1   **A Systematic Analysis of Read-Across Adaptations in**

2   **Testing Proposal Evaluations by the European Chemicals Agency**

3

4   Hannah M. Roe[1], Han-Hsuan D. Tsai[1],

5   Nicholas Ball[2], Fred A. Wright[3], Weihsueh A. Chiu[1], Ivan Rusyn[1,*]

6

7   [1]Department of Veterinary Physiology and Pharmacology, College of Veterinary Medicine and

8   Biomedical Sciences, Texas A&M University, College Station, TX, USA

9   [2]Dow Chemical Company, Horgen, Switzerland

10   [3]Departments of Statistics and Biological Sciences, North Carolina State University, Raleigh, NC,

11   USA

12

13   * Corresponding author: Ivan Rusyn, MD, PhD, e-mail: irusyn@tamu.edu

Abstract

An important element of the European Union's "Registration, Evaluation, Authorisation and Restriction of Chemicals" (REACH) regulation is the evaluation by the European Chemicals Agency (ECHA) of testing proposals submitted by the registrants to address data gaps in standard REACH information requirements. The registrants may propose adaptations, and ECHA evaluates the reasoning and issues a written decision. Read-across is a common adaptation type, yet it is widely assumed that ECHA often does not agree that the justifications are adequate to waive standard testing requirements. From 2008 to August 2023, a total of 2,630 Testing Proposals were submitted to ECHA; of these, 1,538 had published decisions that were systematically evaluated in this study. Each document was manually reviewed, and information extracted for further analyses. Read-across hypotheses were standardized into 17 assessment elements (AEs); each submission was classified as to the AEs relied upon by the registrants and by ECHA. Data was analyzed for patterns and associations. Testing Proposal Evaluations (TPEs) with adaptations comprised 23% (353) of the total; analogue (168) or group (136) read-across adaptations were most common. Of 304 read-across-containing TPEs, 49% were accepted; the odds of acceptance were significantly greater for group read-across submissions. The data was analyzed by Annex (i.e., tonnage), test guideline study, read-across hypothesis AEs, as well as target and source substance types and their structural similarity. While most ECHA decisions with both positive and negative decisions on whether the proposed read-across was adequate were context-specific, a number of significant associations were identified that influence the odds of acceptance. Overall, this analysis provides an unbiased overview of 15 years of experience with testing proposal-specific read-across adaptations by both registrants and ECHA. These data will inform future submissions as they identify most critical AEs to increase the odds of read-across acceptance.

Keywords: ECHA, Read-across, OECD test guideline studies, adaptations, REACH

Introduction

Testing Proposals are a critical element of the European Union's (EU) "Registration, Evaluation, Authorisation and Restriction of Chemicals" (REACH) regulation [Article 40(2)] (European Council, 2006) insofar they are a mechanism to eliminate unnecessary testing in animals and other model organisms and ensure that the most appropriate tests are performed. These submissions are prepared by the registrants where they identify data gaps in complying with the standard information requirements for their registration type (ECHA, 2011). They provide the European Chemicals Agency (ECHA) the opportunity to comment on the proposed studies and to suggest refinements, so the information obtained from any new testing is most informative for hazard and risk characterization. Annex XI of REACH provides a range of possible adaptations to the standard information requirements (European Council, 2006). Among potential adaptations, read-across is one of the major methods used to fulfil information requirements in REACH (ECHA, 2023a). The Organisation for Economic Co-operation and Development (OECD) and ECHA published guidance on read-across (ECHA, 2015, 2017; OECD, 2017) and there are many authoritative commentaries form diverse stakeholders on ways to improve read-across in regulatory submissions (Ball et al., 2016; Pestana et al., 2022; Patlewicz et al., 2014; Blackburn and Stuard, 2014; Rovida et al., 2020).

Few studies exist that systematically evaluated regulatory decisions and the reasoning that regulators used to accept or reject read-across hypotheses. A review of published decisions by ECHA that were available as of July 2015 was a product of multi-stakeholder collaboration (Ball et al., 2016). Both compliance checks (CCH, 524 documents) and testing proposal evaluations (TPE, 388 documents) were examined with regards to the relative successes and pitfalls of different scientific arguments used in proposed read-across hypotheses. The timing of that publication coincided with the publication of the first edition of ECHA's read-across assessment framework (RAAF) (ECHA,

62   2015);   additional   guidance   (ECHA,   2017,   2022,   2020)   and   recommendations

63   (https://echa.europa.eu/recommendations-to-registrants) has been provided by ECHA. The study by

64   Ball et al (2016) summarized the state of the art in read-across based on their analysis of the data

65   available   then,   and   highlighted   the   areas   where   improvements   were   needed   in   improving

66   justifications and informing the registrants about best practices and successful cases. It was also

67   acknowledged that RAAF guidance was likely to have a major impact on read-across adaptations

68   and that additional systematic analyses will be needed.

69   Another   example   is   a   recent   systematic   analysis   of   read-across   in   REACH   registration

70   dossiers, data was extracted for target-source analogue pairs for mono-constituent substances from

71   the information in IUCLID (International Uniform Chemical Information Database) (Patlewicz et

72   al., 2024). The authors looked only at the data provided in the submission dossier and did not

73   consider if ECHA has evaluated these through the process TPE or CCH. They also focused on the

74   entries where read-across was used to satisfy information requirements for repeated dose toxicity or

75   developmental toxicity studies – standard test requirements for high-tonnage substances that require

76   the use of a large number of animals (Taylor et al., 2014b; Rovida et al., 2023). The analyses were

77   restricted   to   substances   with   defined   organic   structures,   the   final   dataset   comprised   270   target

78   substances and 259 source substances. The study focused on examining physicochemical, structural

79   and metabolic similarity between source and target substances, as well as on the analysis of dose-

80   response data from the animal studies – comparing the data from IUCLID to predictions using the

81   United   States   Environmental   Protection   Agency   (US   EPA)   generalized   read-across   (GenRA)

82   approach (Helman et al., 2019). This study concluded that identification of suitable analogues for

83   read-across is not only a challenge with respect to defining a similarity cutoff, but also with respect

84   to finding the substances that already have data that satisfy standard test requirements. Collectively,

85   the study found that low structural similarity was a common occurrence in REACH submissions

86 based on read-across. They also concluded that GenRA provides more conservative estimates for

87 dose-response analysis of chemical effects.

88 Overall, the discussions on the best practices to perform and evaluate read-across are critical

89 to increase the familiarity with this approach, and for strengthening the justifications for read-across

90 adaptations in regulatory submissions. Active discourse between the industry and regulators

91 continues and progress is being made to increase regulatory acceptance of read-across submissions.

92 Concomitantly, it is important to determine what do successful examples of accepted read-across

93 adaptations look like and whether there are common themes in accepted and rejected submissions

94 to ECHA. The extensive database of ECHA decisions on both CCH and TPE submissions now exists

95 because nearly 15 years elapsed since the first implementation of review of submissions under

96 REACH regulation. The current study focused on TPE decisions that were public as of August 2023

97 – a total of 2,630 Testing Proposals were submitted to ECHA from 2008 to 2023; of these, 1,538

98 had published decisions and were evaluated herein. Each document was manually reviewed, and

99 information extracted in a systematic approach. Read-across hypotheses were standardized into 17

100 assessment elements (AEs), and each submission was classified based on these AEs. Data was

101 analyzed for patterns, including Annex (i.e., tonnage), test guideline study, read-across hypothesis

102 AEs, and the structural similarity of target and source substances.

103

104 **Methods**

105 *Review of ECHA TPE Decisions and Data Extraction*

106 ECHA publishes its decisions on the testing proposal evaluation (TPE) on its website

107 (https://echa.europa.eu/information-on-chemicals/dossier-evaluation-status). On August 11, 2023,

108 we exported information on the identities of the substances listed (i.e., the target substance), stage

109 of their evaluation process, and the date the decision (if any) was issued. Among all downloaded

5

110   records, 2,636 records were identified as TPEs, of which 1,538 included a link to a publicly available

111   decision. This information is listed as part of **Supplemental Table 1**.

112   All TPEs with available decisions (n=1,538) were downloaded from the abovementioned

113   website in a PDF (portable document format) and manually evaluated for information on the

114   registration and missing data required to address data gaps for evaluation under REACH. It should

115   be noted that while PDF files are machine-readable, about half of the files we examined were

116   scanned images of varying quality. Attempts to perform optical character recognition processing of

117   these files resulted in poor text quality that would make machine reading difficult to impossible. In

118   addition, the terminology used by ECHA, as well as document formats, have changed over time.

119   Collectively, attempts to automatically process these documents for information retrieval were

120   vacated in favor of manual examination of each document. Information extracted from each PDF at

121   this stage included Annex, the proposed tests to fulfill the data gap, and ECHA's decision on the

122   proposal (for more detail, see **Supplemental Table 1**). This step included identifying documents in

123   which adaptations (e.g., read-across) were proposed; ECHA's decision(s) on the proposed adaptation

124   was also recorded. Of the 1,538 TPEs examined, 310 documents were identified as containing read-

125   across adaptations. However, when evaluating testing proposal with read-across, some documents

126   were found to include more than one read-across (either a hypothesis or end-point). In total, 314

127   records were included in the final analysis.

128   Next, the documents that included read-across adaptations were evaluated in greater detail

129   and additional information was extracted. First, we recorded the type of read-across, analogue or

130   group/category, as well as the EC number for each source substance(s). The EC numbers for the

131   target and source substances were used to search the ECHA database to identify the type of each

132   substance from the registration dossier. ECHA classifies substances into three primary types –

133   mono-constituent, multi-constituent, or substances of "unknown or variable composition, complex

6

134   reaction products, or biological" (UVCB). There were instances where the substance type could not

135   be ascertained from the TPE because it was redacted, or the registrant had ceased manufacturing

136   and the registration dossier was no longer publicly accessible. We also recorded separately what

137   OECD test(s) were proposed by the registrants to be adapted through read-across and ECHA's

138   decision on each adaptation. Second, to enable evaluation of the read-across hypotheses in each

139   TPE, we defined 17 assessment elements (AEs, **Table 1**) based on (i) information in ECHA's RAAF

140   guidance documents (ECHA, 2015, 2017) and (ii) additional considerations presented in TPEs that

141   were not part of RAAF. The AEs were grouped into three categories; those based on toxicodynamic

142   considerations, on toxicokinetic considerations, and other assessment considerations. Additional

143   details on each AE, including corresponding explanation from RAAFs and/or TPE decisions (where

144   applicable) can be found in **Supplemental Table 2**. For each TPE, AEs were recorded separately as

145   those proposed by the registrant as part of their read-across hypothesis, and those used by ECHA in

146   justifying their decision. Third, for each read-across decision, we noted whether submission was

147   based on analogue or group/category, and whether the proposed standard test requirement adaptation

148   was found by ECHA to be satisfactory or not.

149   *Statistical Data Analysis*

150   Statistical analyses were performed using *R* 4.40 and GraphPad Prism (v. 10.2.2, GraphPad

151   Software, La Jola, CA). Using predictors such as OECD test and the proposed AEs in testing

152   proposals and decisions, we constructed 2×2 tables (Acceptance/Rejection of a testing proposal vs.

153   other binary predictors). These tables were analyzed using Fisher's exact test (`fisher.test`) for

154   *p*-values, odds ratios, and associated confidence intervals. As these predictors were considered to be

155   of independent interest, nominal confidence intervals and *p*-values are reported. Multiple logistic

156   regression using `glm()` function in R was performed using submission year, decision year, group

157   vs. analogue-based read-across, and each AE. For several AEs, the small number of associated

7

158   testing proposals could produce unstable estimates resulting in infinite standard errors, and these

159   AEs were removed in a single iterative step. Reported *p*-values for the multiple logistic regression

160   were corrected using the Benjamini-Hochberg *q*-value via `p.adjust` and *q*<0.05 was considered

161   as significant.

162   *Analysis of Chemical Similarity*

163   For testing proposals with analogue read-across for mono-constituent target and source

164   substances, the EC number of each substance was used to obtain "simplified molecular input line

165   entry system" (SMILES) identifiers for subsequent analysis of structural similarity. For each

166   substance, SMILES were used to obtain Morgan Fingerprint information using the "*rcdk*" R package

167   (Guha, 2007) to calculate the similarity between source and target substances. The Jaccard

168   Similarity Score (Chung et al., 2019) was derived for each source and target compound pair. A t-test

169   was then performed to compare the Jaccard similarity scores between accepted and rejected cases.

170   To account for the possibility that acceptance rates might differ for scores greater than a threshold,

171   a rigorous cutpoint analysis was performed to find an "optimal" threshold, while accounting for the

172   implicit multiple testing in such an analysis. Here permutation-based significance testing was

173   employed as follows. First, Jaccard similarity scores were screened to identify an optimal cutpoint.

174   Each possible cutpoint (defined as an observed Jaccard similarity score in the dataset) was evaluated,

175   and the cutpoint that returned the smallest nominal *p*-value using Fisher's exact test

176   (acceptance/rejection vs. Jaccard similarity < or ≥ cutpoint) was identified as optimal. Second, a

177   null population was constructed by permuting the accepted and rejected labels 100,000 times, and

178   for each permutation these labels were combined with the observed Jaccard Similarity Scores.

179   Finally, an empirical *p*-value was determined by computing the proportion of permuted minimum

180   nominal *p*-values that were ≤ the minimum nominal *p*-value from the first step, and the final odds

181   ratio using the associated optimal cutpoint.

182

183   **Results**

184   We first examined all publicly accessible TPE decisions by ECHA (1,538 were public as of

185   August 11, 2023) with respect to whether any adaptations were proposed (**Figure 2**). We found that

186   over ¾ of TPEs (1,185 or 77% of all TPEs) did not include adaptations, i.e., the registrants

187   acknowledged that they need to perform additional tests as a remedy to addressing the data gap(s)

188   in standard information requirements for their substance(s). The remaining 353 testing proposal

189   submissions (23% of total) contained one or more adaptations, most of these (86% or 304) contained

190   some form of read-across reasoning for one or more tests. Among the remaining 49 (14%) testing

191   proposals, 39 contained adaptations other than read-across, such as QSAR or Weight of Evidence;

192   these were not evaluated because they were not based on read-across. Another 10 submissions

193   contained read-across adaptations but were not examined herein because ECHA rejected these

194   submissions as redundant (e.g., required data was already available in another registration or

195   endpoint proposed not necessary based on tonnage band) without weighing in on the registrant-

196   proposed read-across arguments (for details, see **Supplemental Table 3).** Among 304 examined

197   TPEs with read-across adaptations, more than half (55%) were of the analogue type – ECHA

198   accepted read-across justifications for 39% of these submissions. Among testing proposals that

199   proposed the group type of read-across, a larger fraction (62%) was found to be acceptable. Overall,

200   based on the published TPE decisions from 2008 to 2023, the odds that a testing proposal with read-

201   across hypothesis would be found adequate by ECHA for group read-across submissions were 2.6

202   times as large as that for analogue read-across submission ($p < 0.05$, Fisher's exact test). It should be

203   noted, however, that this result should be interpreted with caution because if a read-across

204   hypothesis/justification was found to be satisfactory for a group of similar substances, then several

9

205 "positive" decisions might ensue thus amplifying the numbers as compared to analogue-type read

206 across submissions.

207   The process of evaluation of testing proposals by ECHA has been characterized as "lengthy

208 and bureaucratic" (Taylor et al., 2014b), yet ECHA had to rapidly implement a substantial new

209 regulatory framework, as well as establish internal expertise and competencies for evaluating a large

210 number of testing proposals and registrations. **Figure 3A** shows the time trends in the public release

211 of TPE decisions (bars) and the fraction of TPEs with read-across adaptations (line). Starting in

212 2012, when the output became more uniform, between 58 and 194 (126 ± 42, mean $\pm$ S.D.) TPE

213 decisions were published by ECHA each calendar year. The fraction of TPEs with read-across

214 adaptations varied even more widely, from 6.2% to 42% ($22.5 \pm 11.2\%$). **Figure 3B** shows the data

215 for published TPEs with read-across adaptations. The stacked bars show the numbers of proposals

216 with analogue and group read-across types. The line shows the fraction of read-across-containing

217 submissions that were deemed acceptable by ECHA. There appears to be a large difference in the

218 rate of acceptance with much higher acceptance rate in 2012-2015, before ECHA published final

219 guidance on read across (ECHA, 2015). For example, in 2014, ECHA released decisions on a group

220 of "Higher Olefins" substances which contained 21 substances. The decrease in acceptance of

221 testing proposals between 2015 and 2017 may reflect the time needed to standardize the evaluation

222 process according to RAAF. Overall, the time-dependent trends for the rate of testing proposal

223 acceptance when either proposal submission year or decision publication year is considered, are

224 slightly negative (slopes of -.078 and -0.101, respectively) but not significant.

225   There are three main substance types recognized under REACH – mono-constituent, multi-

226 constituent, and UVCBs [unknown or variable composition, complex reaction products or of

227 biological materials] (ECHA, 2023b). **Figure 4A** shows that over 90% of published TPE decisions

228 were on mono-constituent and UVCB substances. While the relative proportion of accepted read-

across adaptations is about 50% for both, analogue submissions were most common for mono-constituent substances, while for UVCBs it was the group-type read-across that dominated. While the numbers for multi-constituent substances are relatively small, most submissions contained the analogue type read-across and the majority of these were not accepted. **Figure 4B** shows what source substance types were used by the registrants in read-across adaptations for each target substance type. While it is not surprising that most (79% for mono-constituent and 75% for UVCB) of the time the target and source substances were of the same type, it is curious that in some instances the registrants attempted to read-across from a UVCB to a mono-constituent substance. It also appears that when a UVCB substance was used as a source for another UVCB, more instances of the read-across adaptation were deemed inadequate as opposed to when a mono-constituent substance was used as the source; however, the differences were not significant between source/target pairs for either substance type.

The REACH regulation established information requirements for substance registration (ECHA, 2011); these are based on the annual tonnage produced or imported into the European Union – the higher the tonnage, the greater number of studies that must be done (Botham et al., 2023). **Figure 5** shows the number of read-across adaptations across tonnage bands, from Annex VII (1-10 tons) to Annex X (over 1,000 tons). Recent registration data from ECHA (ECHA, 2023a) shows that among ~12,500 registered substances, most (39%) are of the lowest tonnage/data requirements. Substances that are subject to most comprehensive testing are in Annexes IX and X, these comprise about 19% of the total for each Annex (2,346 and 2,335, respectively). When substances with testing proposals are considered (**Figure 5A**), the trends are largely reversed – few Annex VII (3.1%) and VIII (9.8%) substances had published TPE decisions, while the bulk of the evaluations were for Annex IX (52.5%) and Annex X (34.6%) substances. All animal tests required by Annexes IX and X but not yet available require a testing proposal to be submitted, while Annexes VII and VIII only

253   do under certain circumstances (ECHA, 2011). Most TPEs with read-across adaptations were for

254   Annex IX; however, the relative proportion of such submissions is similar across four tonnage

255   bands. When TPEs with read-across adaptations are compared by tonnage band, read-across type,

256   and ECHA decision (**Figures 5B-C**), it is evident that most rejected testing proposals with read-

257   across adaptations for substances in Annexes VIII and IX were of the analogue type. The analysis

258   of the odds that a proposed read-across type would favor a positive outcome shows that indeed, the

259   odds of acceptance are significantly greater for group read-across adaptation for Annex VIII and IX

260   substances (**Figure 5D**). While the odds were highest for Annex VIII substances, the 95%

261   confidence interval was also wide, owing to a relatively smaller number of observations.

262   We also stratified the data by the type of a guideline test that was considered a data gap and

263   where a read-across adaptation was proposed to fulfill data requirement(s) for registration under

264   REACH. Most published TPE decisions concerned test of health effects in mammalian systems

265   (**Figure 6A**). Among the so-called OECD test guideline (TG) 400 series assays which are designed

266   to evaluate "health effects", two assays predominated – TG 414 (Prenatal Developmental Toxicity

267   Study) and TG 408 (Repeated Dose 90-Day Oral Toxicity Study in Rodents). **Figure 6B** shows the

268   types of tests in OECD Guidelines for the Testing of Chemicals that are listed in sections 2 (Effects

269   on Biotic Systems) and 3 (Environmental fate and behavior). Among these, the largest number of

270   submissions was for TG 211 (Daphnia magna Reproduction Test). The fraction of submissions that

271   involved read-across adaptations for tests with the greatest number of substances ranged from 14%

272   to 29%. **Figures 6C-D** show the proportions of accepted and rejected TPEs for tests with read-across

273   adaptations; most accepted read-across adaptations were of the group type. When the odds of

274   acceptance for group vs analogue read-across were calculated, the significantly greater odds were

275   for any test and TG 414 and TG 408 – a group read-across argument favors acceptance of an

276   adaptation for these tests (for details, see **Supplemental Table 4**).

12

277   The type of read-across adaptation (analogue or group) may not be a matter of choice for

278   each testing proposal depending on what source substance(s) with appropriate data are available.

279   However, the registrants have options in building scientific support for their read-across hypothesis

280   and can present a number of justifications based on ECHA guidance (ECHA, 2017, 2015, 2022,

281   2020) or other considerations (Ball et al., 2016; Beal et al., 2022; Berggren et al., 2015). It is

282   generally accepted that a number of assessment elements, based on the considerations for how

283   "similarity" is established between the target and source compounds, are needed to build the overall

284   argument and justify the proposed adaptation. Because the decisions we evaluated spanned almost

285   15 years during which the best practices for read-across justifications have been evolving, we

286   decided to craft "assessment elements" based on both formal ECHA guidance, as well as the

287   arguments that were used in the decisions. The short list of assessment elements we used in the

288   systematic analysis is shown in **Table 1**; additional information and example text from ECHA

289   guidelines and decisions can be found in **Supplemental Table 2**. Each TPE decision document with

290   read-across adaptation was examined with respect to what assessment elements were used by the

291   registrants and separately by ECHA in their decision. Such analysis allows us to determine not only

292   the frequency of use for each assessment element, but also the patterns of whether successful

293   proposals have relied on a particular element or combination thereof.

294   **Figures 7A-B** show that among 17 possible assessment elements (AEs) we used in this

295   evaluation, several have been used far more often than others. For example, arguments in support

296   of structural/physicochemical similarity between target and source substances (AE 1) was included

297   in virtually every TPE submission with read-across adaptation; however, there was no difference

298   between successful and unsuccessful TPEs (for details, see **Supplemental Table 4**). Likewise, a

299   large proportion of submissions relied on the reasoning that bridging studies from assays other than

300   OECD test guidelines (AE 16) could aid in rationalization of the similarity argument; however, both

13

301   successful and unsuccessful TPEs had them at the same rate. However, what is also evident from

302   these figures is that the relative proportion of analogue vs group read-across was not identical

303   between accepted and rejected TPEs. When odds of acceptance were analyzed for each assessment

304   element with sufficient data (**Figure 7C**), we found that group-based read-across hypotheses that

305   included AE 1, AE 8 (bringing studies that are OECD TG-type), and AE 15 (characterization of

306   constituents for multi-constituent or UVCB substances) were significantly more likely to be

307   accepted. **Figure 7D** shows the results from multiple logistic regressions, where submission year,

308   decision year, group vs. analogue-based RA, and AEs were included in the same model. Overall, for

309   a testing proposal with read-across to be successful, inclusion of AE 3 (qualitative identification of

310   common biological targets without common metabolites) and AE15 significantly increased the odds.

311   By contrast, inclusion of the information in support of AE 13 (identification of the metabolites of

312   source and target) significantly increased the odds that the TPE will not be accepted. When the

313   similar analyses were done separately for either mono-constituent substances, or UVCB and multi-

314   constituent substances combined, the patterns were mostly similar but the significance for each

315   assessment element was different. For mono-constituent substances, the overall direction was

316   mostly the same as for all substances combined. However, the odds of rejection were considerable

317   for UVCBs when AE 13 and AE 16 were part of a read-across hypothesis. Inclusion of the data

318   showing quantitative evidence for common biological targets without common metabolites (AE 4)

319   appeared to have a significant positive impact on the odds of acceptance, but this result is based on

320   extremely few observations (only two within UVCB TPEs).

321   We also considered whether in unsuccessful TPEs with read-across adaptations (**Figure 7B**),

322   the assessment element reasoning was presented by the registrants but not accepted, or if ECHA

323   pointed out that certain types of data/reasoning were needed to accept the proposed adaptation.

324   **Figure 8** shows the outcome of this analysis (for details, see **Supplemental Table 5**). In many

325 instances, ECHA did not agree with the argumentation for AE 1 even though submissions contained

326 information on this topic. The largest fraction of rejected submissions needed additional guideline

327 bridging studies (AE 8), characterization of constituents (AE 15), or additional characterization of

328 metabolites and/or impurities (AE 14). Characterization of the bioavailability of the parent molecule

329 (AE 11) and quantitative analysis of common biological targets (AE 4) were frequently desired by

330 ECHA but not provided by the registrants.

331 Because AE 1 is the core element of building a similarity argument, we examined structural

332 similarity between target and source compounds in TPEs with read-across adaptations. This analysis

333 was restricted to mono-constituent substances for which chemical descriptors can be calculated,

334 similar to the approach in (Patlewicz et al., 2024). In these analyses (**Figure 9**) we used structural

335 similarity, based on Morgan fingerprints, and calculated Jaccard distance between target and source

336 compounds as a metric for chemical similarity. Similar analyses were conducted using another type

337 of chemical descriptors, the extensible chemistry-aware substructures called Saagar (Sedykh et al.,

338 2021), and the results were nearly identical; therefore, we present here only the results of Morgan

339 fingerprint analyses. For chemical pairs in TPEs with accepted read-across adaptations (**Figure 9A**),

340 the greatest number were highly similar; however, there were many compounds that spanned the

341 entire range. For example, the compounds with lowest similarity scores, but with accepted read-

342 across (**Supplemental Table 6**), were metal-organic compounds such as Co^{2+} or Sr^{2+} 2-

343 ethylhexanoate, Ba^{2+} or Co^{2+} carbonate, Pa^{2+} acetate ammoniate (1:2:4) and Co^{2+} acetate.

344 For substances in TPEs with rejected read-across adaptations, the distribution was much

345 wider, but still right-skewed (**Figure 9B**). When the overall distributions in Jaccard similarity were

346 compared between TPEs with accepted and rejected read-across adaptations, the difference in means

347 was not significant (data not shown). However, when the analyses comparing Jaccard similarity

348 between groups was restricted to substances with nearly perfect similarity based on Morgan

15

349   fingerprints (score=1) and those with scores that were <1, then the chance of acceptance of the read-

350   across adaptation was over 75% with a highly significant odds ratio of 7.0 (p=0.0003 via rigorous

351   cutpoint analysis). When the similarity was not as high, the chance of acceptance was only 31%.

352

353   **Discussion**

354   The goals of this study were three-fold – (i) to examine overall success rate of testing

355   proposals that contained read-across adaptations to standard information requirements, (ii) to

356   determine whether particular AE(s) made proposed read-across more or less successful, and (iii) to

357   improve future read-across submissions by providing a structured database of information based on

358   the actual ECHA decisions so that future proposals can identify relevant examples to "learn" from.

359   As we mentioned in the Introduction, there have been many suggestions on how to "improve" read-

360   across, often lamenting rigidities and conservatism of the decision-makers (Pestana et al., 2022;

361   Patlewicz et al., 2014; Blackburn and Stuard, 2014; Rovida et al., 2020). However, relatively few

362   studies (Ball et al., 2016) attempted to analyze the actual regulatory decisions on read-across

363   submissions and identify specific areas where the registrants may need to do a better job in

364   articulating their read-across justifications.

365   The study presented herein is a follow-up on the latter analysis, but almost a decade later

366   and after numerous guidelines and recommendations were issued by ECHA in an attempt to interpret

367   REACH and explain what justification are necessary for read-across hypotheses (ECHA, 2017,

368   2022, 2020, 2015). As our time-trend analysis showed, it does not appear that the success rates for

369   read-across-based TPE submissions have been improving; therefore, the availability of guidance

370   documents alone may not be sufficient to achieve success of successful adaptations. Similarly, it is

371   also noteworthy that most for 77% of TPEs, the registrants did not propose adaptations and opted to

372   just perform studies that were required. This can be either because there was no viable source

16

373　compound(s), or because the registrants chose certainty with respect to REACH compliance, as

374　opposed to the considerations of cost/time of doing such studies and/or animal welfare

375　considerations. It is possible that the registrants were not interested in "trying" the adaptation route

376　only to find that their submission was not acceptable, and that additional testing will still be

377　necessary. Therefore, we reason that by examining the decisions and cataloging successful and failed

378　read-across hypothesis it is possible to identify factors associated with positive decisions on read-

379　across adaptations and to improve the outcomes of such submissions in the future.

380　***Overall Trends in Read-Across Adaptations in Testing Proposal Submissions and Decisions***

381　There appears to be distinct trends in TPEs submissions with read-across – a rapid growth

382　of read-across-containing adaptations as a fraction of all evaluated testing proposals before 2015,

383　when the first RAAF guidance was published (ECHA, 2015), a precipitous decline to near zero in

384　2015, and another steep incline in the 5 years since 2015. On the decision side, a high proportion

385　(60-90%) of read-across-containing TPEs submitted before 2014 was accepted; this was followed

386　by a precipitous drop in 2016 and since. There are likely many reasons for these remarkable swings,

387　most of these can be attributed to the registration deadlines, internal considerations at ECHA to deal

388　with a large number of initial submissions under REACH, mutual learning of what acceptable read-

389　across is, and intense advocacy from the industry to provide clear guidance on interpretation of

390　vague language in REACH regulation. While certainly interesting from a historical perspective, and

391　clear evidence that RAAF guidance did make an impact, perhaps not as intended, it is unlikely that

392　a deeper dive into potential specifics of these trends will yield instructive learnings.

393　Our observation that the overall number of proposed adaptations was considerably greater

394　for Annex IX and X substances is likely a direct result of increase in standard testing requirements

395　with higher tonnage (Botham et al., 2023). Similarly, the finding that most common data gaps pertain

396　to a few tests (e.g., TG 414 [Prenatal developmental toxicity study] and 408 [Repeated dose 90-day

397   oral toxicity study in rodents]) is not surprising because these studies carry considerable cost and

398   use a large number of animals (**Table 2**). For Annex X substances, the main human health

399   information requirements are the TG 443 (Extended one generation reproduction study) and the TG

400   414 in a second species. Where a registrant required both, as well as the studies in Annex IX, it was

401   common for several years for the TPEs to be split, with a decision on the TG 408 and first species

402   in TG 414 and then follow up decision on the remaining studies once the results of the TG 408 and

403   414 were provided. This may explain why in general there were not as many TPEs covering these

404   additional higher tier studies – because decisions on the follow up studies may still be under

405   consideration and not publicly available decisions were available before this study's cutoff of August

406   2023.

407   It is noteworthy that the apparent "success" rate of read-across adaptations was significantly

408   different between analogue and group/category submissions. The odds for a successful testing

409   proposal adaptation through group read-across were almost 3-times as great as those for the

410   analogue read-across. These odds were most pronounced for the substances in Annexes VIII and IX,

411   and for TG 414 and TG 408 studies. Even though this finding indicates that if the registrant had a

412   choice between analogue and group, they may wish to opt into a group approach, there will be many

413   instances where the choice of group or analogue read-across is dictated by the type of a substance

414   and availability of the standard test data to which to read-across. In addition, one needs to keep in

415   mind that for an analogue approach, one argument to use read-across covers one substance and so

416   the ECHA decision covers just one substance. For the groups/categories, the same argument may

417   cover multiple substances, but there will be a separate decision for each one, hence the appearance

418   of a larger number of successful submissions may be misleading. Analogue-type submissions must

419   prove that a source substance represents equal or worse case than target, while group-type

420   submissions need to demonstrate a trend in the effect; therefore, it is generally more intuitive that

421 read-across within a group/category would be more acceptable as compared to an analogue approach

422 as there are more substances, and often more data from which to derive trends.

423 ***Read-Across Assessment Elements – Examples***

424 Similarity in structure/physicochemical properties (AE 1) is widely regarded as the

425 foundational element for any read-across hypothesis and it is not surprising that this AE was part of

426 virtually every read-across-containing testing proposal submission. For example, ECHA's RAAF

427 states that "*structural similarity is a pre-requisite for any grouping and read-across approach under*

428 *REACH*" (ECHA, 2015). Similarly, this principle is central to the application of read-across to

429 decisions by other agencies (Helman et al., 2019; Lizarraga et al., 2023); studies of the utility of

430 various descriptors of chemical's structure for prediction and characterization of chemical toxicity

431 span several decades (McKinney et al., 2000). While structure-activity predictions are very useful,

432 they also suffer from a number of pitfalls (Zvinavashe et al., 2008) and often need other data to

433 increase confidence (Rusyn et al., 2012).

434 Our analysis of ECHA TPEs found that AE 1 was included in both accepted and rejected

435 testing proposals, essentially to the same degree; however, very high chemical similarity (i.e.,

436 Jaccard similarity score of 1 based on Morgan fingerprints) was a strong predictor of acceptable

437 read-across. Still, testing proposals with read-across for substances with very low structural

438 similarity were also accepted. For example, there were several metalorganic compounds with

439 acceptable read-across, but very low structural similarity based on Morgan fingerprints. One

440 example is a group of "Cobalt-containing compounds" that included Co^{2+} 2-ethylhexanoate, Co^{2+}

441 carbonate, and Co^{2+} diacetate. The registrants reasoned that a common metal cation, not the organic

442 counterions, was the driver of any adverse health effects; the metal cation rapidly dissociates from

443 the organic counterion when encountering the biological fluids. These submissions also included

444 AE 9 (formation of common/identical compounds) and AE 12 (formation and impact of non-

19

445   common compounds/exposure to other compounds than those linked to the prediction) as part of

446   their read-across hypothesis. The overall rationale presented in these testing proposals was found

447   acceptable by ECHA. It is interesting that decisions on other metalorganic compounds that used

448   similar arguments were published at different times between 2013 to 2023, demonstrating that well-

449   rationalized reasoning based on AE 9 and 12 can overcome low similarity in AE 1.

450   It should be noted that while considerations of toxicokinetics (AE 9 through 14) are widely

451   regarded as important for "good" read-across (Ball et al., 2016; Rovida et al., 2020), there were

452   relatively few submissions that included AEs other than 9 and 12, as discussed above. In fact,

453   inclusion of AE 13 (Metabolites of source and target have been identified) to argue for exposure to

454   structurally similar metabolites or different compounds that cause the same effect appeared to

455   greatly diminish the odds of acceptance. AE 13 was present in 4% of accepted and 26% of rejected

456   testing proposals with read-across. ECHA used AE 13 as a reason to reject proposed read-across in

457   4% of all unfavorable decisions – 3% were a difference of opinion on how the data was interpreted

458   or whether the data are supportive of the read-across hypothesis, and 1% was the lack of discussion

459   of this element by the registrant. For example, when computational predictors of metabolism were

460   used without corresponding analytical evidence, ECHA did not find those arguments satisfactory.

461   ECHA also frequently noted the lack of discussion of other metabolites that could be formed (i.e.,

462   AE 12) and whether those may be associated with adverse health effects.

463   The second most common element in proposed read-across hypotheses was AE 16

464   (Toxicodynamic similarity based on the data from a bridging (not a guideline) study). This AE

465   includes any non-TG data submitted to support a read-across, such as *in vitro* methods, QSAR, and

466   non-guideline animal studies. It was proposed in 77% of accepted and 74% of rejected testing

467   proposals with read-across. It is even more noteworthy that ECHA used this AE to reject a read-

468   across adaptation in 86% of all unfavorable decisions. Among these, ECHA disagreed with the

20

469   strengths of the registrant's reasoning based on the available data in 66% of decisions; in 20% of

470   decisions, ECHA pointed out that such data would be needed to strengthen the read-across

471   hypothesis. For example, the use of short-term studies to justify read-across adaptation of chronic

472   or prenatal developmental toxicity studies or presenting the data for only the target or source

473   substance, or for a different endpoint, was not deemed to be a satisfactory justification. Similarly,

474   when target and source substances did not demonstrate similar effects in non-guideline studies, or

475   when observations in non-animal (e.g., *in vitro* test) studies were not substantiated by data from *in*

476   *vivo* studies, the read-across hypotheses were found to be not acceptable. For example, the data from

477   *in vitro* mouse lymphoma assay, without support from *in vivo* studies, is not adequate as a

478   justification for adaptations of higher-tier endpoints (e.g., TG 408).

479   A related read-across hypothesis element was AE 8 (Bridging [guideline] study), for which

480   the registrant submitted data derived from a test conducted using an OECD TG protocol. AE 8 was

481   part of 39% of accepted and 43% of rejected testing proposals with read-across; odds of acceptable

482   read-across adaptation were 4.3 for group-type submissions that had this AE as part of their

483   hypothesis. ECHA decisions discussed AE 8 in 76% of unfavorable decisions; half of these cases

484   were a disagreement with the registrant as to whether such data were supportive of the read-across

485   hypothesis, and another half were cases where ECHA pointed out that such data would be needed

486   to justify the proposed adaptation. For example, when testing proposals for higher-tier endpoints

487   (TG 408 and TG 414) included data from TG 422 (Combined repeated dose toxicity with the

488   reproduction/developmental toxicity screening) in support of read-across and showing similar

489   effects (e.g., same target organ, magnitude of effects, etc.) for both source and target substance,

490   ECHA generally agreed with the registrant's hypothesis. It is also worth noting that AE 8 was

491   frequently a focus of "third party" comments on testing proposal submissions. Specifically, "third

492   parties" reasoned that the findings of low toxicity in a 28-day oral study (TG 407) should be used

493   as "bridging" evidence in support of read-across for adaptation of the TG 408 (90-day study) (Taylor

494   and Andrew, 2017; Taylor et al., 2014a). ECHA considered these arguments in their decisions;

495   however, the Agency stated in their decisions that it is the registrant's (i.e., and not "third party")

496   responsibility to consider such reasoning and any other justifications when proposing adaptations.

497   Two AEs pertain to the argument that compounds may have common biological effects even

498   if they are not structurally similar or do not form identical metabolites. Specifically, AE 3 (Common

499   Biological Targets without Common Metabolites (Qualitative)) and AE 4 (Common Biological

500   Targets without Common Metabolites (Quantitative)) address this type of a rationale in read-across.

501   While AE 3 was included in a fairly large number of testing proposals with read-across, AE 4 was

502   only found in 2 submissions. While it is impossible to extrapolate from only 2 submissions with AE

503   4, both of those were accepted by ECHA, it very well may indicate that quantitative arguments (i.e.,

504   the magnitude of the effect) are welcome. The latter conclusion is also supported by the fact that

505   ECHA pointed out the lack of information pertaining to AE 4 in many unfavorable decisions on

506   submitted testing proposals with read across (13%). Reasoning related to AE 3 was included in 29%

507   of accepted and 17% of rejected testing proposals with read-across. Among these, ECHA disagreed

508   with the strengths of the registrant's reasoning based on the available data in 11% of decisions; in

509   5% of decisions, ECHA pointed out that such data would be needed to strengthen the read-across

510   hypothesis. For example, submissions based on large groups (e.g., higher olefins and resin acids)

511   relied on the arguments pertaining to AE 3 when reasoning that "*different compounds have the same*

512   *effect.*" The inclusion of AE 3 significantly increased the odds of acceptance for group-type

513   submissions.

514   Substances that are classified as multi-constituent or UVCB have several additional

515   challenges with respect to the need for establishing both substance identity, and to characterize the

516   chemical composition to support read-across. ECHA published separate guidance for these

517   substances (ECHA, 2017); in addition, the chasm between the regulator's expectations and the

518   realities of analytical chemistry solutions for identifying and quantifying the constituents in highly

519   complex substances has been documented (Roman-Hubers et al., 2023). It is not surprising,

520   therefore, that AE 15 (Characterization of multi-/UVCB substances) was commonly included in

521   testing proposals with read-across adaptations for these types of substances. In submissions

522   concerning multi-constituent/UVCB substances, AE 15 was present in 76% of accepted and 32% of

523   rejected testing proposals with read-across. ECHA used this AE in 43% of unfavorable decisions;

524   in 15% it was a difference of opinion on the interpretation of the available data and 28% because

525   the registrant did not address this AE. The odds of acceptance were significantly higher (4.2) for

526   group-type submissions when this AE was part of the read-across hypothesis. In the cases of

527   unfavorable decision, ECHA reasoned that insufficient characterization was provided by the

528   registrant for target or source (or both) substances, which means that there was no way to compare

529   their similarity or lack thereof. While this reasoning is similar to that of AE 1, it is far less clear how

530   to define "*broad similarity*" for substances that have too many constituents and when their

531   composition is expected to be variable. Case examples of petroleum substances have been recently

532   published with respect to the use of other supporting data types from *in vitro* studies to justify

533   grouping (Tsai et al., 2023; House et al., 2022; House et al., 2021); however, ECHA did not find

534   these non-guideline "bridging" studies (AE 16) satisfactory for a number of reasons. Many of these

535   related to chemical characterization of the substances and/or their extracts that were used for *in vitro*

536   testing, challenges that need to be addressed by additional research (Roman-Hubers et al., 2022;

537   Cordova et al., 2022).

538   ***Using a Database of TPE Read-Across Adaptations to Construct Future Submissions***

539   Even though our analysis showed that about 50% of testing proposal submissions that relied

540   on read-across adaptations were found to be acceptable by ECHA, improvements are needed to

541   increase success. We found that over the past decade, the fraction of accepted proposals stayed

542   relatively constant and the efforts by all parties involved in the registration process have yet to

543   produce a measurable improvement. Therefore, we reason that the data we extracted from the

544   decisions may provide instructive examples of successful submissions that could be emulated in the

545   near future. It is also clear that by improving success rate of testing proposal submissions with read-

546   across, measurable impact can be made in terms of reduction in animal use for chemical registrations

547   under REACH. **Table 2** shows that the number of animals that will be required to meet information

548   requirements if new testing is performed is substantial. Similarly, if more registrants will consider

549   read-across-based or other adaptations, rather than defaulting to new testing, the reduction in animal

550   use will be even more pronounced.

551   While it is without a doubt that the registrants and ECHA are committed to sharing

552   encouraging examples and mutually develop best practices for read-across, we reason that the data

553   collected in this project will prove useful to both registrants and regulators (for full database, see

554   **Supplemental Tables 1, and 3-5**). On the one hand, to achieve greater chance of success, registrants

555   will be able to identify specific aspects of read-across justifications that merit most attention and

556   improve these in the testing proposals they are working on. Our results also likely to encourage

557   consideration of a category versus analogue approach, although the latter may not be possible for

558   many submissions. We also found that read-across between different substance types should be

559   discouraged, particularly reading-across from UVCBs to mono-constituent substances. On the other

560   hand, the trends identified in our study could indicate changes in ECHA's approach to interpretation

561   of read-across and suggest that additional efforts may be needed by ECHA to improve consistency

562   in the application of their guidance over time. In addition, ECHA and other agencies may use our

563   results and the database to identify areas where they could provide more granular advice to improve

564   how read-across arguments are presented and justified.

565   *Limitations of the Study*

566   While our analyses and interpretation of the findings may prove instructive, we also

567   acknowledge that our results may not be taken as definitive and deterministic. First, we note that

568   our dataset may have limitations as it had a cutoff date of August 2023 and the trends in most recent

569   submissions, and thus decisions, may be different given the attention given to the efforts to improve

570   read-across (Pestana et al., 2022; Patlewicz et al., 2014; Blackburn and Stuard, 2014; Rovida et al.,

571   2020). Second, we relied on the robust summaries of the testing proposals as included in ECHA's

572   decisions, we did not examine the testing proposal submissions. While it is possible that some mis-

573   interpretation could have occurred, we reason that robust summaries may be eventually more

574   informative in terms of categorizing the elements of the proposed read-across into standardized

575   categories (i.e., AEs, **Table 1**). Even though the AEs were defined by the authors, we reason that no

576   single guidance, or combination of different documents released by ECHA over time, clearly defines

577   read-across AEs. While this is another limitation, we point out that every effort was made to assure

578   that the AEs used herein were comprehensive and consistent with both RAAF and the actual

579   decisions (**Supplemental Table 2**).

580   Third, a large difference in the success rate of read-across adaptations between testing

581   proposals and compliance checks (Ball et al., 2016) indicates that the "bar" may be lower on the

582   former because the decision to allow for the use of read-across as part of a testing proposal is

583   provisional. Whether the use of read-across is ultimately acceptable or not for the registration can

584   only be decided once the registration dossiers are submitted. Consequently, there may be cases

585   where the use of read-across has been accepted for the testing proposal but was subsequently

586   challenged or rejected. For example, the registrants of decan-4-olide (EC: 211-892-8) submitted a

587   testing proposal with read-across adaptation for TG 211 (long-term toxicity to aquatic invertebrates)

588   and this submission was found to be acceptable by ECHA in 2015. However, a compliance check

25

589   decision by ECHA (published in 2023) concluded a read-across adaptation to standard information

590   requirement was insufficient and that required TG 211 data were needed to justify proposed

591   classification. Even though this caveat does not invalidate the results presented in this manuscript,

592   it does highlight that even "*final*" decisions on read-across can be subject to change.

593   Finally, we need to point out that although our analysis has identified some specific areas

594   that were influential in read-across acceptance, it is rare that ECHA decisions focus on just one

595   aspect; often, multiple reasons may lead to the ultimate decision to accept or not accept.

596   Furthermore, the submission may be accepted not because ECHA agreed with the rationale

597   presented in a testing proposal, but because different combination of the information included in the

598   submission was deemed to be sufficient. For example, for the Asphalt UVCB (EC: 232-490-4)

599   submission, the registrant proposed a read-across adaptation of TG 414 to another substance in the

600   group (Residues (petroleum), thermal cracked vacuum; CAS: 92062-05-0) on which this test is to

601   be performed. The principal reasoning by the registrant was that the target and source substances

602   belong to a group formed based on the refining process and on similarity in carbon number

603   distribution and the hydrocarbon class profiles. Furthermore, the registrant hypothesized that one of

604   the hydrocarbon classes – polycyclic aromatic hydrocarbons containing 4 or more aromatic rings –

605   is the putative reproductive toxicant. While ECHA found that the proposed justification for the

606   overall group was insufficient, it agreed that a read-across to the proposed source substance selected

607   based on the PAH with 4+ aromatic rings was "appropriate." While such cases complicate the overall

608   analysis as they make it difficult to be confident that one or two specific AEs had an impact on each

609   decision, our data still identified several "critical" AEs as detailed above. Ultimately, we encourage

610   the registrants to focus their efforts on improving the rationales for their read-across hypotheses to

611   stand a better chance of acceptance. Even though each read-across submission is unique in terms of

612   the type of a substance, availability of the data and the endpoint, we show that there are several

26

613   general trends and that the registrants can rely on case examples in the database when crafting their

614   submissions.

615   **Table 1.** Assessment Elements (AE) used in Testing proposal Evaluation (TPE) submissions and

616   decisions. See **Supplemental Table 2** for a detailed explanation of each AE.

Assessment Elements Based on <u>Toxicodynamic</u> and Other Considerations		Assessment Elements Based on <u>Toxicokinetic</u> Considerations	
AE.1	Structure/Physicochemical Properties	AE.9	Formation of Common (Identical) Compounds
AE.2	Common Biological Targets with Common Metabolites	AE.10	Exposure of Biological Targets to Common Compounds
AE.3	Common Biological Targets without Common Metabolites (*Qualitative*)	AE.11	Extent of the Bioavailability of the Parent Compound
AE.4	Common Biological Targets without Common Metabolites (*Quantitative*)	AE.12	Formation and Impact of Non-common Compounds/Exposure to Other Compounds than those Linked to the Prediction
AE.5	Environmental Degradation to Non-common Compounds	AE.13	Metabolites of Source and Target have been Identified
AE.6	Environmental Bioaccumulation of Potential Non-common Compounds	AE.14	Potential Presence of Uncharacterized Metabolites or Impurities
AE.7	Common Environmental Degradation Pathways	**Other Assessment Elements**	
AE.8	Toxicodynamic Similarity based on the Data from a Bridging (OECD Test Guideline) Study	AE.15	Characterization of Constituents (for Multi-Constituent or UVCB Substances)
		AE.16	Toxicodynamic Similarity based on the Data from a Bridging (Not a Guideline) Study
		AE.17	Lack of Observed Adverse Effects

617

28

618 **Table 2.** Animal count for representative tests in animals that had proposed read-across adaptations

619 in testing proposals to satisfy REACH substance registration requirements.

OECD Test Guideline Studies That Were the Subject of Read-Across Adaptations in Testing Proposals	Accepted Testing Proposals with Read-Across: Number (% of 149 total)	Not Accepted Testing Proposals with Read-Across: Number (% of 165 total)	Number of Animals Required for Each Test		
			Minimum Number According to each OECD TG	Average Number According to (Taylor, 2018)	Average Number According to (Knight et al., 2023)
TG 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents (OECD, 2018a)	74 (49.7%)	69 (41.8%)	100	100	122
TG 414: Prenatal Developmental Toxicity Study [rat] (OECD, 2018b)	88 (59.1%)	101 (61.2%)	100*	900	1,459
TG 443: Extended One-Generation Reproductive Toxicity Study (OECD, 2018c)	10 (7.1%)	19 (11.5%)	580	960	2,733/1,830[#]
TG 489: *In Vivo* Mammalian Alkaline Comet Assay (OECD, 2016)	6 (4.0%)	4 (2.4%)	25	-	50

620 *, Not including the number of animals in each litter.

621 [#], Reflecting the use of mated/unmated animals.

622 -, Information was not included in the publication.

623   References
624

625   Ball, N., Cronin, M. T., Shen, J. et al. (2016). Toward Good Read-Across Practice (GRAP)
626   guidance. *ALTEX 33*, 149-166. http://dx.doi.org/10.14573/altex.1601251

627   Beal, M. A., Gagne, M., Kulkarni, S. A. et al. (2022). Implementing in vitro bioactivity data to
628   modernize priority setting of chemical inventories. *ALTEX 39*, 123-139.
629   http://dx.doi.org/10.14573/altex.2106171

630   Berggren, E., Amcoff, P., Benigni, R. et al. (2015). Chemical Safety Assessment Using Read-
631   Across: Assessing the Use of Novel Testing Methods to Strengthen the Evidence Base for
632   Decision Making. *Environ Health Perspect 123*, 1232-1240.
633   http://dx.doi.org/10.1289/ehp.1409342

634   Blackburn, K. and Stuard, S. B. (2014). A framework to facilitate consistent characterization of
635   read across uncertainty. *Regul Toxicol Pharmacol 68*, 353-362.
636   http://dx.doi.org/10.1016/j.yrtph.2014.01.004

637   Botham, P., Cronin, M. T. D., Currie, R. et al. (2023). Analysis of health concerns not addressed by
638   REACH for low tonnage chemicals and opportunities for new approach methodology. *Arch
639   Toxicol 97*, 3075-3083. http://dx.doi.org/10.1007/s00204-023-03601-5

640   Chung, N. C., Miasojedow, B., Startek, M. et al. (2019). Jaccard/Tanimoto similarity test and
641   estimation methods for biological presence-absence data. *BMC Bioinformatics 20*, 644.
642   http://dx.doi.org/10.1186/s12859-019-3118-5

643   Cordova, A. C., Ford, L. C., Valdiviezo, A. et al. (2022). Dosing Methods to Enable Cell-Based In
644   Vitro Testing of Complex Substances: A Case Study with a PAH Mixture. *Toxics 11*, 19.
645   http://dx.doi.org/10.3390/toxics11010019

646   ECHA (2011). Guidance on information requirements and chemical safety assessment.
647   https://echa.europa.eu/guidance-documents/guidance-on-information-requirements-and-
648   chemical-safety-assessment Acessed on: July 21, 2024

649   ECHA (2015). Read-Across Assessment Framework (RAAF).
650   http://echa.europa.eu/documents/10162/13628/raaf_en.pdf
651   http://dx.doi.org/10.2823/546436, Acessed on: July 21, 2024

652   ECHA (2017). Read-Across Assessment Framework (RAAF) - considerations on multi-constituent
653   substances and UVCBs.
654   https://echa.europa.eu/documents/10162/13630/raaf_uvcb_report_en.pdf/3f79684d-07a5-
655   e439-16c3-d2c8da96a316 Acessed on: August 25, 2020

656   ECHA (2020). Dossier Evaluation Process - principles applied by ECHA in revieweing read-
657   across and category approaches.
658   https://echa.europa.eu/documents/10162/17220/echa_testing_strategies_under_reach_en.p
659   df/6ffd4ed1-8f14-6326-8142-2c881846fdac Acessed on: July 26, 2024

660 ECHA (2022). Advice on using read-across for UVCB substances - obligations arising from
661   Commission Regulation 2021/979, amending REACH annexes.
662   https://www.echa.europa.eu/documents/10162/11395738/advice_uvcb_read-
663   across_en.pdf/ac1f64a6-9ee5-441e-cf1c-92914b843b4e?t=1651665130365 Acessed on:
664   May 19, 2022

665 ECHA (2023a). The use of alternatives to testing on animals for the REACH Regulation: Fifth
666   report under Article 117(3) of the REACH Regulation. http://dx.doi.org/10.2823/805454,
667   Acessed on: July 21, 2024

668 ECHA (2023b). Guidance for identification and naming of substances under REACH and CLP.
669   https://echa.europa.eu/documents/10162/2324906/substance_id_en.pdf/ee696bad-49f6-
670   4fec-b8b7-2c3706113c7d Acessed on: July 21, 2024

671 European Council (2006). Regulation (EC) No 1907/2006 of the European Parliament and the
672   Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and
673   Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending
674   Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and
675   Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and
676   Commission Directives 91/155/EEC, 93/105/EC and 2000/21/EC. 1-849. https://eur-
677   lex.europa.eu/eli/reg/2006/1907/oj

678 Guha, R. (2007). Chemical Informatics functionality in R. *J Stat Softw 18*, 1 - 16.
679   http://dx.doi.org/10.18637/jss.v018.i05

680 Helman, G., Shah, I., Williams, A. J. et al. (2019). Generalized Read-Across (GenRA): A
681   workflow implemented into the EPA CompTox Chemicals Dashboard. *ALTEX 36*, 462-
682   465. http://dx.doi.org/10.14573/altex.1811292

683 House, J. S., Grimm, F. A., Klaren, W. D. et al. (2021). Grouping of UVCB substances with new
684   approach methodologies (NAMs) data. *ALTEX 38*, 123-137.
685   http://dx.doi.org/10.14573/altex.2006262

686 House, J. S., Grimm, F. A., Klaren, W. D. et al. (2022). Grouping of UVCB substances with dose-
687   response transcriptomics data from human cell-based assays. *ALTEX 39*, 388-404.
688   http://dx.doi.org/10.14573/altex.2107051

689 Knight, J., Hartung, T. and Rovida, C. (2023). 4.2 million and counting... The animal toll for
690   REACH systemic toxicity studies. *ALTEX 40*, 389-407.
691   http://dx.doi.org/10.14573/altex.2303201

692 Lizarraga, L. E., Suter, G. W., Lambert, J. C. et al. (2023). Advancing the science of a read-across
693   framework for evaluation of data-poor chemicals incorporating systematic and new
694   approach methods. *Regul Toxicol Pharmacol 137*, 105293.
695   http://dx.doi.org/10.1016/j.yrtph.2022.105293

31

McKinney, J. D., Richard, A., Waller, C. et al. (2000). The practice of structure activity relationships (SAR) in toxicology. *Toxicol Sci 56*, 8-17. http://dx.doi.org/10.1093/toxsci/56.1.8

OECD (2016). *Test No. 489: In Vivo Mammalian Alkaline Comet Assay*. https://www.oecd-ilibrary.org/content/publication/9789264264885-en http://dx.doi.org/doi:https://doi.org/10.1787/9789264264885-en

Organisation for Economic Co-operation and Development (2017). Guidance on Grouping of Chemicals, Second Edition, OECD Series on Testing and Assessment, No. 194. http://dx.doi.org/https://doi.org/10.1787/9789264274679-en, Acessed on: July 21, 2024

OECD (2018a). *Test No. 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents*. https://www.oecd-ilibrary.org/content/publication/9789264070707-en http://dx.doi.org/doi:https://doi.org/10.1787/9789264070707-en

OECD (2018b). *Test No. 414: Prenatal Developmental Toxicity Study*. https://www.oecd-ilibrary.org/content/publication/9789264070820-en http://dx.doi.org/doi:https://doi.org/10.1787/9789264070820-en

OECD (2018c). *Test No. 443: Extended One-Generation Reproductive Toxicity Study*. https://www.oecd-ilibrary.org/content/publication/9789264185371-en http://dx.doi.org/doi:https://doi.org/10.1787/9789264185371-en

Patlewicz, G., Ball, N., Becker, R. A. et al. (2014). Read-across approaches--misconceptions, promises and challenges ahead. *ALTEX 31*, 387-396. http://dx.doi.org/http://dx.doi.org/10.14573/altex.1410071

Patlewicz, G., Karamertzanis, P., Friedman, K. P. et al. (2024). A systematic analysis of read-across within REACH registration dossiers. *Comput Toxicol 30*, 1-15. http://dx.doi.org/10.1016/j.comtox.2024.100304

Pestana, C., Enoch, S. J., Firman, J. W. et al. (2022). A strategy to define applicability domains for read-across. *Comput Toxicol 22*, 100220. http://dx.doi.org/https://doi.org/10.1016/j.comtox.2022.100220

Roman-Hubers, A. T., Cordova, A. C., Rohde, A. M. et al. (2022). Characterization of Compositional Variability in Petroleum Substances. *Fuel 317*, 123547. http://dx.doi.org/10.1016/j.fuel.2022.123547

Roman-Hubers, A. T., Cordova, A. C., Barrow, M. P. et al. (2023). Analytical chemistry solutions to hazard evaluation of petroleum refining products. *Regul Toxicol Pharmacol 137*, 105310. http://dx.doi.org/10.1016/j.yrtph.2022.105310

Rovida, C., Barton-Maclaren, T., Benfenati, E. et al. (2020). Internationalization of read-across as a validated new approach method (NAM) for regulatory toxicology. *ALTEX 37*, 579-606. http://dx.doi.org/10.14573/altex.1912181

732   Rovida, C., Busquet, F., Leist, M. et al. (2023). REACH out-numbered! The future of REACH and
733   animal numbers. *ALTEX 40*, 367-388. http://dx.doi.org/10.14573/altex.2307121

734   Rusyn, I., Sedykh, A., Low, Y. et al. (2012). Predictive modeling of chemical hazard by integrating
735   numerical descriptors of chemical structures and short-term toxicity assay data. *Toxicol Sci*
736   *127*, 1-9. http://dx.doi.org/10.1093/toxsci/kfs095

737   Sedykh, A. Y., Shah, R. R., Kleinstreuer, N. C. et al. (2021). Saagar-A New, Extensible Set of
738   Molecular Substructures for QSAR/QSPR and Read-Across Predictions. *Chem Res Toxicol*
739   *34*, 634-640. http://dx.doi.org/10.1021/acs.chemrestox.0c00464

740   Taylor, K., Andrew, D. J. and Rego, L. (2014a). The added value of the 90-day repeated dose oral
741   toxicity test for industrial chemicals with a low (sub)acute toxicity profile in a high quality
742   dataset. *Regul Toxicol Pharmacol 69*, 320-332.
743   http://dx.doi.org/10.1016/j.yrtph.2014.04.008

744   Taylor, K., Stengel, W., Casalegno, C. et al. (2014b). Experiences of the REACH testing proposals
745   system to reduce animal testing. *ALTEX 31*, 107-128.
746   http://dx.doi.org/10.14573/altex.1311151

747   Taylor, K. and Andrew, D. J. (2017). The added value of the 90-day repeated dose oral toxicity test
748   for industrial chemicals with a low (sub)acute toxicity profile in a high quality dataset: An
749   update. *Regul Toxicol Pharmacol 90*, 258-261.
750   http://dx.doi.org/10.1016/j.yrtph.2017.09.018

751   Taylor, K. (2018). Ten years of REACH - An animal protection perspective. *Altern Lab Anim 46*,
752   347-373. http://dx.doi.org/10.1177/026119291804600610

753   Tsai, H. D., House, J. S., Wright, F. A. et al. (2023). A tiered testing strategy based on in vitro
754   phenotypic and transcriptomic data for selecting representative petroleum UVCBs for
755   toxicity evaluation in vivo. *Toxicol Sci 193*, 219-233.
756   http://dx.doi.org/10.1093/toxsci/kfad041

757   Zvinavashe, E., Murk, A. J. and Rietjens, I. M. (2008). Promises and pitfalls of quantitative
758   structure-activity relationship approaches for predicting metabolism and toxicity. *Chem*
759   *Res Toxicol 21*, 2229-2236.

760

761

Conflict of interest

N. Ball is an employee of Dow Chemical Company that submitted several registrations and testing proposals for evaluation by ECHA. Other authors declare no conflicts of interest.

Data availability

Data extracted from publicly available testing proposal evaluations by ECHA are included as Supplemental Tables.
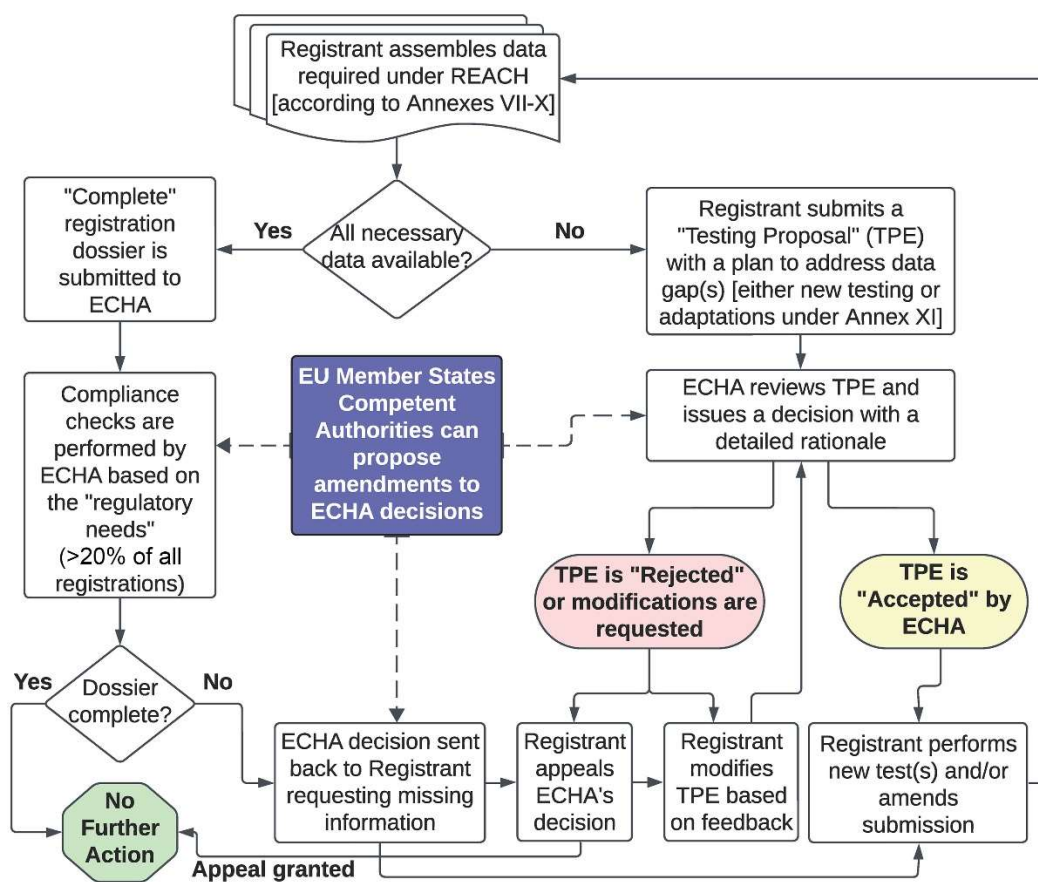
Funding

Acknowledgements

781 **Figure 1**. A simplified workflow of substance registration under REACH.

782

783

784 **Figure 2**. A diagram illustrating categorization of the publicly available TPE (Testing Proposal
785 Examination) documents with and without adaptations, as well as the acceptance and rejection rates
786 of TPEs with read-across (RA) adaptations. Numbers indicate the number of submissions with a
787 published TPE as substances may have submitted multiple TPs throughout the registration period.
788

789

790
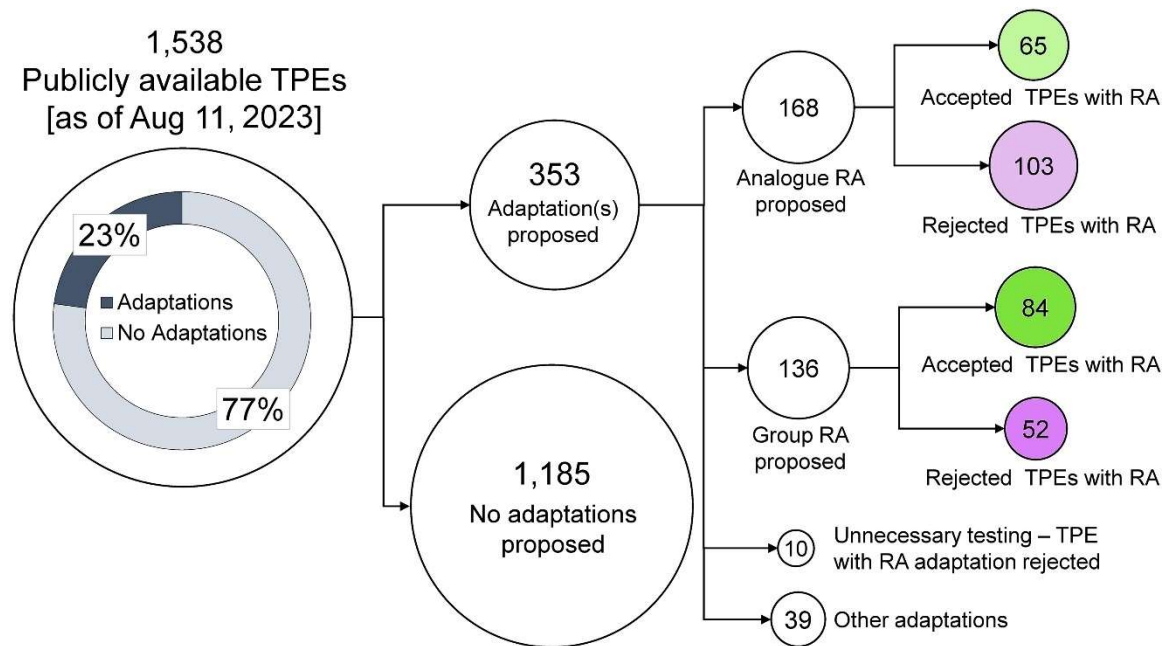
Figure 3. Time-trend plots for when TPE decisions were published. **(A)** The number of all published TPE decisions per calendar year (bars, left y-axis) and a fraction of TPEs that contained read-across (RA) adaptations (line, right y-axis). **(B)** The number of published TPE decisions per calendar year indicating the type of read-across (stacked bar plots, left y-axis) and a fraction of these that was accepted (line, right y-axis).
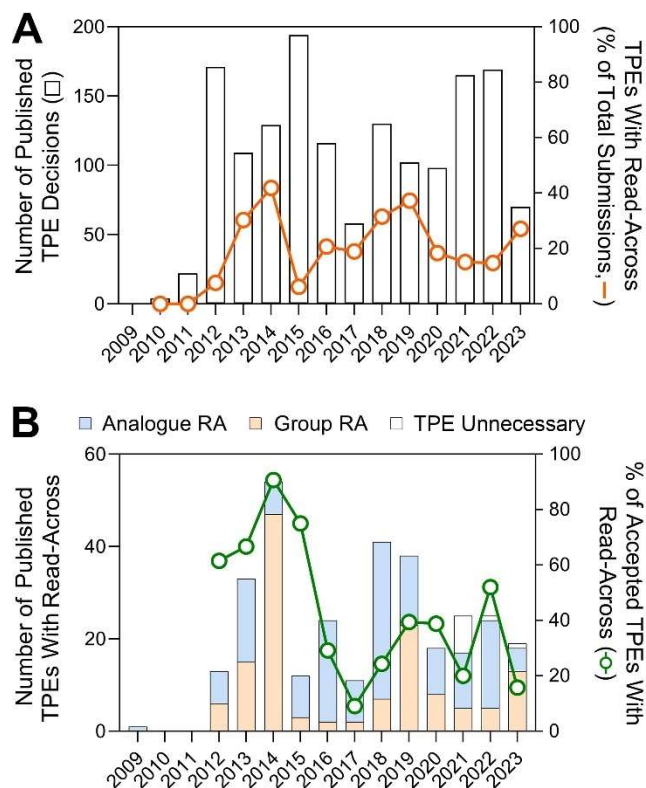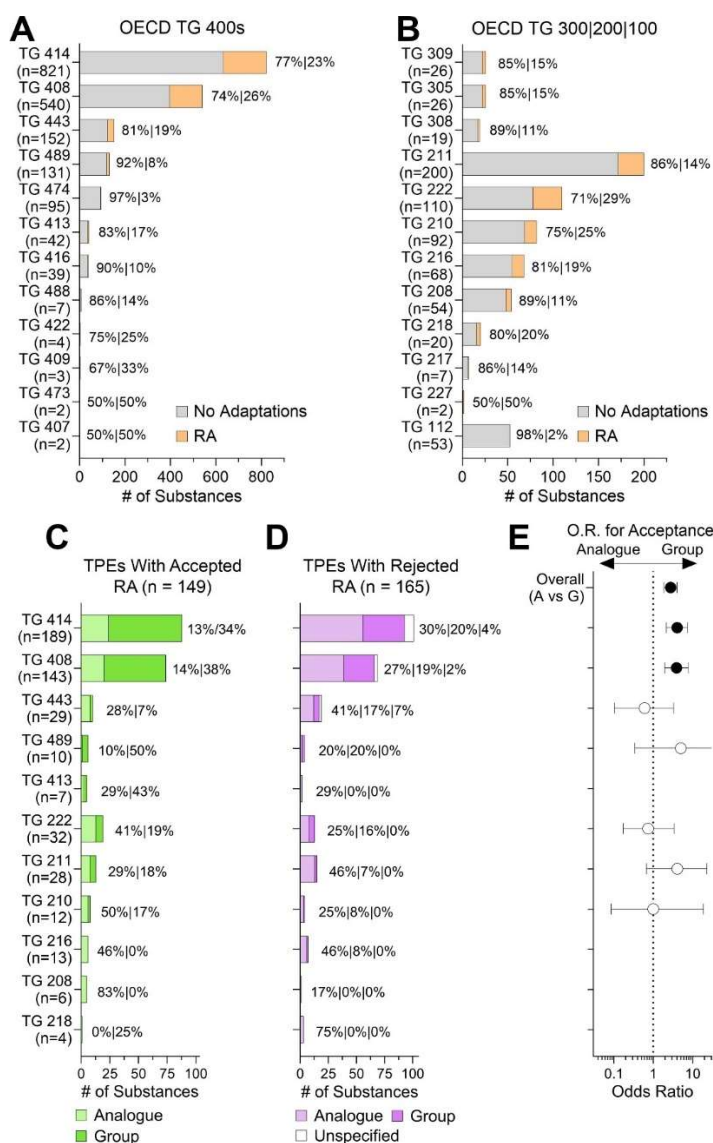
Figure 4. Analysis of the publicly available TPE decisions by substance type. **(A)** Stacked bar plots show the number of substances for which read-across adaptations were accepted or rejected, separated into substance categories. The total number of substances in each category is shown. Within each stacked bar plot, submissions that used analogue (lighter shade) or group (darker shade) read-across are shown. In some instances of rejected read-across, the type of read-across could not be determined (white). **(B)** For each target substance type, different types of source substances were used as indicated. The outcome of read-across evaluation is shown by the adjacent bars.
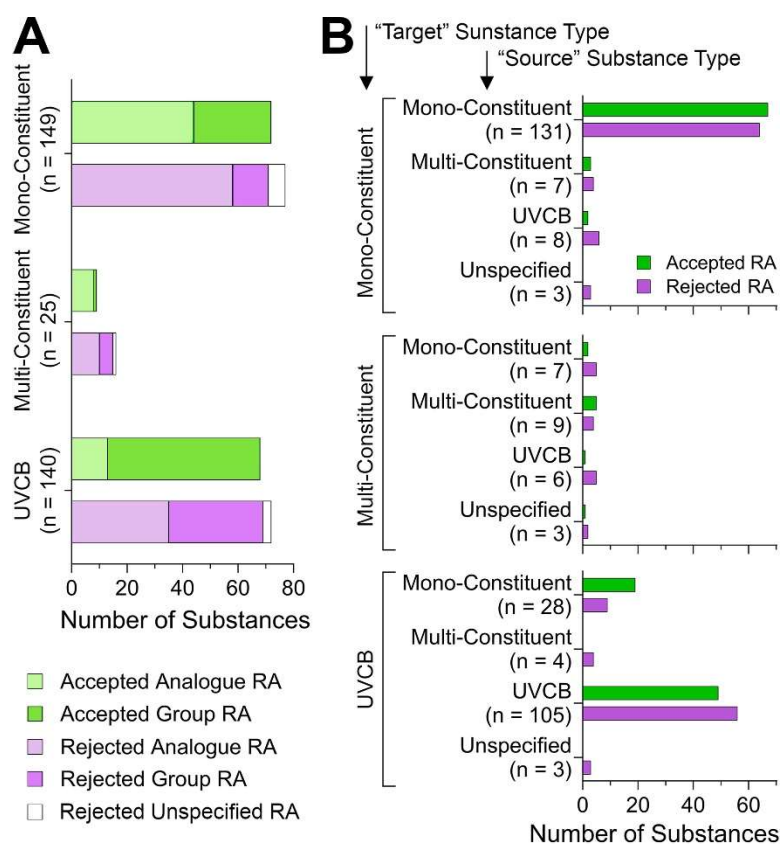
Figure 5. Analysis of the publicly available TPE decisions by REACH Annex number (substance tonnage). **(A)** Stacked bar plots show the number of substances in each Annex category for which publicly available TPEs were examined. Gray color represents TPEs without read-across adaptations and orange represents those with read-across adaptations. (**B-C**) The number of substances with accepted (B) and rejected (C) read-across adaptations. Within each stacked bar plot, submissions that used analogue (lighter shade) or group (darker shade) read-across are shown. In some instances of rejected read-across, the type of read-across could not be determined (white). **(D)** For each Annex number, odd ratios (OR) and 95% confidence intervals for Acceptance in Group vs. Analogue (i.e. an OR>1.0 corresponds to greater odds of Acceptance for Group). The intervals for Annex #IX and VIII do not contain the null OR=1.0, corresponding to a significantly greater odds of Acceptance for Group ($p<0.05$).
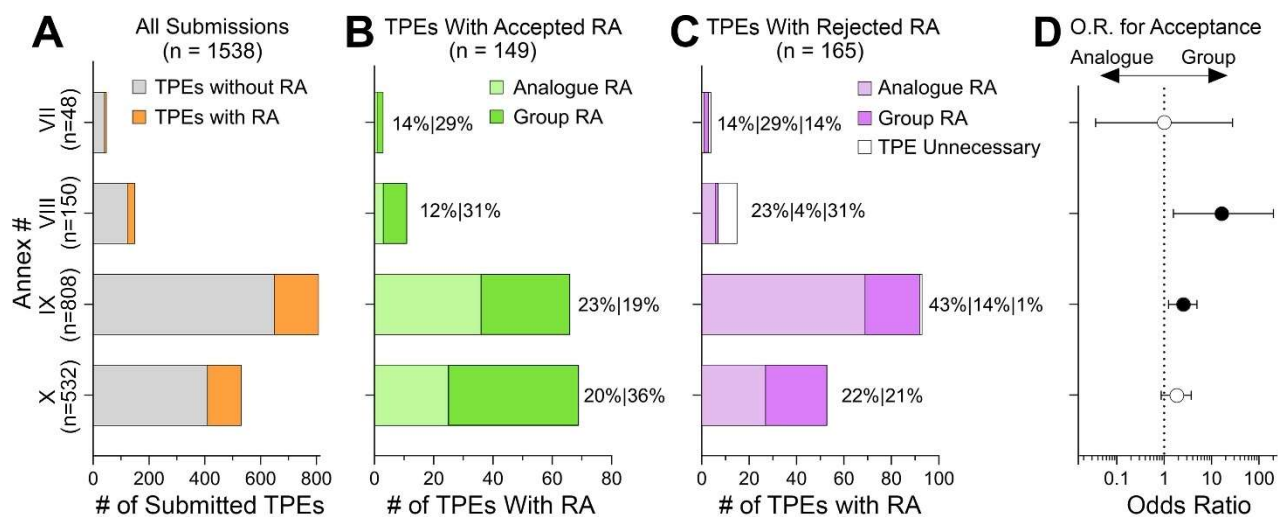
Figure 6. Analysis of the publicly available TPE decisions by OECD test guideline study type. **(A-B)** Stacked bar plots show the number of substances for each guideline test, separated into "human health" (A), and other (B) OECD test categories. Gray color represents TPEs without read-across adaptations and orange represents those with read-across adaptations. **(C-D)** The number of substances with accepted (C) and rejected (D) read-across adaptations. Within each stacked bar plot, submissions that used analogue (lighter shade) or group (darker shade) read-across are shown. In some instances of rejected read-across, the type of read-across could not be determined (white). **(E)** For the Overall data and when split by OECD guideline type, odd ratios (OR) and 95% confidence intervals for Acceptance in Group vs. Analogue (i.e. an OR>1.0 corresponds to greater odds of Acceptance for Group). The intervals for Overall, TG 414, TG 408, and TG443 do not contain the null OR=1.0, corresponding to a significantly greater odds of Acceptance for Group (*p*<0.05).

Figure 7. Analysis of the publicly available TPE decisions by assessment elements (AEs). **(A-B)** Stacked bar plots show the number of substances Overall and for each AE, separated into TPEs that were Accepted (A), and Rejected (B). For AE15 (marked with *), the fractions shown is for UVCB substances only. **(C)** For the Overall data and for each AE, odd ratios (OR) and 95% confidence intervals for Acceptance in Group vs. Analogue (i.e. an OR>1.0 corresponds to greater odds of Acceptance for Group). The intervals for Overall, AE 1, AE 8, and AE 15 do not contain the null OR=1.0, corresponding to a significantly greater odds of Acceptance for Group ($p<0.05$). **(D)** ORs from multiple logistic regression analyses with predictors submission year, decision year, group vs. analogue-based RA, and all AEs. For each column (all TPEs, Mono and UVCB), the ORs are displayed on a color scale, where red indicates predictor increases the chance of Acceptance, blue indicates predictor decreases chance of Acceptance, and '*' denotes coefficients that have false discovery rates of $q<0.05$.
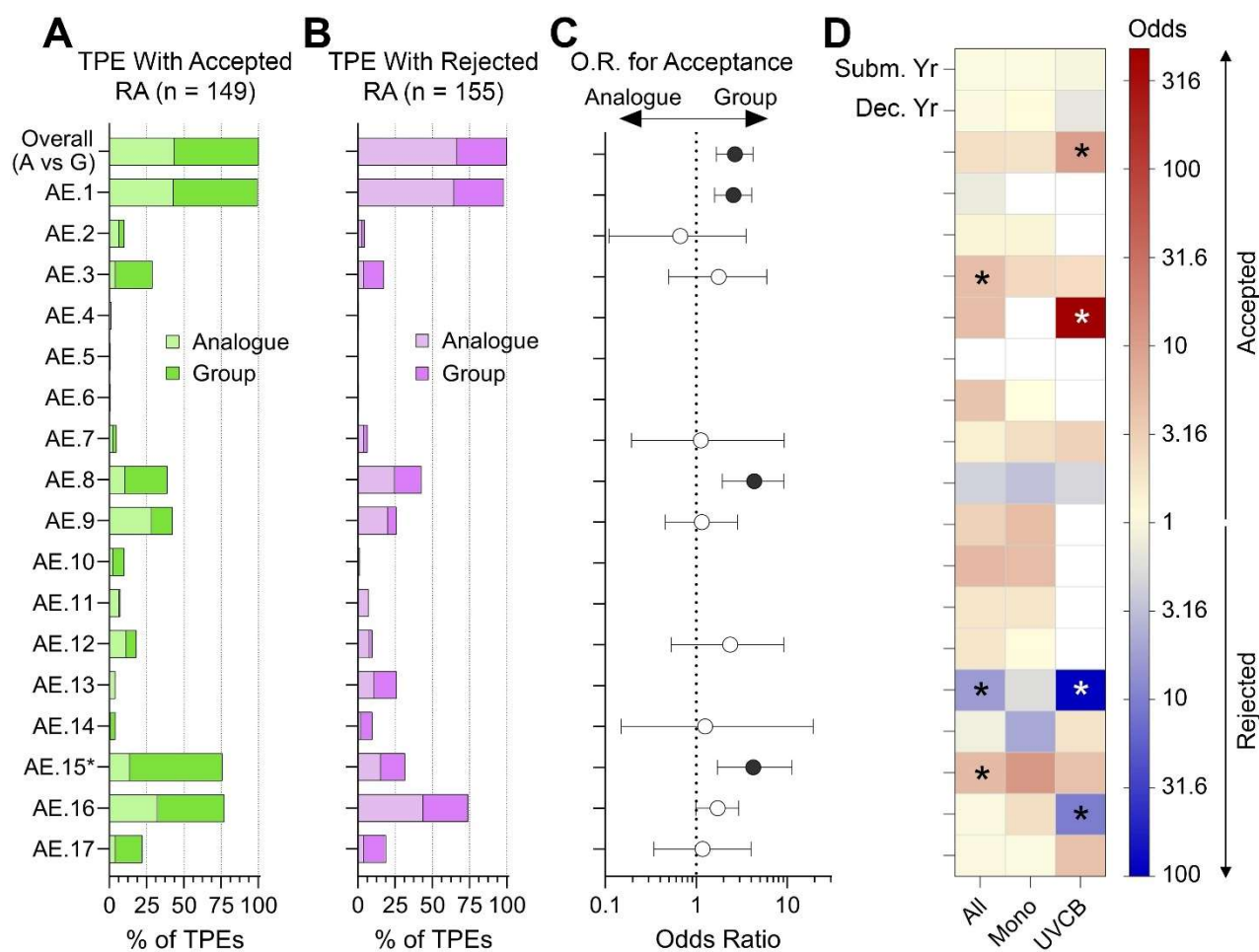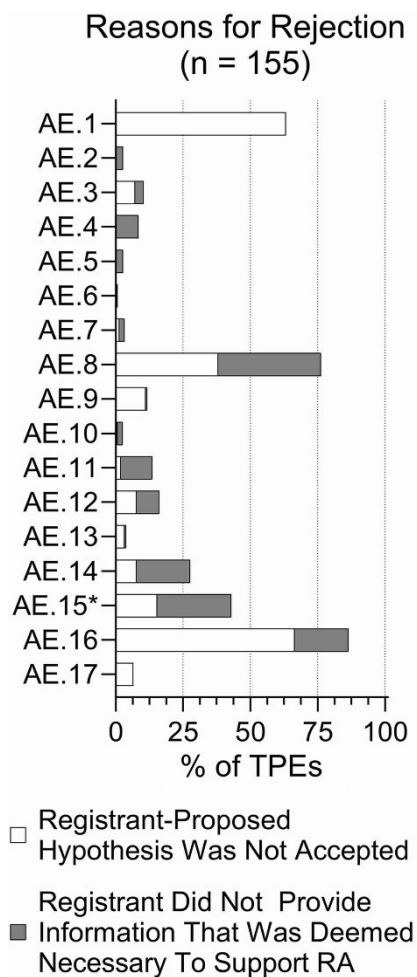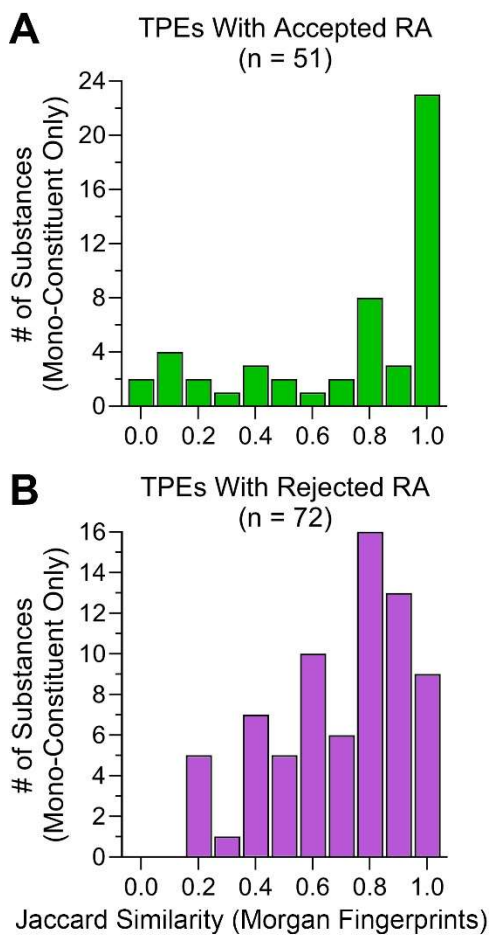
853   **Figure 8.** Reasons that were stated by ECHA for the rejection of a TPE with read-across. The light
854   bar represents when a registrant proposed an element that ECHA disagreed with and/or interpreted
855   differently. The dark bar represents the registrant failing to take an element into account in the read-
856   across hypothesis.
857

858

859

860 **Figure 9**. Source to target Jaccard similarity values for read-across proposals. **(A)** Accepted and **(B)**
861 Rejected TPEs. Mean similarity values were not significantly different. However, a rigorous
862 cutpoint analysis revealed that TPEs with very high similarity (>0.9) were significantly more likely
863 to be accepted (p=0.0003, OR=7.0).
864

865

43