	QMRF identifier (JRC Inventory): **To be entered by JRC**
	QMRF Title: **Derek Nexus - HERG (human Ether-à-go-go-Related Gene) channel inhibition**
	Printing Date: **22 May 2020**

1.QSAR identifier

1.1.QSAR identifier (title):

Derek Nexus - HERG (human Ether-à-go-go-Related Gene) channel inhibition

1.2.Other related models:

Derek Nexus contains alerts for multiple endpoints, including mutagenicity, chromosome damage, carcinogenicity, hepatotoxicity, teratogenicity and skin irritation

1.3.Software coding the model:

Derek Nexus v6.1 contains 5 alerts for HERG channel inhibition, together with reasoning rules encoding physicochemical descriptors.

2.General information

2.1.Date of QMRF:

7 January 2016

2.2.QMRF author(s) and contact details:

Lilia Fisk Lhasa Limited Granary Wharf House, 2 Canal Wharf, Leeds, LS11 5PS, UK

2.3.Date of QMRF update(s):

22 May 2020

2.4.QMRF update(s):

Rachael Tennant, Lhasa Limited, Granary Wharf House, 2 Canal Wharf, Leeds, LS11 5PS, UK

1.3, 2.3, 2.4, 2.6, 3.4, 4.6

2.5.Model developer(s) and contact details:

Lhasa Limited Granary Wharf House, 2 Canal Wharf, Leeds, LS11 5PS

2.6.Date of model development and/or publication:

Derek Nexus 6.1 was released on 11 June 2020

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

[1]Sanderson DM & Earnshaw CG (1991). Computer prediction of possible toxic action from chemical structure; The DEREK system. Human and Experimental Toxicology 10, 261-273.

[2]Judson PN, Marchant CA & Vessey JD (2003). Using argumentation for absolute reasoning about the potential toxicity of chemicals. Journal of Chemical Information and Computer Sciences 43, 1364-1370.

[3]Marchant CA, Briggs KA & Long A (2003). In silico tools for sharing data and knowledge on toxicity and metabolism: Derek for Windows, Meteor, and Vitic. Toxicology Mechanisms and Methods 18, 177–187.

[4]Judson PN, Stalford SA & Vessey J (2013). Assessing confidence in predictions made by knowledge-based systems. Toxicology Research 2, 70-79.

2.8.Availability of information about the model:

Derek Nexus is a proprietary, rule-based expert system for the prediction of toxicity. Its knowledge base is composed of alerts, examples and reasoning rules which may each contribute to the

predictions made by the system. Each alert in Derek describes a chemical substructure believed to be responsible for inducing a specific toxicological outcome (often referred to as a toxicophore). Alerts are derived by experts, using toxicological data and information regarding the biological mechanism of action. Where relevant, metabolism data may be incorporated into an alert, enabling the prediction of compounds which are not directly toxicity but are metabolised to an active species. The derivation of each alert is described in the alert comments along with supporting references and example compounds where possible. By reporting this information to the user, Derek provides highly transparent predictions. The use of structural alerts for the prediction of toxicity is both widely understood and the subject of many publications.

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

No

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Predictions are made for the domain of bacteria and can be broken down into species (e.g. mouse, human, guinea pig).

3.2.Endpoint:

3.3.Comment on endpoint:

The Derek Nexus model for HERG channel inhibition is developed from several sources of data. Primary data used for alert development are in vitro whole-cell patch clamp electrophysiology assays in a variety of cell types that either have native (AT-1) or transfected (HERG) ion channels. Positive calls are made based on IC50s cut off.

3.4.Endpoint units:

Derek Nexus makes predictions for and against toxicity through reasoning. For the endpoint of HERG channel inhibition, predictions for toxicity decrease in confidence in the following order: certain> probable>plausible>equivocal. Predictions against toxicity increase in confidence in the following order: doubted<improbable. These likelihood levels have been shown to correlate with predictivity [Judson et al, 2013]. Multiple data sources (e.g. toxicity data from multiple assays and knowledge of ligand binding interactions with the ion channel) are synthesised into the structure-activity relationships that underpins Derek Nexus predictions. Units are considered by the alert writers when building the alert training set, however, as predictions are made using data from multiple assays these do not include units as default.

3.5.Dependent variable:

Data from whole-cell patch clamp electrophysiology assays in a variety of cell types and knowledge of ligand binding interactions are synthesised to arrive at an expert conclusion of whether compounds within the model training set is likely to be a HERG channel inhibitor.

3.6.Experimental protocol:

The model is primary based on whole-cell patch clamp electrophysiology assays in a variety of cell types.

3.7.Endpoint data quality and variability:

Alert writers use all available and relevant information in the public domain (and proprietary data, where available) for alert development. Wherever possible, primary references are used as data sources: (i) the data are subject to expert assessment prior to inclusion in the alert training set and (ii) the references themselves are cited in the alert comments enabling users to conduct their own expert assessments on data quality.

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

Expert derived structural alerts for HERG channel inhibition (2D SARs), physicochemical properties and associated reasoning.

4.2.Explicit algorithm:

Structural alerts

logic of argumentation

4.3.Descriptors in the model:

[1]Markush structures encoding activating and deactivating features (known as patterns in the Derek Nexus knowledge base)

[2]ClogP

4.4.Descriptor selection:

There is an a priori assumption that patterns and associated reasoning will be used to model toxicity within Derek Nexus.

4.5.Algorithm and descriptor generation:

Alert writers design the patterns to describe the activating and deactivating features found during expert assessment of the alert training set.

4.6.Software name and version for descriptor generation:

Alert writers use the Derek Knowledge Editor (v2.0) for the implementation of patterns. ClogP predictions generated using the BioByte model (v5.9).

4.7.Chemicals/Descriptors ratio:

This is not applicable to structural alerts as these are knowledge-based rather than statistically based.

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

The scopes of the structure-activity relationships describing the HERG channel inhibition endpoint are defined by the developer to be the applicability domain for the model. Therefore, if a chemical activates an alert describing a structure-activity for HERG channel inhibition it can be considered to be within the applicability domain. If a compound

does not activate an alert or reasoning rule in Derek, a result of 'nothing to report' is presented to the user. This can be interpreted as a negative prediction or that the query compound is outside the domain of the model. Which of these is more appropriate may depend on the endpoint of interest.

5.2.Method used to assess the applicability domain:

The applicability domain of each alert is defined by the alert developer on the basis of the training set data and expert judgement on the chemical and biological factors which affect the mechanism of action for each alert.

5.3.Software name and version for applicability domain assessment:

This is not applicable.

5.4.Limits of applicability:

Limits for individual alerts are mainly defined by restrictions in the scope of the alerts which are available for inspection within the software.

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

NanoMaterial: 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:

Non-proprietary elements of the training set are available through the references, and illustrated by the examples, within Derek Nexus. The illustrative examples are not available, due to the proprietary nature of Derek Nexus.

6.6.Pre-processing of data before modelling:

This is not applicable.

6.7.Statistics for goodness-of-fit:

This is not applicable.

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

This is not applicable.

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

This is not applicable.

6.10.Robustness - Statistics obtained by Y-scrambling:

This is not applicable.

6.11.Robustness - Statistics obtained by bootstrap:

This is not applicable.

6.12.Robustness - Statistics obtained by other methods:

This is not applicable.

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

NanoMaterial: 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:

External validation is carried out on each knowledge base release. Three proprietary data sets have been used for alert validation. The data sets used for validation are not available in the public domain, so are not made available. Two proprietary data sets and one published data set (Doddareddy et al) have been used for alert validation. Further, the relationship between likelihood levels and prediction accuracy has been assessed [Judson et al 2013].

7.6.Experimental design of test set:

Proprietary data sets were sought.

7.7.Predictivity - Statistics obtained by external validation:

The software reports the number of positive and negative compounds from the validation data sets that activate each alert and calculates positive predictivity using this data.

7.8.Predictivity - Assessment of the external validation set:

Data derived from patch clamp electrophysiology assays. Published data set contains 607 and two proprietary data sets > 1000 compounds. The compounds in the data sets are primarily small and medium-sized chemicals and so are representative of the structures used to build the model.

7.9.Comments on the external validation of the model:

This is not applicable.

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

All alerts describing structure-activity relationships for the HERG channel inhibition endpoint are based on the known structure of the HERG channel and likely ligand binding interactions. This information is detailed in the alert comments.

8.2.A priori or a posteriori mechanistic interpretation:

The mechanistic basis of the model was developed a priori.

8.3.Other information about the mechanistic interpretation:

All references supporting the mechanistic basis of an alert are detailed and available for inspection within the software.

9.Miscellaneous information

9.1.Comments:

Derek Nexus may be used to assess the toxicity of a wide range of chemical classes, including food, drug, cosmetic, and industrial chemicals, and the system provides predictions for over 50 toxicological endpoints, including mutagenicity, chromosome damage, carcinogenicity, skin sensitisation and reproductive toxicity.

9.2.Bibliography:

[1]Sanderson DM & Earnshaw CG (1991). Computer prediction of possible toxic action from chemical structure; The DEREK system. Human and Experimental Toxicology 10, 261-273.

[2]Judson PN, Marchant CA & Vessey JD (2003). Using argumentation for absolute reasoning about the potential toxicity of chemicals. Journal of Chemical Information and Computer Sciences 43, 1364-1370.

[3]Marchant CA, Briggs KA & Long A (2003). In silico tools for sharing data and knowledge on toxicity and metabolism: Derek for Windows, Meteor, and Vitic. Toxicology Mechanisms and Methods 18, 177–187.

[4]Judson PN, Stalford SA & Vessey J (2013). Assessing confidence in predictions made by knowledge-based systems. Toxicology Research 2, 70-79.

[5]Doddareddy MR, Klaasse EC, Shagufta, Ijzerman AP & Bender A (2010). Prospective Validation of a Comprehensive In silico hERG Model and its Applications to Commercial Compound and Drug Databases. ChemMedChem, 5, 716–729.

9.3.Supporting information:

Training set(s)Test set(s)Supporting information

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

To be entered by JRC

10.2.Publication date:

To be entered by JRC

10.3.Keywords:

To be entered by JRC

10.4.Comments:

To be entered by JRC