



## RESEARCH ARTICLE

## A New QSAR Model for Acute Fish Toxicity based on Mined Structural Alerts

Giuseppina Gini<sup>1\*</sup>, Thomas Ferrari<sup>1</sup>, Anna Lombardo<sup>2</sup>, Antonio Cassano<sup>2</sup> and Emilio Benfenati<sup>2</sup>

<sup>1</sup>Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Italy

<sup>2</sup>Mario Negri Institute for Pharmacological Research, Italy

\*Corresponding author: Giuseppina Gini, Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milano, Italy, Tel: +39-0223993626



### Abstract

In this paper we discuss the problem of predicting the fish toxicity property of chemical compounds, and show how this can be approached using a computational intelligence method. There are two views on assessing toxicities: One says that such properties can be derived from the whole molecular structure, the other that some specific functional substructures, called Structural Alerts (SA), are able to explain the toxicity. In this work, a new Structure-Activity Relationship (SAR) approach is proposed to mine molecular fragments that act like SAs for the biological activity. We apply our data mining method, called SARpy, to a dataset about LC<sub>50</sub> for the fathead minnow, and build a multiclass classifier in the categories defined by the legislation. We test the model on an external test set of data about trout toxicity. The new model shows marked prediction skills and, more interestingly, it is based on mined structural alerts. Discovering new knowledge about substructures statistically strongly connected to toxicity opens to other future in-silico methods. The model is freely available in the VEGA hub<sup>a</sup> among other models for aquatic toxicity.

### Keywords

Structure activity relationships (SAR), Data mining, Knowledge discovery, Fish toxicity, Structural alerts

### Introduction

In our industrialized society, a huge amount of chemical substances are used and produced every day. This increasing number of chemicals around us raises the problem of characterization, prediction and evaluation of their consequences to the human health and to the environment.

<sup>a</sup><https://www.vegahub.eu>

Toxicology provides the knowledge of mechanisms, rules, and data characterized by quality levels and defines the limits of safety of chemical agents. Chemistry offers knowledge on chemical descriptors and physico-chemical properties. Biology studies the mechanisms of the action of the chemicals on the animals and the other organisms used for tests.

Predictive Toxicology is a multi-disciplinary science that requires the close collaboration among toxicologists, chemists, biologists, statisticians and Artificial Intelligence researchers. Statistics and Machine Learning integrates all these items to analyze the existing data and, especially, to extract new knowledge from the data, or at least to generate reliable toxicity predictions for chemical compounds [1-4]. In this way, the main drawbacks in toxicity studies, like high costs of experiments, duration of the tests, using animals in scientific experiments [5], can be surmounted.

The goal of toxicity prediction is to describe the relationship between chemical properties and biological and toxicological processes. SAR relates features of a chemical structure to a property, effect or biological activity associated with the chemical. There are two kinds of SAR: qualitative and quantitative. A qualitative relationship is a general rule, which provides a class label. Quantitative structure-activity relationships (QSARs) can be developed using continuous (quantitative) data, mostly through a regression process. This area was developed in the last forty years to assess the value of drugs and more recently to assess general toxicity.

Three “postulates” of modeling QSARs are accepted:

P1: The molecular structure is responsible for all the activities shown.

P2: Similar compounds have similar biological and chemo-physical properties.

P3: QSAR is applicable only to similar compounds.

In toxicity prediction various machine learning techniques were proposed, from expert systems to Artificial Neural Networks (ANN), from decision trees to neuro-fuzzy models. There were also local versus global models, and different combinations of classifiers [6].

Human experts usually estimate toxicity through the detection of particular structural fragments, already known to be responsible for the toxic property under investigation. In the literature such fragments are referred to as structural alerts (SAs) and can be derived by human-experts, from knowledge of the biochemical mechanism of action (such as the activation of an enzyme cascade or the opening of an ion channel, which leads to a biological response); these mechanisms are still poorly understood and largely unknown.

Only a few approaches have been developed to assist experts in directly extracting such knowledge from data. Some are based on inductive logic programming (ILP) [7]; they cannot be directly applied to standard chemical formats for molecule representation and require extra computation. Other, including NIKE [8], uses a neuro-fuzzy approach to integrate explicit and implicit knowledge; however they use chemical descriptors derived after complex simulations and not directly the chemical 2D structure.

In the following we introduce the problem of our investigation, i.e. acute fish toxicity, the data available, our method for knowledge extraction, the obtained model and its statistics.

Among the available data sets with experimental results, the acute toxicity database to fathead minnow (*Pimephales promels*) is considered a “gold standard” for quality of toxicity measurement. The chemicals were selected from the Toxic Substances Control Act (TSCA) inventory of chemicals to represent a cross-section of industrial organic chemicals. Thus, the fathead minnow database is chemically heterogeneous and the substances in it represent a large spectrum of mechanisms of toxic action. In fact it includes compounds acting by different modes and mechanisms such as narcosis (type I, II and III), uncouplers of the oxidative phosphorylation, reactive electrophiles/proelectrophiles, acetylcholinesterase inhibitors, and central nervous system (seizure) agents.

Recent work is expanding the set of chemicals for which more testing results are available. Recently [9] analyzed the effect of some antibiotics on fish larvae (*Tilapia nilotica*) toxicity. Another important area of inves-

tigation is the effect on fishes of pollutants derived from agriculture. Recent papers [10,11] report about tests on fish of the toxicity of chlorpyrifos, carbaryl, diuron, nematicur, and malathion, showing how their presence in water poses a concern, in particular for chlorpyrifos (for which the experimental  $LC_{50}$  was  $0.08 \mu\text{ mol/L}$ ).

The QSAR models so far developed on those data predict either the  $LC_{50}$  value or a class, using the classification adopted at international levels. We shortly present this topic in the next section. Predicting a class is more useful for regulatory purposes, and this is our choice.

## Materials and Methods

The endpoint considered in this paper is fish toxicity, an important test for water quality assessment. All toxicity data were retrieved from Russom, et al. [12]. Tests were conducted at EPA (Environmental Protection Agency) using Lake Superior water at  $25 \pm 1 \text{ }^\circ\text{C}$ . Aqueous toxicant concentrations were measured in all tests with quality assurance criteria requiring 80% agreement between duplicate samples and 90% to 100% spike recovery. Flow-through exposures were conducted using cycling proportional, modified Benoit or electronic diluters. Median lethal concentrations ( $LC_{50}$ , in mg/L) were calculated using the Trimmed Spearman-Kärber method, with 95% confidence intervals being calculated when possible.  $LC_{50}$  (96 h), i.e. the lethal concentration for 50% of a population within 96 hours. For the purposes of this study the logarithm of their reciprocal values ( $\log(1/LC_{50})$ ) was used for QSAR modeling.

The data set contains 568 different compounds. Data are quite representative for most industrial chemicals, but they are still a very small percentage of the commercialized chemicals, and an even minor part of all possible chemicals humans can be exposed to. Nevertheless, they represent a unique collection of data. As typical in data from studies in the life sciences, the cost of experimental data is high. Thus, the number of experiments is quite low.

Since these values are widely spread and to take into account the regulations, the results were also transformed into the classification for toxicity to fish as provided by Directive 92/32/EEC of the EU for dangerous substances [13] (Table 1).

According to EPA the data set can be studied also considering the chemical classes included, which are indicated in (Table 2).

**Table 1:** EU classification for fish toxicity.

Class	LC50	Damage
1	< 1 mg/L	Very toxic to aquatic organisms
2	1 - 10 mg/L	Toxic to aquatic organisms
3	10 - 100 mg/L	Harmful to aquatic organisms
4	> 100 mg/L	May cause long term adverse effects to aquatic organisms

**Table 2:** Chemical classes in the data set.

Chemical classes	Number of compounds
Hydrocarbons	26
Ethers	24
Alcohols	60
Aldehydes	44
Ketones	39
Acids	68
Nitriles, Sulfur compounds	33
Amines	74
Benzenes	33
Phenols	49
Heterocyclics	48
Carbamates, Other pesticides	28
Various classes	42

In a previous work on toxicity prediction in the fathead minnow Russom, et al. [12] encoded human expert knowledge in the definition of modes of action (MOA). They classified eight mechanisms involved in aquatic toxicity. Then specific chemical fragments have been identified, and a heuristic has been developed to find the MOA on the basis of these fragments. Finally, quite simple QSAR models have been used to predict toxicity for specific MOA.

Another interesting QSAR model that uses a low number of chemical 3D descriptors and a linear regression is described in [14].

The fish dataset contains a diversity of compounds, with a diversity of structures. Because of the lack of homogeneity it is hard for a single technique to model this data and to obtain good results. One attempt to overcome the diversity of the data, dividing the set using the official chemical classification from EPA, was developed in [15]. The results obtained in prediction are quite good for some models, with the coefficient of determination on the test set  $R^2_{\text{test}}$  ranging from 0.98 for ethers to 0.60 for aldehydes. The accuracy prediction of the combined model has  $R^2_{\text{test}}$  of about 0.8.

Since the classification in chemical classes is sometimes ambiguous because more than one functional group could be inside one molecule, an alternative approach grouped the chemicals together with a clustering technique in the input space (the descriptors space) using Euclidian distance. In fact the chemicals with similar biological and physico-chemical properties have similar descriptors and are grouped in the same cluster; moreover, for similar descriptors is more likely to have similar behavior. This work, as reported in [16], obtained accuracy very similar to the previous one. Using an unsupervised method it built ten clusters and ten models (one for each). All the models were tested on an external test set. The performance of every local model and of the global model on the test set was about 0.77 in terms of  $R^2_{\text{test}}$ .

Other approaches have been reported, using Genetic Algorithms, Support Vector Machines, and integration

of them [17-19]. The EC project ANTARES<sup>b</sup> listed about 250 implemented QSAR models, including many on fish toxicity, with  $R^2$  similar to those above reported.

To evaluate the performance of the models, a test set of 351 compounds with toxicity data on Rainbow trout (*Oncorhynchus mykiss*) was used. This dataset was compiled from two sources: the beta version of the OECD toolbox<sup>c</sup> selecting the OECD-HPV inventory, and the pesticide dataset used to develop the Demetra<sup>d</sup> model for fish. The data were checked to eliminate duplicates, mixture and inorganic compounds. Salt were considered in their neutral form.

## Extracting Toxicity Rules as SAS

Instead of computing chemical descriptors and extracting a good subset for modeling, our approach relies on extracting fragments and combining them into rules. Our method has been implemented in SARpy [20,21], a Python script employing the open source OpenBabel 2.2.3 library. SARpy also is public and available at <https://sourceforge.net/projects/sarpy/>.

The model is built in SARpy starting from the training dataset given in input as a matrix of chemical structures and their experimental activity label, written as a Comma Separated Values (CSV) table, with structures expressed in Simplified Molecular Input Line Entry Specification (SMILES) notation. SMILES are ASCII strings obtained by printing the symbol nodes (chemical elements) encountered in a depth-first tree visit of the chemical graph.

From the training set of molecular structures along with their experimental activity binary labels, SARpy generates every substructure of the chemicals in the set, and mines correlations between the incidence of a particular molecular substructure and the activity of the molecules that contain it. Such task is carried out in three subsequent steps:

1. Fragmentation: A recursive algorithm that considers every combination of bond breakages working directly on the SMILES string and computes every substructure of the molecules in the input set.
2. Evaluation: Each substructure is validated as potential SA on the training set.
3. Rule set extraction: From the huge set of substructures collected, a reduced set of rules is extracted in the form: "IF contains <SA> THEN <activity label>".

## Fragmentation

The aim of this phase is to detect all the chemical substructures present in the training set of chemicals. The substructures are identified by recursively applying

<sup>b</sup><http://www.antares-life.eu/index.php?sec=modellist>

<sup>c</sup>[http://www.oecd.org/document/23/0,3746,en\\_2649\\_34379\\_33957015\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/23/0,3746,en_2649_34379_33957015_1_1_1_1,00.html)

<sup>d</sup>[https://cordis.europa.eu/project/rcn/67471\\_en.html](https://cordis.europa.eu/project/rcn/67471_en.html)

a simple fragmentation algorithm, which at each step breaks a bond in the molecule and collects the two fragments that result. Applying the next fragmentation step to the output of the previous, all the possibilities of a second bond breakage are explored; and so on, until no more new fragments could be extracted. The fragmentation of chemical structures is performed directly on their SMILES strings. Thus the problem of processing a two-dimensional molecular graph is reduced to a fast ASCII strings processing. Considering that parsing with context-free grammars has worst-case time complexity cubic, the complexity of SARpy in enumerating the substrings is polynomial. In addition it has been decided to consider rings as single entities during the fragmentation phase: the idea is that the same substructures contained in a ring might be found as open skeleton in other compounds in the training data, otherwise, if always embedded in a ring, then only the whole ring itself has to be taken into account.

A further consideration concerns the length, in terms of number of atoms, of relevant fragments. Several experiments gave evidences that fragments longer than 18 atoms don't affect the final model.

## Evaluation

Once all the substructures have been collected, the next phase consists in their individual evaluation as potential SAs using the training set of chemicals. Here the binary case of a "positive" or "negative" experimental activity label, associated to each structure, is considered, so to focus on the search for SAs for positive activity.

First of all, each substructure is matched against every molecular structure in the training data. Such matches can be divided into matches against positive structures, called True Positives (TP), and matches against negative structures, called False Positives (FP). The structural comparison is carried out by the OpenBabel SMARTS matching function after being optimized by the use of fingerprints and by taking advantage of the hierarchical organization of the fragmentation process. In fact, the search for structures potentially containing a given fragment can be restricted to the ones that contain a substructure of the fragment itself. Therefore, the evaluation process is performed backward in the hierarchy of fragments.

Then every fragment is paired with its TPs and FPs matches. From these two values, the likelihood ratio, which is a measure of precision intrinsic to the test (does not depend on prevalence of activity labels in the training set) is computed:

$$\text{likelihood ratio} = (TP / FP) \times (\text{negatives} / \text{positives})$$

The likelihood ratios are used to rank all the potential SAs; also sensitivity is used as a secondary sorting key. The evaluation is aimed at identifying the substructures

that show high precision and good sensitivity; likelihood ratio will be used in the next phase to dynamically extract the best set of SAs.

## Rule extraction

The final goal is to obtain, from the huge list of potential SAs, a reduced set able to predict the target class with the best precision. Such requirement is implemented as follows:

1. Order the list of potential alerts by likelihood ratio.
2. Select the top ranked one, add it at the rule set and remove it from the list of potential alerts.
3. Update TP and FP values of the remaining potential SAs, by discarding TPs and FPs in common with the alert just selected.
4. Update likelihood ratios of potential SAs.
5. Return to 1.

To avoid rules with irrelevant behavior, a lowest bound on the TP value is considered; furthermore, rules that show a precision worse than the prevalence of the target class in the training set (i.e., likelihood ratio < 1) are removed, since they predict worse than a "random rule". Such rules pruning is dynamically re-executed before the extraction of every rule, using the actual TP values and likelihood ratios.

With this procedure it is possible to select the next SA keeping into account the effect of the SAs already in the rule set, The definition of a termination condition affects the behavior of the rule set, making it more sensitive (late stop) or more specific (early stop). The output is presented to the user as an ordered set of rules in the form "IF contains <SA> THEN <apply label>".

The resulting set of rules can be checked on an external test set or many-folds cross-validated.

## Results on Predicting Fish Toxicity

As already said, the definition of the toxicity in case of fish acute toxicity follows thresholds defined on a conventional basis, as reported in [Table 1](#). It means that the toxic effect is defined according to opportunistic values, and there are different thresholds used by regulators, but not necessarily connected to a unique effect (see the discussion about MOA).

Another point to underline is that the robustness of the model depends on the number of chemicals in each class. Thus, for instance, in case of the class of chemicals with the highest toxicity (meaning toxic at a concentration < 1 mg/L), the number of chemicals is quite low, and thus the model has to work on unbalanced data.

Due to the relatively low number of chemicals with experimental data on the fathead minnow, we decided to use all of them for modeling purposes, and then to

**Table 3:** Chemical classes in the data set.

Predicted				Observed
< 1 mg/L	1 - 10 mg/L	10 - 100 mg/L	> 100 mg/L	
<b>35</b>	23	14	0	< 1 mg/L
6	<b>102</b>	31	8	1- 10 mg/L
4	27	<b>134</b>	21	10 - 100 mg/L
0	4	38	<b>123</b>	> 100 mg/L

**Table 4:** Results for the external test set (TROUT).

Predicted				Observed
< 1 mg/L	1- 10 mg/L	10 - 100 mg/L	> 100 mg/L	
<b>24</b>	62	35	19	< 1 mg/L
5	<b>53</b>	40	10	1- 10 mg/L
1	14	<b>20</b>	21	10 - 100 mg/L
2	8	14	<b>23</b>	> 100 mg/L

use a second set of compounds for testing purposes. However, the second set of compounds is for another fish, trout, and includes a high number of pesticides (more than 50%). These two factors should be carefully considered, because they represent differences compared to the population of the values used to train the model.

We applied SARpy to identify the fragments related to toxicity, according to the different thresholds of [Table 1](#). Thus, SARpy discovers a series of fragments which serves to label chemicals as toxic with the threshold of 100 mg/L, another series of fragments for the threshold at 10 mg/L, and so on. The overall model is composed by three individual models done with SARpy.

The overall model has been then implemented within the VEGA platform (<http://www.vega-qsar.eu>) [22], where the user can freely get the predicted class simply introducing the SMILES code of the chemical(s) of interest.

[Table 3](#) shows the numbers of predicted molecules in each class as obtained for the four toxicity classes. Computing the accuracy on the training set, the accuracy is 87.5% the highest threshold of 100 mg/L, is 85% for the threshold of 10 mg/L, and it is 92% for the threshold at 1 mg/L. The average accuracy is so greater than 88%.

Thus, the models for the three thresholds provide good results.

To check the generalization capability of the model we tested it with an external test set. This test set contains compounds for which results on trout toxicity were available. [Table 4](#) shows the results for this test set. As before stated, there is difference in the fish, which may get different toxicity values. Furthermore, this data set contains in most of the cases pesticides, and thus there is quite a lot of chemicals which are labeled as very toxic, and may contains SAs not covered within the training set. The accuracy for the threshold at 100 mg/L is still quite good (79%), but it is lower for the thresholds at 10 mg/L (63%) and at 1 mg/L (65%). The average accuracy is so 69% on the trout test set.

The reasons are quite probably both the diversity in the chemicals used and the fish kind, as expected. In fact applying a QSAR model to another species has to consider the inter-species differences too.

SARpy has been able to extract fragments that can provide useful insights on the understanding of the toxic mode of action. For instance, SARpy has found that fragments related to toxicity have Sn, dithiophosphates, and halogenated aromatic rings, which are known to be associated to aquatic toxicity. Using the same tool those fragments would be eventually correlated to MOA, an area where there is not sufficient structural knowledge. The full list of the alerts found by SARpy and applied by the model to make predictions according to the three thresholds is in [Appendix 1](#).

This new model can be compared to other available models.

In a recent exercise [23], researchers tested five available models for fish toxicity, namely ECOSA, KATE, TEST, ADMET, and CADRE-AT, all based on logP, against a new test set, not used in developing any of the models. This new set contains 83 chemicals, distributed in many chemical classes. The output of the models was one of the 4 categories. CADRE-AT (a model based on neural nets, and not freely available) obtained the high total accuracy of 83%, missing only 3 predictions. The other models had a total accuracy lower than 58% and many missing predictions. For the majority of compounds it was impossible to classify them into MOA.

The same EPA Fathead Minnow Acute Toxicity database has been used to create a model as reported in the EU project OpenTox<sup>e</sup>. The models were developed using descriptors and regression on a training set of 90% (522 compounds) and test set of 10% (58 compounds). The reported statistical parameters are:  $R^{2\text{test}}$  about 0.7, and  $Q^{2\text{train}}$  about 0.6 [24].

The use of (Q)SAR models is expected to increase in various regulations. The European Plant Protection Products Regulation requires that registrants establish

<sup>e</sup><http://www.opentox.org>

whether pesticide metabolites pose a risk to the environment. Fish acute toxicity is one of the required endpoints. EFSA recently published a “Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters” in which it outlines the opportunity to apply non-testing methods, such as SAR and QSAR. Considering this issue Burden and coworkers developed a retrospective analysis about the state of the art in predicting fish toxicity. They extracted experimental fish LC<sub>50</sub> values for 150 metabolites from the Pesticide Properties Database<sup>f</sup>. Then predicted them using the US EPA’s ECOSAR software, and compared the predictions to the experimental data. They observed a significant correlation between predicted and experimental fish LC<sub>50</sub> values and concluded that for 91% of the substances the predictions were sufficiently predictive. They concluded that the applicability of QSAR models in the metabolite assessment could be recommended [25].

It is worth discussing our results in perspective, considering other available approaches and future developments.

Recent research has individuated toxicokinetic models as tools to help in assessing the environmental risk. For aquatic species, and in particular for fish and for some chemical classes, some PBTK (physiologically based toxicokinetic) models are available [26]. This software can be of great interest in studying the acute toxicity of fish.

Another recent direction in filling the data-gap in toxicity prediction is providing new software to improve read-across. In our experience, after the results on a round-robin exercise on read across [27] involving about 20 experts, we found that the results were largely user-dependent on assessing fish toxicity (spanning from non toxic to high toxic for the same chemical), even considering that most of the experts used tools as OECD toolbox. In case of other toxicological endpoints, when experts used a property tailored system as ToxRead [28], the answers were more consistent with each other. ToxRead presents in a structured and hierarchical way the structural alerts present in the target and the similar molecules.

It is our future task to extend ToxRead by inserting new properties, as fish toxicity, and the structural alerts so far collected. We are aware that those alerts have now only a statistical meaning (the full process of transformations not being available for most of them); however we expect that they will help experts in the assessment.

## Conclusions

Alongside classical methods as *in vivo* and *in vitro* experiments, the use of computational tools is gaining

more and more interest in the scientific community and in the industrial world as accompaniment or replacement of existing techniques. For regulatory purposes it is important to obtain satisfactory classification accuracy on new chemical families not well studied. In this area it is important to develop models that can take advantage of statistical analysis on great numbers and can be further refined to improve or confirm the results and give more insights into the domain.

To this end we have developed a new model for fish toxicity using SARpy, a system able to focus on the important structural features hidden in the database. The difference of SARpy with respect to other (Q)SAR approaches is its ability to extract relevant knowledge in the form of structural alerts during the learning stage. The models show their capability to predict the fish toxicity, even though they suffer in some cases for the unbalanced distribution of the data and chemical nature.

## References

1. Benfenati E, Gini G (1997) Computational predictive programs (expert systems) in toxicology. *Toxicology* 119: 213-225.
2. Giuseppina C Gini, Alan R Katritzky (2000) Predictive toxicology of chemicals: Experience and impact of AI tools. *AI Magazine* 21: 81-84.
3. Gini G, Benfenati E (2007) E-modelling: Foundations and cases for applying AI to life sciences. *International J on Artificial Intelligence Tools* 16: 243-268.
4. Benfenati E, Mazzatorta P, Neagu CD, Gini G (2002) Combining classifiers of pesticides toxicity through a neuro-fuzzy approach. *Proc 3<sup>rd</sup> International Workshop on Multiple Classifier Systems, MCS 2002, Cagliari (Italy)*, 293-303.
5. Hansch C, Hoekman D, Leo A, Zhang L, Li P (1995) The expanding role of quantitative structure-activity relationship (QSAR) in toxicology. *Toxicol Lett* 79: 45-53.
6. Neagu CD, Benfenati E, Gini G, Mazzatorta P, Roncaglioni A (2002) Neuro-fuzzy knowledge representation for toxicity prediction of organic compounds. *15<sup>th</sup> European Conference on Artificial Intelligence (ECAI02), Lyon, France*, 498-550.
7. Dehaspe L, Toivonen H, King RD (1999) Finding frequent substructures in chemical compounds. In: Gini G, Katritzky, A predictive toxicology of chemicals: Experiences and impact of ai tools, AAAI-SS-99, AAAI Press, Menlo Park, CA.
8. Neagu CD, Gini G (2003) Neuro-fuzzy knowledge integration applied in toxicity prediction. In: Jain R, Abraham A, Faucher C, Ven Der Zwaag BJ, *Innovations in knowledge engineering*. Advanced Knowledge International, Magill, South Australia 311-342.
9. El-Nahhal Y, El-Dahdouh N (2015) Toxicity of amoxicillin and erythromycin to fish and mosquitoes. *Ecotoxicol Environ Contam* 10: 13-21.
10. El-Nahhal Y, El-Najjar S, Affi S (2015) Impact of organic contamination on some aquatic organisms. *Toxicol Int* 22: 45-53.
11. El-Nahhal Y (2018) Toxicity of some aquatic pollutants to fish. *Environ Monit Assess* 190: 449.

<sup>f</sup><http://sitem.herts.ac.uk/aeru/ppdb/en/atoz.htm>

12. Russom CL, Bradbury SP, Hammermeister DE, Drummond SJ (1997) Predicting modes of toxic action from chemical structure: Acute toxicity in the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 16: 948-967.
13. Directive 92/32/ECC, the seventh amendment to Directive 67/548/EEC, UJL 154 of 5.VI.92, 1992, [www.eurunion.org/1legislat1chemical.htm](http://www.eurunion.org/1legislat1chemical.htm).
14. Netzeva TI, Aptula AO, Benfenati E, Cronin MTD, Gini G, et al. (2005) Description of the electronic structure of organic chemicals using semi-empirical and Ab initio methods for development of toxicological QSARs. *J Chem Inf Model* 45: 106-114.
15. König C, Gini G, Benfenati E, Crăciun M (2004) Multi-class classifier from a combination of local experts: Toward distributed computation for real-problem classifiers. *Int J of Pattern Recognition and Artificial Intelligence* 18: 801-817.
16. Gini G, Craciun M, Benfenati E (2004) Combining unsupervised and supervised artificial neural networks to predict aquatic toxicity. *J Chem Inf Comput Sci* 44: 1897-1902.
17. Pintore M, Piclin N, Benfenati E, Gini G, Chretien JR (2003) Predicting toxicity against the fathead minnow by adaptive fuzzy partition. *QSAR Comb Sci* 22: 210-219.
18. Toropova P, Toropov AA, Lombardo A, Roncaglioni A, Benfenati E (2012) Coral: QSAR models for acute toxicity in fathead minnow (*Pimephales promelas*). *J of Computational Chemistry* 33: 1218-1223.
19. Wang Y, Zheng M, Xiao J, Lu Y, Wang F, et al. (2010) Using support vector regression coupled with the genetic algorithm for predicting acute toxicity to the fathead minnow. *SAR QSAR Environ Res* 21: 559-570.
20. Ferrari T, Gini G, Golbamaki Bakhtyari N, Benfenati E (2011) Mining toxicity structural alerts from SMILES: A new way to derive structure-activity relationships. *Proc IEEE SSCI, CIMD, Paris*, 120-127.
21. Gini G, Ferrari T, Cattaneo D, Bakhtyari NG, Manganaro A, et al. (2013) Automatic knowledge extraction from chemical structures: The case of mutagenicity prediction. *SAR QSAR Environ Res* 24: 365-383.
22. Benfenati E, Manganaro A, Gini G (2013) VEGA-QSAR: AI inside a platform for predictive toxicology. *Workshop Popularize Artificial Intelligence, Torino* 5: 21-28.
23. Melnikov F, Kostal J, Voutchkova-Kostal A, Zimmerman JB, Anastas PT (2016) Assessment of predictive models for estimating the acute aquatic toxicity of organic chemicals. *Green Chem* 18: 4432-4445.
24. <http://www.way2drug.com/PASSOnline/downloads/articles/Druzilovsky-abstract.pdf>.
25. Burden N, Maynard SK, Weltje L, Wheeler JR (2016) The utility of QSARs in predicting acute fish toxicity of pesticide metabolites: A retrospective validation approach. *Regul Toxicol Pharmacol* 80: 241-246.
26. Grech A, Brochot C, Dorne JL, Quignot N, Bois FY, et al. (2017) Toxicokinetic models and related tools in environmental risk assessment of chemicals. *Sci Total Environ* 578: 1-15.
27. Benfenati E, Belli M, Borges T, Casimiro E, Cester J, et al. (2016) Results of a round-robin exercise on read-across. *SAR QSAR Environ Res* 27: 371-384.
28. Gini G, Franchi AM, Manganaro A, Golbamaki A, Benfenati E (2014) ToxRead: A tool to assist in read across and its use to assess mutagenicity of chemicals. *SAR QSAR Environ Res* 25: 999-1011.

## APPENDIX 1: Structural Alerts as SMARTS strings

### SAs for class1toxicity

```

C(OCCCC)c1cccc(c1)
c1cc(c(O)c(c1)C(C)C)
Oc1ccc(cc1C)Cl
O=Cc1cccc(Oc2ccc(cc2))c1
Nc1ccc(cc1)CCCCCCCC
O=[N+][([O-])c1c(cc(c1)[N+](=O)[O-])Cl)Cl
NCCCCCCCCCCC
O(CC)P(=S)(OCC)SCS
CC[Sn](CC)(CC)CC
c1cc(ccc1CCl)
O=CC=CC(=O)
N#CCC#N
[c]([Cl,Br,F])[c]([Cl,Br,F])[c]([Cl,Br,F])
CCC(=O)OC[a]

```

### SAs for class2toxicity

```

Oc1ccc(cc1)Cl
C(C)CCCCCCCCC
c1ccc(Oc2ccccc2)cc1
C(=O)Oc1ccccc1

```

C(OCC=C)CC  
c1c(cccc1CC)CC  
O(CC)P(=S)(O)  
C=CC=C  
C(=O)OCCCC  
O=[N+][[O-]]c1cc(cc(c1)C)  
c1cc(c(cc1C))C  
SCC  
O(c1ccccc1)CCCC  
c1cccc2ccccc12  
c1cc(ccc1O)Br  
c1ccccc1c2ccccc2  
O=Cc1c(F)cccc1  
O=C(OC)CC  
[\*;D1]#C[C;!D4][!C;D1]  
\*[C;D2][C;D2][C;D2][C;D2][C;D2][C;D2]\*  
[s;R]  
[\$([c]([Cl,Br,F])[c]([Cl,Br,F])),\$([c]([Cl,Br,F])[c][c]([Cl,Br,F]))]

### SAs for class 3 toxicity

C(OC)c1ccc(cc1)  
NCCCCC  
c1cc(c(cc1)C)C  
Fc1ccccc1  
O=C(OCCC)C  
c1ccc(cc1)Br  
O=[N+][[O-]]c1cc(N)ccc1  
c1ccc(cc1)CCCC  
C=CCCC  
c1ccc(cc1)Cl  
Oc1ccc(cc1)C  
c1cc(ccc1N(C)C)  
c1cc(N)ccc1O  
CCCCCCCC  
C(C(Cl))Cl  
O=P(OCC)(OCC)OCC  
N(CCC)(CCC)C  
o1c(ccc1)  
C#CCC(O)  
C(CCCl)C  
OCC#CC  
O=[C;D2][C;D2]  
C=C  
c1ccccc1