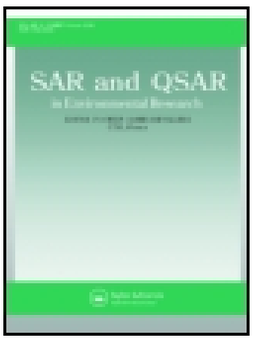

SAR and QSAR in Environmental Research

ISSN: 1062-936X (Print) 1029-046X (Online) Journal homepage: http://www.tandfonline.com/loi/gsar20

Integrating QSAR and read-across for environmental assessment

E. Benfenati, A. Roncaglioni, M.I. Petoumenou, C.I. Cappelli & G. Gini

To cite this article: E. Benfenati, A. Roncaglioni, M.I. Petoumenou, C.I. Cappelli & G. Gini (2015): Integrating QSAR and read-across for environmental assessment, SAR and QSAR in Environmental Research, DOI: 10.1080/1062936X.2015.1078408

To link to this article: http://dx.doi.org/10.1080/1062936X.2015.1078408

SAR and QSAR in Environmental Research, 2015
http://dx.doi.org/10.1080/1062936X.2015.1078408

Integrating QSAR and read-across for environmental assessment[$]

E. Benfenati[a]*, A. Roncaglioni[a], M.I. Petoumenou[a], C.I. Cappelli[a] and G. Gini[b]

[a]*IRCCS - Istituto di Ricerche Farmacologiche Mario Negri, Milano, Italy;* [b]*Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milano, Italy*

(*Received 22 June 2015; in final form 28 July 2015*)

Read-across and QSAR have different traditions and drawbacks. We address here two main questions: (1) How do we solve the issue of the subjectivity in the evaluation of data and results, which may be particularly critical for read-across, but may have a role also for the QSAR assessment? (2) How do we take advantage of the results of both approaches to support each other? The QSAR model starts from the training set. The presence of similar chemicals with property values close to that predicted can support the result. The approach in read-across is the opposite. The assessment is focused on the few substances similar to the target. The data quality of the similar chemicals is fundamental. A risk is poor standardization in the definition of 'similarity', because different approaches may be applied. Inspired by the principles of high transparency and reproducibility, a new program for read-across, called ToxRead, has been developed and made freely available (www.toxgate.eu). The output of ToxRead can be compared and integrated with the output of QSAR, within a weight-of-evidence strategy. We discuss the evaluation and integration of ToxRead and QSAR with examples of the assessment of bioconcentration factors of chemicals.

Keywords: read-across; BCF; REACH; QSAR; log *P*

1. Introduction

The EC regulation Registration, Evaluation, Authorisation and restriction of Chemicals (REACH) (EC Regulation 1907/2006) requires that the registrant assess if alternative methods are available, to provide valid results. Among the alternative methods, the so-called non-testing methods represent a possibility. Non-testing methods include (quantitative) structure–activity relationship ((Q)SAR) and read-across. However, a main drawback of read-across is that the evaluation, typically done by the human expert, is quite subjective. The evaluation of similar compound(s) and the reasoning about similarity are often done manually, on the basis of a mental process that is often difficult to replicate.

The issue of chemical-to-chemical similarity is not directly present in the case of QSAR models. In the case of QSAR models the target chemical is in some way compared with the whole population of chemicals as the basis of the model, and this is addressed within the so-called applicability domain of the model. Thus, the comparison is done not between one chemical and another, or a few others, as in the read-across case, but with the whole set of compounds used for the model.

*Corresponding author. Email: emilio.benfenati@marionegri.it
[$]Presented at the 8th International Symposium on Computational Methods in Toxicology and Pharmacology Integrating Internet Resources, CMTPI-2015, June 21–25, 2015, Chios, Greece.

However, in QSAR models we should distinguish between models which are statistically based, for which the above is true, and SAR models which are based on rules, usually called structural alerts. In the latter case, the overall structure of the model is like a collection of read-across models, with the special situation that rules for similarity are unequivocally defined by the presence of the structural alert. Thus, if the target chemical has a certain structural alert, it belongs to the family characterized by that alert, and thus it is labelled as toxic. (Within some models there are exception alerts, thus the presence of the exception alert overrules the toxic effect; this applies, for instance, to the presence of a sulfonic group which can cancel the genotoxic activity of the aromatic amine.) The structural alert is the ontology which defines the family, but in some cases the structural alert is supported by one single chemical, thus this is conceptually very similar to the one-to-one read-across model. This is the case, for instance, in the rule for genotoxicity of SA_1: Acyl halides [1].

Thus, SAR models have similarities with read-across models. However, the read-across approach is typically not so strictly formalized, and thus different results may be obtained based on expert judgment.

Furthermore, several QSAR include tools to visualize the chemicals used to build up the model. These tools are used to evaluate the reliability of the model; however, they can be also used for read-across. For example, VEGA and TEST [2,3], a couple of freely available platforms for QSAR models, show the most similar compounds to the target chemical, and this information can be easily used for read-across. These considerations indicate that there is a certain overlap between QSAR and read-across.

In the case of QSAR, the 'similarity' is usually addressed with measurements related to the chemical structure, such as the presence of certain fragments, or using descriptors as done within VEGA, TEST and AMBIT [2–4].

However, the way to assess similarity within read-across exceeds the chemical similarity. Annex XI within REACH specifies some other features which may be addressed, such as a common functional group. Recently the ECHA published a document with a framework to address read-across [5]. One of the scenarios described in this document covers the analogue approach for which the read-across hypothesis is based on different compounds which have the same type of effect, and the mutagenicity module of ToxRead can be applied to address this case. Other scenarios cover the category approach, and the bioconcentration factor (BCF) module of ToxRead can be used for these scenarios.

The multiple choices offer more powerful strategies using read-across, but the drawback is an increased level of subjectivity. Each expert is guided by his or her past experience; pieces of information may escape her or his knowledge, the weight assigned to each element of evidence and value may be different, and expressed in a subjective way, such as likely, plausible, reasonable, level of concern, etc. Typically, this evaluation is not associated with a quantitative value, and thus the integration between different results, such as different QSAR models and read-across, is quite difficult.

Thus, even though examples of integration of results of QSAR and read-across models have appeared very recently [6], the way to assess read-across is through manual, individual expert assessment.

We show here the use of non-testing methods in the case of BCF. BCF has a significant contribution in environmental assessments of the aquatic compartment [7]. BCF expresses the likelihood of a chemical to concentrate and consequently accumulate in an organism, increasing the risk of adverse effects [8]. REACH requires BCF data, which are usually used within the Globally Harmonized System. BCF is typically done using the OECD test No. 305 [9].

The disadvantages of experimental studies for determining the BCF of each chemical are the high cost and the requirement of hundreds of vertebrate test animals [10]. The duration of the test is at least 1 month. Considering the methods used by registrants to address BCF in REACH, this endpoint is one of those with greater use of alternative methods, which have increased in comparison with the first evaluation deadline [11]. The most used alternative method to deal with BCF was weight of evidence (where also QSAR estimations are sometimes nested), followed by read-across (whose use has, however, decreased) and QSAR.

We studied an approach that offers the expert a systematic and reproducible way to analyse the reasons for similarity in a guided way, showing data of good quality on similar chemicals. In order to mitigate the subjectivity problems we developed a program for read-across, called ToxRead [12,13]. It has been already described for mutagenicity. Here we extend its use for BCF, and present the integration of its results with QSAR results.

2. Materials and methods

2.1 *Experimental BCF data*

Data for ToxRead have been derived from the set of compounds available within VEGA [14]. The source of these data has been described in the literature, referring to the VEGA models on BCF [15,16]; basically they derive from data considered of good quality, such as the EURAS database [17], with the addition of further compounds from the literature [8,18]. When multiple data for the same compounds were available the arithmetic mean value was used in the database for the read-across. In total there are 860 chemicals in the database supporting the read-across.

2.2 *The log* P *values*

The log *P* values derive from two sources. One source is the experimental data obtained from the EPI Suite KOWWIN model [19]. When the experimental value is not available, it is calculated using the VEGA KOWWIN model [3].

2.3 *The algorithm for visualization*

The algorithm for the visualization of similar compounds has been adapted from the ToxRead model for mutagenicity, with some modifications necessary to deal with a continuous endpoint as the BCF rather than a categorical one, as described within the results section. The algorithm for measuring similarity is the same used within VEGA [20].

2.4 *The rules for BCF*

For the BCF endpoint, four sets of rules have been used. The first set of rules already exists within the CAESAR model of the VEGA platform [3]; these rules, obtained with SARpy software [21], specify if a chemical has a high probability of having a low BCF value, despite its high log *P* value. The second set is a couple of new rules, also obtained with SARpy, which identify compounds with high log *P* and BCF values. The third set of rules has been obtained with the istChemFeat software developed by KODE within the CALEIDOS project [22,23]; istChemFeat identifies if, in a given molecule, there is an atom/fragment, from a list of about 700, representing chemical classes and atom-centred fragments. ToxRead finds if the

target chemical has any of these atom-centred fragments, and on this basis it shows the expected range of BCF value. Thus, families of compounds with the same fragment can be evaluated. In this case, ToxRead indicates if the BCF value of chemicals within a family is homogeneous, and this family is characterized by its mean value and standard deviation. Finally, the fourth set of rules is related to empirical observation and reasoning, such as the correlation among molecular weight, log *P* and BCF values, which is higher within the range of log *P* values between 2 and 6.

2.5 *Using the ToxRead software*

The user has to choose the property of interest, in this case BCF. In the current version there are only two properties. A previous paper discussed the use of the ToxRead software for mutagenicity [13]. The chemical structure is introduced as SMILES. The user has to choose the number of similar chemicals to be shown. This is the only option allowed.

3. Results

3.1 *The visualization scheme of ToxRead for BCF*

The target chemical is shown as a circle at the centre of the graph. The chemical structure (with the SMILES: O=[N+]([O-])c1c(c(cc(c1Cl)Cl)Cl)) of the target chemical is shown in a box on the right (see Figure 1). The target chemical shown in blue is linked to three other circles, which represent the most similar chemicals, in a number chosen by the user. The typical number is three, but depending on the chemical it may be useful to explore a larger number.

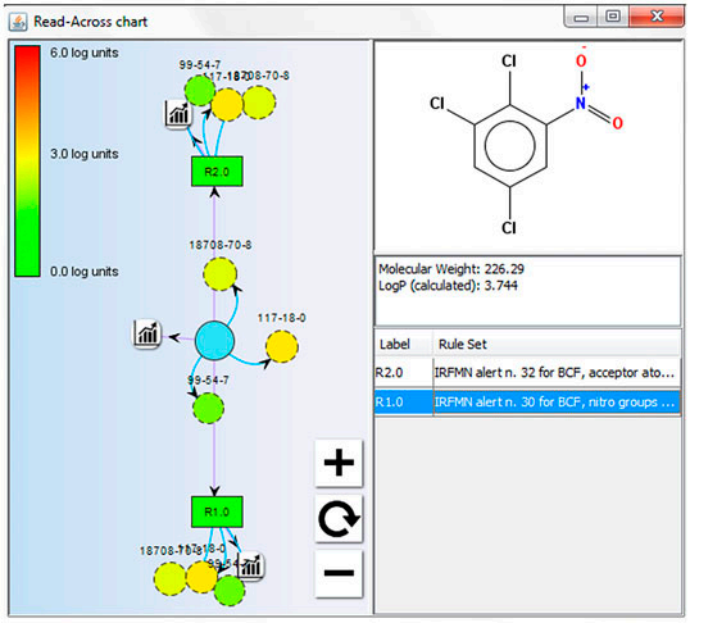

Figure 1. ToxRead screen showing the similar molecules and the rules found in the analysis of the molecule: O=[N+]([O-])c1c(c(cc(c1Cl)Cl)Cl).

Below we provide an example of using different number of similar chemicals. The similarity is measured as described in the material and methods section. This provides only a first level of evidence. Further information is obtained through the analysis of the rules. The rules are endpoint specific, and allow more insights to be obtained on the likely property value, depending on the endpoints. In other words, the first level of evaluation through chemical similarity is purely on a structural basis, and it is the same approach applied to all properties. It allows a rough assessment of the likely property value, on the basis of property values of the most similar compounds. It is conceptually a kind of kNN approach, purely done on chemometric, general features. This approach is suitable in case of poor availability of specific rules, which are endpoint specific, and provides simple information related to similar compounds assessed on the general chemical structure.

In addition to this level of information there is a second level of information, which is endpoint specific. The rules are specific features that can be useful for the evaluation of the property value of the target compound; the rules are endpoint specific. For BCF the rules indicate that within the chemical's structure the software identified a special residue, which can be useful for the evaluation of the property value. Associated with each rule there are again similar chemicals, in this case the ones most similar to the target compound, and in addition containing the residue at the basis of the rule.

The colours of the rule and of the similar chemicals are associated to a BCF scale, from red to green, which is reported in the output graph (see Figure 1). Figure 2 summarizes the graphical symbols used within the BCF module of ToxRead. Circles always represent chemicals. The size of the circle is proportional to the similarity of the chemical; the largest is the most similar.

The colour of the circle refers to the same scale of BCF values used for the rules. Yellow means that the chemical has a BCF value of about 3, and thus it is close to the 3.3 threshold. Red means a BCF value exceeding the 3.3 threshold, while green means a BCF value far from that threshold. Of course different thresholds apply depending on the regulation, and the exact experimental value of the similar chemicals is given by clicking on the circle.

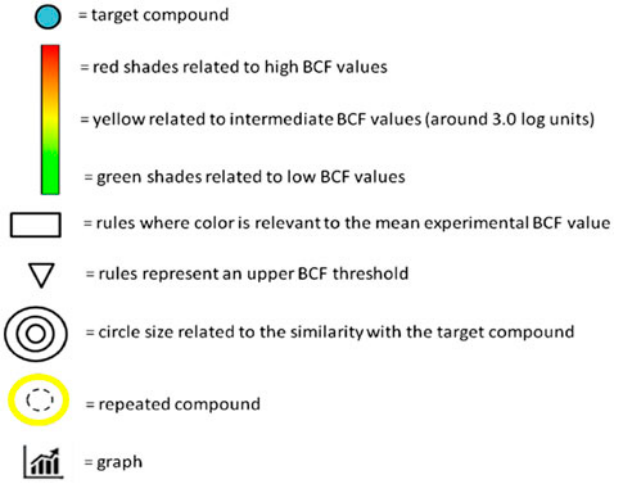

Figure 2. The graphical elements used by ToxRead.

Moving the mouse on the circles, if the same chemical appears more than once, the software identifies all the occurrences with a yellow ring around the chemical. Indeed, the same similar chemical may be similar simply from a chemical point of view, but it may also contain one or more of the rules applicable to the target compound, and thus the same chemical may appear several times. For instance, in Figure 1 the chemical with the CAS number 18708-70-8 appears three times.

Figure 2 also shows that there are two different symbols for rules. Indeed, some rules, indicated by rectangles, show that the target chemical has a residue that usually is present in chemicals with a certain average BCF value. Again, the colour of the rule represents the BCF value of the rule according to the scale. The rules drawn as triangles represent the empirical set of rules dealing with physico-chemical property values associated to BCF value. A log *P*/BCF interpolation chart is available for the target molecule and for each rule.

More information is available by clicking on the symbols. Clicking on the rule, the software shows the average BCF value of chemicals containing that rule, the number of chemicals with that rule within the software database, and the standard deviation of the BCF values for the defined family of substances (see Figure 3). For instance, for the rule reported in Figure 3 the software tool indicates that there is an aromatic nitro group, that there are 62

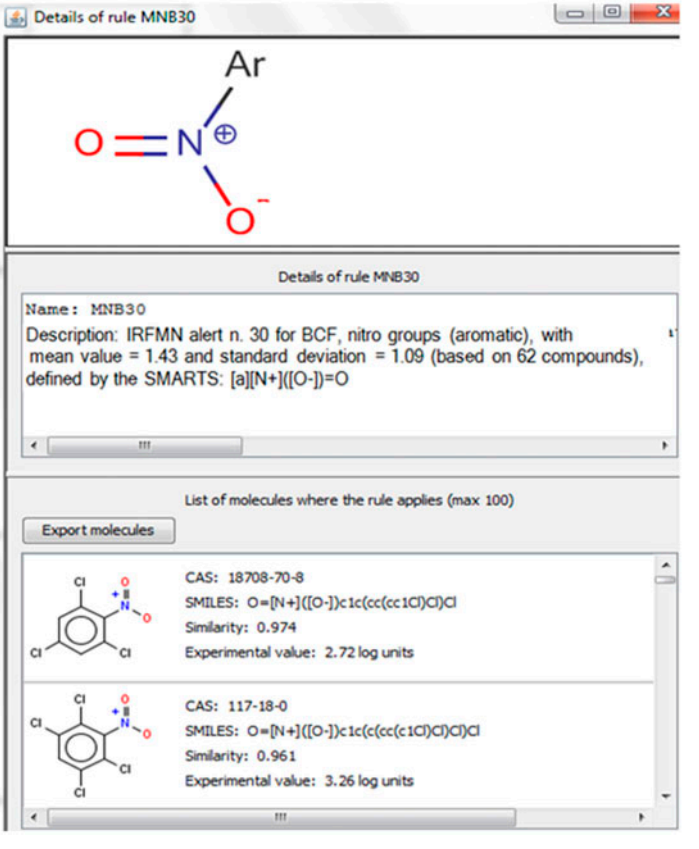

Figure 3. An example of the rules found for the target molecule (upper part of the figure) and the first two of the similar molecules containing the rule (lower part of the figure).

compounds with that rule, and that the average value is 1.43 and the standard deviation is 1.09. There is also a button to click, if the user wants to see or download the 100 most similar compounds containing that rule.

Clicking on the similar chemical (circle), the software shows in a box its structure (see Figure 4), with its experimental values (not only BCF, also other property values are reported if available). The similarity value is also indicated; it ranges from 1 (identity to the target compound) to 0. A good similarity value is higher than 0.85. If the similarity is lower than 0.7, there are many atoms and fragments which are different between the target compound and the similar one.

Clicking on the interpolation chart, a figure reports the most similar molecules with their log *P* value (experimental, if available, otherwise predicted with the VEGA model; if the value is experimental the colour of the chemical is dark in the chart) versus their experimental BCF value (see Figure 5). A dotted vertical line is drawn in correspondence with the log *P* value of the target compound. The area between the highest and lowest BCF value is white, indicating that the BCF value of the target compound should be within this area. Such a chart can be useful to evaluate the relationship and the trend between the log *P* and BCF values for the most similar compounds. There is also a table reporting the log *P* and BCF values of the similar compounds.

3.2 *Using the information given by the ToxRead software*

The figures of the similar compounds give an example of how the software can help in the assessment based on read-across. In Figure 1 the similar compounds have values 'green' or 'yellow', i.e. close to 3 or lower than 3. There are two rules associated with the target compound, and both are 'green', indicating that generally this kind of chemical is safe. One rule, rule 1.0, is related to aromatic nitro compounds, and its explanation is shown in Figure 3. Rule 2.0 identifies if there is an atom that can accept H-bonds (N, O, F).

This kind of overview provides the first framework to assess the nature of the target substance. In this case there is no major concern (no 'red' rule, no 'red' similar compound), but some chemicals may have a BCF value close to 3. Thus, a closer look at the similar compounds allows a more detailed view. This is offered by the graph, as in Figure 5. The target compound has an aromatic ring with a nitro group and three chlorine atoms linked to it. The

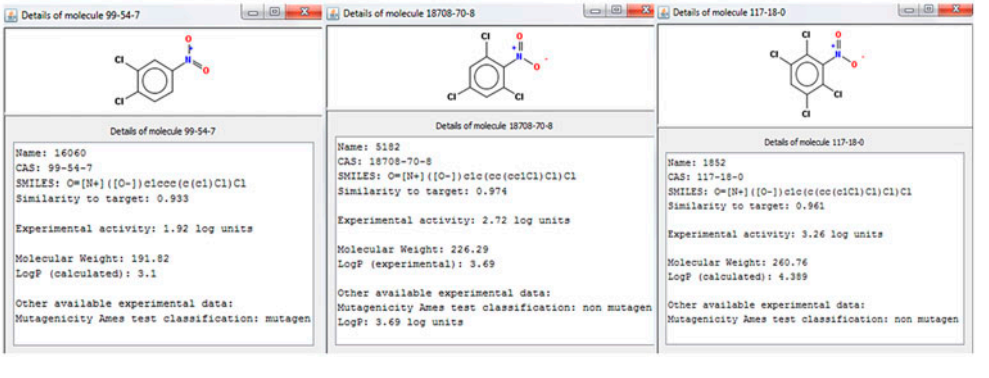

Figure 4. The three most similar molecules to the target chemical (as in Figure 1).

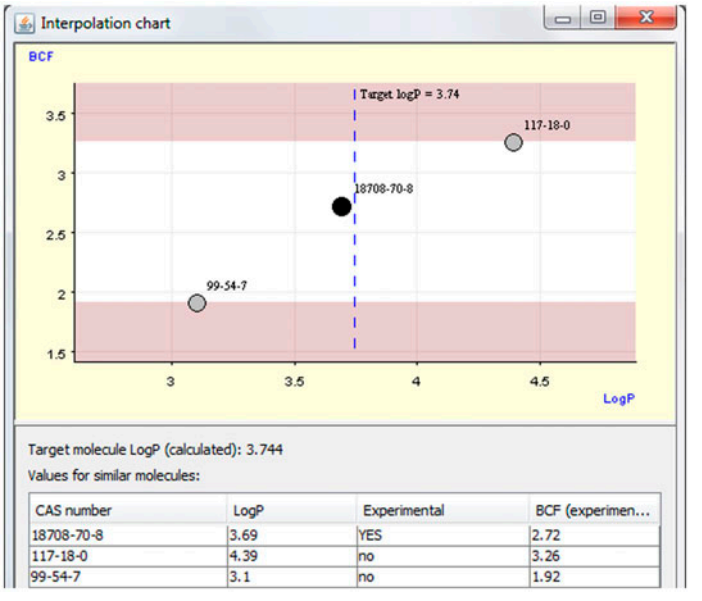

Figure 5. Interpolation chart representing the three most similar molecules to the target molecule as in Figure 1, with their log *P* versus log BCF values.

most similar compound has the same kind of substituent groups, only in different positions. The most similar compound has 2.72 log units as experimental BCF value. If we look at the other two most similar compounds, one of them has only two chlorine atoms, while the second one has four chlorine atoms. Chlorine brings a hydrophobic contribution, and indeed we see that a BCF of dichloronitrobenzene has the BCF of 1.92 log units, while the tetrachloro analogue has the BCF of 3.26 log units. If we look at the graph in Figure 5 there is quite a nice trend with log *P*.

Figure 6 shows the same picture analysing a larger number of similar compounds. In this case there are more compounds, and the trend with log *P* is confirmed. The new most similar compounds now introduced show lower BCF values; indeed, one of them has an amine group, which introduces hydrophilicity. The second new similar compound is a phenol, with an expected lower BCF value. The third new similar compound has only one chlorine atom, and as we have already commented this reduces the BCF value, also compared with the dichloro analogue. Thus, the new similar compounds, all of them with a nitro group, confirm the line of reasoning we had before. The new similar compound confirms the trend, but their similarity is lower, compared with the first ones.

If we further extend our view to a larger set of similar compounds, as shown in Figure 7, we see again a kind of trend, but more noise is present. In this case there are similar chemicals which do not contain the nitro group.

3.3 *Integrating results from QSAR models*

We used the VEGA software, which provides three QSAR models for BCF: one is the Meylan model, as developed within EPI Suite, the second is the CAESAR model [15],

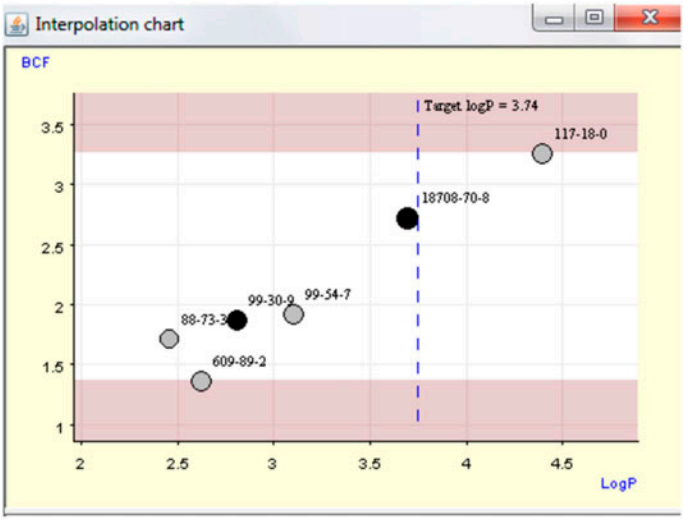

Figure 6. Interpolation chart representing the six most similar molecules, with their log *P* versus log BCF values.

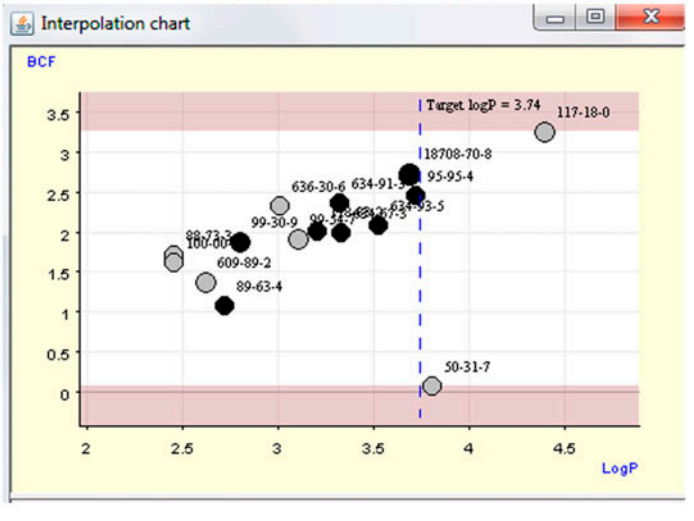

Figure 7. Interpolation chart representing the 15 most similar molecules, with their log *P* versus log BCF values.

and the third is a kNN model. VEGA provides an assessment of the applicability domain in a quantitative manner, through the Applicability Domain Index (ADI), whose value ranges from 1 to 0, and is based on a series of sub-factors. Briefly, the software evaluates the similarity of the most similar compounds, and compares the predicted BCF value with the experimental BCF values of the most similar compounds. Furthermore, it evaluates the

correctness of the predictions of the most similar values. Other checks are done, considering the occurrence of unusual fragments in the molecule, and (for the CAESAR model) perturbing the values of the descriptors, checking if these changes of the descriptor values (up to 10%) provoke a much higher BCF value change. All these checks are used within the ADI, and offer a good way to assess the model reliability for the specific chemical and property.

In our first example the predictions of the three models are in a good agreement: values range from 1.6 to 2.47 log units. The ADI values of the three models are good: they range from 0.7 to 1. Thus, overall the likely BCF value is lower than 3, and is in line with the conclusions derived from ToxRead. The most similar compound, CAS number 18708-70-8, a trichloronitrobenzene, has a BCF of about 2.7.

Thus, overall, these two parallel approaches, read-across and QSAR, can be used within a weight-of-evidence strategy, and they reinforce each other. This is the ideal situation, because the uncertainty of the overall evaluation is reduced, through a consensus process.

3.4 *The example of a chemical with more uncertain and conflicting evidences*

Another example is presented in Figure 8 (with the SMILES: CCCNCc1cc(Cl)c(cc1Cl)-c1cc(Cl)c(Cl)cc1Cl), where there are two red rules and one green. Similarly, some chemicals are red while others are green. Thus, there is a higher level of uncertainty, compared with the previous example. The most similar compounds are those connected to the red rules; similarity value: 0.863 for the three most similar compounds (Figure 9). Their experimental BCF values vary from 4.02 to 5.69 log units. The compounds connected to the green rule are less similar, with similarity value that varies from 0.784 to 0.849. These chemicals have a lower number of chlorine atoms, and more polar groups. Thus, it may be more reasonable to refer to the chemicals that are more similar.

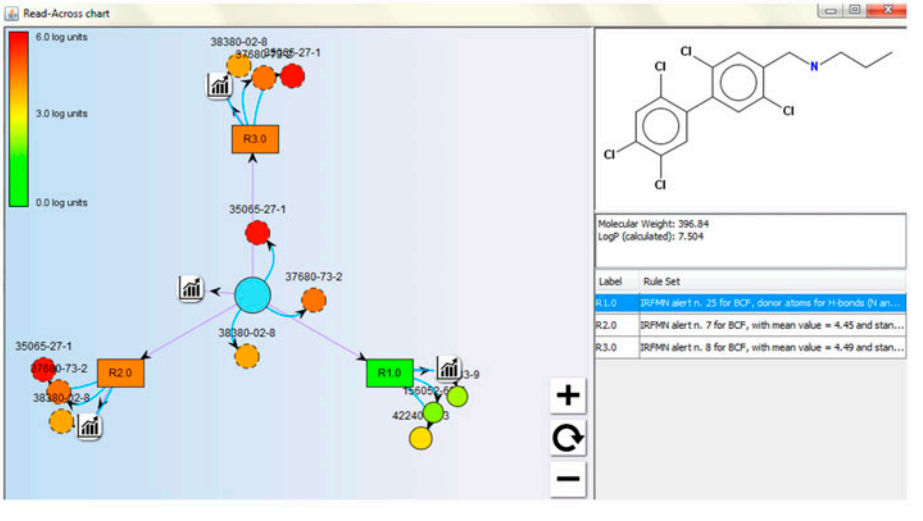

Figure 8. ToxRead screen showing the similar molecules and the rules found in the analysis of the molecule: CCCNCc1cc(Cl)c(cc1Cl)-c1cc(Cl)c(Cl)cc1Cl.

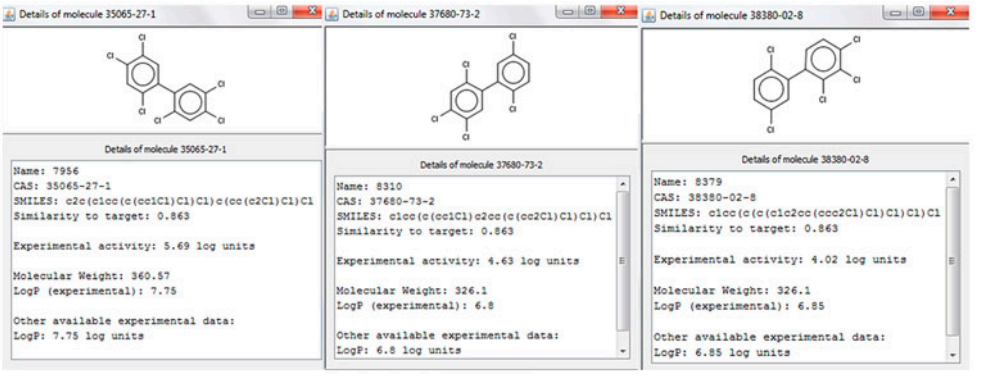

Figure 9. The three most similar molecules to the target molecule of the Figure 8.

3.5 *Integrating read-across and QSAR for the second case study*

The three models within VEGA provide values close to or higher than 4, so the prediction is very bioaccumulative. The ADI of the three models are acceptable. Overall, the main pieces of evidence for each QSAR model indicate that the chemical has a BCF value in the range 4–4.5. Combining the overall QSAR results, they support the conclusion that the rule for low BCF value (Rule 1, related to donor atoms for H-bonds) in this case is overruled by the two other rules, both associated with the biphenyl structure linked to two chlorine atoms (in different positions). The BCF value is probably above 4, but there is a higher level of uncertainty, compared with the previous example, for the higher uncertainty associated with the read-across evaluation. In this case the expert may get conflicting results from the read-across evaluation provided by ToxRead, but the QSAR model helps in supporting the preference for rules 2 and 3, which indicate the high log *P* value.

3.6 *Discussion*

As we have shown, the availability of tools like ToxRead opens the way to a deeper integration between the results of QSAR models and read-across. Thus, ToxRead increases the possible integration of these two approaches. A first level of integration is the extent of agreement between the results. Of course, if there is good agreement, it is easy to get a decision, but problems may arise in the case of conflicting results. In our case the user can easily verify this possible agreement. What we suggest is to arrive at an assessment on the basis of the QSAR result(s) (each model should be assessed individually, if more than one is available, and then the consistency of the results should be evaluated), separately to obtain a conclusion for the read-across, and only finally to evaluate the agreement or disagreement between read-across and QSAR(s).

The transparency and documentation offered by VEGA and ToxRead can provide valid support to integrate the results of the models, evaluating the reliability of individual results and thus facilitating the decision in case of conflicting results. We have shown here how to proceed for the read-across case, reasoning about conflicting evidence. The relevance assigned to each result can be established on that basis. This increases the flexibility and robustness of the integration of the results of the two approaches. The integration, as we have seen, is not

the simple mathematical average of different results. Conversely, the chemical information on the similarity of the chemicals and the presence of specific rules improve the possibility to improve the overall assessment. The user has presented different pieces of evidence, which can help in the process of reasoning (currently the user should save the preferred graph with the chosen number of similar substances and related information. A future version of ToxRead will simplify the reporting process). The same rules are shown to all users, which do not select them a priori. The evaluation is based on elements that are automatically shown, and this increases the transparency and reproducibility of the process. Furthermore, the software is based on parameters that are processed through simple mathematical equations. This opens the possibilities to further integration of the results within the same platform, in an automatic way, which may process and integrate results from QSAR and read-across. This is not implemented yet, but this objective is greatly facilitated by the fact that each result is based on parameters, which are produced by the software through quantitative parameters, with a possible use within a unified program.

4. Conclusions

There are more and more cases of read-across done for practical use to register chemicals for regulations such as REACH and the Cosmetics Directive. However, papers discussing the use of read-across are less frequent, and even less frequent are papers on software for read-across. In most cases, human experts perform the evaluation using their personal experience. This opens the question of the possible subjectivity of the evaluation.

In order to offer a more systematic way to address read-across, we developed the publicly available tool ToxRead, which has already been downloaded by hundreds of users. The approach we adopted offers a general overview of possible reasons of concern, or the expected behaviour of the target compound, on the basis of a collection of rules. These rules can be analysed, and the most similar chemicals that are associated with the rules can be seen and evaluated, also through the use of graphs summarizing the different scenarios. In this way the multiple factors to be evaluated are represented in an overall visual scheme, and the user should navigate and analyse them, assessing the levels of uncertainty (also through the quantitative similarity scores of the related compounds) and the trend analysis representations; greater care should be given to cases with conflicting elements (similar chemicals or rules with large differences in values).

Within ToxRead it is clear that the read-across strategy is endpoint specific, since it refers to specific rules which apply for BCF for instance, as shown here, and is not based on a simple chemical similarity on the related compounds.

QSAR is often criticized as a black-box approach, since it operates at the abstract level of descriptors, not always easily related to simple reasoning. ToxRead addresses this kind of need, to represent reasons of concern directly related to chemical structures, which are organized according to rules. Combining the two non-testing methods may reciprocally support the individual assessment.

ToxRead is based on a diagnostic philosophy, while the QSAR approach is typically prognostic. In other words, in the current version of ToxRead no prediction of the property for the target compound is attempted. Prediction is left to the QSAR models, which are present for instance within VEGA (www.vega-qsar.eu).

Acknowledgements

The authors gratefully acknowledge the financial support of the LIFE+ projects CALEIDOS and PROSIL [23,24].

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] R. Benigni, C. Bossa, N.G. Jeliazkova, T.I. Netzeva, and A. Worth, *The Benigni/Bossa rulebase for mutagenicity and carcinogenicity – A module of Toxtree*, EUR-Scientific and Technical Report Series, EUR 23241 EN, Office for the Official Publications of the European Communities, Luxembourg, 2008.
[2] VEGA, platform available at http://www.vega-qsar.eu/.
[3] TEST, software available at http://www.epa.gov/nrmrl/std/qsar/qsar.html#TEST.
[4] AMBIT, Version 2.4.11, 2014; software available at http://ambit.sourceforge.net/.
[5] European Chemicals Agency (ECHA), Read-Across Assessment Framework (RAAF), ECHA/PR/15/07, ECHA, Helsinki, 2015.
[6] N. Price and Q. Chaudhry, *Application of in silico modelling to estimate toxicity of migrating substances from food packaging*, Food Chem. Toxicol. 71 (2014), pp. 136–141.
[7] J.A. Arnot and F.A.P.C. Gobas, *A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms*, Environ. Rev. 14 (2006), pp. 257–297.
[8] M. Pavan, T.I. Netzeva, and A.P. Worth, *Review of literature-based quantitative structure-activity relationship models for bioconcentration*, QSAR Comb. Sci. 27 (2008), pp. 21–31.
[9] Organisation for Economic Co-operation and Development (OECD), Guideline for testing of chemicals, No. 305. Bioaccumulation in fish: Aqueous and dietary exposure, OECD, 2012.
[10] N. Rodriguez-Sanchez, M.T.D. Cronin, A. Lillicrap, J.C. Madden, P. Piechota, and K.E. Tollefsen, *Development of a list of reference chemicals for evaluating alternative methods to in vivo fish bioaccumulation tests*, Environ. Toxicol. Chem. 33 (2014), pp. 2740–2752.
[11] European Chemicals Agency (ECHA), *The Use of Alternatives to Testing on Animals for the REACH Regulation*, Second report under Article 117(3) of the REACH Regulation, ECHA, Helsinki, 2014.
[12] ToxRead, software available at http://www.toxgate.eu/.
[13] G. Gini, A.M. Franchi, A. Manganaro, A. Golbamaki, and E. Benfenati, *ToxRead: A tool to assist in read across and its use to assess mutagenicity of chemicals*, SAR QSAR Environ. Res. 25 (2014), pp. 999–1011.
[14] E. Benfenati, A. Manganaro, and G. Gini, *VEGA-QSAR: AI inside a platform for predictive toxicology*. Proceedings of the Workshop PAI 2013, 5 December 2013, Turin, Italy, pp. 21–28.
[15] C. Zhao, E. Boriani, A. Chana, A. Roncaglioni, and E. Benfernati, *A new hybrid system of QSAR models for predicting bioconcentration factors (BCF)*, Chemosphere 73 (2008), pp. 1701–1707.
[16] A. Lombardo, A. Roncaglioni, E. Boriani, C. Milan, and E. Benfenati, *Assessment and validation of the CAESAR predictive model for bioconcentration factor (BCF) in fish*, Chem. Cent. J. 4 (2010), (Suppl 1):S1.
[17] EURAS, database available at http://cefic-lri.org/lri_toolbox/bcf-database/.
[18] S. Dimitrov, N. Dimitrova, T. Parkerton, M. Comber, M. Bonnell, and O. Mekenyan, *Base-line model for identifying the bioaccumulation potential of chemicals*, SAR QSAR Environ. Res. 16 (2005), pp. 531–554.

[19] Estimation Programs Interface Suite (EPI Suite) platform, Version 4.11, available at http://www.epa.gov/oppt/exposure/pubs/episuite.htm.
[20] M. Floris, A. Manganaro, O. Nicolotti, R. Medda, G.F. Mangiatordi, and E. Benfenati, *A generalizable definition of chemical similarity for read-across*, J. Cheminform. 6 (2014), p. 39.
[21] T. Ferrari, D. Cattaneo, G. Gini, N.G. Bakhtyari, A. Manganaro, and E. Benfenati, *Automatic knowledge extraction from chemical structures: The case of mutagenicity prediction*, SAR QSAR Environ. Res. 24 (2013), pp. 365–383.
[22] KODE, available at http://kode-solutions.net/en/index.php.
[23] CALEIDOS LIFE+ project, available at http://www.caleidos-life.eu/.
[24] PROSIL LIFE+ project, available at http://www.life-prosil.eu/.