The Use of Computational Methods for the Assessment of Chemicals in REACH

Tsakovska I.[1], Worth A.[2]*

[1] *Centre of Biomedical Engineering, Bulgarian Academy of Sciences*
105 Acad. G. Bonchev Str., 1113 Sofia, Bulgaria
[2] *Institute for Health and Consumer Protection*
European Commission Joint Research Centre, Italy
E-mail: andrew.worth@ec.europa.eu

Summary: This article provides an overview of the use of computational methods in chemicals hazard and risk assessment under the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) legislation. The key aspects of the REACH guidance on the assessment of chemicals are discussed that treat the possible stepwise (tiered) approach combining multiple computational methods in assessing chemicals. Several publicly accessible software tools for the computer-based estimation of chemical hazard, developed by the European Commission's Joint Research Centre (JRC), are described.

Keywords: QSAR, Read-across, Grouping methods, Software tools, Chemical hazard, Regulatory assessment of chemicals, REACH

1. INTRODUCTION

The current European chemicals legislation REACH (Registration, Evaluation, Authorisation and Restriction of CHemicals) came into force on 1 June 2007 [10]. One of the aims of the new legislation is to harmonise the information requirements for New and Existing Chemicals, thereby improving the information base for assessing the risks of Existing Chemicals. It has been estimated that approximately 30,000 Existing Chemicals (phase-in chemicals under REACH) for which environmental and toxicological datasets are currently incomplete will need additional information. This raised understandable financial and animal welfare concerns. To address these concerns REACH calls for utilizing where possible computational methods such as (Q)SARs (Quantitative Structure – Activity Relationships), read-across and grouping methods. In particular, there is an obligation to carry out vertebrate testing only as a last resort, and to consider all other options before industry

---

* *Corresponding author*

performs (or the authorities require) testing, as stipulated in REACH Article 25 [10]. This includes the need to gather all existing information on physico-chemical, toxicological and ecotoxicological properties of a substance, including information generated by (Q)SARs and chemical grouping methods. In the REACH Guidance on Information Requirements and Chemical Safety Assessment it is recommended that these and other tools should be used in a stepwise (tiered) approach [9]. This means the hazard of chemicals can be assessed by using multiple computational tools within the overall framework of a tiered non-testing approach.

With the overall aim to promote the regulatory use of (Q)SAR methods [26] and particularly to promote the availability of reliable computer-based estimation methods for the regulatory assessment of chemicals under the scope of REACH, the European Commission's Joint Research Centre (JRC) has been developing, in collaboration with various partners, a range of user-friendly and publicly accessible software tools, such as Toxtree, Toxmatch, DART, QSAR Model Database [19].

Toxtree predicts various kinds of toxic effects and modes of toxicological action by applying decision tree approaches. Toxmatch generates quantitative measures of chemical similarity that can be used to compare datasets and to calculate pairwise similarity between compounds. Toxmatch can thus be used to support the formation of chemical groups (categories) and the application of read-across between analogues within groups. Toxtree and Toxmatch were developed under contract to the JRC by Ideaconsult (Sofia, Bulgaria). DART (Decision Analysis by Ranking Techniques was developed, under JRC contact by Talete srl (Milan, Italy) to make ranking methods available to scientific researchers. DART is designed to support the ranking of chemicals according to their environmental and toxicological concern. Finally, the JRC is developing a web-based inventory of (Q)SAR models (the JRC QSAR Model Database) to help to identify relevant (Q)SARs for chemicals undergoing regulatory review. The JRC QSAR Model Database provides publicly-accessible information on QSAR models and will enable any developer or proponent of a (Q)SAR model to submit this information by means of a QSAR Model Reporting Format (QMRF). These *in silico* tools are freely available from the JRC website (http://ecb.jrc.ec.europa.eu/qsar/).

2. A STEPWISE APPROACH FOR USING NON-TESTING DATA IN REACH

In the REACH guidance on the assessment of chemicals, a number of computational tools are recommended to be used in a stepwise (tiered) approach [9]. A structured workflow for the generation and use of non-testing data is proposed [2, 9]. It comprises a sequence of operations exploiting the functionalities of a wide array of computational tools and databases. Some of these tools are already available, whereas others need to be developed. The workflow is summarized in Fig. 1 and includes a number of steps.

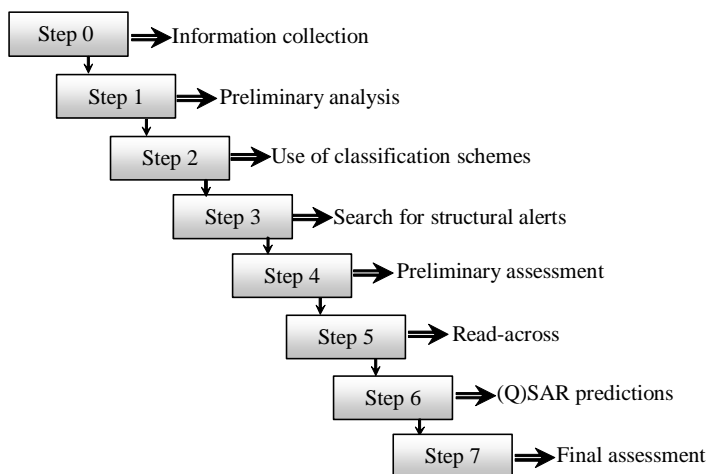| Step 0 | ⟹ Information collection |
Step 1	⟹ Preliminary analysis
Step 2	⟹ Use of classification schemes
Step 3	⟹ Search for structural alerts
Step 4	⟹ Preliminary assessment
Step 5	⟹ Read-across
Step 6	⟹ (Q)SAR predictions
Step 7	⟹ Final assessment

Fig. 1 Workflow of the stepwise approach for the use of non-testing data in regulatory assessment of chemicals

In the starting step (Step 0), information on experimentally estimated properties is collected. It is is used as a growing summary of the workflow process. Thus, it is possible to identify information gaps by comparing the REACH information requirements and the collected information. An endpoint for which non-testing data is needed and which can be generated by means of (Q)SAR methods and category/read-across approaches is then selected, and one or more of Steps 1-7 are followed to obtain the non-testing data. Step 1 involves a preliminary analysis of the reactivity, uptake and fate profile expected for both the substance of interest and its

transformation products. It is based on existing information and/or physico-chemical data. Available information can be retrieved from the peer-reviewed literature and from existing software tools and databases. Step 2 includes usage of classification schemes (where available) for the endpoint of interest to collect further information on the biological activity of the compound. For example, the classification scheme developed by Cramer et al [6] is useful for evaluating the likely systemic toxicity of a compound. Step 3 involves an investigation for the presence of structural alerts within the chemical(s) of interest (e.g. the BfR (Bundesinstitut für Risikobewertung) rules for eye irritation/corrosion and skin irritation/corrosion). Several commercial software programs are available for performing this kind of analysis - Derek (Lhasa Ltd, UK), MCASE (MultiCASE Inc, USA), Leadscope (Leadscope Inc, USA). Step 4, which requires expert judgement, involves a preliminary assessment of the expected uptake, toxicity and fate profile of the query compound, using the outcomes of Steps 1-3. In Step 5 the endpoint information for one or more source chemicals (analogues of the query compound), is used to make a prediction of the endpoint for the target chemical. Read-across is based on the identification of similar compounds. In Step 6 predictions of biological activity of the chemical are generated by using (Q)SAR models or expert systems that incorporate such models. Finally an overall assessment is carried out in Step 7, based on the information from Steps 1-6 for the chemical and endpoint(s) of interest. The workflow is designed as a flexible that means, certain steps may be omitted, or performed in a different order depending on the endpoint of interest and regulatory purpose.

3. COMPUTATIONAL TOOLS POTENTIALLY USEFUL FOR REGULATORY PURPOSES

A large and increasing number of freely available and commercial computational tools have been developed that could be applicable in the regulatory assessment of chemical risk. They are based on different (Q)SAR approaches and statistical methods and could be used for different purposes as descriptor generation, molecular similarity analysis, analogue searching, hazard assessment, etc. Commercially available software tools have been described else where [for example, 7, 13]. Examples of freely available tools include the OECD QSAR Application Toolbox [20], AMBIT [1] and

tools developed in the CAESAR [5] and OpenTox [15]. In the section below an overview of several freely available open-source tools as developed within the JRC are presented (for more details see [19]).

Toxtree
Toxtree is an open-source software application capable of estimating different types of toxic effects and modes of action by applying decision tree approaches (Fig. 2) [23].

Fig. 2 Example of using Toxtree functionalities

Currently it includes the following decision trees:

- Cramer decision tree to estimate toxic hazard [6]. Recently a new module has been developed that has in addition five new rules;
- Verhaar scheme for predicting mode of action for ecotoxicity [24];

- A decision tree for estimating skin irritation and corrosion potential [25];
- A decision tree for estimating eye irritation and corrosion potential [12];
- A decision tree for estimating carcinogenicity and mutagenicity [3];
- START (Structural Alerts for Reactivity in Toxtree) biodegradation and persistence plug-in based on a compilation of structural alerts for environmental persistence and biodegradability;
- Structural alerts for the in vivo micronucleus assay in rodents [4];
- Structural alerts for identification of Michael Acceptors [21].

The program implements a decision tree editor that allows existing rulebases to be adapted and new rulebases to be easily developed.

Toxmatch
Chemical similarity is an important concept in the field of computational toxicology. It is based on the hypothesis that similar compounds have similar properties [14]. This is a relative concept, rather than an absolute one. Hence two chemicals cannot be similar in absolute terms, but only with respect to some property. The practical realisation of the concept is based on numerical representation of the compound and a measure (similarity index) between these representations.

Toxmatch is an open-source software application that encodes several chemical similarity indices to facilitate the grouping of chemicals into categories and read-across (Fig. 3) [22].

The core functionalities include: (i) comparison of datasets based on various structural and descriptor-based similarity indices; (ii) calculation of pairwise similarity between compounds; (iii) calculation of aggregated similarity of a compound to a set. The datasets from .mol, .sdf files, or from a list of SMILES codes arranged in txt, .csv and .xls files could be loaded. Descriptors can be calculated on the fly or imported from separate files. A range of similarity indices are implemented, including Euclidean Distance Index, Hodgkin-Richards index, Tanimoto index, Cosine-like

similarity index or Carbó index. The applicability of Toxmtach is reported elsewhere [11, 17].
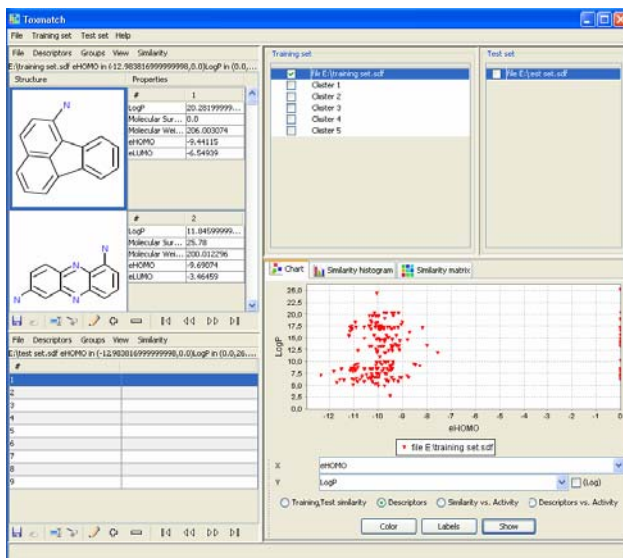

Fig. 3 Example of using Toxmach functionalities

DART
A way to perform analysis of the order relationship among elements (chemical substances, chemical processes, etc.) is by ranking methods. They can be roughly classified as total and partial-order ranking methods, according to the specific order they provide. In the complex systems evaluated by ranking strategies, elements are described by several attributes, also referred to as criteria. Thus the system must be analysed by more than one criterion, and decisions must be made by taking several criteria into account simultaneously. A detailed description of the application of these methods is given in [18].

DART is an open-source application that implements seven total ranking methods and one partial ranking method, the Hasse diagram technique (Fig. 4) [8]. It also implements several types pre-processing analysis, like cluster analysis by k-Means, Principal Component Analysis, the bin partition, reduction of significant digit.
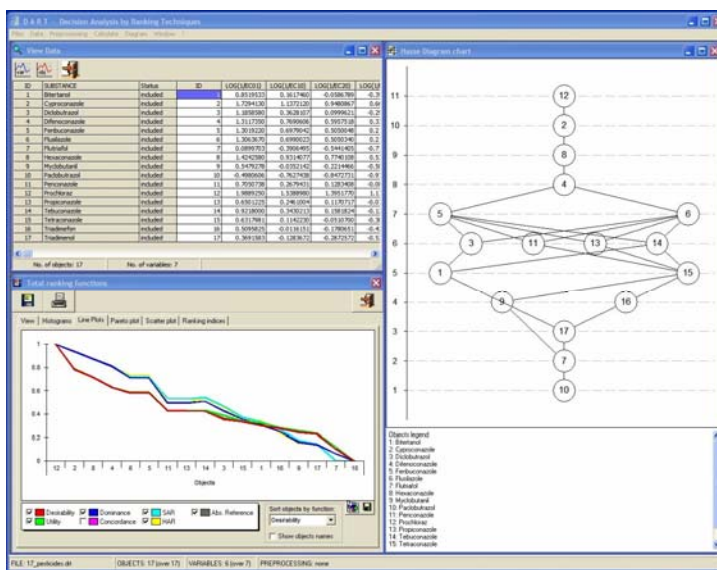
Fig. 4 Example of using DART functionalities

DART can be used for the ranking of chemicals according to their environmental and toxicological concern.

JRC QSAR Model Database
Taking into consideration the huge amount of computational models for chemicals assessment available and in order to facilitate their use under REACH, the JRC is developing QSAR Model Database, that is a web-accessible and searchable inventory of (Q)SAR models. Each model is documented in the form of a QMRF, which is a robust summary of the model. The structure of the QMRF reflects the OECD principles for the validation of (Q)SAR models for regulatory purposes [16], which are collectively intended to provide a systematic and transparent approach to describing QSAR models that will facilitate their regulatory acceptance. The QMRF format can be downloaded from the JRC site. There is also an offline QMRF editor that the developer or proponent of a model can use to prepare a model-specific QMRF before submission to the JRC for inclusion in the QSAR database. It is important to note that there is no official adoption procedure for QSARs, and the publication of a QMRF by

the JRC does not imply its endorsement by the JRC or the European Commission.

4. CONCLUSIONS

REACH foresees the use (Q)SARs to indicate the presence or absence of a certain dangerous property when testing is not considered necessary. However (Q)SAR results can only be used for regulatory purposes if they are considered relevant, reliable and adequate for the purpose, and if they are documented in an appropriate manner. In general, the integrated use of multiple estimation methods is considered to provide a more robust assessment than the use of any method on its own. To facilitate the integrated use of non-testing methods, a stepwise approach integrating (Q)SAR and grouping approaches has been proposed. To support the practical implementation of non-testing methods, the JRC has been developing a number of freely available open-source computational tools to help the use of *in silico* methods in the regulatory assessment of chemicals. A major challenge for the near future lies not so much in the development of new software tools, but in developing guidance on how to use the outputs of existing computational models in a range of regulatory contexts, including classification and labeling, risk assessment, and the identification of substances of very high concern (e.g. persistent, bioaccumulative and toxic chemicals; carcinogenic, mutagenic and reprotoxic chemicals; endocrine disruptors). Another challenge, in the longer term, will be to embed non-testing methods in the context of wider testing strategies including biological information from *in vitro* methods and novel high-throughput screening and "*omics*" methods.

ACKNOWLEDGEMENTS

REFERENCES

1. AMBIT v2.0. http://ambit.sourceforge.net/ (accessed 30 October 2009).
2. Bassan A., A. P. Worth, The Integrated Use of Models for the Properties and Effects of Chemicals by means of a Structured Workflow, *QSAR & Comb. Sci.*, 2008, 27, 6-20.
3. Benigni R., C. Bossa, N. Jeliazkova, T. Netzeva, A. Worth, The Benigni / Bossa rulebase for mutagenicity and carcinogenicity – a module of Toxtree, European Commission report EUR 23241 EN, 2008.
4. Benigni R., C. Bossa, O. Tcheremenskaia, A. Worth, Development of structural alerts for the in vivo micronucleus assay in rodents, European Commission report EUR 23844 EN, 2009.
5. CAESAR project. http://www.caesar-project.eu/ (accessed 30 October 2009).
6. Cramer G. M., R. A. Ford, R. L. Hall, Estimation of Toxic Hazard - A Decision Tree Approach, *J. Cosmet. Toxicol.*, 1978, 16, 255-276.
7. Cronin M T. D., G. M. Maggiora, Predicting Chemical Toxicity and Fate. CRC Press, New York, 2004.
8. DART v2.05. http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=DART (accessed 30 October 2009).
9. European Chemicals Agency, Guidance on information requirements and chemical safety assessment, Helsinki, 2008. Available at http://guidance.echa.europa.eu/docs/guidance_document/information_requirements_en.htm (accessed 30 October 2009).
10. European Commission, Regulation (EC) No. 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No. 793/93 and Commission Regulation (EC) No. 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC, Off. J. Eur. Union L396 (2006), pp. 374–375.

11. Gallegos Saliner A., A. Poater, N. Jeliazkova, G. Patlewicz, A. P. Worth, Toxmatch - A Chemical Classification and Activity Prediction Tool based on Similarity Measures. *Regul. Toxicol. Pharm.*, 2008, 52, 77-84.

12. Gerner I., M. Liebsch, H. Spielmann, Assessment of the eye irritating properties of chemicals by applying alternatives to the Draize rabbit eye test: the use of QSARs and in vitro tests for the classification of eye irritation, *ATLA*, 2005, 33, 215-237.

13. Greene N., Computer systems for the prediction of toxicity: an update, *Adv. Drug Deliv. Rev.*, 2002, 54, 417-431.

14. Johnson A. M., G. M. Maggiora, Concepts and Applications of Molecular Similarity. John Willey & Sons, New York, 1990.

15. OpenTox project. http://opentox.org/ (accessed 30 October 2009).

16. Organisation for Economic Cooperation and Development, The report from the expert group on (quantitative) structure activity relationship ([Q]SARs) on the principles for the validation of (Q)SARs, Series on Testing and Assessment No. 49 (ENV/JM/MONO(2004)24), Paris, 2004.

17. Patlewicz G., N. Jeliazkova, A. Gallegos Saliner A. A. P. Worth AP, Toxmatch – A new software tool to aid in the development and evaluation of chemically similar groups. *SAR QSAR Environ. Res.*, 2008, 19, 397–412.

18. Pavan M., R. Todeschini, Multicriteria Decision Making Methods. In: Walczak B., R. Taulér, S. Brown (Eds.), Comprehensive Chemometrics, Elsevier, Amsterdam, 2009, 591-629.

19. Pavan M, A. P. Worth, Publicly-accessible QSAR software tools developed by the Joint Research Centre, *SAR QSAR Environ. Res.*, 2008, 19, 785-799.

20. QSAR Application Toolbox v1.1. http://www.oecd.org/document/54/0,3343,en_2649_34379_42923638_1_1_1_1,00.html (accessed 30 October 2009).

21. Schultz T. W., J. W. Yarbrough, R. S. Hunter, A. O. Aptula, Verification of the Structural Alerts for Michael Acceptors. *Chem. Res. Toxicol.*, 2007 20, 1359–1363.

22. Toxmatch v1.06. http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXMATCH (accessed 30 October 2009).

23. Toxtree v1.60. http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE (accessed 30 October 2009).
24. Verhaar H.J.M., C. J. van Leeuwen and J. L. M. Hermens, Classifying environmental pollutants. 1. Structure-activity relationships for prediction of aquatic toxicity. *Chemosphere*, 1992, 25, 471-491.
25. Walker J. D., I. Gerner, E. Hulzebos, K. Schlegel, The Skin Irritation Corrosion Rules Estimation Tool (SICRET), *QSAR Comb. Sci.*, 2005, 24, 378-384.
26. Worth A. P., A. Bassan, J. de Bruijn, A. Gallegos Saliner, T.I. Netzeva, G. Patlewicz, M. Pavan, I. Tsakovska, S. Eisenreich, The role of the European Chemicals Bureau in promoting the regulatory use of (Q)SAR methods, *SAR QSAR Environ. Res.*, 2007, 18, 111–125.