The Acceptance of Computational Methods for the Regulatory Assessment of Chemicals

Andrew Worth

European Commission, Joint Research Centre, Institute for Health & Consumer Protection[1]

US Food & Drug Administration, Center for Food Safety & Applied Nutrition, Office of Food Additive Safety[2]

CIR workshop on SARs, Washington DC, 5 March 2012

[1] Permanent affiliation; [2] Current temporary affiliation

European Commission

Overview

1. The EU Cosmetics legislation – the challenge for alternative methods
2. Historical perspective on the regulatory use of QSAR
3. How to document QSAR models and their predictions
4. How to assess model predictions
5. Computational tools developed by the JRC
6. Some EU research initiatives to promote alternatives for cosmetics
7. Background reading
8. Appendix – illustration of the framework for assessing model predictions

Challenge set by the EU Cosmetics Directive

Testing Ban

Prohibits animal testing in the EU to meet requirements of Cosmetics Directive

Finished products: September 2004

Ingredients: March 2009

Marketing Ban

Prohibits marketing of cosmetics when the product or ingredients have been tested on animals to meet requirements of Cosmetics Directive

Gradual prohibition (ingredients)

1st Deadline: March 2009

2nd Deadline: March 2013
(sensitisation, carcinogenicity, repeat dose toxicity, reproductive toxicity, toxicokinetics)

7th Amendment to Council Directive 76/768/EEC

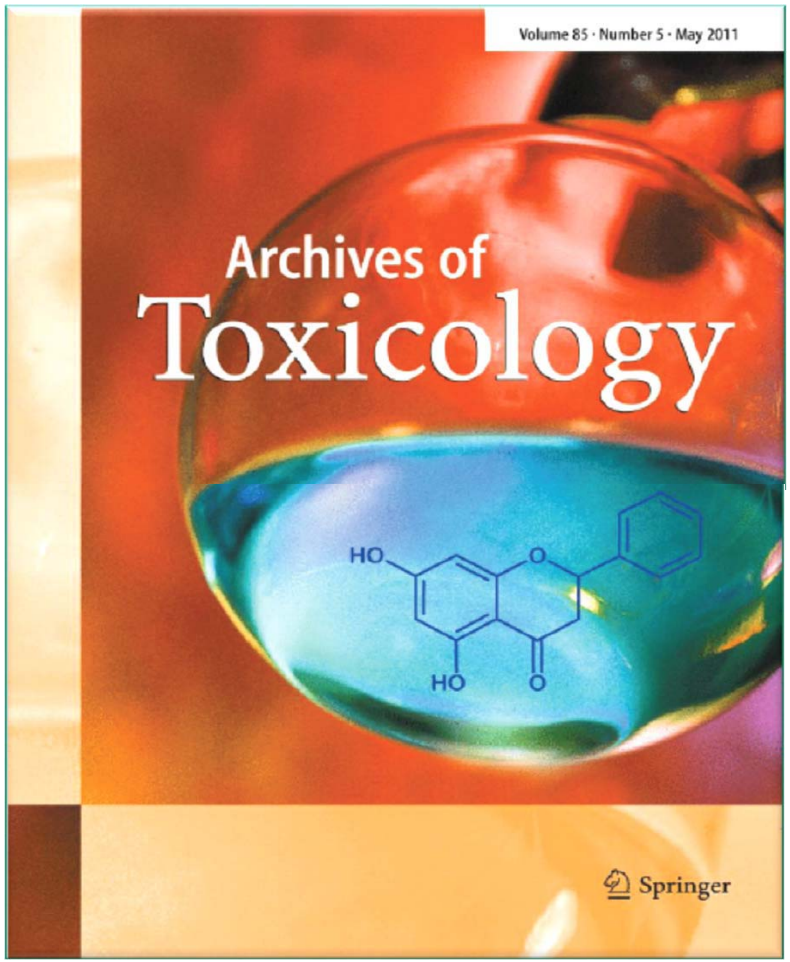
Toxicological endpoints relevant to cosmetic ingredients

- Acute toxicity (oral, dermal, inhalation)
- (Skin and eye) Irritation and corrosion
- Skin sensitisation
- Dermal absorption
- Repeated dose toxicity
- Mutagenicity / genotoxicity
- Carcinogenicity
- Reproductive toxicity
- Toxicokinetics
- Photo-induced toxicities

Specific Information requirements / recommendations are context-dependent

Current status of alternative methods

Alternative (non-animal) methods for cosmetics testing: current status and future prospects 2010



- Another 7– 9 years for full replacement of animal tests for skin sensitisation
- Due to underlying scientific challenges, no specific timelines estimated in the areas of:
  - toxicokinetics
  - repeated dose toxicity
  - carcinogenicity
  - reproductive toxicity

Regulatory use of (Q)SAR – before & during REACH

- Industrial chemicals - extensive use of grouping and read-across, generally without documented rationale
- Industrial chemicals - occasional use of QSARs in risk assessment, PBT assessment, and classification & labelling, generally without documented rationale
- QSARs and read-across usually for hazard identification, not hazard characterisation
- ECHA's first evaluation (June 2011) of the use of alternative methods under REACH: limited use of *in vitro* methods and QSAR/read-across

Regulatory use of (Q)SAR – Cosmetics

- Possible use of QSAR for cosmetic ingredients is recognised in the SCCS *Notes of Guidance for the Testing of Cosmetic Ingredients and their Safety Evaluation (December 2010)*

- ”At present, safety evaluation of cosmetic ingredients is carried out by the SCCS using data obtained from animal studies (in vivo), in vitro experiments, QSAR (quantitative structure activity relationship) calculations, clinical studies, epidemiological studies and accidents. The physical and chemical data of the compounds under investigation are also taken into consideration.”

- “ …This risk assessment procedure is subdivided in 4 parts:

  1) Hazard identification: based on the results of in vivo tests, in vitro tests, clinical studies, accidents, human epidemiological studies and, when available, quantitative structure activity relationship (QSAR) studies. The intrinsic physical, chemical and toxicological properties of the molecule under consideration are studied to identify whether the substance has the potential to damage human health …”

Barriers to acceptance

- Industry assessors and regulators may not be familiar with QSAR methodology and therefore not comfortable using the results
- Modellers and regulators do not speak the same language
- Models might not be relevant to the regulatory question
- Models might not be sufficiently transparent
- Models might not be reproducible or readily available
- Insufficient practical guidance on how to use models in a regulatory context

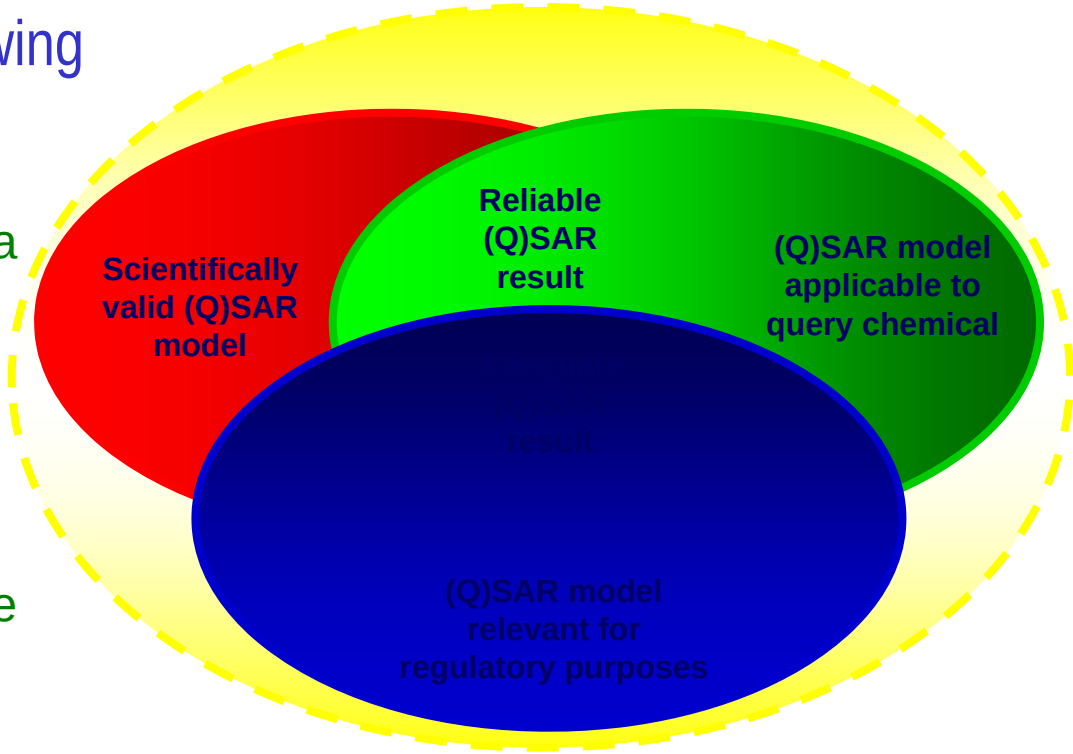
Overcoming the barriers to acceptance

- Consistent and internationally accepted reporting formats on QSAR models and their predictions: QMRF and QPRF
- Freely accessible software tools
- Training on how to use the tools
- Guidance on how to interpret the prediction results

Adequacy of (Q)SAR prediction

In order for a (Q)SAR result to be adequate for a given regulatory purpose, the following conditions must be fulfilled:

- the prediction should be generated by a valid model
- the model should be applicable to the chemical of interest with the necessary level of reliability
- the prediction should be relevant for the regulatory purpose
- adequate and reliable documentation should be provided



ECHA guidance on Information Requirements & Chemical Safety Assessment

http://guidance.echa.europa.eu/docs/guidance_document/information_requirements_en.htm

Standardised (Q)SAR Reporting Formats

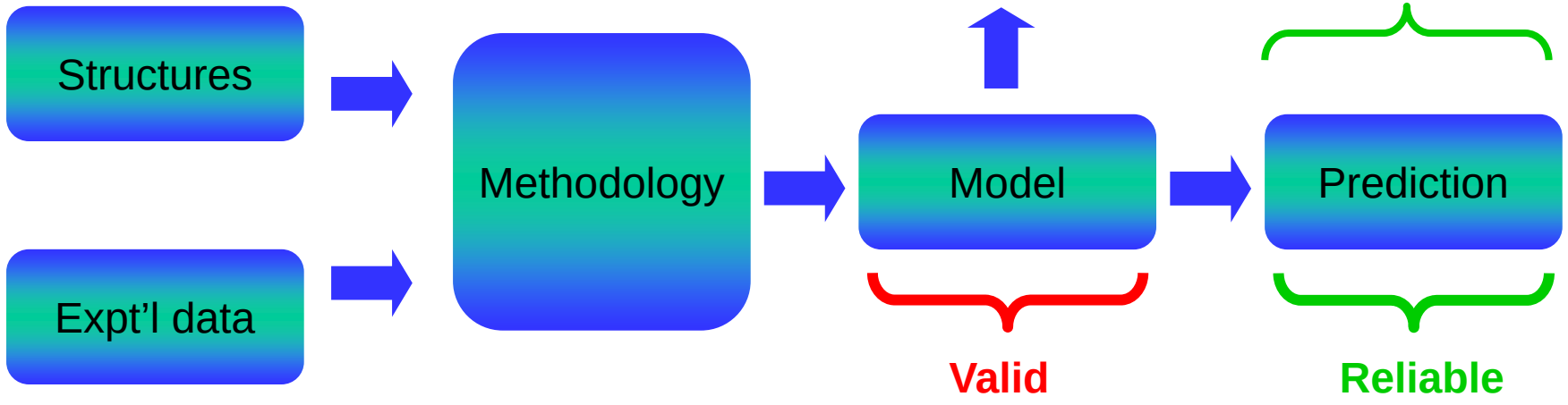
The need for “adequate and reliable” documentation is met by using standardised reporting formats:

QMRF

Robust summary of a (Q)SAR model, which reports key information on the model according to the 5 OECD validation principles.

QPRF

Description and assessment of the prediction made by given model for a given chemical



(Q)SAR Reporting Formats: QMRF

QMRF captures information on fulfilment of OECD validation principles, but no judgement or “validity statement” is included

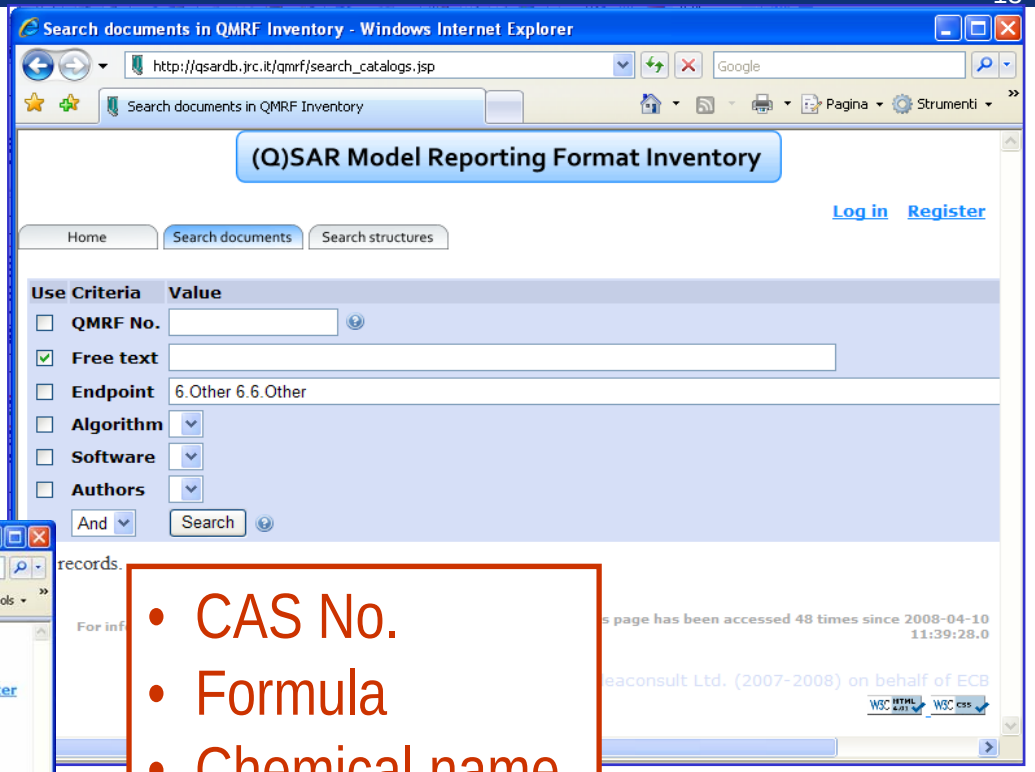
A (Q)SAR should be associated with the following information:

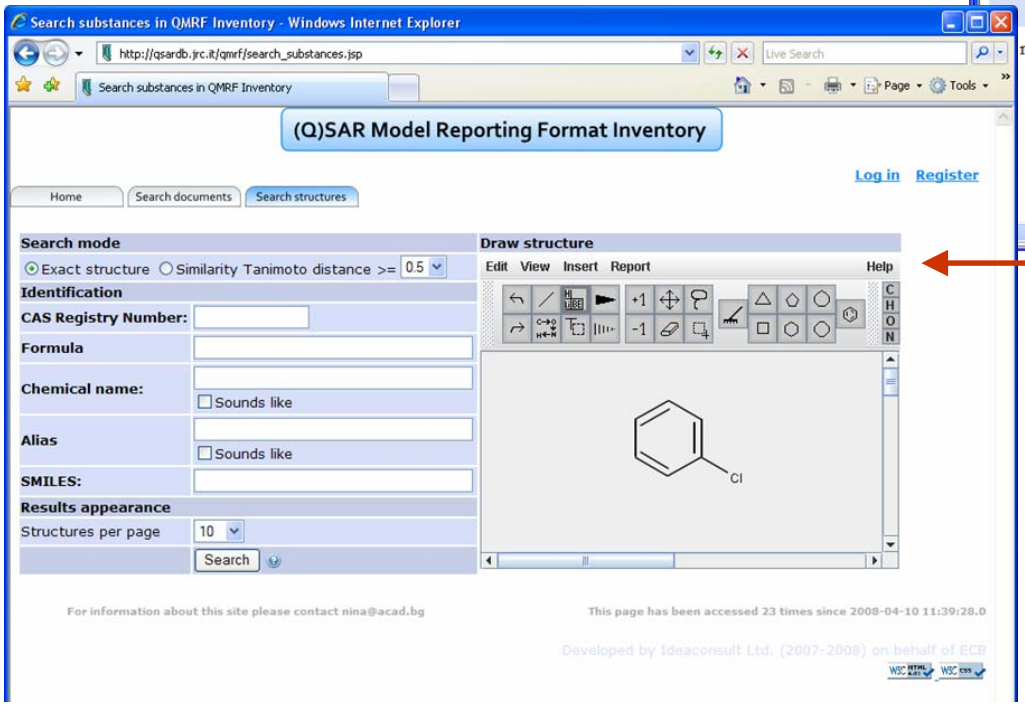
1. a defined endpoint
2. an unambiguous algorithm
3. a defined applicability domain
4. appropriate measures of goodness-of-fit, robustness and predictivity
5. a mechanistic interpretation, if possible

- Principles adopted by 37th Joint Meeting of Chemicals Committee and Working Party on Chemicals, Pesticides & Biotechnology; 17-19 Nov 2004
- ECB preliminary Guidance Document published in Nov 2005
- OECD Guidance Document published in Feb 2007
- OECD Guidance summarised in REACH guidance (IR and CSA) 2008

JRC QSAR Model Database

- QMRF No.
- Free text
- Endpoint
- Algorithm
- Software
- Authors



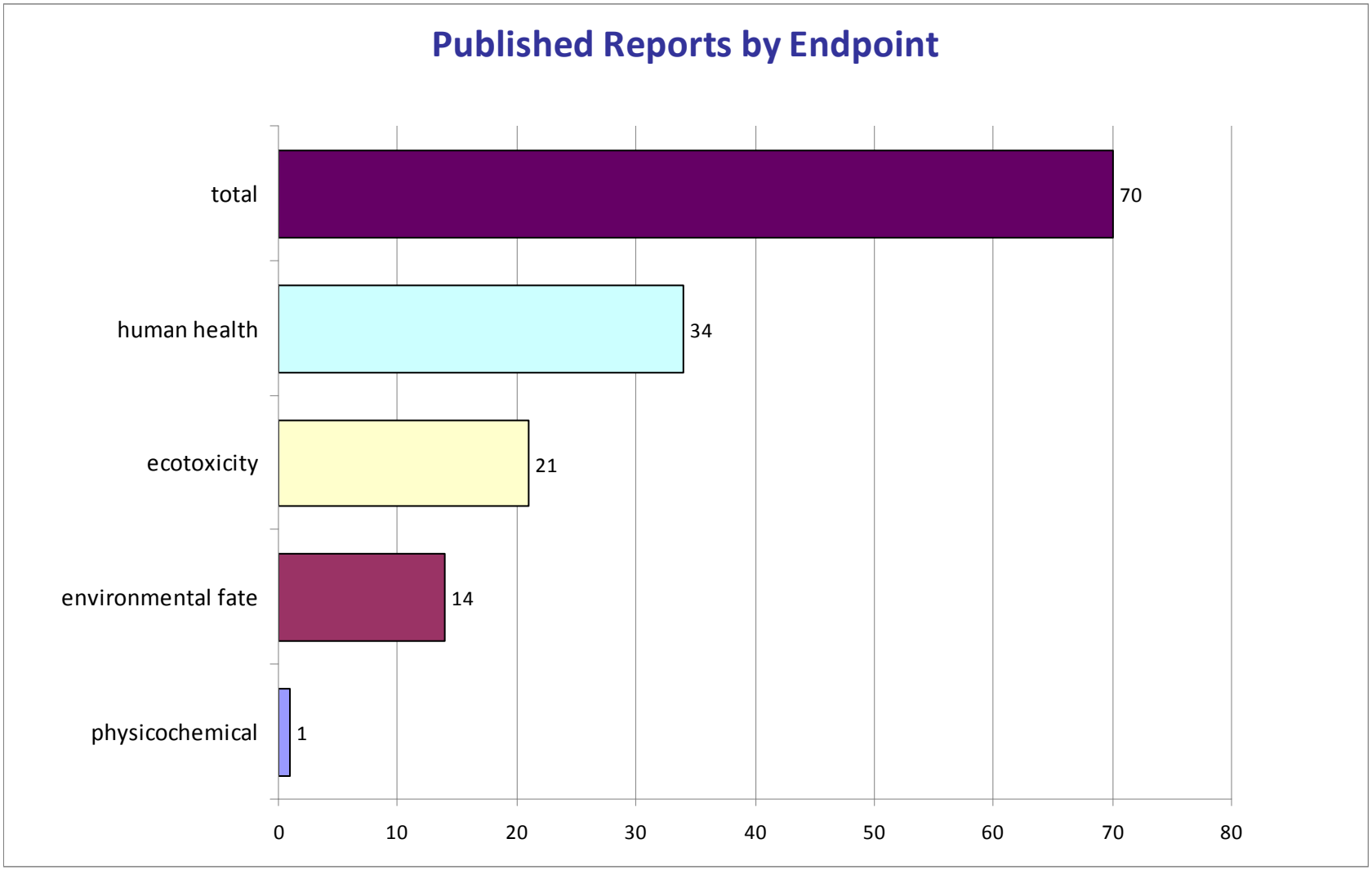


- CAS No.
- Formula
- Chemical name
- SMILES

QMRF editor

http://qsardb.jrc.ec.europa.eu

QMRFs in JRC database



http://qsardb.jrc.ec.europa.eu

1 March 2012: 70 QMRFs published

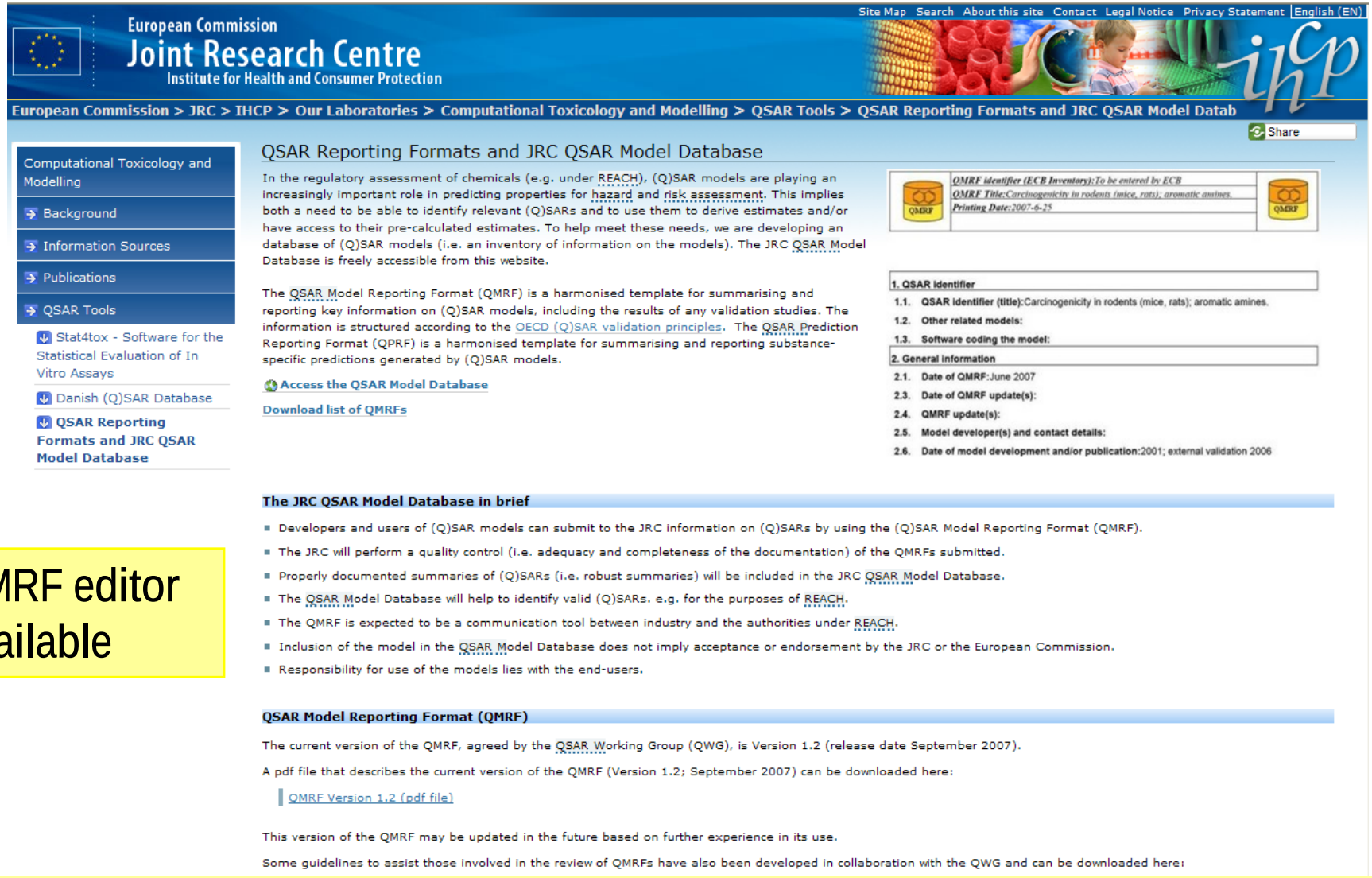
(Q)SAR Reporting Formats: QPRF

QPRF captures information on the substance and its prediction, and is intended to facilitate considerations of the adequacy of a prediction

1. Substance information
2. General (administrative) information on QPRF
3. Information on prediction (endpoint, algorithm, applicability domain, uncertainty, mechanism)
4. Adequacy (includes judgement and indicates whether additional information is needed for WoE assessment)

- Assessment of adequacy depends on reliability and relevance of prediction, but also on the availability of other information, and the consequence of being wrong
- Not just a scientific consideration, but also a policy decision

Information on Reporting Formats



European Commission
Joint Research Centre
Institute for Health and Consumer Protection

Site Map Search About this site Contact Legal Notice Privacy Statement English (EN)

European Commission > JRC > IHCP > Our Laboratories > Computational Toxicology and Modelling > QSAR Tools > QSAR Reporting Formats and JRC QSAR Model Datab

Share

Computational Toxicology and Modelling
- Background
- Information Sources
- Publications
- QSAR Tools
  - Stat4tox - Software for the Statistical Evaluation of In Vitro Assays
  - Danish (Q)SAR Database
  - **QSAR Reporting Formats and JRC QSAR Model Database**

QSAR Reporting Formats and JRC QSAR Model Database

In the regulatory assessment of chemicals (e.g. under REACH), (Q)SAR models are playing an increasingly important role in predicting properties for hazard and risk assessment. This implies both a need to be able to identify relevant (Q)SARs and to use them to derive estimates and/or have access to their pre-calculated estimates. To help meet these needs, we are developing an database of (Q)SAR models (i.e. an inventory of information on the models). The JRC QSAR Model Database is freely accessible from this website.

The QSAR Model Reporting Format (QMRF) is a harmonised template for summarising and reporting key information on (Q)SAR models, including the results of any validation studies. The information is structured according to the OECD (Q)SAR validation principles. The QSAR Prediction Reporting Format (QPRF) is a harmonised template for summarising and reporting substance-specific predictions generated by (Q)SAR models.

Access the QSAR Model Database

Download list of QMRFs

QMRF identifier (ECB Inventory):To be entered by ECB
QMRF Title:Carcinogenicity in rodents (mice, rats); aromatic amines.
Printing Date:2007-6-25

1. QSAR identifier
- 1.1. **QSAR identifier (title):**Carcinogenicity in rodents (mice, rats); aromatic amines.
- 1.2. **Other related models:**
- 1.3. **Software coding the model:**

2. General information
- 2.1. **Date of QMRF:**June 2007
- 2.3. **Date of QMRF update(s):**
- 2.4. **QMRF update(s):**
- 2.5. **Model developer(s) and contact details:**
- 2.6. **Date of model development and/or publication:**2001; external validation 2006

The JRC QSAR Model Database in brief

- Developers and users of (Q)SAR models can submit to the JRC information on (Q)SARs by using the (Q)SAR Model Reporting Format (QMRF).
- The JRC will perform a quality control (i.e. adequacy and completeness of the documentation) of the QMRFs submitted.
- Properly documented summaries of (Q)SARs (i.e. robust summaries) will be included in the JRC QSAR Model Database.
- The QSAR Model Database will help to identify valid (Q)SARs. e.g. for the purposes of REACH.
- The QMRF is expected to be a communication tool between industry and the authorities under REACH.
- Inclusion of the model in the QSAR Model Database does not imply acceptance or endorsement by the JRC or the European Commission.
- Responsibility for use of the models lies with the end-users.

QSAR Model Reporting Format (QMRF)

The current version of the QMRF, agreed by the QSAR Working Group (QWG), is Version 1.2 (release date September 2007).

A pdf file that describes the current version of the QMRF (Version 1.2; September 2007) can be downloaded here:

QMRF Version 1.2 (pdf file)

This version of the QMRF may be updated in the future based on further experience in its use.

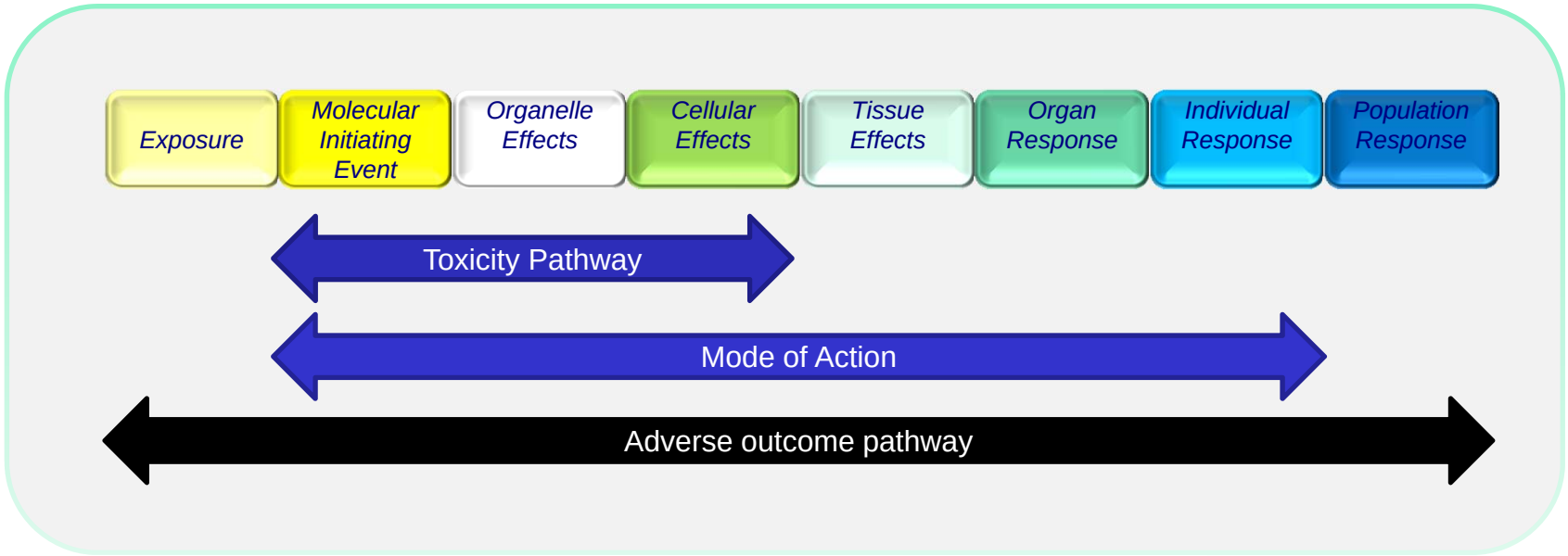
Some guidelines to assist those involved in the review of QMRFs have also been developed in collaboration with the QWG and can be downloaded here:

QMRF editor available

http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/qsar_tools/QRF

Towards a template for Intermediate Effects

- Trend towards assessment based on Toxicity Pathways, Mode-of-Action (MoA) and Adverse Outcome Pathway (AOP)
- Development of OECD Harmonised Template (OHT) 201
- OECD project, led by JRC, and in collaboration with ECHA
- Compatibility with IUCLID, OECD Toolbox, Effectopedia (http://www.effectopedia.org/go/) and other tools

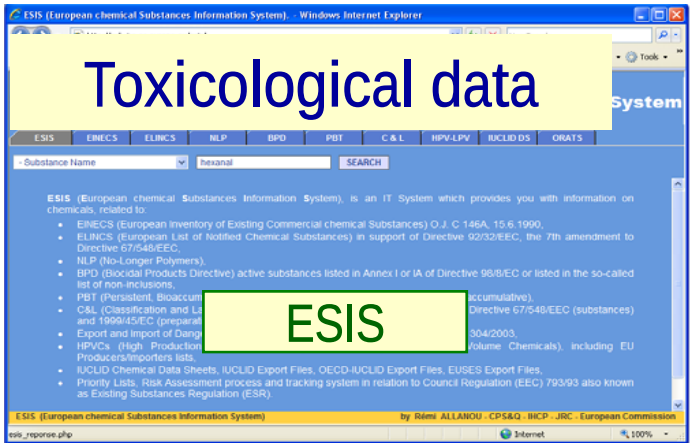


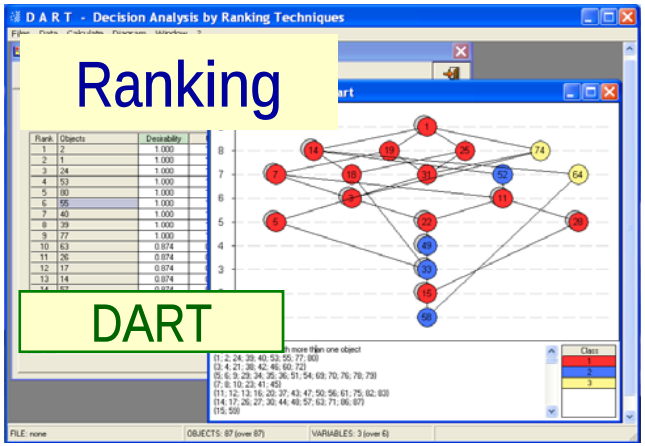
Adequacy: 10 questions for assessing model predictions

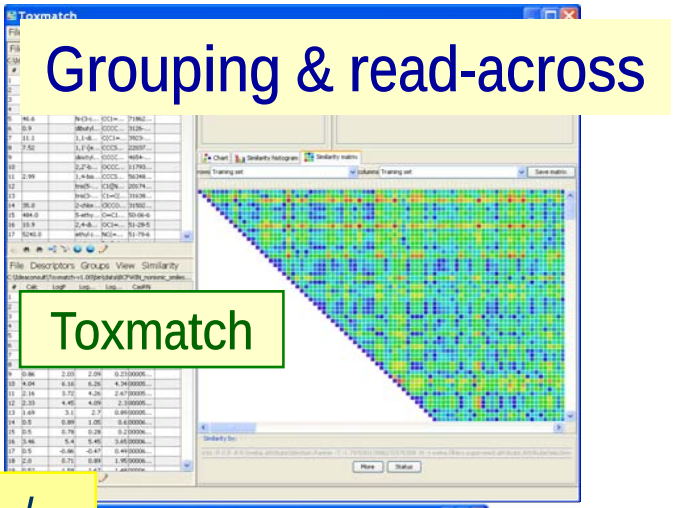
1. Is the predicted endpoint clearly defined?
2. Is the predicted endpoint a *direct* information requirement?
3. Is the model training set fully available (for statistical models)?
4. Is the method used to develop the model well documented?
5. Is information available concerning the performance of the model?
6. In the case of a statistical model, is there evidence of overfitting?
7. Does the model training set contain the chemical of interest ?
8. Does the model make reliable predictions for analogues of the chemical of interest?
9. Is the prediction substantiated with argumentation based on the applicability domain of the model?
10. Can the prediction be easily reproduced?

Worth et al (2011). A Framework for assessing in silico Toxicity Predictions: Case Studies with selected Pesticides. JRC report EUR 24705 EN.

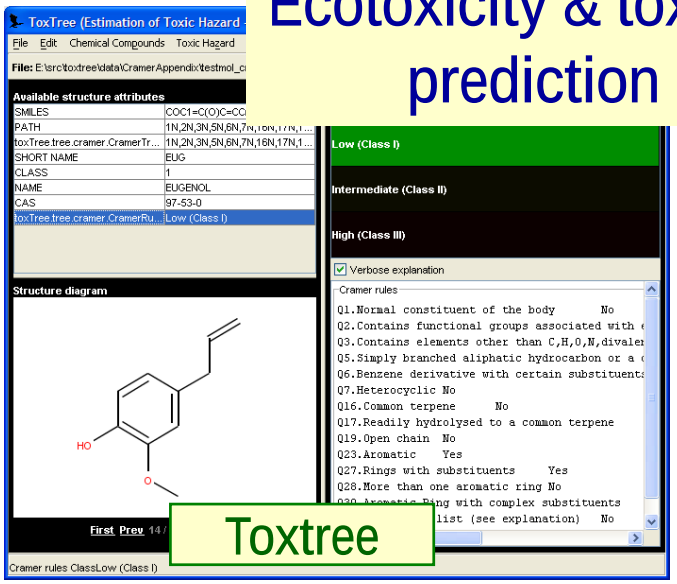
Computational tools developed by the JRC

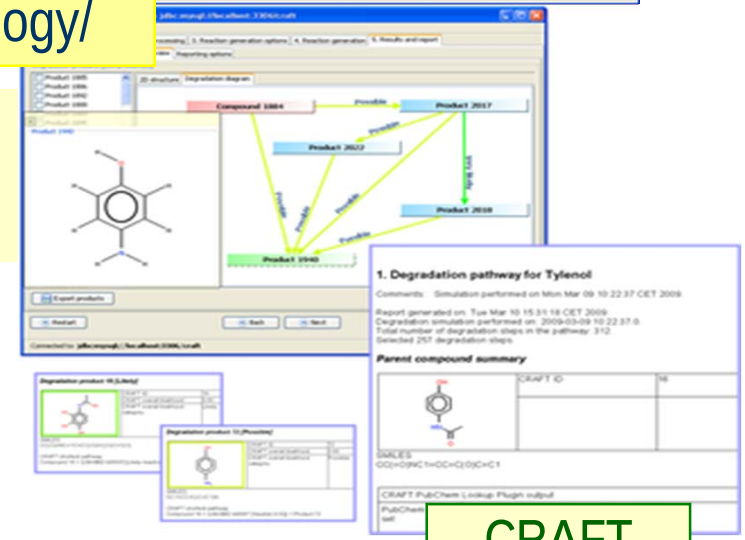


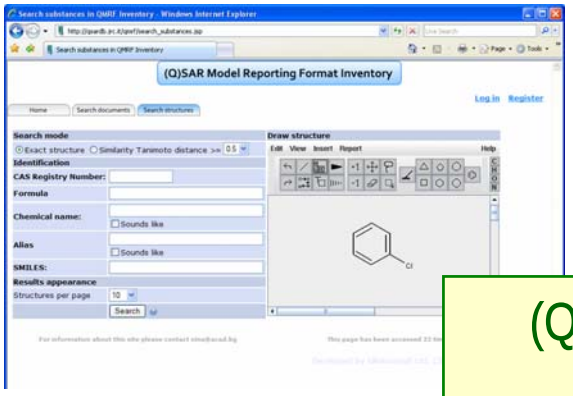




http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/

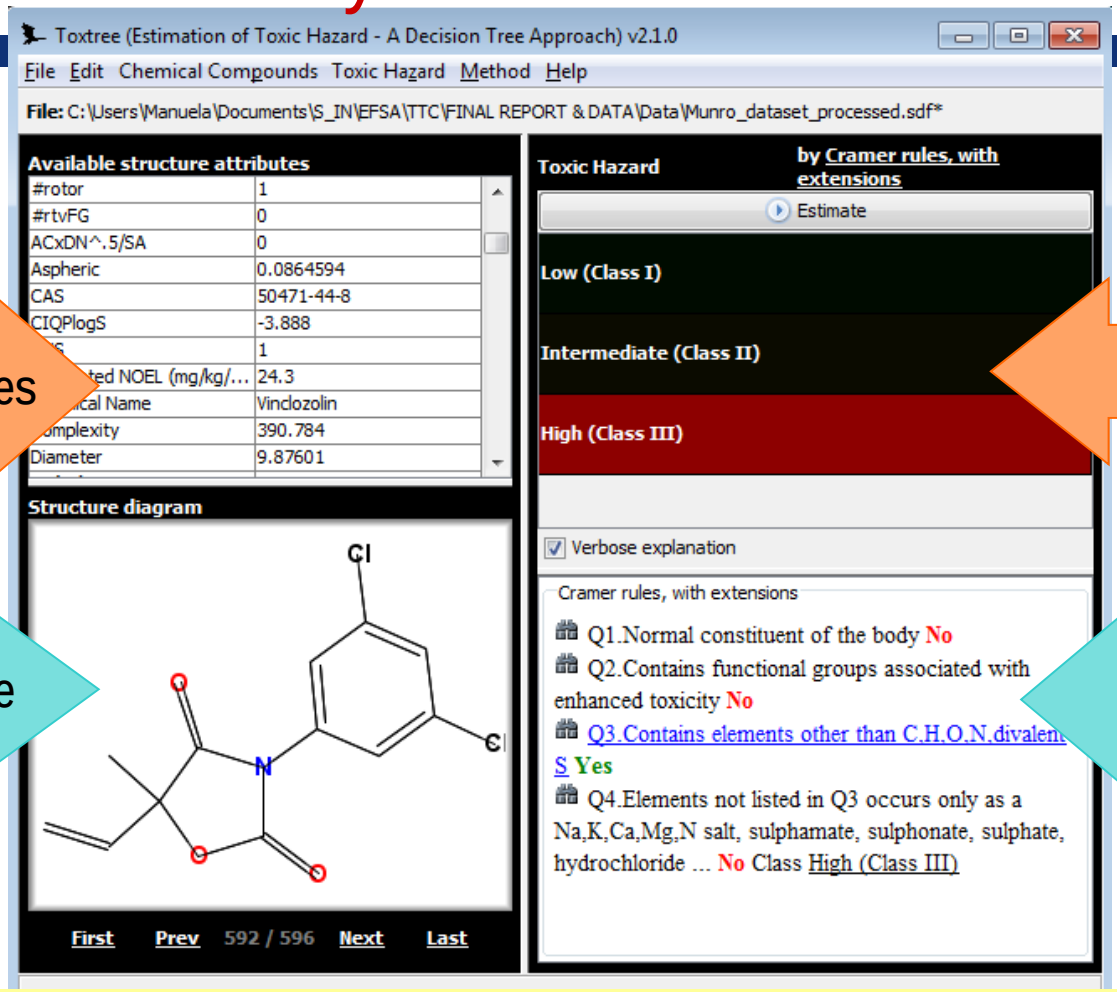






METIS

Toxicity Estimation Tool: Toxtree



Compound properties

Compound structure

Prediction

Reasoning

- Downloadable versions from JRC and Sourceforge (http://toxtree.sourceforge.net)
- Online version: OpenTox - ToxPredict (http://www.opentox.org/)
- Current version 2.5.0 (August 2011) includes skin and eye irritation, systemic toxicity (Cramer), genotoxicity and carcinogenicity, sensitisation, sites of CYP metabolism

7th Framework Programme: SEURAT -1 Cluster

Joint funding by European Commission and Cosmetics Europe: 2011-2015

Computational Toxicology

The SEURAT vision is to fundamentally change the way we assess the safety of chemicals, by superseding traditional animal experiments with a predictive toxicology that is based on a complete understanding of how chemicals can cause adverse effects in humans.

COSMOS – key facts and figures

- Full title: Integrated In Silico Models for the Prediction of Human Repeated Dose Toxicity of Cosmetics to Optimise Safety
- Coordinator: Prof. Mark Cronin, Liverpool John Moores University (LJMU)
- Duration: 01/01/11 – 31/12/15 (60 months)
- Partners: 15

  LJMU, JRC, US FDA (CFSAN), Henkel, Merck, Ineris, ILSI Europe, Altamira, Insilico Biotech, Knime, Molecular Networks, Soluzioni Informatiche, Bulgarian Academy of Sciences, National Chemicals Institute (Slovenia), Bradford University

http://www.cosmostox.eu

COSMOS: overall aim and scope

To develop an integrated suite of open source and open access computational models to assist in the prediction of human repeated dose toxicity for cosmetics

- Collation, curation and quality control of **toxicological data**
- Development and application of novel **Threshold of Toxicological Concern** (TTC) approaches to cosmetics
- Development of novel **in silico methods** (QSAR, grouping, ranking) for predicting repeat dose toxicity
- **Multi-scale modelling**: in vitro fate, cell growth, uptake & toxicity, molecular network models, 2D liver, PBPK
- **Integration** of open source and open access tools (database and models) into adaptable and flexible **in silico workflows**

Alternatives to Animal Testing (AAT) extended program

A further 8M EUR to drive pre-validation & acceptance of alternative assays

- Pre-validation of 'promising' toolbox test methods for skin sensitization and data integration activities

- Finalizing development and conduct pre-validation of the already developed 3D-model for genotoxicity, and promote regulatory acceptance in this field

- Refinement of eye irritation assays to address last remaining gaps

Summary & conclusions

- In principle, (Q)SAR estimates can be used as direct replacements for test data, but in practice, their use in Integrated Testing Strategies is more likely
- To promote the harmonised use of QSARs, standardised templates for reporting the validity of QSAR models, and the adequacy of QSAR predictions, have been developed
- No reporting format (yet) for Integrated Testing Strategies, but a template for intermediate effects (OHT 201) is under development
- No formal validation and adoption procedures for (Q)SAR models
- Each organisation needs to develop its own criteria / guidelines for assessing the adequacy of (Q)SAR predictions. Illustrative examples useful
- Acceptable models for certain endpoints should be achievable in the short-term: topical toxicities, skin penetration, sensitisation, genotoxicity

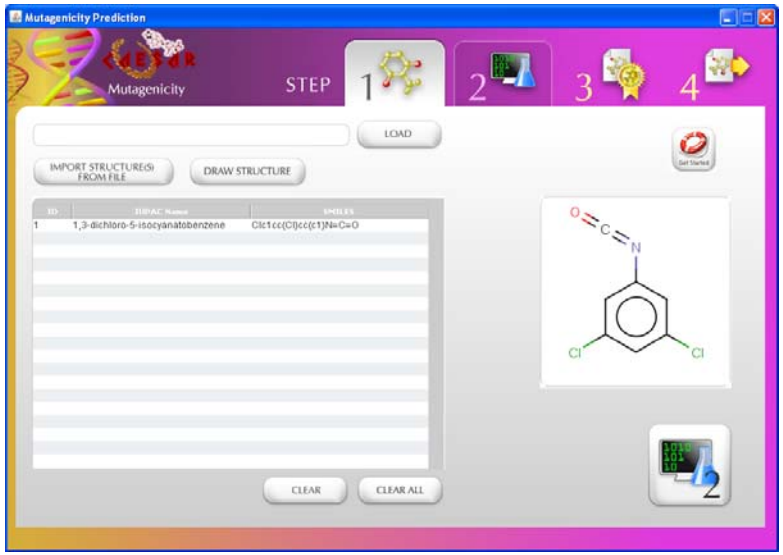
Key references

- Adler et al (2011). Alternative (Non-Animal) Methods for Cosmetics Testing: Current Status and Future Prospects – 2010.
  http://www.springerlink.com/content/y33r3u3854246277
- ECHA 2011 status report on use of alternatives under REACH: http://echa.europa.eu/documents/10162/17231/alternatives_test_animals_2011_en.pdf
- OECD Guidance on QSAR validation (2007): http://www.oecd.org
- REACH Guidance on ITS and use of QSARs (2008)
  http://guidance.echa.europa.eu/docs/guidance_document/information_requirements_en.htm
- QSAR reporting formats (QMRF and QPRF) and QMRF Editor
  http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/qsar_tools/QRF
- QSAR Model Database: http://qsardb.jrc.ec.europa.eu/qmrf
- Worth A, Lapenna S, Lo Piparo E, Mostrag-Szlichtyng A & Serafimova R (2011). A Framework for assessing in silico Toxicity Predictions: Case Studies with selected Pesticides. JRC report EUR 24705 EN. http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/

Appendix

Assessment of QSAR predictions

Example of the CAESAR mutagenicity model



http://www.caesar-project.eu/

Example: CAESAR mutagenicity predictions

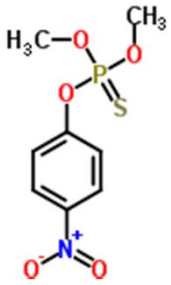
Q1	Is the predicted endpoint clearly defined?
A1	Yes, the endpoint is Ames (*S. Typhimurium*) mutagenicity
Q2	If the predicted endpoint is clearly defined ("yes" to Q1), does it represent a direct information requirement under the legislation of interest, or is it related to one of the information requirements?
A2	Yes, genotoxicity test data are required under most types of chemicals legislation (e.g. industrial chemicals, pesticides, biocides)
Q3	If the model is statistically based (as opposed to knowledge-based), is the model training set fully available?
A3	Yes, the training and test set are published (http://www.caesar-project.eu)
Q4	Is the method used to develop the model documented or referenced (e.g. in a scientific paper or QMRF)
A4	Yes, a QMRF is in preparation, based on the following publications: Ferrari T, Gini G & Benfenati E (2009). Support vector machines in the prediction of mutagenicity of chemical compounds. Proc NAFIPS 2009, June 14-17, Cincinnati, USA, p 1-6. Ferrari T & Gini G (2010). A new multistep model to predict mutagenicity from statistic analysis and relevant structural alerts. Central Chemistry 4, Suppl 1, S2.

Example: CAESAR mutagenicity predictions

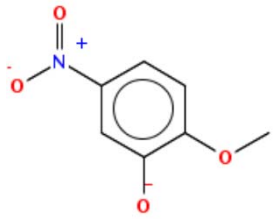
Q5	Is information available (in terms of statistical properties) concerning the performance of the model, including its goodness-of-fit, predictivity, robustness and error of prediction (uncertainty)?
A5	Yes. Information on the accuracy (82.1%), sensitivity (90.6%) and specificity (71.4%) are provided.
Q6	If the model is statistically based (as opposed to knowledge-based), does examination of the available statistics indicate that the model may have been overfitted?
A6	The model is statistically based but should not be overfitted because the ratio of chemicals (3380) to descriptors (42) is 80.5.

Q7	Does the model training set contain the chemical of interest?
A7	The model training set includes some pesticides including parathion-methyl but not sodium nitroguaiacolate.

Methyl parathion
Dimethoxy-(4-nitrophenoxy)-thioxo-phosphorane
CAS 298-00-0
S=P(Oc1ccc(cc1)[N+]([O-])=O)(OC)OC
Mutagen
Correctly predicted by CAESAR



Sodium Nitroguiacolate
2-methoxy-5-nitro-phenolate
CAS 67233-85-6
[Na+].[O-]c1cc(ccc1OC)[N+]([O-])=O
Non mutagen
Incorrectly predicted as mutagen by CAESAR



Example: CAESAR mutagenicity predictions

Q8	Does the model make reliable predictions for analogues of the chemical structure of interest?
A8	Yes, the Caesar software gives the chance to examine, for each compound submitted, the six most similar compounds found in the model training set. For these compounds the experimental value for the selected endpoint is shown, together with the prediction made by the model. The similarity measure employed by the Caesar software takes into account functional group similarity, constitutional similarity, ring similarity and fingerprint similarity. For parathion methyl (correctly predicted by the software), the similar structures obtained are: parathion methyl (input structure), aminofenitrothion, 1-ethenoxy-4-nitro-benzene, fenitrooxon, o-nitroanisole, N-hydroxy-N-(4-nitrophenyl)acetamide. All of them are predicted correctly by the software. For nitroguaiacolate (wrongly predicted by the software) the similar structures obtained are: o-nitroanisole, 1-ethoxy-3-nitro-benzene, 2,5-dinitrophenol, p-nitrosoanisole, 2-methoxy-1,3,5-trinitro-benzene, 1-ethenoxy-4-nitro-benzene. All of them are predicted correctly by the software.

Example: CAESAR mutagenicity predictions

Q9	Is the model prediction substantiated with argumentation based on the applicability domain of the model?
A9	Yes, Caesar addresses the applicability domain in several ways, namely by: a) checking whether the compound of interest falls in the descriptor space – if the compound is out of domain, this is noted in the output; b) providing a similarity score (1=identity) for the structure-based comparison with analogues; c) visual representation of the most similar compounds; d) by revealing the known and predicted toxicities for the analogues, thereby indicating the prediction error. Thus Caesar provides an assessment based on both the input (descriptor) space and the output (toxicological endpoint) space.
Q10	Can the model prediction be easily reproduced?
A10	Yes, the software is accessible in the form of a freely accessible web platform (http://www.caesar-project.eu) The software is easy to use, even for non-specialists.