Leadscope Enterprise model for maximum recommended daily dose (MRDD) in humans

1. QSAR identifier

1.1 QSAR identifier (title)

Leadscope Enterprise model for maximum recommended daily dose (MRDD) in humans, Danish QSAR Group at DTU Food.

1.2 Other related models

MultiCASE CASE Ultra model for maximum recommended daily dose (MRDD) in humans, Danish QSAR Group at DTU Food.

SciMatics SciQSAR model for maximum recommended daily dose (MRDD) in humans, Danish QSAR Group at DTU Food.

1.3. Software coding the model

Leadscope Predictive Data Miner, a component of Leadscope Enterprise version 3.1.1-10.

2. General information

2.1 Date of QMRF

January 2015.

2.2 QMRF author(s) and contact details

QSAR Group at DTU Food;

Danish National Food Institute at the Technical University of Denmark;

http://qsar.food.dtu.dk/;

qsar@food.dtu.dk

Sine Abildgaard Rosenberg;

National Food Institute at the Technical University of Denmark;

Trine Klein Reffstrup;

National Food Institute at the Technical University of Denmark;

Eva Bay Wedebye;

National Food Institute at the Technical University of Denmark;

Nikolai Georgiev Nikolov;

National Food Institute at the Technical University of Denmark;

Marianne Dybdahl;

National Food Institute at the Technical University of Denmark.

2.3 Date of QMRF update(s)

2.4 QMRF update(s)

2.5 Model developer(s) and contact details

Jay Russel Niemelä;

National Food Institute at the Technical University of Denmark;

Eva Bay Wedebye;

National Food Institute at the Technical University of Denmark;

ebawe@food.dtu.dk

Nikolai Georgiev Nikolov;

National Food Institute at the Technical University of Denmark;

nign@food.dtu.dk

Danish QSAR Group at DTU Food;

National Food Institute at the Technical University of Denmark;

http://qsar.food.dtu.dk/;

qsar@food.dtu.dk

2.6 Date of model development and/or publication

January 2014.

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

Roberts, G., Myatt, G. J., Johnson, W. P., Cross, K. P., and Blower, P. E. J. (2000) LeadScope: Software for Exploring Large Sets of Screening Data. *Chem. Inf. Comput. Sci.*, 40, 1302-1314.

Cross, K.P., Myatt, G., Yang, C., Fligner, M.A., Verducci, J.S., and Blower, P.E. Jr. (2003) Finding Discriminating Structural Features by Reassembling Common Building Blocks. *J. Med. Chem.*, *46,* 4770-4775.

Valerio, L. G., Yang, C., Arvidson, K. B., and Kruhlak, N. L. (2010) A structural feature-based computational approach for toxicology predictions. *Expert Opin. Drug Metab. Toxicol.*, 6:4, 505-518.

2.8 Availability of information about the model

The training set is non-proprietary and was compiled from the Maximum Recommended Daily Dose (MRDD) Database which is publically available at the FDA/CDER Webpage (http://www.epa.gov/comptox/dsstox/sdf_fdamdd.html, accessed 9th of July 2013). The model algorithm is proprietary from commercial software.

2.9 Availability of another QMRF for exactly the same model

3. Defining the endpoint

3.1 Species

Human (phase 1 clinical trial).

3.2 Endpoint

QMRF 4. Human Health Effects

QMRF 4.14. Repeated dose toxicity

3.3 Comment on endpoint

The Maximum Recommended Daily Dose (MRDD) for a pharmaceutical is an estimated upper dose limit beyond which a drug's efficacy is not increased and/or undesirable adverse effects begin to outweigh beneficial effects. The MRDD is related to the No Observed Effect Level (NOEL) for non-pharmaceuticals (NOEL equals 1/10 MRDD), a dose at which there are no statistically or biologically significant increases in the frequency or severity of any effect between the exposed population and its appropriate control. Because of this relation this model can be used to estimate both the MRDD and NOEL values for a given compound.

Data for this model was compiled from FDA's Center for Drug Evaluation and Research, Office of Pharmaceutical Science, Informatics and Computational Safety Analysis Staff's Maximum Recommended Daily Dose (FDAMDD) database. Most of the MRDD values in the FDAMDD database were determined from pharmaceutical phase 1 human clinical trials that employed an oral route of exposure and daily treatments, usually for 3 - 12 months. The pharmaceuticals were given as single or divided dose treatment regimens to achieve desired pharmacological effects. In contrast, roughly 5% of the pharmaceuticals in the FDAMDD database were anti-neoplastics and anesthetics and these were administered intravenously and/or intramuscularly. When separate MRDDs were reported for different routes of exposure, only the oral MRDD was included in the database. In addition, some pharmaceuticals have different MRDD values for male and female adults, children, or elderly patients. In this situation only MRDD values for the average adult patient were used.

Pharmaceuticals that are administered orally are usually tested over a limited range of doses and have MRDDs reported as mg/day. The MRDDs were converted from the mg/day unit to mg/kg body weight (bw)/day based upon an average adult weighing 60 kg. In contrast, the dose unit for most antineoplastic drug MRDDs is reported as mg/m^2 which was converted to mg/kg bw/day using the formula mg/kg bw/day = $mg/m^2/37$ for an average adult. Additionally, a few drugs had MRDDs reported in parts per million (ppm) which were converted to mg/kg bw/day on the basis that 1000 ppm equals 25 mg/kg bw/day for an average 60 kg adult. MRDD values for the 1,220 chemicals in this training set range from 0.00001 to 1000 mg/kg bw/day (Matthews *et al.* 2004).

As data for this model is derived directly from human data it can be argued that the model predictions can give a more accurate estimate of human MRDD than data derived from repeat-dose tests in rodents.

To make a categorical model compounds with a MRDD value between 0.0167-2.69 mg/kg bw/day were defined as positive and compounds with MRDD values between 5.00-1000 mg/kg bw/day were defined as negative. Intermediate compounds were defined as marginal.

3.4 Endpoint units

No units, 1 for positives and 0 for negatives.

3.5 Dependent variable

Maximum recommended daily dose (MRDD) in humans, positive or negative.

3.6 Experimental protocol

Data originate from pharmaceutical phase 1 human clinical trials that employed an oral route of exposure and daily treatments, usually for 3 - 12 months. The pharmaceuticals were given as single or divided dose treatment regimens to achieve desired pharmacological effects. In contrast, roughly 5% of the pharmaceuticals in the FDAMDD database were anti-neoplastics and anesthetics and these were administered intravenously and/or intramuscularly (Matthews *et al.* 2004).

3.7 Endpoint data quality and variability

According to (http://www.epa.gov/comptox/dsstox/sdf_fdamdd.html) "Several features of DSSTox FDAMDD have the potential to impact on SAR analysis and should be taken into account in any future use of these data. Most prominent among these is the imprecise nature of the reported MRDD value, both in terms of the wide range of adverse or toxic effects that would be considered in assigning the MRDD, and in terms of the ambiguous chemical structure association with this dose measure. In DSSTox FDAMDD and the corresponding Source FDA MRDD database, there are several cases where a single Dose_MRDD_mg value is assigned to multiple related structural derivatives of a pharmaceutical, i.e., the same activity is assigned to multiple Structure/CASRN records in the database. In theory, an MRDD value will reflect the lowest dose of a drug producing adverse effects but for the FDA MRDD database this value has been derived from pooled clinical reports where more than one form of a drug may have been administered. When MRDD mg mass units are converted to mmol units for SAR analysis, a single Dose_MRDD_mg is converted to a range of mmol doses, taking into account the different molecular weights of the various drug derivatives. Assuming that these various drug derivatives have similar or equal molar potencies, the reported Dose_MRDD_mg could be presumed to reflect the dose of the smallest STRUCTURE_MolecularWeight derivative that would register as the highest molar content and, therefore, most potent for a given mass dose."

4. Defining the algorithm

4.1 Type of model

A categorical (Q)SAR model based on structural features and numeric molecular descriptors.

4.2 Explicit algorithm

This is a categorical (Q)SAR model made by use of partial logistic regression (PLR). The specific implementation is proprietary within the Leadscope software.

4.3 Descriptors in the model

structural features,

aLogP,

polar surface area,

number of hydrogen bond donors,

Lipinski score,

number of rotational bonds,

parent atom count,

parent molecular weight,

number of hydrogen bond acceptors

4.4 Descriptor selection

Leadscope Predictive Data Miner is a software program for systematic sub-structural analysis of a chemical using predefined structural features stored in a template library, training set-dependent generated structural features (scaffolds) and calculated molecular descriptors. The feature library contains approximately 27,000 pre-defined structural features and the structural features chosen for the library are motivated by those typically found in small molecules: aromatics, heterocycles, spacer groups, simple substituents. Leadscope allows for the generation of training set-dependent structural features (scaffold generation), and these features can be added to the pre-defined structural features from the library and be included in the descriptor selection process. It is possible in Leadscope to remove redundant structural features before the descriptor selection process and only use the remaining features in the descriptor selection process. Besides the structural features Leadscope also calculates eight molecular descriptors for each training set structure: the octanol/water partition coefficient (alogP), hydrogen bond acceptors (HBA), hydrogen bond donors (HBD), Lipinski score, atom count, parent compound molecular weight, polar surface area (PSA) and rotatable bonds. These eight molecular descriptors are also included in the descriptor selection process.

Leadscope has a default automatic descriptor selection procedure. This procedure selects the top 30% of the descriptors (structural features and molecular descriptors) according to X^2-test for a binary variable, or

the top and bottom 15% descriptors according to *t*-test for a continuous variable. Leadscope treats numeric property data as ordinal categorical data. If the input data is continuous such as IC_{50} or cLogP data, the user can determine how values are assigned to categories: the number of categories and the cut-off values between categories. (Roberts *et al.*2000).

When developing this model, intermediate models with application of different modelling approaches in Leadscope were tried:

1. 'Single model' using only the Leadscope pre-defined structural features, i.e. no scaffolds, and calculated molecular descriptors for descriptor selection.
2. 'Single model' using both the Leadscope pre-defined structural features and the training set dependent features (scaffolds generation) as well as the calculated molecular descriptors in the descriptor selection.
3. 'Single model' using both Leadscope pre-defined structural features and the training set dependent features (scaffolds generation), with subsequent removal of redundant structural features, and calculated molecular descriptors for descriptor selection.
4. 'Composite model' using only the Leadscope pre-defined structural features, i.e. no scaffolds, and calculated molecular descriptors in the descriptor selection.
5. 'Composite model' using both Leadscope pre-defined structural features and the training set dependent features (scaffolds generation) as well as the calculated molecular descriptors in the descriptor selection.

Based on model performance as measured by a preliminary cross-validation the model developed using approach number 2. was chosen.

For this model scaffolds were generated by Leadscope for the training set structures and added to the Leadscope library of structural features. Descriptors were then automatically selected among the structural features and the eight molecular descriptors.

4.5 Algorithm and descriptor generation

For descriptor generation see 4.4.

After selection of descriptors the Leadscope Predictive Data Miner program performs partial least squares (PLS) regression for a continuous response variable, or partial logistic regression (PLR) for a binary response variable, to build a predictive model. By default the Predictive Data Miner performs leave-one-out or leave-groups-out (in the latter case, the user can specify any number of repetitions and percentage of structures left out) cross-validation on the training set depending on the size of the training set. In the cross-validation made by Leadscope the descriptors selected for the 'mother model' are used when building the validation submodels and they therefore have a tendency to be overfittet and give overoptimistic validation results.

In this model, because of the categorical outcome in the response variable, PLR was used to build the predictive model. For this model 279 descriptors were selected to build the model. These include 8 Leadscope calculated molecular descriptors, 155 hierarchy features, and 116 scaffolds. The 279 descriptors were distributed on 6 PLS factors.

4.6 Software name and version for descriptor generation

Leadscope Predictive Data Miner, a component of Leadscope Enterprise version 3.1.1-10.

4.7 Descriptors/chemicals ratio

In this model 279 descriptors were used and distributed on 6 PLS factors. The training set consists of 1106 compounds. The descriptor/chemical ratio is 1:4.0 (279:1106).

5. Defining Applicability Domain

5.1 Description of the applicability domain of the model

The definition of the applicability domain consists of two components; the definition of a structural domain in Leadscope and the in-house further probability refinement algorithm on the output from Leadscope to reach the final applicability domain call.

1. Leadscope

For assessing if a test compound is within the structural applicability domain of a given model Leadscope examines whether the test compound bears enough structural resemblance to the training set compounds used for building the model (i.e. a structural domain analysis). This is done by calculating the distance between the test compound and all compounds in the training set (distance = 1 - similarity). The similarity score is based on the Tanimoto method. The number of neighbours is defined as the number of compounds in the training set that have a distance equal to or smaller than 0.7 with respect to the test compound. The higher the number of neighbours, the more reliable the prediction for the test compound. Statistics of the distances are also calculated. Effectively no predictions are made for test compounds which are not within the structural domain of the model or for which the molecular descriptors could not be calculated in Leadscope.

2. The Danish QSAR group

In addition to the general Leadscope structural applicability domain definition the Danish QSAR group has applied a further requirement to the applicability domain of the model. That is only positive predictions with a probability equal to or greater than 0.7 and negative predictions with probability equal to or less than 0.3 are accepted. Predictions within the structural applicability domain but with probability between 0.5 to 0.7 or 0.3 to 0.5 are defined as positives out of applicability domain and negatives out of applicability domain, respectively. When these predictions are weeded out the performance of the model in general increases at the expense of reduced model coverage.

5.2 Method used to assess the applicability domain

Leadscope does not generate predictions for test compounds which are not within the structural domain of the model or for which the molecular descriptors could not be calculated.

Only positive predictions with probability equal to or greater than 0.7 and negative predictions with probability equal to or less than 0.3 are accepted.

5.3 Software name and version for applicability domain assessment

Leadscope Predictive Data Miner, a component of Leadscope Enterprise version 3.1.1-10.

5.4 Limits of applicability

The Danish QSAR group applies an overall definition of structures acceptable for QSAR processing which is applicable for all the in-house QSAR software, i.e. not only Leadscope. According to this definition accepted structures are organic substances with an unambiguous structure, i.e. so-called discrete organics defined as: organic compounds with a defined two dimensional (2D) structure containing at least two carbon atoms, only certain atoms (H, Li, B, C, N, O, F, Na, Mg, Si, P, S, Cl, K, Ca, Br, and I), and not mixtures with two or more ‘big components’ when analyzed for ionic bonds (for a number of small known organic ions assumed not to affect toxicity the ‘parent molecule’ is accepted). Calculation 2D structures (SMILES and/or

SDF) are generated by stripping off ions (of the accepted list given above). Thus, all the training set and prediction set chemicals are used in their non-ionized form. See 5.1 for further applicability domain definition.

6. Internal validation

6.1 Availability of the training set

Yes

6.2 Available information for the training set

Yes

6.3 Data for each descriptor variable for the training set

No

6.4 Data for the dependent variable for the training set

Yes

6.5 Other information about the training set

1106 compounds are in the training set: 524 positives and 582 negatives.

6.6 Pre-processing of data before modelling

Data was originally collected from the FDAMDD database. Only compounds for which SMILES codes could be found and that had a structure acceptable for the commercial software were used in the final training set, that is only discrete organic chemicals as described in 5.4 were used. In case of replicate structures, one of the replicates was kept if all the compounds had the same activity and all were removed if they had different activity.

6.7 Statistics for goodness-of-fit

Not performed.

6.8 Robustness – Statistics obtained by leave-one-out cross-validation

Not performed. (It is not a preferred measurement for evaluating large models).

6.9 Robustness – Statistics obtained by leave-many-out cross-validation

A five times two-fold 50 % cross-validation was performed. This was done by randomly removing 50% of the full training set used to make the “mother model”, where the 50% contains the same ratio of positive and negatives as the full training set. A new model (validation submodel) was created on the remaining 50% using the same settings in Leadscope but with no information from the “mother model” regarding descriptor selection etc. The validation submodel was applied to predict the removed 50% (within the defined applicability domain for the submodel). Likewise, a validation submodel was made on the removed 50% of the training set and this model was used to predict the other 50% (within the defined applicability domain for this submodel). This procedure was repeated five times.

Predictions within the defined applicability domain for the ten validation submodels were pooled and Cooper's statistics calculated. This gave the following results for the 52.0% (3179/(5*1222)) of the predictions which were within the applicability domain:

- Sensitivity (true positives / (true positives + false negatives)): 78.6%
- Specificity (true negatives / (true negatives + false positives)): 82.5%
- Concordance ((true positives + true negatives) / (true positives + true negatives + false positives + false negatives)): 80.7%

6.10 Robustness - Statistics obtained by Y-scrambling

Not performed.

6.11 Robustness - Statistics obtained by bootstrap

Not performed.

6.12 Robustness - Statistics obtained by other methods

Not performed.

7. External validation

7.1 Availability of the external validation set

7.2 Available information for the external validation set

7.3 Data for each descriptor variable for the external validation set

7.4 Data for the dependent variable for the external validation set

7.5 Other information about the validation set

7.6 Experimental design of test set

7.7 Predictivity – Statistics obtained by external validation

7.8 Predictivity – Assessment of the external validation set

7.9 Comments on the external validation of the model

External validation not performed for this model.

8. Mechanistic interpretation

8.1 Mechanistic basis of the model

The global model identifies structural features and molecular descriptors which in the model development was found to be statistically significant associated with effect. Many predictions may indicate modes of action that are obvious for persons with expert knowledge for the endpoint.

8.2 A priori or posteriori mechanistic interpretation

A posteriori mechanistic interpretation. The identified structural features and molecular descriptors may provide basis for mechanistic interpretation.

8.3 Other information about the mechanistic interpretation

9. Miscellaneous information

9.1 Comments

The model can be used to predict the human MRDD in a categorical way: Positive means MRDD value between 0.0167-2.69 mg/kg bw/day, negative means MRDD values between 5.00-1000 mg/kg bw/day and marginal means intermediate in between 2.69 and 5.00 mg/kg bw/day. It can be argued that the predictions from this (Q)SAR model give a more accurate estimate of human MRDD/NOEL than those derived from animal toxicity studies, where multiple uncertainty/safety factors are necessary to compensate for incompatibility and uncertainty underlying the extrapolation of animal toxicity to humans.

9.2 Bibliography

Matthews, E.J, Kruhlak, N.L, Benz, R.D and Contrera, J.F. (2004) Assessment of the Health Effects of Chemicals in Humans: I. QSAR Estimation of the Maximum Recommended Daily Dose (MRDD) and No Effect Level (NOEL) of Organic Chemicals Based on Clinical Trial Data. *Current drug Discovery Technologies,* 1, 61-76.

Maximum recommended daily dose (MRDD) Database: http://www.epa.gov/comptox/dsstox/sdf_fdamdd.html. Accessed 9th of July 2013. According to which the data originates from:

- Blacow, N. W. *Martindale: The Extra Pharmacopoeia, 26th ed.*; The Pharmaceutical Press: London, **1972**.
- Wade, A. *Martindale: The Extra Pharmacopoeia, 28th ed.* ; The Pharmaceutical Press: London, **1982**.
- Reynolds, J. E. F. *Martindale: The Extra Pharmacopoeia, 30th ed.;* The Pharmaceutical Press: London, **1993**.
- Sweetman, S. C. *Martindale: The Complete Drug Reference*; Ed.; The Pharmaceutical Press: London. Electronic version, Thomson
- MICROMEDEX: Greenwood Village, CO. Edition expires **2003**.
- Arky, R. *The Physicians' Desk Reference, 49th ed.*; Medical Economics Company: New Jersey, **1995**.

Arky, R. *The Physicians' Desk Reference, 53rd ed.*; Medical Economics Company: New Jersey, **1999**.

9.3 Supporting information