	QMRF identifier (JRC Inventory): **Q13-66-0044**
	QMRF Title: **Derek for Windows - Chromosome damage**
	Printing Date: **Dec 11, 2019**

1.QSAR identifier

1.1.QSAR identifier (title):

Derek for Windows - Chromosome damage

1.2.Other related models:

The Derek for Windows model for Chromosome damage is the only model and invokes no submodels.

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):

21st February 2011

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 chromosome damage endpoint.

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

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

[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. http://pubs.acs.org/doi/abs/10.1021/ci020272g

2.8.Availability of information about the model:

The alerts are available for inspection within the software and representative examples are provided to illustrate a given alert if available. Alerts are developed against a set of compounds which are generally gleaned from publicly available data, though a formal training set is not recorded for all alerts

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

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Mammal (primarily hamster, mouse and human)

3.2.Endpoint:

6.Other 6.6.Other Chromosome damage

3.3.Comment on endpoint:

The model predicts potential chromosome damage. It is primarily based on data from in vitro chromosome aberration test and predicts for the chromosome damage endpoint, with the majority of predictions being for in vitro chromosome damage and includes additional data from in vitro micronucleus and L5178Y TK+/- assays. In vivo chromosome aberration test and in vivo micronucleus test data have also been considered to develop a limited number of alerts which predict for in vivo chromosome damage.

3.4.Endpoint units:

The percentage of cells displaying chromosome aberrations (excluding gaps) is used to assign a compound as causing chromosome damage in vitro or in vivo: a level of 10% or above classifies a compound as positive.
The percentage of cells with micronuclei is used to classify compounds as positive or negative in the in vitro and in vivo micronucleus assay.
In all cases the authors call is also taken into consideration.

3.5.Dependent variable:

3.6.Experimental protocol:

The model is based principally on data from the standard Chinese hamster cell lines in the in vitro chromosome aberration test (Chinese hamster lung, Chinese hamster ovary) conducted following standard test protocol, although data from human lymphocyte cells is also considered in a number of alerts. Alerts predicting for in vivo chromosome damage are typically based on in vivo micronucleus assay data in the mouse. If activity is observed in a non-standard test protocol or cell line, this will be mentioned in the comments. Activity observed in additional assays will also be mentioned in the comments.

3.7.Endpoint data quality and variability:

Alerts are developed against data generated following standard test protocols. The data forming the basis of each alert is fully referenced within the alert.

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

Expert system

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:

4.5.Algorithm and descriptor generation:

4.6.Software name and version for descriptor generation:

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:

The scope of the structure-activity relationships describing the chromosome damage endpoint are defined by the developer to be the applicability domain for the model. Therefore, if a chemical matches an alert describing a structure-activity for chromosome damage it can be considered to be within the applicability domain.

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:

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

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:

The training set for each alert is not available. The number of compounds upon which an alert is built can vary substantially from alert to alert depending on the amount information available in the public domain or donated proprietary data. However an alert will not be built against a single compound. No internal validation has been performed.

6.6.Pre-processing of data before modelling:

Data is not processed before an alert is developed, though an alert will not be built against a single compound, and data generated from standard

test protocols is always used unless the alert comments specifically mention otherwise.

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: 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:

Three external validation sets exist and have been used for the external validation of the alerts within the chromosome damage model. Two are publicly available but not attached, the other contains proprietary data.

1) Revised Edition 1998 Data Book of Chromosomal Aberration Test in Vitro, Sofuni T (editor), Life-Science Information Center, Tokyo, 1999.

2) A collection of in vitro chromosome aberration test data for 1620 compounds compiled from the US Food and Drug Administration (FDA) SAR Genetox Database (extracted 2 September 2010).

3) Mohr J, Jain B, Sutter A, Ter Laak A, Steger-Hartmann T, Heinrich N and Obermayer K. A maximum common subgraph kernel method for predicting the chromosome aberration test. Journal of Chemical Information and Modeling, 2010.

7.6.Experimental design of test set:

In addition to the publicly available dataset, proprietary datasets were sought for robust external evaluations.

7.7.Predictivity - Statistics obtained by external validation:

The positive predictivity for each alert for the three datasets is available within the software.

7.8.Predictivity - Assessment of the external validation set:

The total number of compounds in the validation datasets is 1088, which is sufficiently large to validate the model (although there may be some compounds which are common to both datasets).

The compounds in the datasets are primarily small chemicals and so are representative of the structures used to build the model.

7.9.Comments on the external validation of the model:

The three datasets used for external validation were tested against each alert in the chromosome damage model and the positive predictivity calculated for each alert and each dataset.

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

All alerts describing structure-activity relationships for the chromosome damage endpoint have a mechanistic basis wherever possible. Mechanistic information is detailed in the comments associated with an alert and can include information on both the mechanism of action and biological target.

8.2.A priori or a posteriori mechanistic interpretation:

The mechanistic basis of the model was developed a priori by examining the active and inactive structures before developing the structure-activity relationship.

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 for Windows is an knowledge-based expert system containing mechanistically-based rules which are built using all the underlying evidence available to the SAR developer. Therefore, there is no defined training or test set, and there are no internal validation statistics to report. The model may be used to consider the likelihood of chromosome damage when preparing a regulatory submission (e.g. for REACH).

9.2.Bibliography:

[1]Sofuni T, Ed (1999). Revised Edition 1998 Data Book of Chromosomal Aberration Test in Vitro. Life-Science Information Center, Tokyo.

[2]A collection of in vitro chromosome aberration test data for 376 compounds compiled from US Food and Drug Administration (FDA) Center for Drug Evaluation and Research (CDER) and Center for Food Safety and Applied Nutrition (CFSAN) 2009 Genetox Databases.

[3]Mohr J, Jain B, Sutter A, Ter Laak A, Steger-Hartmann T, Heinrich N and Obermayer K (2010) A maximum common subgraph kernel method for predicting the chromosome aberration test. Journal of Chemical Information and Modeling 50(10), 1821–1838 http://dx.doi.org/10.1021/ci900367j

9.3.Supporting information:

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

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

Q13-66-0044

10.2.Publication date:

2013-06-28

10.3.Keywords:

Lhasa Limited;chromosome damage;Derek for Windows;mammal;

10.4.Comments:

former Q13-34-36-313