Comparison of Seven in Silico methods for Evaluating of Ecotoxicological Acute Toxicity of Daphnia magna and Pimephales promelas: Case study on Chinese Priority Controlled Chemicals

Current Status: UNDER REVIEW

BMC Bioinformatics

Linjun Zhou
Nanjing Tech University

Deling Fan
Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment

Wei Yin
Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment

Wen Gu
Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environmental

Zhen Wang
Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment

Jining Liu
Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment

Yanhua Xu yanhuaxu18@hotmail.com
Nanjing Tech University
Corresponding Author

Lili Shi
Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment

Mingqing Liu
Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment
Abstract

Background: The acute toxicity on aquatic organisms are indispensable parameters in the ecological risk assessment priority chemical screening process (e.g. persistent, bioaccumulative and toxic chemicals). Currently, a number of predictive models for aquatic toxicity are available, however, the accuracy of in silico tools in priority assessment and risk assessment still remains to be further studied. Herein, this study evaluated the performance of seven Quantitative Structure–Activity Relationship (QSAR) in silico methods (Danish QSAR Database, Ecological Structure Activity Relationships, KAshinhou Tool for Ecotoxicity on PAS, Toxicity Estimation Software Tool, QSAR Toolbox, Read Across, and Virtual models for property Evaluation of chemicals within a Global Architecture) for assessing acute aquatic toxicity to Daphnia magna and Pimephales promelas using the first batch list of Priority Controlled Chemicals in China.

Results: Based on the values for the median lethal dose and the US Environmental Protection Agency’s acute aquatic toxicity categories of concern, the acute toxicity grade was classified into six categories. According to the comparative prediction results, the accuracy of the Daphnia magna toxicity categories prediction was 25%–56%, the correlation coefficient ranged from 0.1236 to 0.6349, and the correlation coefficients of the applicability domain were 0.040 and 0.5148. The corresponding values for the Pimephales promelas toxicity categories prediction were 22%–44%, 0.1495–0.4144, 0.2156 and 0.6793.

Conclusion: As the structure of chemicals of first batch list of Priority Controlled Chemicals in China are complex, the accuracy of model prediction is low, which depends on the quality of the constructed model and application domain. Although
in silico methods can be used to preliminarily estimate aquatic toxicity, experimental data validation is still required for prioritizing environmental hazards assessments and risk assessments.

Background

Global regulations have called for systematic testing of potential environmental contaminants to protect human health and the environment from exposure to anthropogenic chemicals, such as industrial chemicals and pharmaceuticals [1]. Considering the ever-increasing number of chemicals, more than 140000 chemicals used in the global market currently, are presenting challenges to traditional ecotoxicity testing strategies for in vivo experiments, which are expensive, time-consuming, and reliant on large number of animal subjects. Therefore, it is virtually impossible to determine acute toxicity for all the chemicals used globally, especially in light of new EU legislation to phase out animal testing[2]. The National Research Council (NRC) and global regulations, including the European Chemical Agency’s REACH initiative, the U.S. Toxic Substances Control Act (TSCA), and the Canadian Environmental Protection Act (CEPA), are encouraging increased reliance on in silico approaches to mitigate the challenges associated with in vitro and in vivo toxicity testing approaches[3-6].

With regard to the risk assessment of chemicals, the use of Quantitative Structure-Activity Relationship (QSAR) models, which relate chemical molecular structures with their physicochemical properties and environmental behavioural parameters, in the absence of toxicity data, toxicities relating to hazard identification can be assessed at minimal computational costs[7]. The cost-benefit advantages of in silico methods and regulatory support for them have led to the development of a number
of ecotoxicity assessment tools [8]. Such tools include the Ecological Structure Activity Relationships (ECOSAR), Kashinhou Tool for Ecotoxicity (KATE) and Toxicity Estimation Software Tool (T.E.S.T.), Virtual models for property Evaluation of chemicals within a Global Architecture (VEGA), Danish QSAR Database, and OECD QSAR Toolbox. ECOSAR and T.E.S.T. were developed by the United States Environmental Protection Agency (US EPA) and the Syracuse Research Corporation [9], KATE is a product of the Japanese Ministry of the Environment and the Japanese National Institute for Environmental Studies. The Danish QSAR Database, which comprises more than 200 (Q)SAR models covering 600,000 chemicals, was developed by the National Food Institute, Technical University of Denmark, with support from the Danish Environmental Protection Agency and the Nordic Council of Ministers and the European Chemicals Agency. To increase the regulatory acceptance of QSAR methods, the Organisation for Economic Co-operation and Development (OECD) developed the QSAR Toolbox. VEGA, developed by the Mario Negri Institute for Pharmacological Research, allows comprehensive evaluation of the applicability domain (AD) of the prediction results [10].

In 2007, the OECD guidelines on the development and validation of QSAR models were issued. They proposed that a QSAR model for practical application should be associated with an unambiguous algorithm[11], a defined endpoint, an AD, appropriate goodness-of-fit measures, robustness as well as predictive ability, and a mechanistic interpretation, if possible [12, 13]. A number of studies developed QSAR models for the endpoint of acute toxicity to Daphnia magna and Pimephales promelas[14–18]. Based on models for specific chemical classes and different classes of substances, Golbamaki, et al and Moore, et al have compared the performance of some QSAR models for acute toxicity [19, 20], but not systematic
and overall. Despite these guidelines, lack of external validations and model performances of the test sets, model overfitting, and poor AD definitions remain major concerns [21-24]. A clear AD definition would ensure that the model assumptions are met [25, 26]. Compared to physicochemical QSARs, previous validation efforts neglected to conduct a strictly external validation, relied on small data sets, or evaluated one tool at a time. The results of such efforts suggested that model accuracy for aquatic toxicity endpoints decreases during validation [27]. Thus, to ensure practical utility, we validated specific acute toxicity in silico models using an external testing set [28].

To implement the regulatory requirements of the “Action Plan for Prevention and Control of Water Pollution,” the Ministry of Ecological Environment, Ministry of Industry and Information Technology, and Health Planning Commission (China) jointly issued the List of Priority Controlled Chemicals (the first batch) at the end of 2017 [29, 30]. This list is presented in Table 1.

Table 1 chemicals of first batch priority controlled chemicals in China
| No. | Chemicals                                      | CAS                  | Predict? |
|-----|-----------------------------------------------|----------------------|----------|
| 1   | 1,2,4-Trichlorobenzene                        | 120-82-1             | Yes      |
| 2   | 1,3-butadiene                                  | 106-99-0             | Yes      |
| 3   | Musk xylene                                    | 81-15-2              | Yes      |
| 4   | 1,4-Benzenediamine,N1,N4-bis(methylphenyl)-   | 27417-40-9           | Yes      |
| 5   | Chloroalkanes C10-13                          | 85535-84-8/68920-70-7/71011-12-6/85536-22-7/85681-73-8/108171-26-2 | Yes |
| 6   | Dichloromethane                                | 75-09-2              | Yes      |
| 7   | Formaldehyde                                   | 50-00-0              | Yes      |
| 8   | Hexachlorocycloptadiene                       | 77-47-4              | Yes      |
| 9   | Hexabromocyclododecane                        | 25637-99-4/3194-55-6/134237-50-6/134237-51-7/134237-52-8 | Yes |
| 10  | Naphthalene                                    | 91-20-3              | Yes      |
| 11  | Perfluoroctyl sulfonic acid salts&perfluorocetyl sulfonyl chloride | 1763-23-1/307-35-7/2795-39-3/29457-72-5/29081-56-9/70225-14-8/56773-42-3/251099-16-8 | Yes |
| 12  | Nonylphenol and Nonylphenol Polyoxyethylene Ethyl | 25154-52-3/84852-15-3/9016-45-9 | Yes |
| 13  | Chloroform                                     | 67-66-3              | Yes      |
| 14  | Trichloroethylene                              | 79-01-6              | Yes      |
| 15  | Decabromodiphenyl oxide                       | 1163-19-5            | Yes      |
| 16  | Tetrachloroethylene                            | 127-18-4             | Yes      |
| 17  | Acetaldehyde                                   | 75-07-0              | Yes      |
| 18  | Cadmium                                       | 7440-43-9            | No       |
| 19  | Mercury                                       | 7439-97-6            | No       |
| 20  | Chromium VI                                    | No                   |          |
| 21  | Pb                                            | No                   |          |
| 22  | AsH₃                                          | 7440-38-2            | No       |

In view of the possible uses of QSAR tools, regulators often use predictions from multiple QSAR tools to arrive at a decision. However, the precision of predictions by different QSAR tools oriented for Chemical management is unknown. In this study, selected in silico tools, namely ECOSAR, T.E.S.T., Danish QSAR Database, VEGA, KATE, OECD Toolbox, and Read Across, were used to predict acute aquatic toxicity to *Daphnia magna* and *Pimephales promelas*. To provide insight into the applicability accuracy and ease of use of these QSAR tools, this study presents a systematic assessment using the correlation coefficient ($R^2$), correlation coefficient...
of the AD ($R^2_{AD}$), and acute toxicity accuracy within 1 log unit (which was classified into six categories). The test set used in this evaluation comprised the first list of priority controlled chemicals, which are representative of those evaluated by in silico tools.

Methods

Validation dataset

Systematic and rigorous model evaluation requires reliable experimental data. As such, acute aquatic toxicity experimental data ($LC_{50}$) were first obtained from ECHA’s risk assessment report, GLP, or standard test methods, and then converted into log values. Based on the US EPA’s acute aquatic toxicity categories, chemical toxicity categories were defined based on $LC_{50}$ values for the following six categories: Category 1 - High hazard I ($\log LC_{50} < -2$), Category 2 - High hazard II ($-2 < \log LC_{50} < -1$), Category 3 - High hazard III ($-1 < \log LC_{50} < 0$), Category 4 - Moderate hazard ($0 < \log LC_{50} < 1$), Category 5 - Low hazard ($1 < \log LC_{50} < 2$), and Category 6 - No hazard ($\log LC_{50} > 2$).

Predictive tools

The following seven in silico methods were evaluated for predicting acute aquatic toxicity to Daphnia magna and Pimephales promelas: ECOSAR, T.E.S.T., Danish QSAR Database, VEGA, KATE, QSAR Toolbox, and Read Across. A brief description of each program is provided below, and the pertinent details are summarized in Table 2.

Table 2. Summary of the predictive tools
Predictive tools | AD definition | Structural similarity | Algorithm | AD judge | Training set size
--- | --- | --- | --- | --- | ---
ECOSAR | log K\textsubscript{OW} range | Structural debirs | Linear regression | No | 1000/1000
T.E.S.T | Molecular descriptor space | Structural similarity coefficient | Linear regression | No | 823/353
Danish QSAR Database | Molecular descriptor space | Structural debirs and similarity search | Battery algorithm | Yes | 565/626
VEGA | Molecular descriptor space | Structural and atom centered fragments similarity | Read Across | Yes | > 5/>5
KATE | log K\textsubscript{OW} range | Structural debirs | Linear regression | Yes | 535/562
QSAR Toolbox (QSAR) | logK\textsubscript{OW} range & Molecular descriptor space | Structural debirs and similarity search | Linear regression | Yes | 823/353
QSAR Toolbox (Read Across) | Molecular descriptor space | Structural similarity coefficient | Read Across | Yes | > 5/>5

Ecological Structure Activity Relationships (ECOSAR) estimates acute aquatic toxicity via the Mayer–Overton relationship for chemicals within structurally similar classes[31]. ECOSAR is trained on a large data set of ecotoxicity studies from the ECOTOX database that follow the USEPA Office of Chemical Safety and Pollution Prevention (OCSPP) guidelines, which comprise 130 structural classes. The linear regression models between the LC\textsubscript{50} toxicity estimates and log K\textsubscript{OW} were developed for substances in each class, assuming the following limitations: maximum value log K\textsubscript{OW} = 5 and maximum molecular weight = 1000 amu. Chemicals that do not meet the latter two criteria or are structurally dissimilar from every QSAR model are considered to lie outside the AD. However, this tool suffers from the lack of external validations and model performance data of the test sets.

KAshinhou Tool for Ecotoxicity (KATE) on PAS 2011 estimates acute aquatic toxicity via the Mayer–Overton relationship for chemicals within a total of 40 structural chemical classes[32]. KATE is trained on the US EPA fathead minnow (Pimephales promelas) and the Japanese Ministry of Environment Oryzias latipes datasets [33].
The tool is available as a standalone application or as a web plug-in. The log $K_{OW}$ value of the test chemical, which is obtained from an internal experimental database or is estimated with the alternative forced choice method, is compared to the range of log $K_{OW}$ values in each structural class of the training set, and it internally defines the Ads.

Toxicity Estimation Software Tool (T.E.S.T.) v.4.1 estimates acute aquatic toxicity by Read Across among structural analogues or via multivariate regression (e.g. hierarchical clustering, single model (a single multiple linear regression model), the Food and Drug Administration method, nearest neighbour method, and consensus-based methods). T.E.S.T. is trained on the endpoint from the EPA ECOTOX database [34]. The optimal descriptor set, determined by the genetic algorithm, is used to characterize the toxicity of the chemicals[35]. The data from the suitable cluster is used to make predictions for test compounds. Each Read Across or regression model has a specific AD and structural similarity coefficient. The program provides estimated LC$_{50}$ thresholds based on each model’s prediction as well as the most accurate estimate of the component model [36].

VEGA estimates physicochemical, environmental, ecotoxicological, and toxicological properties using the Read Across method. The six most similar compounds are identified using the chemical similarity algorithm[37]. VEGA provides a simple way for Read Across using a similarity search, which is also used to evaluate the reliability of the model prediction. VEGA performs a sophisticated procedure to evaluate the reliability of the model; it refers to aspects of the AD of the model (such as the similarity of the most similar compounds, descriptor space, descriptor sensitivity, outliers based on specific fragments, identification of the presence of
rare fragments, and accuracy of the prediction assessment). The prediction results and applicability domain assessment are outputted.

Danish QSAR Database provides estimates for more than 600,000 chemicals in over 200 QSAR models by sorting with regard to chemical similarity to facilitate Read Across groupings. The database was developed at the Technical University of Denmark. The endpoints are modelled in three software systems (Leadscope, CASE Ultra, and SciQSAR), and an overall battery prediction is made to reduce “noise” from the individual model estimates and thereby improve accuracy and broaden the AD. The prediction results and AD judgment are outputted [38].

OECD QSAR Toolbox v4.2 finds structurally and mechanistically defined analogues and chemical categories, which serve as sources for Read Across and QSAR for filling in data gaps. The prediction results and AD assessment are outputted. Toolbox has multiple functions, such as identifying analogues of a chemical, retrieving the existing experimental results of those analogues, and filling in data gaps through Read Across or QSAR, classifying a large number of existing chemicals using the mechanism or behaviour model, using a QSAR model to fill in the data gaps for chemicals, evaluating the robustness of a potential analogue with Read Across, evaluating the applicability of a (Q)SAR model for a target compound to fill in the missing data, and establishing the QSAR model.

Results

External validation results for acute toxicity of

Daphnia magna

The 36 organic chemicals assessed in this study represent a diverse array of commercial substances. They include olefins, halides, nitrobenzene, perfluorooctane
sulfonate, phenols, aldehydes, diphenyl ether, biphenyl, and phenylamine. Heavy metals are not included in the QSAR prediction. The external validation results pertaining to the acute toxicity of Daphnia magna are shown in Table 3.
Table 3
Results of predicted toxicities and categories of Daphnia magna for the validation set

| CAS Exp No. | Predict data (LC50 mg/L) | Ecosar | T.E.S.T | QSAR Toolbox | Danish | VEGA | Read Across | KATE |
|-------------|---------------------------|--------|--------|--------------|--------|------|-------------|------|
|             | Cat. LC50 | Cat. LC50 | Cat. LC50 | Cat. LC50 | Cat. LC50 | Cat. LC50 | Cat. LC50 | Cat. LC50 | Cat. LC50 | Cat. LC50 |
| 120-82-1    | 120-82-1 | 120-82-1 | 120-82-1 | 120-82-1 | 120-82-1 | 120-82-1 | 120-82-1 | 120-82-1 | 120-82-1 | 120-82-1 |
| 01-79-66-9  | 01-79-66-9 | 01-79-66-9 | 01-79-66-9 | 01-79-66-9 | 01-79-66-9 | 01-79-66-9 | 01-79-66-9 | 01-79-66-9 | 01-79-66-9 | 01-79-66-9 |
| 81-15-2     | 81-15-2 | 81-15-2 | 81-15-2 | 81-15-2 | 81-15-2 | 81-15-2 | 81-15-2 | 81-15-2 | 81-15-2 | 81-15-2 |
| 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 |
| 007-5-09-2  | 007-5-09-2 | 007-5-09-2 | 007-5-09-2 | 007-5-09-2 | 007-5-09-2 | 007-5-09-2 | 007-5-09-2 | 007-5-09-2 | 007-5-09-2 | 007-5-09-2 |
| 50-00-0     | 50-00-0 | 50-00-0 | 50-00-0 | 50-00-0 | 50-00-0 | 50-00-0 | 50-00-0 | 50-00-0 | 50-00-0 | 50-00-0 |
| 77-47-3     | 77-47-3 | 77-47-3 | 77-47-3 | 77-47-3 | 77-47-3 | 77-47-3 | 77-47-3 | 77-47-3 | 77-47-3 | 77-47-3 |
| 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 | 855-35-84-8 |
| 91-20-3     | 91-20-3 | 91-20-3 | 91-20-3 | 91-20-3 | 91-20-3 | 91-20-3 | 91-20-3 | 91-20-3 | 91-20-3 | 91-20-3 |
| 176-3-23-1  | 176-3-23-1 | 176-3-23-1 | 176-3-23-1 | 176-3-23-1 | 176-3-23-1 | 176-3-23-1 | 176-3-23-1 | 176-3-23-1 | 176-3-23-1 | 176-3-23-1 |
| 307-35-7    | 307-35-7 | 307-35-7 | 307-35-7 | 307-35-7 | 307-35-7 | 307-35-7 | 307-35-7 | 307-35-7 | 307-35-7 | 307-35-7 |
| 279-5-39-3  | 279-5-39-3 | 279-5-39-3 | 279-5-39-3 | 279-5-39-3 | 279-5-39-3 | 279-5-39-3 | 279-5-39-3 | 279-5-39-3 | 279-5-39-3 | 279-5-39-3 |
| 251-54-23   | 251-54-23 | 251-54-23 | 251-54-23 | 251-54-23 | 251-54-23 | 251-54-23 | 251-54-23 | 251-54-23 | 251-54-23 | 251-54-23 |
| 848-52-15   | 848-52-15 | 848-52-15 | 848-52-15 | 848-52-15 | 848-52-15 | 848-52-15 | 848-52-15 | 848-52-15 | 848-52-15 | 848-52-15 |
| 901-6-45    | 901-6-45 | 901-6-45 | 901-6-45 | 901-6-45 | 901-6-45 | 901-6-45 | 901-6-45 | 901-6-45 | 901-6-45 | 901-6-45 |
| 67-66-3     | 67-66-3 | 67-66-3 | 67-66-3 | 67-66-3 | 67-66-3 | 67-66-3 | 67-66-3 | 67-66-3 | 67-66-3 | 67-66-3 |
| 79-01-6     | 79-01-6 | 79-01-6 | 79-01-6 | 79-01-6 | 79-01-6 | 79-01-6 | 79-01-6 | 79-01-6 | 79-01-6 | 79-01-6 |

R²: 0.3223  0.2377  0.3729  0.1236  0.5766  0.6349
R² AD: 0.3165  0.2277  0.1605  0.0400  0.5148  0.4063

Accuracies:

- 1. U. R. R. A. D. 0.5  0.4  0.3  0.2  0.1  0.0
- 2. 0.5  0.4  0.3  0.2  0.1  0.0
- 3. 0.5  0.4  0.3  0.2  0.1  0.0
- 4. 0.5  0.4  0.3  0.2  0.1  0.0

* indicated the chemical is out the applicability domain; NA indicated not available for prediction.
Total accuracy measures the fraction of chemicals correctly placed in the toxicity category with a log scale error within 1, while missing predictions are excluded from the analysis. The correlation coefficient (R^2) and the correlation coefficient in the AD (R^2_AD) between the experimental and predictive data were calculated for these models. Based on the predictive power of classification of the entire data set into the six toxicity categories, the tested tools for Daphnia magna can be ranked in the following order, from the highest to the lowest performers: T.E.S.T. > Danish QSAR Database = QSAR Toolbox > KATE > ECOSAR = Read Across > VEGA. The total accuracy of T.E.S.T. is relatively high at 56%, and R^2 of the experimental and predicted values is 0.2095. The accuracies of both the Danish QSAR Database and the QSAR Toolbox are 44%, while the corresponding R^2 values are 0.3729 and 0.2277, and the R^2_AD values are 0.1605 and 0.2277. A few of the chemicals, namely, 2, 8, 3, 2, and 5, are outside the ADs of the Danish QSAR Database, VEGA, Read Across, ECOSAR, and KATE. The accuracy, R^2, and R^2_AD of KATE are 37%, 0.6349, and 0.4063, respectively; while those of ECOSAR are 31%, 0.3223, and 0.1615, respectively. The corresponding values for Read Across are 31%, 0.5766, and 0.5148; and those for VEGA are 25%, 0.1236, and 0.04.

The predicted deviations (Fig. 1) of T.E.S.T., Read Across, and VEGA for hexachlorocyclopentadiene (CAS: 77-47-4) and hexabromocyclododecane (CAS: 25637-99-4) exceed 1 log unit, whereas those of ECOSAR and QSAR Toolbox for perfluorooctyl sulfonic acid salts and perfluorooctyl sulfonyl chloride (CAS: 307-35-7) exceed 1 log unit. The values for the other substances fall within the 1 log unit deviation. Hexabromocyclododecane is considered to be outside the ADs of five in silico methods, and perfluorooctyl sulfonic acid salts and perfluorooctyl sulfonyl
chloride are outside the ADs of two in silico methods.

External validation results for acute toxicity of Pimephales promelas

The external validation results for the acute toxicity of Pimephales promelas are shown in Table 3. Based on predictive power of classification into the six toxicity categories of the entire data set, the tested tools for Pimephales promelas can be ranked in the following order from the highest to the lowest performers: QSAR Toolbox > ECOSAR > Read Across > KATE > T.E.S.T. = Danish QSAR Database > VEGA. The total accuracy of QSAR Toolbox for the six categories is 44%, whereas the values of $R^2$ and $R^2_{AD}$ of the experimental and predicted values are 0.318 and 0.2156, respectively. The accuracies of ECOSAR and Read Across are 44% and 39%, whereas the corresponding $R^2$ and $R^2_{AD}$ values are 0.2166 and 0.4144, and 0.4229 and 0.4430. Moreover, 2, 3, 6, 2, 2, and 4 of the chemicals are outside the AD for the Danish QSAR Database, QSAR Toolbox, VEGA, Read Across, ECOSAR, and KATE, respectively. The accuracy, $R^2$, and $R^2_{AD}$ of KATE are 38%, 0.2714, and 0.6793, respectively, while the corresponding values of the Danish QSAR Database are 27%, 0.1495, and 0.2381.
Table 4
Results of predicted toxicities and categories of *Pimephales promelas* for the validation set

| CAS No. | Exp data LC₅₀ (mg/L) | Predict data (LC₅₀ mg/L) | ECOSAR | T.E.S.T | QSAR Toolbox | Danish | Vega | Read Across | KATE |
|---------|----------------------|--------------------------|--------|---------|--------------|--------|------|-------------|------|
|         | LC₅₀ Cat. | LC₅₀ Cat. | LC₅₀ Cat. | LC₅₀ Cat. | LC₅₀ Cat. | LC₅₀ Cat. | LC₅₀ Cat. | LC₅₀ Cat. | LC₅₀ Cat. | LC₅₀ Cat. | LC₅₀ Cat. |
| 120-82-1 | 0.7 | 3 | 2.82 | 4 | 1.84 | 4 | 1.81 | 4 | 1.36 | 4 | 2.67 | 4 | 15.8 | 5 | 2.4 | 4 |
| 81-15-2 | 0.2 | 3 | 0.23 | 3 | 0.07 | 2 | 0.19 | 3 | 0.28 | 3 | 1.06 | 4 | 0.27 | 3 | 1.4 | 4 |
| 855-35-84-8 | 100 | 6 | 0.13 | 3 | 6.44 | 4 | 127 | 6 | NA | NA | NA | NA | 192 | 6 | 0.19 | 3 |
| 007-5-09-2 | 330 | 6 | 249 | 6 | 316 | 6 | 284 | 6 | 249 | 6 | 311 | 6 | 31 | 5 | 180 | 6 |
| 50-00-0 | 23.95 | 12.5 | 5 | NA | NA | 17.95 | 5 | 12.5 | 6 | 0.77 | 3 | 31 | 5 | 11 | 5 |
| 71-47-4 | 0.001 | 6 | 0.34 | 3 | 0.33 | 3 | 0.09 | 2 | 0.34 | 3 | 0.24 | 3 | 0.35 | 3 | 0.03 | 2 |
| 256-37-99-4 | 100 | 6 | 0.004 | 5 | 0.04 | 2 | 40 | 8 | 0.001 | 4* | 0.001 | 4* | 9.62 | 4 | 40.0 | 5 |
| 91-20-3 | 0.96 | 5 | 9.24 | 9 | 7.27 | 4 | 6.31 | 4 | 3.19 | 4 | 6.12 | 4 | NA | NA | 8.8 | 4 |
| 307-35-7 | 4.7 | 4 | 0.04 | 2 | 0.24 | 3 | 0.04 | 2 | 77 | 6 | 102 | 6 | 654 | 6 | 0.08 | 2 |
| 279-5-39-3 | 9.5 | 4 | 23.6 | 5 | 0.57 | 3 | 74.5 | 5 | 0.00 | 1 | NA | NA | 654 | 6 | NA | NA |
| 251-54-52-3 | 0.12 | 3 | 0.03 | 2 | 0.34 | 3 | 0.45 | 4 | 0.49 | 4 | 0.13 | 3 | 0.14 | 3 | 0.08 | 2 |
| 848-5 | 0.13 | 3 | 0.05 | 2 | 1.16 | 4 | 0.5 | 3 | 0.08 | 7 | 2 | 0.01 | 2 | 0.17 | 3 | 0.08 | 2 |
| 901-645-9 | 5.00 | 4 | 2.23 | 4 | 9.95 | 4 | 0.93 | 3 | 0.02 | 2 | 16.2 | 5 | 0.21 | 5 | 3.5 | 4 |
| 67-66-3 | 121 | 6 | 242 | 6 | 72.2 | 4 | 41.9 | 5 | NA | NA | 54.1 | 5 | 24.3 | 5 | 68.0 | 5 |
| 79-01-62 | 44.5 | 5 | 11.2 | 5 | 30.4 | 9 | 45.6 | 5 | 9.95 | 4 | NA | NA | 52.3 | 5 | 23.0 | 5 |
| 116-319-5 | 0.18 | 3 | 40.1 | 3 | 40.1 | 3 | 9.45 | 5 | 40.1 | 4* | 0.79 | 3 | 5.37 | 4 | 7.3E-4 | 4 |
| 127-184 | 8.4 | 4 | 5.4 | 4 | 15.6 | 5 | 7.44 | 4 | 2.86 | 4 | 11.6 | 4 | 0.24 | 3 | 7.1 | 4 |
| 75-070 | 30.85 | 3 | 33.8 | 3 | 36.9 | 5 | 34.3 | 5 | 37.4 | 9 | 31.9 | 5 | 27.4 | 5 | 36.0 | 5 |
| r² | 0.21 | 66 | 0.29 | 1 | 0.31 | 80 | 0.14 | 95 | 0.41 | 23 | 0.41 | 44 | 0.41 | 44 | 0.27 | 14 |
| R² AD | 0.42 | 29 | 0.29 | 10 | 0.21 | 56 | 0.23 | 81 | 0.19 | 43 | 0.29 | 10 | 0.67 | 93 | 0.67 | 93 |
| Six categories accuracy | 44% | 27% | 44% | 27% | 22% | 39% | 35% |

* indicated the chemical is out the applicability domain; NA indicated not available for prediction.

The deviation (Fig. 2) of the in silico methods for hexabromocyclododecane, hexachlorocyclopentadiene, and perfluoroctyl sulfonic acid salts and perfluoroctyl sulfonyl chloride are above 1 log unit, similar to the predictions of acute toxicity of...
Daphnia magna. Hexabromocyclododecane is outside the ADs of five in silico methods, while perfluorooctyl sulfonic acid salts and perfluorooctyl sulfonyl chloride are outside the ADs of one in silico method.

Uncertainty of the predictions

To improve regulatory confidence and acceptance of toxicity predictions made by in silico methods, it is crucial to assess the uncertainty of the predictions and determine to what extent the uncertainty is acceptable. The sources of uncertainty are as follows: quality of the data, mechanisms of action, descriptors (experimentally measured or calculated properties; e.g. log $K_{OW}$ and relevance of the descriptors), statistical method, and AD[39]. Currently, no coherent mapping or definition of uncertainty exists with regard to QSAR models[40]. The predictions of acute toxicity in the regulatory data set containing chemicals such as hexabromocyclododecane, hexachlorocyclopentadiene, and perfluorooctyl sulfonic acid salts and perfluorooctyl sulfonyl chloride are not used in the training set as some amount of uncertainty is associated with them. Thus, they inapplicable. T.E.S.T. and VEGA do not provide any correct toxicity classifications for LC$_{50}$ (median lethal dose) < 0.01 mg/L, 0.01 mg/L < LC$_{50}$ < 0.1 mg/L, LC$_{50}$ > 100 mg/L for Daphnia magna, and the same is true for predictions of Pimephales promelas by T.E.S.T. and VEGA for LC$_{50}$ < 0.01 mg/L. Moreover, none of the tools provide correct toxicity classifications for 0.01 mg/L < LC$_{50}$ < 0.1 mg/L for Pimephales promelas. This shows that some chemicals are mispredicted as being extremely toxic (0.01 mg/L < LC$_{50}$ < 0.1 mg/L and LC$_{50}$ < 0.01 mg/L). Notably, the tools provide predictions regarding the acute toxicity of the chemicals to Daphnia magna for LC$_{50}$ values ranging from less than 0.01 mg/L to greater than 100 mg/L. The tools provide
poor predictions of acute toxicity of the chemicals to Pimephales promelas when \( \text{LC}_{50} \) ranges from 0.01 mg/L to 0.1 mg/L. The number of correct classifications of acute toxicity for Daphnia magna and Pimephales promelas are shown in Figs. 3 and 4. The results show that the predictions for perfluorooctyl sulfonic acid salts, perfluorooctyl sulfonyl chloride, hexachlorocyclopentadiene, decabromodiphenyl oxide, 1,2,4-trichlorobenzene, and musk xylene, all of which are composed of special structures, are poor.

Discussion

Chapter R.6 of the “Guidance on Information requirements and Chemical Safety Assessment” is devoted to QSAR and the grouping of chemicals [41] does not provide a list of QSAR tools that have regulatory acceptance. However, it does provide the criteria that must met before a QSAR tool can be accepted in the regulatory context. These criteria refer to the OECD Setubal Principles for (Q)SAR Validation[42, 43]. The results show that the accuracies of four categories of acute toxicity are significantly higher than six categories of acute toxicity (see Table 5). Our study is unique in that Priority Controlled Chemicals grouped in six defined toxicity categories have been used to predict acute toxicities using commercially available tools and compare these values with their experimentally derived counterparts. Currently, a number of predictive in silico models for aquatic toxicity are available, but most face challenges in producing accurate predictions across a wide variety of functional chemical classes. These results are affected by the information on which the QSAR is based, including the quality, consistency, reporting, and reliability of the toxicological data, chemical structures, and properties descriptors, as well as the development and nature of the algorithm that
forms the basis of the QSAR. The toxicological data of most in silico tools (ECOSAR, T.E.S.T., Danish QSAR Toolbox, and KATE) are sourced from ECOTOX, which comprises peer-reviewed literature, and the test results are identified through comprehensive searches of the open literature. However, preferably high quality data can be modelled with the OECD Test Guideline or Good Laboratory Practice (GLP). The prediction results are also closely related to the types of chemicals and the amount of data. Thousands of physicochemical property data (e.g. boiling point, log $K_{\text{OW}}$, melting point, and water solubility) are trained for in silico methods (and used in ECOSAR, T.E.S.T., Danish QSAR Toolbox, and KATE)[44]. By contrast, assessing toxicity mechanisms is complex with relatively few experimental data[45]. Moreover, lack of information about the quality of the prediction, external validation, and ADs increases the uncertainty of the model accuracy (e.g. ECOSAR and KATE). VEGA, Read Across, and the Danish QSAR Database provide predictions that fall inside or outside the AD of the models. There is no single and absolute AD for a given model. Generally, the broader the definition of the AD, the lower the predictivity. The AD should be clearly defined, and the validation results should correspond to this defined domain, which is used again when the model is applied for the predictions[46, 47]. Although there is no criterion to judge the validity or invalidity of the predicted data, predicted results within the AD are preferred. Measured data are thus valuable for assessing toxicities of priority controlled chemicals. And the alternative test systems or endpoints, for instance using fish embryos, may allow reduction or replacement of the the fish early-life stage test or prioritize compounds for conduction of the FELS test[35].
Conclusion

With regard to ecological risk assessments of organic chemicals, QSAR models play an important role in filling the data gaps of toxicity endpoints, decreasing experimental expenses, reducing and replacing actual testing (especially animal testing), and assessing the uncertainty of experimental data. In this study, the performance of seven in silico methods (Danish QSAR Database, ECOSAR, KATE, T.E.S.T., QSAR Toolbox, Read Across, and VEGA) for acute aquatic toxicity to Daphnia magna and Pimephales promelas were evaluated and compared by using the first batch list of Priority Controlled Chemicals in China. As the structure of chemicals of first batch list of Priority Controlled Chemicals in China are complex, the accuracy of model prediction are low and some chemicals are mispredicted as being extremely toxic, which depends on the quality of the constructed model and application domain. Although in silico methods can be used to preliminarily estimate aquatic toxicity, taking into account the uncertainty that has been evaluated in our study, experimental data validation of the alternative test systems or endpoints is still required for prioritizing environmental hazards assessments and risk assessments.

Declarations

Acknowledgments
Authors' contributions
LJZ compared the data. DLF analyzed the data, and was a major contributor in writing the manuscript. WG researched the in silico software. J JL and MQL revised the manuscript. ZW and WY predicted the chemical using models. YHX and LLS contributed to design of the study. All authors read and approved the final manuscript.

Funding
This work was funded by the National Key Research and Development Program of China (No. 2018YFC1801504 and 2018YFC1801601).

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author or from: https://pan.baidu.com/s/1wBFhArHQ79f7mODP_vnNiA

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1 Nanjing Tech University, Nanjing, 211816, China.
2 Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment, Nanjing, 210042, China

References
1. MacLeod M, McKone TE, Foster KL, Maddalena RL, Parkerton TF, Mackay D: Applications of Contaminant Fate and Bioaccumulation Models in Assessing Ecological Risks of Chemicals: A Case Study for Gasoline Hydrocarbons. *Environmental Science & Technology* 2004, 38(23):6225-6233.

2. Aptula AO, Roberts DW: Mechanistic Applicability Domains for Nonanimal-Based Prediction of Toxicological End Points: General Principles and Application to Reactive Toxicity. *Chemical Research in Toxicology* 2006, 19(8):1097-1105.

3. Commission E: Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. *Official Journal of the European Union* 2010, L276:33-76.

4. Sanderson H, Solomon K: Contaminants of emerging concern challenge ecotoxicology. *Environmental Toxicology and Chemistry* 2009, 28(7):1359-1360.

5. Council NR: A Framework to Guide Selection of Chemical Alternatives. Washington, DC: The National Academies Press; 2014.

6. ECHA: The Use of Alternatives to Testing on Animals for the REACH Regulation, European Chemicals Agency, Third report under Article 117(3) of the REACH Regulation. In. Helsinki, Finland: European Chemicals Agency; 2017.

7. Cronin MTD, Yoon M: Chapter 5.3 - Computational Methods to Predict Toxicity. In: *The History of Alternative Test Methods in Toxicology*. Edited by Balls M, Combes R, Worth A: Academic Press; 2019: 287-300.

8. Vouthkova AM, Osimitz TG, Anastas PT: Toward a Comprehensive Molecular Design Framework for Reduced Hazard. *Chemical Reviews* 2010, 110(10):5845-5882.
9. Mayo-Bean K, Moran K, Meylan B, Ranslow P: Methodology Document for the ECOlogical Structure-Activity Relationship Model (ECOSAR) Class Program. In. Washington DC: US-EPA; 2012.

10. Benfenati E, Manganaro A, Gini GC: VEGA-QSAR: AI Inside a Platform for Predictive Toxicology. In: Proceedings of the Workshop Popularize Artificial Intelligence co-located with the 13th Conference of the Italian Association for Artificial Intelligence (AIxIA 2013): 2013; Turin, Italy. 21-28.

11. Cardoso-Silva J, Papageorgiou LG, Tsoka S: Network-based piecewise linear regression for QSAR modelling. *J Comput Aided Mol Des* 2019.

12. OECD: Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models; 2014.

13. Toropov AA, Raska I, Jr., Toropova AP, Raskova M, Veselinovic AM, Veselinovic JB: The study of the index of ideality of correlation as a new criterion of predictive potential of QSPR/QSAR-models. *Sci Total Environ* 2019, 659:1387-1394.

14. Sheffield TY, Judson RS: Ensemble QSAR Modeling to Predict Multispecies Fish Toxicity Lethal Concentrations and Points of Departure. *Environ Sci Technol* 2019, 53(21):12793-12802.

15. Ding F, Wang Z, Yang X, Shi L, Liu J, Chen G: Development of classification models for predicting chronic toxicity of chemicals to Daphnia magna and Pseudokirchneriella subcapitata. *SAR QSAR Environ Res* 2019, 30(1):39-50.

16. Fan D, Liu J, Wang L, Yang X, Zhang S, Zhang Y, Shi L: Development of Quantitative Structure-Activity Relationship Models for Predicting Chronic Toxicity of Substituted Benzenes to Daphnia Magna. *Bull Environ Contam Toxicol* 2016, 96(5):664-670.
17. Kluver N, Bittermann K, Escher BI: QSAR for baseline toxicity and classification of specific modes of action of ionizable organic chemicals in the zebrafish embryo toxicity test. *Aquat Toxicol* 2019, 207:110-119.

18. Jia Q, Zhao Y, Yan F, Wang Q: QSAR model for predicting the toxicity of organic compounds to fathead minnow. *Environ Sci Pollut Res Int* 2018, 25(35):35420-35428.

19. Moore DR, Breton RL, MacDonald DB: A comparison of model performance for six quantitative structure-activity relationship packages that predict acute toxicity to fish. *Environ Toxicol Chem* 2003, 22(8):1799-1809.

20. Golbamaki A, Cassano A, Lombardo A, Moggio Y, Colafranceschi M, Benfenati E: Comparison of in silico models for prediction of Daphnia magna acute toxicity. *SAR QSAR Environ Res* 2014, 25(8):673-694.

21. Lombardo A, Roncaglioni A, Benfenati E, Nendza M, Segner H, Jeram S, Pauné E, Schüürmann G: Optimizing the aquatic toxicity assessment under REACH through an integrated testing strategy (ITS). *Environmental Research* 2014, 135:156-164.

22. Benfenati E, Diaza RG, Cassano A, Pardoe S, Gini G, Mays C, Knauf R, Benighaus L: The acceptance of in silico models for REACH: Requirements, barriers, and perspectives. *Chemistry Central Journal* 2011, 5(1):58.

23. Feher M, Ewing T: Global or Local QSAR: Is There a Way Out? *QSAR & Combinatorial Science* 2009, 28(8):850-855.

24. Gramatica P: Principles of QSAR models validation: internal and external. *QSAR & Combinatorial Science* 2007, 26(5):694-701.

25. Nendza M, Muller M, Wenzel A: Discriminating toxicant classes by mode of action: 4. Baseline and excess toxicity. *SAR QSAR Environ Res* 2014, 25(5):393-
26. Cronin MTD, Schultz TW: Pitfalls in QSAR. *Journal of Molecular Structure: THEOCHEM* 2003, 622(1):39-51.

27. Cheng F, Li W, Zhou Y, Shen J, Wu Z, Liu G, Lee PW, Tang Y: admetSAR: A Comprehensive Source and Free Tool for Assessment of Chemical ADMET Properties. *Journal of Chemical Information and Modeling* 2012, 52(11):3099-3105.

28. Khan K, Baderna D, Cappelli C, Toma C, Lombardo A, Roy K, Benfenati E: Ecotoxicological QSAR modeling of organic compounds against fish: Application of fragment based descriptors in feature analysis. *Aquat Toxicol* 2019, 212:162-174.

29. MEE, China. List of Priority Controlled Chemicals (The First Batch) http://www.mee.gov.cn/gkml/hbb/bgg/201712/t20171229_428832.htm. Accessed 15 Oct 2019.

30. Gramatica P, Papa E, Sangion A: QSAR modeling of cumulative environmental end-points for the prioritization of hazardous chemicals. *Environ Sci Process Impacts* 2018, 20(1):38-47.

31. Ministry of the Environment in Japan. Results of eco-toxicity tests of chemicals conducted by Ministry of the Environment in Japan (- March 2015). http://www.env.go.jp/chemi/sesaku/02e.pdf. Accessed 15 Oct 2019

32. Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE, Drummond RA: Predicting modes of toxic action from chemical structure: Acute toxicity in the fathead minnow (Pimephales promelas). *Environmental Toxicology and Chemistry* 1997, 16(5):948-967.

33. Furuham A, Toida T, Nishikawa N, Aoki Y, Yoshioka Y, Shiraishi H:
Development of an ecotoxicity QSAR model for the KAshinhou Tool for Ecotoxicity (KATE) system, March 2009 version. *SAR and QSAR in Environmental Research* 2010, 21(5-6):403-413.

34. Gramatica P, Pilutti P: Evaluation of different statistical approaches for the validation of quantitative structure-activity relationships. Ispra, Italy, The European Commission-Joint Research Centre. *Institute for Health and Consumer Protection–ECVAM* 2004.

35. Furuhama A, Hayashi TI, Yamamoto H: Development of QSAAR and QAAR models for predicting fish early-life stage toxicity with a focus on industrial chemicals. *SAR QSAR Environ Res* 2019, 30(11):825-846.

36. Floris M, Manganaro A, Nicolotti O, Medda R, Mangiatordi GF, Benfenati E: A generalizable definition of chemical similarity for read-across. *Journal of Cheminformatics* 2014, 6(1):39.

37. DTU. User Manual for the Danish (Q)SAR Database http://qsardb.food.dtu.dk/DB_user_manual_21_12_2018.pdf. Accessed 15 Oct 2019.

38. Jaworska JS, Comber M, Auer C, Leeuwen CJV: Summary of a workshop on regulatory acceptance of (Q)SARs for human health and environmental endpoints. *Environmental Health Perspectives* 2003, 111(10):1358-1360.

39. Kar S, Ghosh S, Leszczynski J: Single or mixture halogenated chemicals? Risk assessment and developmental toxicity prediction on zebrafish embryos based on weighted descriptors approach. *Chemosphere* 2018, 210:588-596.

40. Cronin MTD, Richarz A-N, Schultz TW: Identification and description of the uncertainty, variability, bias and influence in quantitative structure-activity relationships (QSARs) for toxicity prediction. *Regulatory Toxicology and
41. ECHA: Guidance on information requirements and chemical safety assessment. Chapter R.6: QSAR and grouping of chemicals. 2008.

42. Cronin MTD, Dearden JC, Walker JD, Worth AP: Quantitative structure-activity relationships for human health effects: Commonalities with other endpoints. *Environmental Toxicology and Chemistry* 2003, 22(8):1829-1843.

43. Eriksson L, Jaworska J, Worth AP, Cronin MTD, McDowell RM, Gramatica P: Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. *Environmental Health Perspectives* 2003, 111(10):1361-1375.

44. Olguin CJM, Sampaio SC, Dos Reis RR, Remor MB, Olguin CFA: QSPR modelling of the soil sorption coefficient from training sets of different sizes. *SAR QSAR Environ Res* 2019, 30(5):299-311.

45. Pogodin PV, Lagunin AA, Filimonov DA, Nicklaus MC, Poroikov VV: Improving (Q)SAR predictions by examining bias in the selection of compounds for experimental testing. *SAR QSAR Environ Res* 2019, 30(10):759-773.

46. Melnikov F, Kostal J, Voutchkova-Kostal A, Zimmerman JB, T. Anastas P: Assessment of predictive models for estimating the acute aquatic toxicity of organic chemicals. *Green Chemistry* 2016, 18(16):4432-4445.

47. ECHA. Substance Evaluation – CoRAP [https://echa.europa.eu/information-on-chemicals/evaluation/community-rolling-action-plan/corap-table. Accessed 15 Oct 2019.

Figures
Correlations (deviations) between the predicted and experimental Daphnia magna LC50 values (log units).
Correlations (or deviations) between the predicted and experimental Pimephales
Figure 3

Number of correct classifications for acute toxicity predictions of Daphnia magna.

Figure 4

Number of correct classifications for acute toxicity predictions of Pimephales promelas.
