	QMRF identifier (JRC Inventory):
	QMRF Title: **iSafeRat® High Accuracy QSAR for physicochemical and ecotoxicological endpoints**
	Printing Date: **Dec 11, 2019**

1.QSAR identifier

1.1.QSAR identifier (title):

iSafeRat® High Accuracy QSAR for physicochemical and ecotoxicological endpoints

1.2.Other related models:

1.3.Software coding the model:

iSafeRat® HA-QSAR toolbox v1.3

2.General information

2.1.Date of QMRF:

16/10/2014

2.2.QMRF author(s) and contact details:

Dr. Paul Thomas KREATiS SAS, 23, rue du Creuzat,38080 L'Isle d'Abeau, France +33 (0)4 28 19 01 06 paul.thomas@kreatis.eu www.kreatis.eu

2.3.Date of QMRF update(s):

27/02/2014 – v1 by KREATiS.05/12/2014 – v1 reviewed by JRC, ISPRA, Italy

16/10/2014 – v2 by KREATiS.

2.4.QMRF update(s):

This QMRF refers to the version 1.3 for the iSafeRat® High Accuracy QSAR for physicochemical and ecotoxicological endpoints

2.5.Model developer(s) and contact details:

Dr. Paul Thomas KREATiS SAS, 23, rue du Creuzat,38080 L'Isle d'Abeau, France +33 (0)4 28 19 01 06 paul.thomas@kreatis.eu www.kreatis.eu

2.6.Date of model development and/or publication:

The results presented in this QMRF refers to the current version of the model (iSafeRat holistic HA-QSAR v1.1) validated on 02-01-2015.

2.7.Reference(s) to main scientific papers and/or software package:

iSafeRat® – in Silico Algorithms For Environmental Risk And Toxicity version 1.1

2.8.Availability of information about the model:

The model is proprietary. The validation set for the iSafeRat holistic HA-QSAR will not be made publicly available.However, if required it may be provided (except for any confidential data) under certain conditions by contacting KREATiS directly.

2.9.Availability of another QMRF for exactly the same model:

None to date.

3.Defining the endpoint - OECD Principle 1

3.1.Species:

As the QMRF refers to a holistic approach addressing multiple endpoints, several species were involved on an endpoint by endpoint basis.

Physicochemical endpoints (octanol-water partition coefficient and water

solubility) values were derived using analytical laboratory methods (no test organisms involved). On the other hand, acute aquatic toxicity values were determined for fish, invertebrates and algae. Table 1 (accompanying documents) summarises the species for which relevant studies were included in the training set for various ecotoxicological endpoints.

3.2.Endpoint:

[1]QMRF 1. Physical Chemical Properties QMRF 1. 3. Water solubility
[2]QMRF 1. Physical Chemical Properties QMRF 1. 6. Octanol-water partition coefficient (Kow)
[3]QMRF 3. Ecotoxic effects QMRF 3. 1. Short-term toxicity to Daphnia (immobilisation)
[4]QMRF 3. Ecotoxic effects QMRF 3. 2. Short-term toxicity to algae (inhibition of the exponential growth rate)
[5]QMRF 3. Ecotoxic effects QMRF 3. 3. Acute toxicity to fish (lethality)

3.3.Comment on endpoint:

This QMRF deals with multiple endpoints in a holistic approach. The selection of these endpoints was based on the thermodynamic principles that relate them.
Species relevant to each ecotoxicological endpoint data:

Endpoint: Short-term toxicity to Daphnia
Species: Daphnia magna, Daphnia pulex

Endpoint: Short-term toxicity to algae
Species: Desmodesmus subspicatus, Pseudokirchneriella subcapitata, Scenedesmus quadricauda

Endpoint: Acute toxicity to fish
Species: Danio rerio, Lepomis macrochirus, Pimephales promelas, Oncorhynchus mykiss, Oryzias latipes

The term ‘holistic approach’ implies that the methodology can be used to predict any of the earlier discussed endpoints from the known values on the remaining endpoints. For example, if the solubility of the molecule is known, it can be used to derive the Log Kow as well as ecotoxicological endpoints. In other words, each missing endpoint is treated as a data gap which is filled using the available information on other endpoints. This makes the approach robust and provides transparency in terms of the thermodynamic relationship between different endpoints. This is further discussed in section 8 on mechanistic understanding of the model.

3.4.Endpoint units:

Table 2 (refer to the accompanying documents) provides with the units relevant for each involved endpoint. Note that for certain endpoints, these units were used only to report predictions in QPRFs. For modelling purposes, different scales were used (refer to section 3.5).

3.5.Dependent variable:

Table 3 (refer to the accompanying documents) provides an overview of the dependent variables of all the endpoints involved (and additionally, test durations are provided for all ecotoxicological endpoints).

3.6.Experimental protocol:

3.6.1. Octanol-water partition coefficient:

The experimental Log Kow values were measured using one of the following lab techniques:

- Shake Flask method (OECD 107 protocol)
- Slow Stirring (OECD 123 protocol)

3.6.2. Water solubility:

The experimental solubility values were measured using one of the following lab techniques:

- OECD 105 protocol on water solubility
- Slow stirring method (Modified 105 protocol)

3.6.3. Ecotoxic effects

Table 4 (refer to the accompanying documents) provides an overview of the test duration and study protocol relevant for each acute aquatic toxicity endpoint.

3.7.Endpoint data quality and variability:

One of the essentials to derive a HA-QSAR model is to perform a validation on the study results included within the training set. This validation was carried out using expert judgement by the dedicated KREATiS model development team. If for any reason the quality of the study results was compromised (for instance, due to unacceptable experimental conditions or issues with laboratory protocol), their corresponding results were withdrawn from the training set and the reason for their removal labelled in the internal database.

The training set data comprised of quality results derived from multiple laboratories and as a result, inter-laboratory differences may be expected. In many cases, diverse experimental methodologies were followed for the same endpoint (for instance, HPLC and Shake Flask methods for log Kow studies). For such cases, the results were not simply averaged but the validity of each result was then justified on a case by case basis. Cases with large differences in validated values for the same substance were treated according to good modelling practices (for instance, potential outlier detection, data verification from available literature resources).

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

QSAR

4.2.Explicit algorithm:

QSAR

see equation

The thermodynamically-driven iSafeRat® holistic approach may be used by taking just the chemical structure as an input by the user.

a) Analysing the input chemical structure, its log Kow is predicted using a 'Regression based-Fragment Approach' where linear regression equations for a series of common structures (for example alkanes) have been generated using high quality log Kow data and are included in the iSafeRat database. For a given chemical structure, the contribution values of all its relevant structural fragments are added together to obtain a high accuracy Log Kow prediction.
b) The Log Kow value (derived in step a) is given as an input to generate the corresponding water solubility value using a 'Simple Linear Regression approach'.
c) The solubility value (derived in step b) is then given as an input to predict the corresponding ecotoxicological endpoint values, again following a 'Simple Linear Regression approach'.

The iSafeRat® approach can be thought of a series of interrelated models where the basic structural input from the user generates prediction for the first endpoint, which is then used as input (prediction for first endpoint) to generate prediction for the second endpoint and so on. By default, the Log Kow of the target chemical structure is predicted using the fragment-based approach which is then used for water solubility estimation by means of a linear regression model and finally the predicted solubility value is used to estimate the toxicity levels at different trophic levels. Additional user input might be needed under specific circumstances. For instance, if the Log Kow of a target chemical cannot be determined as a part of the holistic approach, a measured Log Kow can be given as user input, if known.

Figure 1 (refer to the accompanying documents): A simple representation of the iSafeRat® holistic HA-QSAR workflow

This section provides a brief explanation on each prediction model involved to complete the holistic methodology. To allow a better understanding of the algorithm, a schematic representation of the holistic approach is provided at the end of this QMRF.
4.2.1a Fragment contribution based Log Kow Prediction
a) This model is based on the Log Kow contribution values for various fragment groups represented within the iSafeRat® data inventory. For a given chemical structure, the contribution values of all its relevant structural fragments are added together to obtain a high accuracy Log Kow prediction.b) If any of the desired fragment contribution values are missing, this
hinders a high accuracy Log Kow prediction and as a result, the input

structure is considered to be outside the applicability domain of this model.

4.2.1b Water Solubility Prediction: Local regression models for different chemical classes The Log Kow predicted from the Fragment-based approach is then given as an input to the second phase of the holistic approach which then predicts the corresponding water solubility of the input structure. Note that, an experimental Log Kow value (if known and validity of the study results has been justified) can also be used in the absence of a reliably predicted Log Kow. Depending on the mode of action and the structural profile of the input chemical, it will be allocated to one of the following local HA-QSAR solubility models which were created to provide, with a high accuracy, water solubility predictions: a) iSafeRat® Local solubility HA-QSAR for Alkanes b) iSafeRat®General Solubility HA-QSAR for MOA1 substances c) iSafeRat®Local solubility HA-QSAR for Ethers, Esters, Aldehydes and Ketones d) iSafeRat® Local solubility HA-QSAR for Alcohols e) iSafeRat®Local solubility HA-QSAR for Acids
In case, the input structure doesn't fit to any of the above mentioned solubility models, its solubility prediction may not be feasible and it will therefore be considered beyond the scope (chemical domain) of that model. On the other hand, if the chemical fits within the chemical domain of one the above mentioned models, it will then be examined to fall within the descriptor domain of that model. If it doesn't, the structure cannot qualify for a reliable prediction and will be excluded from the applicability domain of the solubility model.
4.2.1c Acute Aquatic toxicity Prediction using Regression based models
The final phase of this approach correlates the solubility of a substance to the acute toxicity it may cause at different trophic levels. The toxicity prediction was based on a thorough understanding of the thermodynamic relationship between the activity and toxicity that is well discussed in the literature. Indeed, the regressions will heavily rely on the MoA and therefore, depending on different series of MoA the obtained slopes and intercepts may vary significantly. The current version of the iSafeRat Ecotox module can handle the following two categories of chemical substances: non-polar narcotics and esters.
4.2.1d iSafeRat®Mixtures Module If the query chemical given as a user input is a mixture/multi-constituent substance, the iSafeRat® holistic approach can predict the resulting mixture toxicity values at different trophic levels with high accuracy.
a) For this the composition of each mono-constituent present within the mixture should be known.b) Moreover, for each constituent, the toxicity values for each mono-constituent (calculated using iSafeRat® Ecotox module/experimentally measured) are needed.
Given these two inputs to the iSafeRat®Mixtures module, the resulting mixture toxicity values can be predicted with high

accuracy at different trophic levels. The methodology implemented within this module is presented in further details as an appendix to the study report. It should be noted that this iSafeRat® module for mixtures is not a QSAR in itself, but a series of thermodynamic-based calculations which determines the mixture toxicity when the toxicity values for individual constituents are known. The reliability in the output from this module heavily relies on the input given (mainly the reliability in ecotoxicological endpoint values for individual constituents).

4.3.Descriptors in the model:

The approach works on the thermodynamic relationship between different endpoints which allow high accuracy predictions to be achieved following the algorithm discussed in section 4.2. For Log Kow prediction, a fragment based approach was chosen which simply adds together the contribution values for relevant fragments to predict Log Kow. No descriptor based methodology was involved here. On the other hand, the water solubility model for each chemical class was developed based on a linear regression approach using the experimental Log Kow values as the sole descriptor (simple linear regression). Finally, the ecotoxicological models were based on the clear understanding of the thermodynamic relationship between ecotoxic effects and water solubility. This set of models were realised based on simple linear regression between measured ecotoxic effects against high quality measured water solubility values.

4.4.Descriptor selection:

The mechanistic understanding of the thermodynamic principles was the driving factor for descriptor selection. No other variable selection approaches were implemented. The whole idea behind a HA-QSAR is to achieve high accuracy without including inexplicable molecular descriptors and any other mechanistically unjustified complexities.

4.5.Algorithm and descriptor generation:

No specific variable selection method was applied. No statistical tool or packages were used to generate a pool of molecular descriptors. Only validated experimental study results were used (for independent and dependent variables) throughout the holistic approach.

4.6.Software name and version for descriptor generation:

As the holistic approach followed simple modelling techniques (fragment contribution method and simple linear regression), no additional packages or tools were involved. All the descriptors (independent variables) and endpoint values (dependent variables) were experimentally derived and retrieved from various literature resources including some publicly disseminated databases as well as some confidential data available within iSafeRat® database.

4.7.Chemicals/Descriptors ratio:

Table 5 (refer to the accompanying documents) provides with the Chemicals/Descriptors ratio for different iSafeRat®modules included in the holistic methodology. Note that for solubility and ecotoxicity models, Melting Point values were considered to convert the solubility of solids to their subcooled liquid solubility values. Since the ratios were quite high (equal to the number of training set compounds as the descriptor was equal to 1 in all the cases), it indicates that the models were not over fitted with a large number of descriptors.

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

5.2.1. Applicability domain of iSafeRat ® Log Kow module:

The iSafeRat® model follows a cascade scheme including a series of predictive models. The iSafeRat LogKow module predicts the LogKow which is then given as input to the following model in the series (ie, water solubility model) and so on. Since multiple interrelated models are involved, each model has to be validated independently. Unless any measured endpoint value has been specified by the user, the approach by default starts predicting the LogKow of the query substance. The Log Kow prediction is carried out using a fragment-based approach as discussed earlier. Since the prediction involves calculating contribution values for all the fragments represented in the query substance, if one or more fragments cannot be covered, the prediction will not be feasible and the substance will be rendered as outside the applicability domain of iSafeRat® Log Kow module. In this case, the experimentally derived Log Kow can be used (if available) to proceed to the next module (provided the study was considered as valid).

5.2.2. Applicability domain of iSafeRat ® Water solubility module:

Once a reliable Log Kow prediction was derived (from section 5.2.1), it is then given as an input to a simple linear regression-based model to predict water solubility of the query substance. The test substance is verified to be within the chemical AD of the model (e.g. if it is a non-polar narcotic, ester, aldehyde or acid). In case, the substance doesn't fall into any of these categories of chemicals represented in the training set, it will be considered to be outside the AD. Next, the chemical is verified to fall within the descriptor domain of the model. For this, the Log Kow of the test substance is checked to fall within the descriptor range defined for the training set (refer to Table 8). In conclusion, the substance beyond the scope of the model (outside the chemical or descriptor domain) will be considered to fall outside the applicability domain of the iSafeRat ® WatSol module.

5.2.3. Applicability domain of iSafeRat ®

Ecotoxicology module: Finally to complete the cascade approach, the predicted water solubility values (from section 5.2.2) are then given input to iSafeRat® ecotoxicology module to predict the toxicity values at different trophic levels. Following the Simple Linear Regression approach in the same way as the Water Solubility module, the iSafeRat ®
Ecotoxicology module follows the above discussed strategy (see 5.2.2) to ensure that the test substance falls within the chemical and descriptor domain of the model. The descriptor domain within which reliable

predictions can be made with this set of models was further evaluated by approximating the descriptor (solubility value in this case) cut-off value, beyond which the accuracy is not enough to derive a reliable prediction. This exercise was simplified taking into account the thermodynamic relationship between the activity and toxicity values. Since the cut-off descriptor values were approximated, they were then verified by implementing these models on an external test set. As expected, the external substances predicted beyond the approximated solubility cut-off values were associated with a higher prediction error. This further justified the derived domain for reliable prediction with this iSafeRat ® module.

5.2.Method used to assess the applicability domain:

Consensus Applicability Domain strategy:

The cascade approach involves several interrelated models and all these models are based on different modelling approaches for instance, fragment-based method (in case of Log Kow) and simple linear regression (in case of water solubility and ecotoxicological endpoints). Since the outcome of one model is the input for the next model in the series, it has to be validated in order to make sure that unreliable predictions were filtered out and were not considered further in the workflow adding to the uncertainty in further predictions. As a result, a consensus Applicability Domain (AD) approach was adopted. For this, each of the three modules within the cascade approach was individually assessed for the AD check and only those substances were considered within the AD of the model which satisfied the reliability criteria for all the three modules. Table 6 provides an overview of this decision rule to define the AD.

Table 6 (refer to the accompanying documents): An overview of the Applicability Domain of the iSafeRat ® cascade approach.
The table provides a schematic representation of the consensus AD decision. To summarise, a substance falling outside the applicability domain of one or more modules is not considered as a reliable under the holistic approach. This decision rule makes the methodology quite conservative in terms of the AD; however this is crucial in order to retain only quality predictions as reliable which are the ultimate aim of this exercise.

It is essential to specify that a substance falling within the AD of one or more but not all the modules will be considered outside the consensus AD of the cascade approach, however, it may still be reliably predicted for modules for which it satisfies the AD criteria. For example, Test substance 4 in Table 6 can be reliably predicted for PhysChem modules. Falling inside the AD of different modules is therefore an independent event and one module cannot have any impact on the other to disqualify the substance from falling within their AD.

5.3.Software name and version for applicability domain assessment:

iSafeRat® HA-QSAR toolbox v1.1.

5.4.Limits of applicability:

5.4.1. For the Log Kow prediction:

Contribution values of the following 30 structural fragments (reported in Table 7, refer to the accompanying documents) were identified. Substances containing fragments beyond this list are considered as extrapolations.

5.4.2. For water solubility prediction:

Table 8 (refer to the accompanying documents) provides an overview of the limits of applicability for water solubility predictions using iSafeRat® holistic model.

5.4.3. For Ecotoxicity predictions:

Table 9 (refer to the accompanying documents) provides an overview of the limits of applicability for ecotox predictions using iSafeRat®holistic model.

6.Internal validation - OECD Principle 4

6.1.Availability of the training set:

No

6.2.Available information for the training set:

CAS RN: No

Chemical Name: No

Smiles: No

Formula: No

INChI: No

MOL file: No

6.3.Data for each descriptor variable for the training set:

No

6.4.Data for the dependent variable for the training set:

No

6.5.Other information about the training set:

All the predictive models integrated into the iSafeRat® Holistic HA-QSAR are proprietary. For any further information or queries about the model or its validity, contact KREATiS SAS.

6.6.Pre-processing of data before modelling:

All the descriptor and endpoint values were converted to their log units for modelling purposes. The solubility values for training compounds in solid state were converted to their corresponding sub-cooled liquid solubility values taking into account the Melting Point as an additional parameter.

6.7.Statistics for goodness-of-fit:

Since the Log Kow predictions were based on a fragment-based approach, each fragment used a specific (local) training set to derive its Log Kow contribution value. The number of training set substances (n) used for all the local models are specified in Table 7 (refer to the accompanying

documents).

For the rest of the iSafeRat®modules based on linear regression, Table 10 (refer to the accompanying documents) provides with the correlation coefficient (R^2) and the Root Mean Squared Error in calculations (RMSE). The closer the $R^2$is to 1 and lower the RMSE values, the better is the goodness-of-fit for the model. Low RMSE values indicate lower errors in calculation/prediction of the training set compounds.

6.8.Robustness - Statistics obtained by leave-one-out cross-validation:

Leave-one-out cross validation is carried out such that each training compound is excluded once from the training set. The excluded compound is then predicted using the model. The derived predictions are then used to calculate the Q^2LOO and SDEP parameters. Ideally, the Q^2LOO and SDEP values approach 1 and 0, respectively. SDEP is similar to RMSE therefore, the lower its value, the better the prediction. The leave-one-out cross validation was carried out for all the regression based models and their resulting statistics are reported in Table 11 (refer to the accompanying documents).

6.9.Robustness - Statistics obtained by leave-many-out cross-validation:

Leave-many-out cross validation is similar to the leave-one-out approach; however multiple training substances are excluded from the training set and the remaining training set is used to predict the excluded substances. Based on the derived predictions, Q^2LMO and RMSE parameters are derived. For this QMRF, the leave-many-out cross validation was carried out dividing the training set into five cross validation sets. Table 12 (refer to the accompanying documents) reports the resulting statistical parameters.

6.10.Robustness - Statistics obtained by Y-scrambling:

This validation approach indicates that the descriptor and response values had no chance correlation. The results provided in Table 13 (refer to the accompanying documents) were derived performing a Y-scrambling validation on different iSafeRat®modules with 500 iterations. The lower values for Y-scrambling parameters throughout the table further justified the stability of the models.

6.11.Robustness - Statistics obtained by bootstrap:

Bootstrap is another form of internal validation in which a selected subset of the original training set forms the model. Each of the samples of this subset are present in repeated number of times in the model such that the newly created model is of the same size as of the original training set and the excluded training samples form the validation set. This process is repeated in several iterations and predictions for the test set are recorded each time. Based on the derived predictions, Q^2boot is calculated. The results provided in Table 14 were derived performing a bootstrap validation on different iSafeRat®modules with 3000 iterations.

6.12.Robustness - Statistics obtained by other methods:

7.External validation - OECD Principle 4

7.1.Availability of the external validation set:

No

7.2.Available information for the external validation set:

CAS RN: No

Chemical Name: No

Smiles: No

Formula: No

INChI: No

MOL file: No

7.3.Data for each descriptor variable for the external validation set:

No

7.4.Data for the dependent variable for the external validation set:

No

7.5.Other information about the external validation set:

The validation set for the iSafeRat® holistic is not publicly available. However, if required it may be provided (except for any confidential data) under certain conditions by contacting KREATiS directly.

7.6.Experimental design of test set:

Test set selection had been a crucial exercise to demonstrate the validity of the presented cascade approach. In order to derive the test set, we collected experimental data for several substances (not present within the training set. The measured values were retrieved from one of the following publicly available data resources:

a) Available experimental results

b) ECHA dissemination database

c) Data from KREATiS inventory

The test set information was subjected to the following data validation procedure. It was made sure that the measured values were derived under appropriate experimental conditions and using suitable methods.

To qualify as a test set compound, the following verification checks were made:

a) The substance should not be a part of the training set for any of the predictive HA-QSARs included in the cascade approach and must have experimentally derived values for all the five endpoints. An unbiased and complete validation of this cascade approach requires a test set for which the available experimental data should cover all the involved endpoints (as of now, two physicochemical and three acute aquatic toxicity endpoints). b) The chemical groups included in the iSafeRat® training set were identified and listed earlier in this report. One of the essential bits finalizing the test set was that each substance in it is falling within the chemical domain of the training set. The test set substances also covered the majority of the chemical groups represented

within the chemical domain of the training set (this is to avoid any bias and making sure that one chemical group, for instance alcohols, alkanes or alkenes did not dominate the training set). Table 7 provides a list of all the fragments for which the contribution values are available. c) Each test set substance was also verified to fall within the descriptor range of each model involved in the cascade approach. For Log Kow, all the fragments associated with a test substance were verified to have an iSafeRat contribution value. In the case of water solubility and acute aquatic toxicity endpoints, the input descriptor values were made sure to fall within the limits reported in Tables 8 and 9. After evaluating the availability of measured data for all the five endpoints and making an applicability domain check, only 20 test set substances qualified to constitute the test set for the model.

7.7.Predictivity - Statistics obtained by external validation:

In external validation, a new set of substances unknown to the training set is used as the validation/test set. All these substances are predicted using the model. The derived predictions are then used to calculate the $Q^2$and RMSEP. In theory, Q^2values closer to 1 and RMSEP closer to 0 indicates that the model is associated with a reliable predictivity. Table 15 provides the results derived applying the iSafeRat® modules on the validation set with 20 substances. No $Q^2$was provided for Log Kow module as different local training sets were used to derive the contribution values for various fragments. The 20 validation substances were divided into three local solubility models for their predictions.

7.8.Predictivity - Assessment of the external validation set:

The criterion to have experimentally derived values for all five endpoints was crucial to justify the validity of the holistic approach in its entirety. Moreover, care was taken that the test set covered the chemical domain of the training data reasonably well to avoid any possible bias in the resulting statistical validation.

7.9.Comments on the external validation of the model:

The validation set will be extended from time to time and the revised validation results will be presented as an updated version of this QMRF.

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

The thermodynamic relationship between surrogates for chemical activity, such as solubility, log Kow and narcosis has been widely reported in the literature (Mackay et al., 2009) but only recently demonstrated to be reliable as a method which can be applied to accurately predict endpoint values for standard regulatory guideline studies (ECETOC, 2014). The holistic approach is advantageous as it validates the predictions using a wider scope than just one experimental method which may be inherently subject to variability. With this method all the parameters are examined simultaneously and the overall validity of the approach is justified.

These methods are now being applied to specific groups of substances that do not demonstrate baseline toxicity but have also been found to provide QSAR relationships which were determined statistically as valid.

8.2.A priori or a posteriori mechanistic interpretation:

The iSafeRat®Holistic HA-QSAR combines several predictive models that were developed on a clear understanding of the thermodynamic principles. The Log Kow prediction was based on the fragment-based approach which had been commonly implemented in the literature. The relationships between Log Kow and water solubility as well as between solubility and aquatic toxicity endpoints has been well discussed in the literature based on the understanding of thermodynamics.On one hand, based on the thermodynamic relationship between different endpoints, the modelling approach and the strategy of combining multiple iSafeRat®modules were planned*a priori*, while on the other hand, the validation results presented in this report (internal and external validation) helped justifying the validity of this methodology.

8.3.Other information about the mechanistic interpretation:

To allow a better understanding of the methodology, no inexplicable molecular descriptors or modelling algorithms were included. Since multiple predictive modules were included, the validity of each module was justified independently and holistically.

9.Miscellaneous information

9.1.Comments:

This QMRF can be used as a reference document for QPRFs providing iSafeRat® predictions for one or more of the five endpoints relevant to the holistic approach.

9.2.Bibliography:

[1]MacKay D, Arnot J, Petkova E, Wallace K, Call D, Brooke L & Veith G (2009). SAR and QSAR in Environmental Research 20 (3-4) 393-414.
http://www.tandfonline.com/doi/full/10.1080/10629360902949153

[2]ECETOC Technical Report no.120. Activity-Based Relationships for Aquatic Ecotoxicology Data: Use of the Activity Approach to Strengthen MoA Predictions. 13 January 2014
http://www.ecetoc.org/index.php?mact=MCSoap,cntnt01,details,0&cntnt01by_category=5&cntnt01template=display_list_v2&cntnt01order_by=Reference%20Desc&cntnt01display_template=display_details_v2&cntnt01document_id=8303&cntnt01returnid=89

9.3.Supporting information:

Tables for QMRF	http://qsardb.jrc.it:80/qmrf/download_attachment.jsp?name=qmrf448_Tables for QMRF Q19-46-41-422 (revised v1.3).pdf

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

10.2.Publication date:

10.3.Keywords:

iSafeRat;HA-QSAR;holistic;physicochemical;Log Kow;Solubility;algae;fish;Desmodesmus subspicatus;Pseudokirchneriella subcapitata;Scenedesmus quadricauda;Daphnia magna;Daphnia pulex;Danio rerio;Lepomis macrochirus;Pimephales promelas;Oncorhynchus mykiss;Oryzias latipes;;

10.4.Comments:

former Q19-46-41-422