	QMRF identifier (JRC Inventory): **Q13-46-0045**
	QMRF Title: **Derek for Windows - Skin sensitisation**
	Printing Date: **Dec 11, 2019**

1.QSAR identifier

1.1.QSAR identifier (title):

Derek for Windows - Skin sensitisation

1.2.Other related models:

1.3.Software coding the model:

Derek for Windows version 13

www.lhasalimited.org/derek

2.General information

2.1.Date of QMRF:

26 July 2010

2.2.QMRF author(s) and contact details:

Kate Langton Lhasa Limited 22-23 Blenheim Terrace, Woodhouse Lane, Leeds, LS2 9HD,UK

kate.langton@lhasalimited.org www.lhasalimited.org

2.3.Date of QMRF update(s):

2.4.QMRF update(s):

2.5.Model developer(s) and contact details:

Lhasa Limited 22-23 Blenheim Terrace, Woodhouse Lane, LS2 9HD, UK

kate.langton@lhasalimited.org www.lhasalimited.org

2.6.Date of model development and/or publication:

Derek for Windows version 13 was released in December 2010 and included updates to the skin sensitisation endpoint.

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

[1]D.M. Sanderson and C.G. Earnshaw; Computer Prediction of Possible Toxic Action from Chemical Structure; The DEREK System; Human and Experimental Toxicology, 1991, 10, 261-273 http://het.sagepub.com/cgi/content/abstract/10/4/261

[2]P.N. Judson, C.A. Marchant,and J.D. Vessey; Using Argumentation for Absolute Reasoning about the Potential Toxicity of Chemicals. JCICS, 2003, 43, 1364-1370 http://pubs.acs.org/doi/abs/10.1021/ci020272g

2.8.Availability of information about the model:

The model contains multiple alerts (version 13 contains 71 skin sensitisation alerts) which each describe an SAR for a specific chemical class and is augmented by a rule to consider LogKp value. The alerts are available for inspection within the software and representative examples are provided to illustrate a given alert if available. The training set underpinning a given alert is proprietary, though generally based on publicly available data. The basis of the rule considering the effect of LogKp is publicly available and referenced within the model.

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

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Mammal (mainly Guinea Pig, Mouse and Human)

3.2.Endpoint:

[1][2]QMRF 4. Human Health Effects QMRF 4. 6. Skin sensitisation

3.3.Comment on endpoint:

The model is primarily based on data from the Guinea Pig Maximisation Test (GPMT) and Local Lymph Node Assay (LLNA). Data from a number of additional assays have also been used to build the model including Human Patch Test results.

3.4.Endpoint units:

A Derek for Windows alert makes a prediction on the likelihood of a query compound causing skin sensitisation and is not restricted to a specific assay, and does not include units. Accordingly, data generated from multiple assays is used to develop an alert and an appreciation of the assay units is required when building the alert training set. The EC3 value in the LLNA is used to assign activity for skin sensitisation. Any compound with an EC3 < 100% is considered positive when developing an alert. For other assays, the author's call is generally accepted.

3.5.Dependent variable:

Data is not processed before an alert is developed, although the model predicts for skin sensitisation and the data used to develop the alerts is assay specific. A positive result in an assay (see section 3.4) is required to develop an alert, though an alert will not be built against a singe compound.

3.6.Experimental protocol:

The model is based primarily on data from Guinea Pig Maximisation Test or Local Lymph Node Assay conducted following standard test protocol (GPMT:OECD Test Guidline 406; LLNA: OECD Test Guideline 429). If activity is observed in a non-standard assay or protocol this will be mentioned in the comments.

3.7.Endpoint data quality and variability:

Alerts are developed against data generated following standard test protocols (predominantly OECD TG 406 and OECD TG 429), the data forming the basis of each alert is fully referenced within the alert. An initial dataset used to develop the endpoint (Cronin MT and Basketter DA. Multivariate QSAR analysis of a skin sensitization database. SAR and QSAR in Environmental Research, 1994, 2, 159-179) is recognised to be of high quality as the tests were run to identical protocols and generated within a single organisation. Subsequent alert (and overall endpoint) development has been against multiple sources of data from which it is hard to draw general conclusions, though all data is generated to standard protocols, and all references are available within the model.

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

Expert system based on multiple structure alerts (2D SARs).

4.2.Explicit algorithm:

Expert system

Expert system based on multiple structure alerts (2D SARs)

4.3.Descriptors in the model:

4.4.Descriptor selection:

LogP (required for the prediction of LogKp) is generated by the reasoning engine within the model using either the Moriguchi equation (Moriguchi I et al, Chemical and Pharmaceutical Bulletin, 1992, 40, 127) or the ClogP model from BioByte. The user can select which model to use or supply their own values.

LogKp is generated using the Potts and Guy equation (Potts RO and Guy RH, Pharmaceutical Research, 1992, 9, 663-669).

4.5.Algorithm and descriptor generation:

The model comprises multiple alerts and associated reasoning rules and one physicochemical-based (LogKp) reasoning rule. The alerts have been built to each describe particular chemical class(es) based on all the available data in the public domain and any donated proprietary data, though an alert will not be written against a single compound. The likelihood and scope of an alert are defined by the developer based on the data collated. An alert is triggered when a query compound matches the scope of the alert as defined by inclusion and exclusion patterns. The LogKp value for the query compound will be considered by the physicochemical-based rule and may augment the overall reasoning level observed. A negative prediction may be generated if a query compound matches no alerts but activated the physicochemical-based reasoning rule.

4.6.Software name and version for descriptor generation:

BioByte Corp

ClogP is used to calculate LogP

None identified

http://www.biobyte.com/

Moriguchi equation

A published equation for the prediction of LogP

n/a

http://www.journalarchive.jst.go.jp/jnlpdf.php?cdjournal=cpb1958&cdvol=40&noissue=1&startpage=127&lang=en&from=jnltoc

4.7.Chemicals/Descriptors ratio:

This is not applicable as the structural alerts 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:

5.2.Method used to assess the applicability domain:

5.3.Software name and version for applicability domain assessment:

5.4.Limits of applicability:

6.Internal validation - OECD Principle 4

6.1.Availability of the training set:

Yes

6.2.Available information for the training set:

CAS RN: Yes

Chemical Name: Yes

Smiles: Yes

Formula: Yes

INChI: Yes

MOL file: Yes

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

All

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

All

6.5.Other information about the training set:

6.6.Pre-processing of data before modelling:

6.7.Statistics for goodness-of-fit:

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

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

6.10.Robustness - Statistics obtained by Y-scrambling:

6.11.Robustness - Statistics obtained by bootstrap:

6.12.Robustness - Statistics obtained by other methods:

7.External validation - OECD Principle 4

7.1.Availability of the external validation set:

Yes

7.2.Available information for the external validation set:

CAS RN: Yes

Chemical Name: Yes

Smiles: Yes

Formula: Yes

INChI: Yes

MOL file: Yes

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

All

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

All

7.5.Other information about the external 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:

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

8.2.A priori or a posteriori mechanistic interpretation:

8.3.Other information about the mechanistic interpretation:

9.Miscellaneous information

9.1.Comments:

9.2.Bibliography:

9.3.Supporting information:

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

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

Q13-46-0045

10.2.Publication date:

2013-06-28

10.3.Keywords:

Lhasa Limited;Derek for Windows;skin sensitisation;mammal;

10.4.Comments:

former Q13-34-36-315