Table A,4, Transformation information and data for the target substance(s) and proposed

1035 source analogue(s) are presented for comparison. The purpose of this information is to provide,

1036 in a transparent manner, assessing similarity in abiotic transformation and/or similarity in the

1037 absorption, distribution, metabolism and elimination information.

1038 **Table A.4: Comparison of Abiotic Transformation and Toxicokinetics**

	Target Substance	Analogue 1	Analogue 2
Name			

	Target Substance	Analogue 1	Analogue 2
Abiotic Transformation			
Toxicokinetics			

1039

1040   In Table A.5, the predictions of potential metabolites derived from *in silico* tools data for the

1041   target substance(s) and proposed source analogue(s) are presented for comparison. A number of

1042   software platforms provide *in silico* predictions of metabolism. These are typically based on

1043   simulations run on the parent compound and initial metabolites using well-studied reactions,

1044   such as oxidation. *Files with name and structure of metabolites should be included for the sake*

1045   *of transparency.*

1046   **Table A.5: Comparison of Potential Metabolic Products**

	Target Substance	Analogue 1	Analogue 2
Name			
Liver metabolism simulator **Toolbox v3.3**			
Other software e.g. MetaPrint2D-React software			
Further software for prediction of metabolites			

1047

1048   In Table A.6, any toxicophore (i.e., toxic endpoint- specific structural alerts) data for the target

1049   substance(s) and proposed source analogue(s) are presented for comparison. A number of

1050   software platforms provide *in silico* predictions based on the presence of toxicophores (e.g.,

1051   OECD QSAR Toolbox, Derek Nexus, MCASE (Computer Automated Structure Evaluation)).

1052   The purpose of this information is to provide, in a transparent manner, any chemical structure

1053   sub-fragments linked to any relevant biological endpoint for assessing similarity.

1054   **Table A.6: Comparison Toxicophores**

	Target Substance	Analogue 1	Analogue 2
Name			

	Target Substance	Analogue 1	Analogue 2
Toxicophores			

1055

1056 In Table A.7, any mechanistic plausibility data including AOP-related, MIE, KEs, KERs or other

1057 mechanistically-relevant endpoints for the target substance(s) and proposed source analogue(s)

1058 are presented for comparison. With few exceptions (e.g., skin sensitisation), there are currently a

1059 limited number of endpoints for which AOPs, MIEs and KEs test methods and data have been

1060 formally developed and causally-linked, especially in the form of a KER. However, in the future,

1061 these pieces of information will become more and more available.

1062 **Table A.7: Comparison of Mechanistic Plausibility and AOP-Related Event Data**

	Target Substance	Analogue 1	Analogue 2
Name			
Mechanistic Plausibility			
Adverse Outcome Pathway or Mode of Toxic Action:			
Molecular Initiating Event:			
Key Event 1 etc.:			
Key Event Relationship 1 etc.:			
Other Mechanistically-Relevant Events			

1063

1064 In Table A.8, any other toxicologically relevant data for the target substance(s) and proposed

1065 source analogue(s) are presented for comparison. In some cases, there is relevant data from other

1066 sources (e.g., alternative species) which can assist in establishing mechanistic similarity.

1067 **Table A.8: Comparison of Other Toxicologically Relevant *In Vivo, In Vitro* and *Ex Vivo***
1068 **Data**

	Target Substance	Analogue 1	Analogue 2
Name			
Endpoint:			
Endpoint:			

1069

1070 **Appendix B: Template for Assessing Uncertainty for Read-Across**

1071 **Table B.1. Data Uncertainty and Weight-of-Evidence Associated with the Fundamentals of**
1072 **Chemical, Transformation/Toxicokinetic and Toxicological Similarity**

Similarity Parameter	Data Uncertainty [a] (empirical, modelled) (low, medium, high)	Strength of Evidence [b] (low, medium, high)	Comment
Substance Identification, Structure and Chemical Classifications			**Example:** All category members have CAS numbers, similar 2D structure and belong to the same chemical class/subclass.
Physio-Chemical & Molecular Properties	Empirical: Modelled:		**Example:** All category members are appropriately similar with respect to key physicochemical and molecular properties. There is a high degree of consistency between measured and model estimated values.
Substituents, Functional Groups, & Extended Structural Fragments			**Example:** Substituents, functional groups and extended structural fragments are consistent across all category members.
Transformation/ Toxicokinetics and Metabolic Similarity	Empirical: In vivo: In vitro: Simulated:		**Example:** Based on *in vivo* and *in vitro* data for multiple category members, there is evidence for similar toxicokinetics and metabolic pathways. Comparison of results from empirical studies and model predictions indicate similar metabolism among all category members.
Potential Metabolic Products			**Example:** Based on *in silico* metabolic simulations, potential metabolic products are similar among all category members.
Toxicophores /Mechanistic alerts			**Example:** Based on *in silico* profilers, all category members contain the same toxicophores.
Mechanistic plausibility and AOP-Related Events			**Example:** Although no AOP is currently available for the hypothesised toxicity pathway, many category members have been tested for what is generally accepted as a mechanistically-relevant event leading to the *in vivo* apical outcome of interest (*a citation could be provided*).
other relevant, *in vivo*, *in vitro* and *ex vivo* endpoints			**Example:** Although not part of the hypothesised toxicity pathway, many category members have been tested for rodent acute oral toxicity and there is general agreement among the reported LC50 values.

Overall uncertainty in similarity of category members: (Low, Moderate, High)

Summary: Key features of chemistry are similar within the category. Key features of transformation toxicokinetics and metabolism are common within the category. Category members are considered mechanistically similar. Category members exhibit a similar toxicological profile with respect to *in vivo* toxicity.

1073 [a] Uncertainty associated with underlying information/data used in the exercise
1074 [b] Consistency within the information/data used to support the similarity rational and prediction

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

Factor	Uncertainty (low, medium, high)	Comment
The problem and premise of the read-across		**Example:** The endpoint to be read across, developmental toxicity, for the category of branched carboxylic acids is well-studied and well-understood, The scenario of the read-across hinges on the inhibition of beta-oxidation of the acid and the subsequent build up of acid in the embryo leading to histone deacetylase inhibitors, increased cell adhesion and concomitant reduced cell motility, prevention of convergent extension during ontogenetic development.
In vivo data read across		
Number of analogues in the source set		**Example:** There are 3 suitable category members with *in vivo* apical endpoint data usable for read-across.
Quality of the *in vivo* apical endpoint data read across		**Example:** High quality empirical data from standard test guidelines for the stated regulatory endpoint exists for 1 category member. Similar non-standard test data of lower quality exists for 2 other category members. All these data are consistent in regards to qualitative description of effects and, where available, similar in quantification.
Severity of the apical *in vivo* hazard		**Example:** Potency data for the *in vivo* apical endpoint (25 mg/kg/day) is limited to a single source substance.
Evidence to biological argument for RA		
Robustness of analogue data set		**Example:** The available data from *in silico*, *in chemico* and *in vitro* studies for the category members were judged to be reliable and conducted under the appropriate conditions.
Concordance with regard to the intermediate and apical effects and potency data		**Example:** There is good agreement between the sequences of biochemical and physiological events leading to the *in vivo* apical outcome. There is consistency and high specificity for the association between the toxicophore and the structural domain of the category. There is general agreement among the dose-response relationships of the tested category members for mechanistically-relevant event(s) which may be assessed *in vitro*.
Weight of Evidence		**Example:** Overall the available information is generally consistent with the stated hypothesis. The sharp structural limitations of the category and narrow range of chemical properties strengthens the WoE. While the toxicokinetics data is limited, the lack of inconsistencies adds to the WoE. While the source substances data is limited, the fact that there is consistent relevant *in vitro* data for 50% of the category members, including the target chemical, strengthens the WoE.
Overall uncertainty of the read across: (Low, Medium, High) Uncertainty associated with the read-across is judged to be low.		

1079 **Summary of Uncertainty**

1080 | **Example:** Overall, the uncertainty in similarity of the analogues or category member is low. The
1081 | key features (i.e., A and B) relevant for toxicity are common within the category. There are only
1082 | minor differences among the analogues or category members with respect to physicochemical
1083 | properties. Analogues or category members are considered chemically similar (i.e., C).
1084 | Analogues or category members are judged to follow the same or similar metabolism. Analogues
1085 | or category members exhibit a similar toxicological profile (i.e., D and E) with respect to the
1086 | endpoint in question. It is concluded that the structural difference between analogues,
1087 | hydrocarbon chain length, has no significant impact on the toxicity being read across.

1088

1089

Appendix C. Work Flow for Reporting a Read-Across Prediction

1. Statement target substance(s) and the regulatory endpoint(s) that is to be read across.

The specific data gap to be filled by the prediction needs to be clearly defined by listing the chemical(s) and toxicity endpoint(s) (i.e., property(s)) for which the read-across prediction is proposed.

2. Description of the analogues or members of the category.

2.1. Premise

A premise for the basis of the analogue or category needs to be presented. This hypothesis should note the relational chemical, toxicokinetic and biological/toxicological features (i.e., structural similarities) which are deemed to be collectively relevant to the endpoint(s) being read across and common to target and source substance or all members of the category.

2.2. Justification

The analogue or category should be justified based on available experimental data, especially for the source substance(s). This is a description of the experimental toxicological data for the analogues or category members, presented in a narrative fashion. Typically, this justification will include endpoint-related mammalian toxicity data via appropriate exposure schemes, toxicokinetic and transformation information, as well as relevant *in vitro* data and structure-activity relationships. These data should demonstrate that the quality and quantity of *in vivo* data to be read across is sufficient to proceed with the exercise. Moreover, these data should be summarised to show the robustness of the read-across and include any indication of data trend(s) within the category for the different endpoints noted.

2.3. Applicability domain

In a category approach, the applicability domain of the category is described by inclusion and/or exclusion rules that identify the extent of values for category members within which reliable predictions can be made. Examples of this are the range of 1-octanol/ water partition coefficients values, functional groups or carbon chain lengths within which the category is appropriate.

2.4. Analogues or category members

Analogues or all members of the category, including target(s) and source substance(s), incorporated in the read-across exercise need to be described in a comprehensive fashion that

1120 takes into account unique substance identifiers such as, names, chemical structures and CAS

1121 numbers.

1122 2.5. Purity/impurities

1123 A purity/impurity profile for each analogue listed in 2.4 needs to be cataloged. The potential

1124 impact of impurities on the endpoint(s) being considered in the adaptation should be identified.

1125 *3. Data matrices for assessing similarity*

1126 Appendix A presented the template for assessing similarity. These data matrices are the central

1127 part of the workflow. They are likely to be the first items examined in any assessed. Data should

1128 be reported clearly, logically and unambiguously. The key study results should be noted and

1129 referenced. The distinction between experimentally measured and model-derived data should be

1130 noted.

1131 *4. Statement of uncertainty*

1132 Appendix B presented the template for assessing uncertainty. This section concludes with a

1133 narrative summary of the uncertainty. Particular consideration needs to be given to pointing out

1134 what are considered to be the weak steps of the read-across argument; why they are considered

1135 weak and how they impact the uncertainty of the read-across prediction.

1136 *5. Statement of the conclusions*

1137 Lastly, an overall concluding statement is made with regard to the category and the read across

1138 prediction relevant to the regulatory decision (e.g., hazard identification, classification and

1139 labelling, risk assessment, etc.) being considered. This should include making the prediction of

1140 toxicity by read-across and fully documenting the prediction to include clarifying the roles of any

1141 endpoint specific and/or endpoint non-specific factors affecting the prediction.