Research Article

Alternative Acute Oral Toxicity Assessment Under REACH Based on Sub-Acute Toxicity Values

Andrea Gissi, Kimmo Louekari, Laurence Hoffstadt, Norbert Bornatowicz and Alberto Martin Aparicio

European Chemicals Agency (ECHA), Helsinki, Finland

Summary

The REACH Regulation requires information on acute oral toxicity for substances produced or imported in quantities greater than one ton per year. When registering, animal testing should be used as last resort. The standard acute oral toxicity test requires use of animals. Therefore, the European Chemicals Agency examined whether alternative ways exist to generate information on acute oral toxicity. The starting hypothesis was that low acute oral toxicity can be predicted from the results of low toxicity in oral sub-acute toxicity studies. Proving this hypothesis would allow avoiding acute toxicity oral testing whenever a sub-acute oral toxicity study is required or available and indicates low toxicity. ECHA conducted an analysis of the REACH database and found suitable studies on both acute oral and sub-acute oral toxicities for 1,256 substances. 415 of these substances had low toxicity in the sub-acute toxicity study (i.e., NO(A)EL at or above the limit test threshold of 1,000 mg/kg). For 98% of these substances, low acute oral toxicity was also reported (i.e., LD_{50} above the classification threshold of 2,000 mg/kg). On the other hand, no correlation was found between lower NO(A)ELs and LD_{50}. According to the REACH Regulation, this approach for predicting acute oral toxicity needs to be considered as part of a weight of evidence analysis. Therefore, additional sources of information to support this approach are presented. Ahead of the last REACH registration deadline, in 2018, ECHA estimates that registrants of about 550 substances can omit the *in vivo* acute oral toxicity study by using this adaptation.

Keywords: acute toxicity, REACH, weight of evidence, alternative methods, adaptation

1 Introduction

1.1 Acute and sub-acute toxicity under REACH

Information on acute oral toxicity is required under the REACH Regulation (EU, 2006) for substances that are manufactured or imported in quantities above 1 ton per year. Usually this information is obtained from an animal study performed according to an OECD Test Guideline (TG) or a corresponding EU test method. The test results are then used for classification under the Classification, Labelling and Packaging (CLP) Regulation (EU, 2008). In addition, information on sub-acute toxicity and screening information on reproductive toxicity are required according to REACH when the tonnage of a chemical exceeds 10 tons per year, and might be already available for chemicals in lower tonnages too.

There have been attempts to apply the 3R principle, i.e., replacement, reduction and refinement of animal experiments (Russell and Burch, 1959), to acute oral toxicity testing. In particular, following the OECD Council decision, TG 401 "Acute Oral Toxicity" was deleted on December 17, 2002 for animal welfare reasons and also because more advanced methods to characterize the acute oral toxicity of chemicals were available. The currently used OECD TGs are: OECD TG 420 (Fixed Dose Procedure), OECD TG 423 (Acute Toxic Class Method), OECD TG 425 (Up-and-down procedure). These test guidelines use fewer animals than the old OECD TG 401, are likely to induce less suffering of the test animals, and provide more information on possible target organs and possible mechanisms of toxicity. Whereas 20 animals were normally used for OECD TG 401, it is now possible to obtain valid results with 3 to 6 animals by following the newer TGs. In addition, TG 423 requires that animals be humanely killed when showing signs of evident toxicity, i.e., "*animals obviously in pain or showing signs of severe and enduring distress*".

Received September 12, 2016;
Accepted November 3, 2016;
Epub November 8, 2016;
doi:10.14573/altex.1609121

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is appropriately cited.

Two OECD Test Guidelines are available to characterize the sub-acute oral toxicity of chemicals: OECD TG 407 (Repeated Dose 28-day Oral Toxicity Study in Rodents) and OECD TG 422 (Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test). Unlike for acute oral toxicity, there are currently no standardized and approved methods that would enable a reduction of the number of test animals used in sub-acute toxicity studies.

Our experience with REACH registrations suggested that the 28-day NO(A)EL (no observed (adverse) effect level) seen in sub-acute toxicity studies (reported in the oral repeated dose toxicity section of REACH dossiers) is usually much lower than the LD_{50} obtained from acute oral toxicity studies. Risk characterization and classification of chemicals are therefore mostly driven by the NO(A)EL values derived from the sub-acute and reproductive toxicity studies and not from the LD_{50}, except in rare cases where high acute toxicity and peak exposures occur. Therefore, the value of acute oral toxicity studies in regulatory decision-making is usually lower than that of the sub-acute toxicity studies or other repeated dose toxicity studies.

1.2 REACH and alternative methods

Under the REACH Regulation (EU, 2006), registrants have obligations to generate information on registered substances by alternative means to vertebrate animal testing (such as *in vitro*, read-across and quantitative structure activity relationships (QSARs)) wherever possible (according to the principle "animal testing as the 'last resort'") as it is the intention of the legislator to apply the 3R principle in the safety assessment of chemicals. Nevertheless, according to EU legislation, the 3R principle must be applied without compromising a high level of human health and environmental protection.

1.3 Predicting acute toxicity

The value of the acute toxicity test already has been questioned previously. Prediction models based on sub-acute toxicity data or *in vitro* cytotoxicity tests have been suggested and developed. According to the respective authors, these models may replace *in vivo* acute toxicity studies partly or – in the future – in total (Creton et al., 2010; Chapman et al., 2010; Indans et al., 1998; Kinsner-Ovaskainen et al., 2013; Robinson et al., 2008; Seidle et al., 2011; Bulgheroni et al., 2009; Luechtefeld et al., 2016; Graepel et.al. 2016).

In 2014, the Joint Research Centre (JRC) of the European Commission published the "*EURL ECVAM strategy to replace, reduce and refine the use of animals in the assessment of acute mammalian systemic toxicity*"[1]. The European Union Reference Laboratory for alternatives to animal testing hosted by the JRC (EURL ECVAM, JRC) considered that efforts should be directed towards a) the reduction and replacement of animal tests for acute systemic toxicity, and b) the refinement of *in vivo* studies, according to the 3Rs principle. Moreover, the JRC suggested that consideration should be given to the mechanistic basis of acute toxicity and validation of any integrated prediction models. The JRC proposed to evaluate promising components of integrated approaches for testing and assessment (IATA), including the better use of existing alternative methods, such as mechanistically relevant *in vitro* assays. Furthermore, according to the JRC, information on repeated dose toxicity might be useful in supporting classification and labelling for acute systemic toxicity.

Luechtefeld et al. (2016) analyzed REACH registration data of 2009-2014 to develop a number of prediction models for oral systemic toxicity. Those authors came to the conclusion that "*a lack of toxicity in a 28-day study is a relatively good predictor that a chemical will be non-toxic acutely*." However, the approach presented in that paper was different compared to the present publication, i.e., the applied toxicity reference values differed, the most recent REACH registration data was not used in that study, and a detailed analysis of outliers was not provided.

On these premises, the European Chemicals Agency (ECHA) examined whether there are alternative ways to generate information on acute oral toxicity. The starting postulate for the proposal presented in this paper was that an acute oral toxicity lethal dose 50% (LD_{50}) above 2,000 mg/kg can be predicted from oral sub-acute toxicity studies when the NO(A)EL is at or above 1,000 mg/kg. The dose thresholds applied in our analysis are those indicated in the CLP Regulation (EU, 2008) for triggering classification, i.e., substances which have an LD_{50} below 2,000 mg/kg need to be classified for acute toxicity (cat. 1 to cat. 4 depending on the exact LD_{50} value), whereas substances with an LD_{50} above 2,000 mg/kg do not need classification for acute toxicity. The threshold applied for sub-acute toxicity is based on the limit test described in OECD TG 407 and 422; this does not require testing with doses higher than 1000 mg/kg. Consequently, 1000 mg/kg is very often the highest dose administered in the studies in the IUCLID database. Therefore, the focus of this analysis has been on whether low acute toxicity could be predicted from sub-acute toxicity test results and other relevant data.

According to the REACH requirement, this approach for predicting acute oral toxicity would need to be reported as part of a Weight of Evidence (WoE) analysis. WoE is the use of "*several independent sources of information leading to the assumption/conclusion that a substance has or not a particular dangerous property, while information from each single source alone is regarded insufficient to support this notion*" (EU, 2006). For acute toxicity, examples of such additional sources are *in vitro* cytotoxicity tests (e.g., Neutral Red Update Assay), QSAR and physico-chemical information.

2 Material and methods

2.1 Analysis of REACH registration data

ECHA's IUCLID database of REACH registrations contains information on all the registered substances in the format of International Uniform ChemicaL Information Database (IUCLID) dossiers. ECHA does not own the REACH registration data, which remain property and responsibility of industry. In these dossiers, information on the endpoints (i.e., properties)

[1] https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-strategy-papers/strategy-acute-mammalian-systemic-toxicity

Tab. 1: Filters applied to obtain a clean dataset of studies

Filter	Comment
Step 1: Filters applied to both acute and repeated-dose ESRs to select experimental data of **sufficient quality**	
"Test material identity same as registered substance" = "yes"	This field is available in IUCLID to distinguish cases where the test was performed on the registered substance (and should be marked as "yes" in these cases) from cases where the study refers to another substance (e.g., read-across, and should be marked "no" in such cases).
"Study type" = "experimental result"	Used to select only experimental data and to exclude other study types such as QSAR results.
Reliability score = "1" or "2"	Used to select studies with uncompromised reliability.
Unit = "mg/kg bw"	For 7.5.1 this includes nominal, actual dose received and total dose.
Step 2: Filters used to select the **relevant studies** performed according to specific OECD/EU guidelines	
Acute toxicity: OECD TG 401, 420, 423, 425, EU Method B.1 (bis and tris)	Only results expressed as LD_{50} values (including results marked as approx. LD_{50}) are selected.
Repeated dose toxicity: OECD TG 407, 422 (i.e., sub-acute toxicity TGs)	Only results expressed as NOAEL or NOEL are selected.
Step 3: **Aggregation rules**[a] applied to obtain one single value per endpoint per substance	
Lowest boundary value selected to represent the toxicity value	Used when the results of a study are given as a range (e.g., 1,000 < LD_{50} < 3,000 mg/kg bw; 1,000 is the selected value).
Lowest LD_{50} value and/or lowest NOAEL value selected	Used when more than one relevant study per endpoint was available for a substance.
Step 4: Selected substances having **both studies**	
Only substances for which relevant studies available in both 7.2.1 and 7.5.1 sections were selected.	Substances for which relevant information was available for only one of the two endpoints were excluded.
Step 5: Confirmation of the test material **identity**	
Tested material identifiers match the registered substance identifiers	Remove "hidden" read-across: There are cases where studies are marked as "experimental results" and where in the field "identity of the test material same as registered substance" is selected "yes". For these cases, providing the test material identifiers was not mandatory because it was assumed that they should match the identifiers of the registered substance. However, there are a few cases where the test material identifiers are provided in the endpoint study record and refer to a substance different from the registered substance. Such studies are therefore called "hidden" read-across, and should not be used for this analysis[b].

[a] E.g., for one substance two endpoint study records are available reporting the following results: ESR 1: LD_{50} = 2,000 mg/kg bw; ESR2: 1,000 < LD_{50} < 3,000 mg/kg bw. The first step is to aggregate the values of the ESR 2, which express the results as a range, to a single value equal to the lowest boundary of the range (1,000 mg/kg bw). The second step is to select the lowest value among the various ESRs, in this case again the value 1,000 mg/kg bw. The single value kept for this substance for section 7.2.1 is 1,000 mg/kg bw.

[b] To avoid the occurrence of such cases, the field "identity of test material same as registered substance" has been removed in the new version of IUCLID (IUCLID 6) and test material identifiers are mandatory for each study.

is stored as Endpoint Study Records (ESRs) and organized into sections. Information on acute oral toxicity and repeated dose oral toxicity is reported in sections 7.2.1 and 7.5.1, respectively. It has to be noted that:

- multiple endpoint study records may be present in a dossier for the same endpoint and section;
- section 7.5.1 also contains data other than sub-acute study data (e.g., 90-day sub-chronic studies);
- ESRs are designed in a generic format and may include studies of different types (not only experimental results but also QSARs, read-across, etc.), reliability (reported as Klimish score, i.e., from 1 to 4), guideline, etc.;
- the results in each ESR can be provided as a range of values and not only as single values.

Given the complex structure of the ESRs, the first step in the data analysis presented in this paper consisted of the filtering of studies to obtain a clean dataset of dossiers (and therefore substances) containing relevant experimental data for both sections 7.2.1 and 7.5.1. Detailed information on the series of filters on different IUCLID fields is provided in Table 1.

Regarding excluded substances, there are many possible reasons for which the pair of relevant studies might not be available, e.g., different routes of administration (often inhalation) were used for the acute and/or sub-acute toxicity tests, or one of these studies was adapted, e.g., by using information on an analogue substance (i.e., read-across adaptation). Such study record(s) were excluded from this analysis.

2.2 Analysis of the results

The results are presented using a confusion matrix and related statistics (Tab. 2). A confusion matrix is a contingency table of size 2x2 widely used in statistical classification to assess the performance of algorithms. In this specific case, where the property to be predicted is LD_{50}, the term negative (i.e., low toxicity) refers to substances with an LD_{50} value > 2,000, while the term positive (i.e., toxic) to substances with an LD_{50} value ≤ 2,000. TN (true negative) and TP (true positive) represent the number of substances that are correctly predicted to be negative and positive, respectively. FP (false positive) are the substances for which the LD_{50} value is predicted to be ≤ 2,000 but the observed value is > 2,000. FN (false negative) are the substances for which the LD_{50} value is predicted to be > 2,000 but the observed values is ≤ 2,000.

The accuracy is then calculated as the number of correct predictions (TN + TP) divided by the total number of data (TN + FN + TP + FP) (Eq. 1).

$$ACC = \frac{TN+TP}{TN+FN+TP+FP} \qquad \text{(Eq. 1)}$$

2.3 Structural analysis of the pool of substances

Some data in the REACH registration database is claimed confidential by the registrants. The full dataset of such structures and results cannot be disclosed. Nevertheless, the QSAR Toolbox v.3.3[2] was used to provide a chemical description of the data set, which can be intended as an applicability domain of the approach. The QSAR Toolbox is a software (co-developed by ECHA and OECD) that allows the retrieval of experimental data, formation of chemical categories based on mechanistic or chemical similarities, and the prediction of properties by read-across, trend analysis or (Q)SAR models.

The software was used to retrieve chemical structures starting from the CAS registry numbers of the registered substances and then characterize their chemical functional groups using the profiler "Organic functional groups (nested)", which can distinguish a total of 499 functional groups.

3 Results and discussion

3.1 Analysis of the REACH registration data

Figure 1 reports the results of each filtering stage:

Starting pool: Data for this analysis were extracted from the IUCLID database in May 2015. At that point in time, the database contained more than 40,000 dossiers for about 10,000 substances. As described above, the extraction was limited to fields in sections 7.2.1 (Acute toxicity: oral) and 7.5.1 (Repeated dose toxicity: oral) of the IUCLID dossiers. In total, 46,948 ESRs were extracted from section 7.2.1 and 37,364 from section 7.5.1.

Step 1: 25,055 studies were retained for acute toxicity and 8,950 for sub-acute toxicity;

Step 2: 18,804 studies were retained for acute toxicity and 3,308 for sub-acute toxicity;

Step 3: multiple studies were aggregated resulting in 4,418 different substances with acute toxicity studies reported and 1,759 substances with sub-acute studies;

Step 4: 1,336 substances with both acute and sub-acute studies were selected;

Step 5: 1,261 substances were selected for the analysis (see supplementary information at doi:10.14573/altex.1609121s for the full list of substances reported as links to ECHA's dissemination website when not confidential). Of these substances:

- 262 substances have an LD_{50} < 2,000 mg/kg bw (i.e., 20.8% of substances are identified as "positive" or "toxic" and must be classified), while 999 have an LD_{50} indicating low toxicity and do not need to be classified;
- 844 substances have a NO(A)EL < 1,000 mg/kg bw (i.e., 66.9% of substances are identified as "toxic" and must be classified), while 417 have a NO(A)EL indicating low toxicity and do not need to be classified.

The results show that the sub-acute toxicity test (and the associated classification threshold) is more sensitive than the acute toxicity test, because the former identifies three times more chemicals requiring classification than the latter. One explana-

Tab. 2: The confusion matrix used to visualise the results

TN stands for true negatives, FN for false negatives, FP for false positives and FN for false negatives.

		Oral LD_{50} **predicted** from the sub-acute toxicity study	
		LD_{50} > 2,000	LD_{50} ≤ 2,000
Oral LD_{50} **observed** *in vivo* in the acute oral toxicity study	LD_{50} > 2,000	**TN**	FP
	LD_{50} ≤ 2,000	FN	**TP**

[2] QSAR Toolbox v.3.3: https://www.qsartoolbox.org

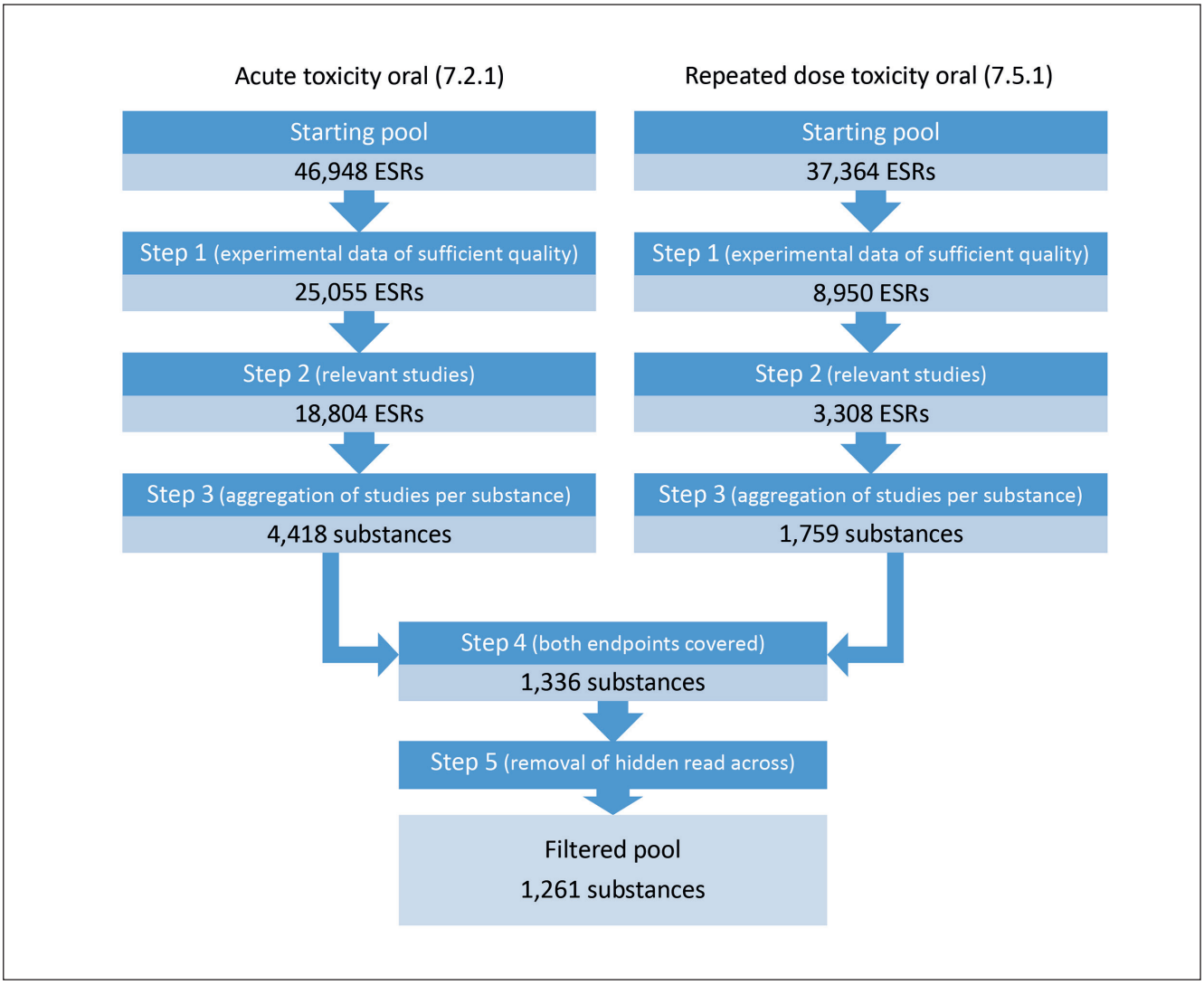

Fig. 1: Results of the extraction and filtering of the studies

tion is that some forms of toxicity only occur upon accumulation of the substance or accumulation of the effect during repeated dose exposure.

The relation between the results of the two tests is visualized in a 2x2 matrix (Tab. 3). The table shows that out of 844 substances having a sub-acute NO(A)EL triggering classification, 253 substances also have an LD_{50} value triggering classification (30.0%). This distribution indicates that sub-acute toxicity NO(A)EL classification cannot be used to predict LD_{50} classification. On the other hand, among the 417 substances with a sub-acute toxicity NO(A)EL indicating low toxicity (i.e., not triggering classification), the vast majority also has an LD_{50} not triggering classification (408, i.e., 97.8%).

This suggests that a correlation can be established and that it is possible to predict no classification for acute toxicity from sub-acute toxicity NO(A)EL ≥ 1,000 mg/kg bw. An LD_{50} value < 2,000 mg/kg bw also is a good predictor for sub-acute toxicity NO(A)EL below 1,000 mg/kg bw.

However, for classification purposes, the prediction of the exact classification based on the exact toxicity value is needed. No correlation was found to predict the exact classification of substances that are acutely toxic (analysis not reported here).

Based on these results, we established a rule that predicts LD_{50} > 2,000 mg/kg for substances with sub-acute toxicity NO(A)EL ≥ 1,000 mg/kg bw. The confusion matrix for such a model is reported in Table 4. The accuracy is: ACC = 408 / (408+9) = 0.978. An accuracy value of 0.978 indicates that predicting all the substances with a NO(A)EL ≥ 1,000 mg/kg bw to have an acute oral toxicity LD_{50} above 2,000 mg/kg bw is correct in 97.8% of the cases.

A second part of the analysis consisted of setting arbitrary lower NO(A)EL thresholds to investigate the possibility of

Tab. 3: Distribution of acute and sub-acute results for substances where both values are available
Results are expressed in mg/kg bw.

	$NO(A)EL_{28d}$ < 1,000	$NO(A)EL_{28d}$ ≥ 1,000
LD_{50} ≤ 2,000	253	9
LD_{50} > 2,000	591	408

Tab. 4: Confusion matrix using a NO(A)EL threshold of 1,000 mg/kg bw

		Oral LD_{50} **predicted** from the sub-acute toxicity study	
		LD_{50} > 2,000	LD_{50} ≤ 2,000
Oral LD_{50} **observed** *in vivo* in the acute oral toxicity study	LD_{50} > 2,000	408	0
	LD_{50} ≤ 2,000	9 [a]	0

[a] outliers discussed later in the text

Tab. 5: False negative, true negative, total number of substances predicted and accuracy using different NO(A)EL values as thresholds

Threshold NO(A)EL	1,000	500	300	200	100	30	Starting pool
Observed LD_{50} ≤ 2,000	9	19	35	44	78	152	262
Observed LD_{50} > 2,000	408	456	536	639	742	888	999
Total (domain)	417	475	571	683	820	1040	1261
Accuracy	**98%**	**96%**	**94%**	**93%**	**90%**	**85%**	**79%**

predicting a lack of acute toxicity classification from NO(A)EL lower than 1,000 mg/kg bw. Table 5 reports the related accuracy values. This analysis shows that using lower sub-acute thresholds to predict lack of acute toxicity classification also would have good performances, at least up to the 300 mg/kg bw limit (94% of the substances with a NO(A)EL higher than 300 mg/kg bw do not need acute classification).

However, to be a good/acceptable adaptation under REACH, any prediction should minimize false negatives to the greatest extent possible to ensure the protection of human safety to the greatest extent possible. Therefore, ECHA only recommends to use the strictest threshold in Table 4, i.e., 1,000 mg/kg bw, which also corresponds to the standard limit dose in sub-acute toxicity studies, as also indicated in the updated ECHA Guidance on Acute Toxicity[3].

3.1.1 Analysis of the outliers

Despite the very good correlation (98%), there were nine cases where the NO(A)EL from the sub-acute studies was higher than 1,000 mg/ kg bw, but the reported LD_{50} values were lower than 2,000 (values between 500 and 1,911 mg/kg bw). These cases represented outliers of the proposed approach. Keeping in mind that the results came from a computerized analysis, these cases were checked manually. In eight out of nine of these cases, the "mismatch" was explained:

- *Case 1:* The NO(A)EL of 1,000 mg/kg bw was reported for a 14-day dose-range finding study. In a subsequent 13-week study, death occurred at 1,000 mg/kg after four weeks. Therefore, the NO(A)EL of a sub-acute toxicity study would be below 1,000 mg/kg.
- *Case 2:* The LD_{50} given in the dossier was expressed as a range: greater than 1,000 mg/kg bw and lower than 3,000 mg/kg bw in a study performed on hamsters. The algorithm picked the lowest value (1,000) as explained above, but the exact value is not available and the range exceeds the 2,000 mg/kg bw threshold. A different study for the substance on rats reported an LD_{50} above 10,000 mg/kg bw, which fits the proposed approach.

[3] see Section R.7.4 of Chapter R.7a of the Guidance on Information Requirements and Chemical Safety Assessment, last update December 2016: https://echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf

– *Cases 3 and 4:* The LD_{50} that was below the threshold did not refer to the registered substance as such (to the anhydrous form of the substance, in one case, and to the metal content in the registered substance, in the other case). Furthermore, in one of these cases, another study reporting the LD_{50} for the registered substances was available and reported a value above 2,000 mg/kg, which fits the prediction model.
– *Case 5:* The LD_{50} indicated in the dossier was 5 mg/kg, while the manual check of the study record indicated a mistake in the reported unit. The correct LD_{50} was 5 g/kg corresponding to 5,000 mg/kg, which fits the proposed approach.
– *Cases 6, 7 and 8:* The results came from tests in which the highest doses tested were below 2,000 mg/kg bw. In all three cases, the registrants indicated an LD_{50} above 2,000 mg/kg bw based on supporting evidence in the summary.

Only one outlier was confirmed. For the only "true outlier", the NO(A)EL was above 1,000 mg/kg bw, but the LD_{50} was 1,410 mg/kg bw, which does not fit the proposed approach. In this case, the substance was administered via gavage in the acute toxicity study but in the diet in the sub-acute toxicity study.

In conclusion, for eight of the nine outliers, the acute oral LD_{50} is considered to be above 2,000 mg/kg.

3.1.2 Structural analysis

801 SMILES were found and retrieved from the QSAR Toolbox for the CAS Registry Numbers (RN) of the registered substances selected for the analysis. For the remaining CAS RN a well-defined structure was not available (e.g., for substances of "unknown or variable composition, complex reaction products or biological material" (UVCB) chemicals which represent around 30% of the substances registered under REACH).

Of a total of 499 chemical functional groups distinguished by the profiler, 227 were represented in the set of substances subject to the analysis, showing a good coverage of the chemical space (see the full list at doi:10.14573/altex.1609121s). As expected, the low toxicity substances do not contain any of the groups commonly associated with high toxicity, such as arsenic, cyanide, chromium, or azo.

3.2 Weight of evidence strategy

To be able to use information from sub-acute toxicity studies in their registration dossier, registrants need to report their finding as part of a WoE approach (see above in 1.3) following the provisions of REACH Regulation Annex XI. Therefore, additional supporting information needs to be submitted to support the evidence resulting from a 28-day NO(A)EL ≥ 1,000 mg/kg bw. Examples of supporting information are results from *in vitro* NRU cytotoxicity tests, physico-chemical properties and QSAR results. A brief review of these methods and their applicability as additional evidence to support the WoE strategy presented in this paper is given below.

3.2.1 *In vitro* NRU cytotoxicity test

Based on the EURL ECVAM validation study to assess the predictive capacity of the 3T3 NRU *in vitro* cytotoxicity test to identify substances not requiring classification for acute oral toxicity (Prieto et al., 2013), ECVAM has issued recommendations concerning the validity and limitations of this *in vitro* test (EURL ECVAM, 2013). Considering the results of that relevant validation study, the 3T3 NRU test method shows a high sensitivity (ca 95%) and, consequently, a low rate (ca 5%) of false negative results when employed in conjunction with a prediction model to distinguish potentially toxic versus non-toxic (i.e., classified versus non-classified) substances. Therefore, substances found to be negative in this test would most likely not require classification for acute oral toxicity based on a cut-off value of > 2,000 mg/kg bw.

The validated 3T3 NRU test method appears to be particularly relevant for the assessment of industrial chemicals since they are not designed to act on specific biological targets and, in general, tend not to be acutely very toxic. Following the provisions of the REACH Regulation (EU, 2006) and in particular its Annex XI, data from the 3T3 NRU test method could be used within a WoE approach to adapt the standard information requirements.

The 3T3 NRU test method is sensitive to hazardous chemicals acting through general mechanisms of toxicity common to most cell types, often referred to as "basal cytotoxicity". Consequently, chemicals which act through (i) mechanisms specific only to certain cell types and tissues (e.g., of the heart or central nervous system) may not be identified as potentially acutely toxic by this method; (ii) metabolic activation to induce toxicity, and therefore not exhibiting significant cytotoxicity, may go undetected, since the 3T3 cell model lacks significant metabolic capacity. Care must be taken therefore in interpreting negative results derived from this assay.

The 3T3 NRU test method has a high false positive rate, therefore positive results cannot be readily used in a meaningful way in characterizing the acutely toxic substances (i.e., acute toxicity classifications Cat 1 to Cat 4). A likely reason is that the test method does not capture important biokinetic processes such as absorption, distribution, metabolism and excretion. Thus, certain chemicals, despite having cytotoxic potential, may not actually be acutely toxic via the oral route.

Due to its limitations, the NRU test should primarily be used to correctly identify and classify substances of low toxicity (i.e., those which are not to be classified for acute toxicity). The 3T3 NRU test method may be a valuable component of a WoE approach for supporting hazard identification and safety assessment in agreement with the EU CLP Regulation implementing the upper threshold of UN GHS Category 4 as the cut-off for non-classification of substances (EURL ECVAM, 2013).

3.2.2 Physico-chemical data

Low reactivity, chemical and biological inertness or very low solubility are examples of physico-chemical properties of the substances that usually suggest that the bioavailability of the substance, and consequently its toxicity, will be low.

In REACH registrations, relevant data on low bioavailability have been provided for some substances, e.g., for substances that have a crystalline structure and extremely low solubility even in aggressive media (hydrogen chloride solution mimick-

ing the gastrointestinal tract). In order to contribute to the WoE, this type of data would normally need to be given as results of reliable and well-documented bioaccessibility or bio-elution tests. Simulated gastric fluid and other relevant biological media need to be used in these tests to be convincing. While the bio-elution method has not been accepted as an OECD TG, a standard protocol is available as ASTM (American Society for Testing and Materials) D-5517 (ASTM, 2007). Under the initiative of Eurometaux, development aiming at an OECD TG project is ongoing.

The rationale of "unreactivity" and lack of bioavailability as indicators of low toxicity, is referred to in the column 2 adaptation in Annex IX of the REACH Regulation (EU, 2006), 8.6.2., fourth indent, according to which "*the sub-chronic toxicity study (90 days) does not need to be conducted if: … the substance is unreactive, insoluble and not inhalable and there is no evidence of absorption and no evidence of toxicity in a 28-day 'limit test', particularly if such a pattern is coupled with limited human exposure*."

Non-reactive substances with very high molecular weight may also have a low bioavailability via the relevant routes of exposure. As for any property of a substance, it has to be considered whether metabolism may also influence reactivity.

3.2.3 QSAR predictions

Several commercial and free QSAR tools include models to predict mammalian acute toxicity. Hydrophobicity is an important descriptor exploited by most of the models, but electronic and steric effects also play a role. In has to be noted that these models for *in vivo* toxicity still cannot cover the complexity of the whole-body phenomena involved in acute toxicity. Predictions need to be critically analyzed and justified to be used. When they can be considered reliable, they can provide an important line of evidence in the overall WoE approach.

4 Conclusions

Information on acute oral toxicity is required under the REACH Regulation for substances registered at tonnages greater than 1 ton per year (EU, 2006). Based on the analysis presented in this paper, ECHA recommends that registrants use results from a 28-day oral toxicity study as part of a WoE approach for covering acute toxicity (instead of the standard animal test) when the substance is anticipated to be non-toxic (as per Annex XI, section 1.2). A NO(A)EL from a sub-acute toxicity study already might be available or, if not, is required for covering REACH Annex VIII requirements. ECHA found it acceptable to rely on studies on sub-acute oral toxicity to predict acute oral toxicity if this approach is undertaken as part of a WoE in combination with other independent sources of information, and if the NO(A)EL obtained in the sub-acute studies is at or above 1,000 mg/kg bw.

There is a high potential for registrants to adapt this standard information requirement before conducting a novel study to meet the acute toxicity information requirement. In light of the results discussed in this paper, ECHA estimates that 550 studies related to the 2018 deadline can be saved by the use of the WoE approach. ECHA provides more regulatory insight on the strategy to cover acute toxicity in the ECHA Guidance on acute oral toxicity[3].

This contributes to the "animal testing as last resort" principle of REACH, where registrants have an obligation to consider alternatives to animal testing prior to undertaking any *in vivo* study in order to avoid unnecessary testing on animals, while they aim to obtain the information necessary to assess the hazards and risks of their substance. However, omitting testing on animals must not compromise the safe use of substances. Registrants also have an obligation to document their considerations when they decide to conduct an *in vivo* assay.

Concerning the last registration deadline in 2018, it is anticipated that for many phase-in substances (65%), an *in vivo* acute oral toxicity study is already available. However, from a forecast number of 5,200 substances (with a tonnage band ≥ 10 tons per year), approximately 35% will require the generation of information for the acute oral toxicity endpoint (i.e., approximately 1,825 studies). In addition, if the newly registered substances show a distribution of toxicity values comparable to the substances already registered and subject to the analysis presented in this paper, approximately 30% of the 5,200 substances may have a 28-day oral NO(A)EL indicating low acute toxicity, and consequently a predicted acute oral LD_{50} higher than 2,000 mg/kg bw/day.

For the WoE approach to be acceptable, registrants need to submit the different pieces of evidence with a sufficient level of detail and documentation. Clear and transparent documentation and argumentation is also essential to allow ECHA to conduct its evaluation. Additional independent pieces of evidence, such as cytotoxicity tests (3T3 NRU Assay in particular), QSAR and physico-chemical information, need to be collected and submitted in the registration dossier to fulfil the legal requirement. ECHA has published a practical guide on "How to use alternatives to animal testing to fulfil your information requirements for REACH registration"[4], which provides additional advice.

References

ASTM (2007). ASTM D5517-07, Standard Test Method for Determining Extractability of Metals from Art Materials, ASTM International, West Conshohocken, PA, 2007. https://www.astm.org

Bulgheroni, A., Kinsner-Ovaskainen, A., Hoffmann, S. et al. (2009). Estimation of acute oral toxicity using the no observed adverse effect level (NOAEL) from the 28 day repeated dose toxicity studies in rats. *Regul Toxicol Pharmacol 53*, 16-19. doi:10.1016/j.yrtph.2008.10.001

Chapman, K., Creton, S., Kupferschmidt, H. et al. (2010). The value of acute toxicity studies to support the clinical management of overdose and poisoning: A cross-discipline con-

[4] https://echa.europa.eu/documents/10162/13655/practical_guide_how_to_use_alternatives_en.pdf

sensus. *Regul Toxicol Pharmacol 58*, 354-359. doi:10.1016/j.yrtph.2010.07.003

Creton, S., Dewhurst, I. C., Earl, L. K. et al. (2010). Acute toxicity testing of chemicals – opportunities to avoid redundant testing and use alternative approaches. *Crit Rev Toxicol 40*, 50-83. doi:10.3109/10408440903401511

EURL ECVAM (2013). Recommendation on the 3T3 NRU Assay for Supporting the Identification of Substances Not Requiring Classification for Acute Oral Toxicity, April 2013, European Commission, Joint Research Centre, Report EUR 25946 EN. http://bit.ly/2pWCpFy

EU – European Union (2006). Regulation (EC) No 1907/2006 of The European Parliament and of The Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). *Off J Eur Union L396*, 10.04.2014. http://bit.ly/1pDFG2I

EU (2008). Regulation (EC) No 1272/2008 of The European Parliament and of The Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending Regulation (EC) No 1907/2006. *Off J Eur Union L353*, 31.12.2008. http://bit.ly/1F3nLbO

Graepel, R., Asturiol, A., Prieto, P. and Worth, A. P. (2016). Exploring waiving opportunities for mammalian acute systemic toxicity tests. *Altern Lab Anim 44*, 271-279.

Indans, I., Fry, T., Parsons, P. and Evans, P. (1998). Classification and labelling of new industrial chemicals for acute toxicity, skin and eye irritation. *Hum Exp Toxicol 17*, 529.

Kinsner-Ovaskainen, A., Prieto, P., Stanzel, S. and Kopp-Schneider, A. (2013). Selection of test methods to be included in a testing strategy to predict acute oral toxicity: An approach based on statistical analysis of data collected in phase 1 of the ACuteTox project. *Toxicol Vitro Int J Publ Assoc BIBRA 27*, 1377-1394. doi:10.1016/j.tiv.2012.11.010

Luechtefeld, T., Maertens, A., Russo, D. P. et al. (2016). Analysis of public oral toxicity data from REACH registrations 2008-2014. *ALTEX 33*, 111-122. doi:10.14573/altex.1510054

Prieto, P., Cole, T., Curren, R. et al. (2013). Assessment of the predictive capacity of the 3T3 neutral red uptake cytotoxicity test method to identify substances not classified for acute oral toxicity (LD50>2000 mg/kg): Results of an ECVAM validation study. *Regul Toxicol Pharmacol 65*, 344-365. doi:10.1016/j.yrtph.2012.11.013

Robinson, S., Delongeas, J.-L., Donald, E. et al. (2008). A European pharmaceutical company initiative challenging the regulatory requirement for acute toxicity studies in pharmaceutical drug development. *Regul Toxicol Pharmacol 50*, 345-352. doi:10.1016/j.yrtph.2007.11.009

Russell, W. M. S. and Burch, R. L. (1959). *The Principles of Humane Experimental Technique*. Johns Hopkins Bloom School of Public Health. http://altweb.jhsph.edu/pubs/books/humane_exp/het-toc

Seidle, T., Prieto, P. and Bulgheroni, A. (2011). Examining the regulatory value of multi-route mammalian acute systemic toxicity studies. *ALTEX 28*, 95-102. doi:10.14573/altex.2011.2.095

Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgements

The authors thank their colleague Pascal Phrakonkham for commenting and reviewing the paper and Pilar Prieto-Peraita for her scientific support and comments.

Correspondence to

Kimmo Louekari
European Chemicals Agency (ECHA)
Annankatu 18
00121 Helsinki, Finland
Phone: +358 9 6861 8459
e-mail: Kimmo.Louekari@echa.europa.eu