JRC QSAR Model Database

EURL ECVAM DataBase service on ALternative Methods to animal experimentation

To promote the development and uptake of alternative and advanced methods in toxicology and biomedical sciences

User Support & Tutorial

2017

DISCLAIMER: Addresses to Internet sites change constantly. Therefore, all links to third parties websites included in the document have been checked last June 2017. We will review it periodically.

CONTENT

Introduction

To address financial and animal welfare concerns in the regulatory assessment of chemicals, computational (*in silico*) models are being increasingly used as a means of filling data gaps and contributing to weight of evidence arguments.

Among *in silico* models are Qualitative and Quantitative Structure-Activity Relationships, collectively referred to as (Q)SARs. These are theoretical models that can be used to predict the physicochemical, biological and environmental fate properties of molecules..

A **structure-activity relationship** (SAR) is a qualitative relationship that relates a (sub)structure to the presence or absence of a property or activity of interest. The substructure may consist of adjacently bonded atoms, or an arrangement of non-bonded atoms that are collectively associated with the property or activity.

A **quantitative structure-activity relationship** (QSAR) is a mathematical model (often a statistical correlation) relating one or more quantitative parameters derived from chemical structure to a property or activity of interest. QSARs are quantitative models yielding a continuous or categorical result.

Read more on (Q)SARs and REACH on the JRC Science Hub

For regulatory purposes, it is important that computational models are properly characterised and documented. For this reason, the JRC developed the JRC QSAR Model Database and the QSAR Model Reporting Format (QMRF).

The QSAR Model Database provides information on QSAR models that have been submitted to the JRC. It is intended to help identify valid QSARs for the purposes of regulatory assessments (e.g. REACH). The database uses a harmonised template (the QMRF) for summarising and reporting key information relating to QSAR models.

This Tutorial introduces the JRC QSAR Model Database and provides guidance how to compile new models, update existing ones and publish them through the JRC QSAR Model database.

Background information on the science and applications of non-testing methods, including (Quantitative) Structure-Activity Relationships and chemical grouping methods is available from the links hereafter.

To facilitate practical application of (Q)SAR approaches in regulatory contexts by governments and industry and to improve their regulatory acceptance, the OECD has adopted outcomes such principles for the validation of (Q)SAR models, has developed guidance on how to apply these principles, as well as the QSAR Toolbox.

In November 2004, the OECD Member Countries agreed on the OECD principles for the validation of QSAR models for their use in the regulatory assessment of chemical safety. The internationally agreed principles provide Member Countries with a consistent and scientifically motivated framework for evaluating the regulatory applicability of QSAR models.

In February 2007, the OECD published a "**Guidance Document on the Validation of QSAR Models**" with the aim of providing guidance on how specific QSAR models can be evaluated with respect to the **OECD principles**.

To facilitate the consideration of a QSAR model for regulatory purposes, it should be associated with the following information:

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

The intent of **Principle 1** (defined endpoint) is to ensure clarity in the endpoint being predicted by a given model, since a given endpoint could be determined by different experimental protocols and under different experimental conditions. It is therefore important to identify the experimental system that is being modeled by the (Q)SAR. Further guidance is being developed regarding the interpretation of "defined endpoint". For example, a no-observed-effect level might be considered to be a defined endpoint in the sense that it is a defined information requirement of a given regulatory guideline, but cannot be regarded as a defined endpoint in the scientific sense of referring to a specific effect within a specific tissue/organ under specified conditions.

The intent of **Principle 2** (unambiguous algorithm) is to ensure transparency in the model algorithm that generates predictions of an endpoint from information on chemical structure and/or physicochemical properties. It is recognized that, in the case of commercially-developed models, this information is not always made publicly available. However, without this information, the performance of a model cannot be independently established, which is likely to represent a barrier for regulatory acceptance. The issue of reproducibility of the predictions is covered by this Principle, and will be explained further in the guidance material.

The need to define an applicability domain (**Principle 3**) expresses the fact that (Q)SARs are reductionist models which are inevitably associated with limitations in terms of the types of chemical structures, physicochemical properties and mechanisms of action for which the models can generate reliable predictions. Further work is recommended to define what types of information are needed to define (Q)SAR applicability domains, and to develop appropriate methods for obtaining this information.

The revised **Principle 4** (appropriate measures of goodness-of-fit, robustness and predictivity) includes the intent of the original Setubal Principles 5 and 6. The wording of the principle is intended to simplify the overall set of principles, but not to lose the distinction between the internal performance of a model (as represented by goodness-of-fit and robustness) and the predictivity of a model (as determined by external validation). It is recommended that detailed guidance be developed on the approaches that could be used to provide appropriate measures of internal performance and predictivity. Further work is recommended to determine what constitutes external validation of (Q)SAR models.

It is recognised that it is not always possible, from a scientific viewpoint, to provide a mechanistic interpretation of a given (Q)SAR (Principle 5), or that there even be multiple mechanistic interpretations of a given model. The absence of a mechanistic interpretation for a model does not mean that a model is not potentially useful in the regulatory context. The intent of **Principle 5** is not to reject models that have no apparent mechanistic basis, but to ensure that some consideration is given to the possibility of a mechanistic association between the descriptors used in a model and the endpoint being predicted, and to ensure that this association is documented.

These principles were agreed by OECD member countries at the 37th Joint Meeting of the Chemicals Committee and Working Party on Chemicals, Pesticides and Biotechnology in November 2004. The principles are intended to be read in conjunction with the associated explanatory notes which were also agreed at the 37th Joint Meeting.

Based on the joint OECD activities to define accepted valid criteria the **QSAR Model Reporting Format (QMRF)** was developed by the **JRC and EU Member State authorities as a harmonised template for summarising and reporting key information on QSAR models**, including the results of any validation studies. The information is structured according to the OECD validation principles.

The **JRC QSAR Model Database** (http://qsardb.jrc.ec.europa.eu/qmrf) is a freely accessible web application that enables users to submit, publish, and search **QSAR Model Reporting Format (QMRF) reports**. Developers and users of QSAR models can submit to the dedicated mailbox information on QSARs by using the QMRF. A downloadable **QMRF editor** (http://sourceforge.net/projects/qmrf/files/QMRF%20Editor/2.0.0) is used for this purpose. The JRC then performs a quality control (i.e. adequacy and completeness of the documentation) of the QMRF submitted. Properly documented QMRFs are included in the JRC QSAR Model Database.

QSAR models are described to such a detail to provide full transparency to any single step of the calculations to allow users to reproduce and apply it. Inclusion of the model does not imply acceptance or endorsement by the JRC or the European Commission, and responsibility for use of the models lies with the end-users.

Reference:
OECD (2007). Guidance Document No. 69 on the Validation of (Quantitative) Structure-Activity Relationship [QSAR] Models.
ENV/JM/MONO(2007)2.
Available from:
http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282007%29 2&doclanguage=en

The **JRC QSAR Model Database** is based on AMBIT2 technology. AMBIT is an open source software for cheminformatics data management, developed within the framework of a CEFIC LRI funded project. The AMBIT system consists of a database and functional modules allowing a variety of evaluations and data mining of the information, stored in the database. AMBIT's database allows storage of chemical structures and their identifiers, experimental data and test descriptions, literature references, information about QSAR models and finally chemical structure attributes such as molecular descriptors.

The database is based on a Relational Database Management System (RDBMS), which allows much faster and convenient access to the data in contrast to flat text files. Data stored in AMBIT can be searched in a number of ways: name, CAS, structure, Smiles, 2D fragment and structural similarity. The unique feature of AMBIT is the ability to store multifaceted information about chemical structures and provide a searchable interface linking these diverse components. Methods for assessing structural similarity, proximity in chemical space or commonality of mechanism of action can be used for QSAR applicability assessment and chemical grouping. The software relies on various open source software libraries and is an open source code to achieve maximum quality, transparency and ease of dissemination.

The AMBIT database consists of several repositories for compounds, descriptors, experimental results, QSAR models, and literature references, as well as several tables containing pre-processed information in order to speed up substructure and similarity queries. The database has been further developed to answer the needs of JRC and enhance its QSAR model repository, which has been adapted to include fields, mirroring the QSAR Model Reporting Format (QMRF), developed by the JRC in collaboration with the former EU QSAR Working Group. A format for importing QMRF reports in the database, as well as a friendly user interface, facilitating data import, have been implemented. The JRC QSAR Model Database is a reference site for retrieving robust summaries of QSAR models.

Data Browsing

Anonymous users can browse the Database for published QMRFs (Qsar Model Reporting Formats) and perform document or substance searching. QMRF documents can be searched by applying a subset of the following criteria:

- Title
- Free text
- Free text (Boolean)
- Endpoint (predefined list, see Endpoint Classification)
- Author
- QMRF Number

The associated structures (e.g. in the training or test set), which have been imported in the QSAR Model Database, can be searched by:

- Exact structure
    - CAS Registry number
    - Formula
    - Chemical name
    - Alias
    - SMILES
- Similarity
- Substructure (Structure drawing)

Either exact or similar structure searching is supported. Similarity can be defined through selectable Tanimoto distance.

Search results can be displayed (view QMRF) and then sorted by different criteria – QMRF Number, Title, Endpoint or Last update.

Documents can be viewed in) in HTML, PDF or Word format (see JRC standard output). The document in XML format can be downloaded and saved for further use.

Step-by Step Guidance on navigating the JRC QSAR Database is provided at page 28.

QMRF Submission

A QMRF Editor (v 2.0.0), which provides a user-friendly way of describing a QSAR model by filling in the QMRF, is accessible as a Java standalone desktop application, available at Source Forge net.

http://qmrf.sourceforge.net/

This can also be downloaded from the QSAR Database home page (left menu).

 Sections from 1 to 9 have to be populated by the author. Section 10 will be filled in by JRC before making the QMRF publicly accessible. Detailed description of the fields in all these sections is provided in this tutorial.

Please send us your models to have them included in the database:

JRC-COMPUTOX@ec.europa.eu

Submission of new QMRF documents to the database is handled entirely by the administrator.

Draft versions of QMRFs, as well as attachments with training and/or test set data, have to be sent to the dedicated mailbox. Supported types of "Training set" and "Test set" attachment file formats include SDF, MOL, CSV and XLS. Any other relevant information can be sent in DOC or PDF format.

The detailed QSAR method reporting format to be used for the model descriptions is provider from page 25 onwards.
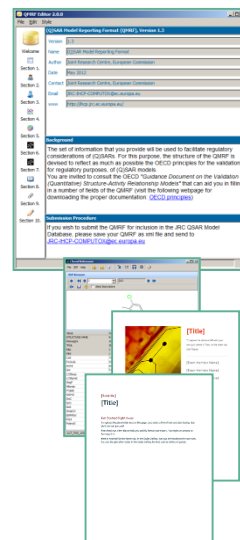
QMRF Reviewing

The entire information content of JRC QSAR Model Database is reviewed by experts in the field considering the following criteria with periodically releases of the revised reports for publication:

- The JRC will perform a quality control, but Inclusion of the model in the JRC 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.

- Properly documented summaries of (Q)SARs (i.e. robust summaries) will be included in the JRC QSAR Model Database.

- The JRC QSAR Model Database will help to identify valid QSARs. e.g. for the purposes of REACH.

Depending on the outcome of the review, the document would be either published in the JRC QSAR Model Database or returned back to the author for further revision(s).

A unique **QMRF Numeric Identifier** is assigned when the QMRF is published. Further details about its format and semantics are provided on page 24.

In case of minor updates already published reports can be republished with the same registration number upon request.

Link to the guidelines for reviewing

From Submission to Publication

Actors

- AUTHOR
- REVIEWER
- EC JRC

- Identification of reviewers
- Storage, completeness check, feedback to author
- Submission of Draft Report
- Review of draft QMRF
- Feed back to authors
- Revision of draft QMRF
- Acceptance
- Final check and edits
- PUBLICATION

Endpoint classification 1

QMRF 1.	**Physical Chemical Properties**	QMRF 1.11	Adsorption/Desorption	
		QMRF 1. 9.	Air- water partition coefficient (Henry's law constant, H)	
		QMRF 1.22.	Auto-Ignition	
		QMRF 1.24.	Average Molecular Weight of Polymers	
		QMRF 1. 2.	Boiling point	
		QMRF 1.12.	Complex Formation Ability in Water	
		QMRF 1.13.	Density	
		QMRF 1.10.	Dissociation constant (pKa)	
		QMRF 1.21.	Explosive Properties	
		QMRF 1.18.	Fat Solubility	
		QMRF 1.20.	Flammability	
		QMRF 1.19.	Flash point	
		QMRF 1.15.	Hydrolysis	
		QMRF 1.26.	Length Weighted Geometric Mean Diameter of Fibres	
		QMRF 1. 1.	Melting point	
		QMRF 1. 8.	Octanol-air partition coefficient (Koa)	
		QMRF 1. 7.	Octanol-water distribution coefficient (D)	
		QMRF 1. 6.	Octanol-water partition coefficient (Kow)	
			EC A.8	Partition Coefficient (EU method includes both shake flask and HPLC)
			OECD 123	Partition Coefficient (n-Octanol/Water): Slow-Stirring Method
			OECD 117	Partition Coefficient (n-octanol/water) HPLC Method
			OECD 107	Partition Coefficient (n-octanol/water); Shake Flask Method
		QMRF 1.23.	Oxidizing Properties	
		QMRF 1.14.	Particle Size Distribution	
		QMRF 1.25.	Solution/Extraction Behaviour of Polymers in Water	
		QMRF 1.16.	Stability	
		QMRF 1. 5.	Surface tension	
		QMRF 1. 4.	Vapour pressure	
		QMRF 1.17.	Viscosity	
		QMRF 1. 3.	Water solubility	

Endpoint classification 2

QMRF 2.	**Environmental fate parameters**	QMRF 2. 8.	Adsorption/Desorption in sediment	
		QMRF 2. 7.	Adsorption/Desorption in soil	
		QMRF 2. 5.a.	Bioaccumulation. BAF fish	
		QMRF 2. 5.b.	Bioaccumulation. BAF other organisms	
		QMRF 2. 4.a.	Bioconcentration . BCF fish	
			EC C.13	Bioconcentration: Flow-through Fish Test
			OECD 305	Bioconcentration: Flow-through Fish Test
		QMRF 2. 4.b.	Bioconcentration . BCF other organisms	
		OECD 228	Determination of Developmental Toxicity of a Test Chemical to Dipteran Dung Flies	
		QMRF 2. 6.	Partition coefficient. Organic carbon-sorption partition coefficient (organic carbon; Koc)	
		QMRF 2. 2.a.	Persistence: Abiotic degradation in air (Phototransformation). Direct photolysis	
		QMRF 2. 2.b.	Persistence: Abiotic degradation in air (Phototransformation). Indirect photolysis (OH-radical reaction, ozone-radical reaction, other)	
		QMRF 2. 1.a.	Persistence: Abiotic degradation in water. Hydrolysis	
		QMRF 2. 1.c.	Persistence: Abiotic degradation in water. Other	
		QMRF 2. 1.b.	Persistence: Abiotic degradation in water. Oxidation	
		QMRF 2. 3.b.	Persistence: Biodegradation. Biodegradation time frame (primary, ultimate degradation)	
		QMRF 2. 3.a.	Persistence: Biodegradation. Ready/not ready biodegradability	
		QMRF 2.10.	Vegetation-air partition coefficient	
		QMRF 2.11.	Vegetation-soil partition coefficient	
		QMRF 2. 9.	Vegetation-water partition coefficient	

Endpoint classification 3

QMRF 3.	Ecotoxic effects	QMRF 3. 3.	Acute toxicity to fish (lethality)	
			EC C. 1.	Acute Toxicity for Fish
			OECD 203	Fish, Acute Toxicity Test
		QMRF 3. 4.	Long-term toxicity to Daphnia (lethality, inhibition of reproduction)	
			EC C.20	Daphnia magna Reproduction Test
			OECD 211	Daphnia magna Reproduction Test
		QMRF 3. 5.	Long-term toxicity to fish (egg/sac fry, growth inhibition of juvenile fish, early life stage, full life cycle)	
		QMRF 3. 6.	Microbial inhibition (activated sludge respiration inhibition, inhibition of nitrification, other)	
			OECD 209	Activated Sludge, Respiration Inhibition Test
			EC C.11.	Biodegradation: Activated Sludge Respiration Inhibition Test
		QMRF 3. 1.	Short-term toxicity to Daphnia (immobilisation)	
			EC C. 2.	Daphnia sp Acute Immobilisation Test
			OECD 202	Daphnia sp Acute Immobilisation Test
		QMRF 3. 2.	Short-term toxicity to algae (inhibition of the exponential growth rate)	
			OECD 201	Alga Growth Inhibition Test
			EC C. 3.	Freshwater Algae and Cyanobacteria, Growth Inhibition Test
		QMRF 3.12.b.	Toxicity to birds. Long-term toxicity (survival, growth, reproduction)	
			OECD 205	Avian Dietary Toxicity Test
		QMRF 3.12.a	Toxicity to birds. Short term toxicity (feeding, gavage, other)	
			OECD 205	Avian Dietary Toxicity Test
		QMRF 3. 8.	Toxicity to earthworms (survival, growth, reproduction)	
		QMRF 3.13.b.	Toxicity to honeybees. Acute contact toxicity	
			EC C.17.	Honeybees-Acute Contact Toxicity
			OECD 214	Honeybees-Acute Contact Toxicity
		QMRF 3.13.a.	Toxicity to honeybees. Acute oral toxicity	
		QMRF 3. 9.	Toxicity to plants (leaves, seed germination, root elongation)	
		QMRF 3.11.	Toxicity to sediment organisms (survival, growth, reproduction)	
		QMRF 3.10.	Toxicity to soil invertebrates (survival, growth, reproduction)	
		QMRF 3. 7.	Toxicity to soil microorganisms (inhibition of C-mineralisation, inhibition of N-mineralisation, other)	

Endpoint classification 4

QMRF 4.	**Human Health Effects**	QMRF 4. 3.	Acute Dermal toxicity	
		QMRF 4. 1.	Acute Inhalation toxicity	
		QMRF 4. 2.	Acute Oral toxicity	
			OECD 423	Acute Oral Toxicity - Acute Toxic Class Method
			OECD 401	Acute Oral Toxicity DELETED
			OECD 420	Acute Oral Toxicity-Fixed Dose Method
			OECD 425	Acute Oral Toxicity: Up-and-Down Procedure
			EC B. 1.	Acute Toxicity (Oral)
			EC B. 1.tris.	Acute Toxicity (Oral) Acute Toxic Class Method
			EC B. 1.bis.	Acute Toxicity (Oral) Fixed Dose Method
		QMRF 4. 5.	Acute photoirritation	
		QMRF 4.12.	Carcinogenicity	
			OECD 451	Carcinogenicity Studies
			EC B.32.	Carcinogenicity Test
			OECD 453	Combined Chronic Toxicity/Carcinogenicity Studies
			EC B.33.	Combined Chronic Toxicity/Carcinogenicity Test
		QMRF 4.18.c.	Endocrine Activity. Other (e.g. inhibition of specific enzymes involved in hormone synthesis or regulation, specify enzyme(s) and hormone)	
		QMRF 4.18.a.	Endocrine Activity. Receptor-binding (specify receptor)	
			OECD 441	The Hershberger Bioassay in rats A Short-term Screening Assay for (Anti) Androgenic Properties
			OECD 440	Uterotrophic Bioassay in Rodents: a short-term screening test for oestrogenic properties
		QMRF 4. 9.	Eye irritation/corrosion	
		QMRF 4.15.	In vitro reproductive toxicity (e.g. embryotoxic effects in cell culture such as embryo stem cells)	
		QMRF 4.17.	In vivo pre-, peri-, post natal development and / or fertility (1 or 2 generation	
		QMRF 4.16.	In vivo pre-natal-developmental toxicity	
		QMRF 4.10.	Mutagenicity	
			OECD 471	Bacterial Reverse Mutation Test
			OECD 482	DNA Damage and Repair, Unscheduled DNA Synthesis in Mammalian Cells in Vitro
			EC B.18.	DNA Damage and Repair- Unscheduled DNA Synthesis - Mammalian Cells in Vitro

Endpoint classification 5

QMRF 4.	**Human Health Effects**		EC B.15.	Gene Mutation Saccharomyces Cerevisiae
			EC B.17.	In Vitro Mammalian Cell Gene Mutation Test
			OECD 476	In Vitro Mammalian Cell Gene Mutation Test
			OECD 473	In Vitro Mammalian Chromosome Aberration Test
			OECD 479	In Vitro Sister Chromatid Exchange Assay in Mammalian Cells
			EC B.21.	In vitro Mammalian Cell Transformation Tests
			OECD 487	In vitro Micronucleous Test
			OECD 474	Mamm Erythrocyte Micronucleus Test
			OECD 475	Mammalian Bone Marrow Chromosome Aberration Test
			EC B.23.	Mammalian Spermatogonial Chromosome Aberration Test
			OECD 483	Mammalian Spermatogonial Chromosome Aberration Test
			EC B.16.	Mitotic Recombination Saccharomyces Cerevisiae
			EC B.25.	Mouse Heritable Translocation
			OECD 485	Mouse Heritable Translocation Assay
			EC B.24.	Mouse Spot Test
			OECD 484	Mouse Spot Test
			EC B.10.	Mutagenicity - In Vitro Mammalian Chromosome Aberration Test
			EC B.12.	Mutagenicity In Vivo Mamm Erythrocyte Micronucleus Test
			EC B.11.	Mutagenicity In Vivo Mammalian Bone Marrow Chromosome Aberration Test
			EC B.13/14.	Mutagenicity: Reverse Mutation Test Using Bacteria
			EC B.22.	Rodent Dominant Lethal test
			OECD 478	Rodent Dominant Lethal test
			OECD 480	Saccharomyces Cerevisiae, Gene Mutation Assay
			OECD 481	Saccharomyces Cerevisiae, Mitotic Recombination Assay
			EC B.20.	Sex-Linked recessive Lethal Test in Drosophila Melanogaster
			OECD 477	Sex-Linked recessive Lethal Test in Drosophila Melanogaster
			EC B.19.	Sister Chromatid Exchange Assay In Vitro
			EC B.39.	Unscheduled DNA Syntesis (UDS) Test with Mammalian Liver Cells In Vivo
			OECD 486	Unscheduled DNA Syntesis (UDS) Test with Mammalian Liver Cells In Vivo

Endpoint classification 6

QMRF 4.	**Human Health Effects**	QMRF 4.19.	Neurotoxicity	
		QMRF 4.13.	Photocarcinogenicity	
		QMRF 4.11.	Photomutagenicity	
		QMRF 4. 8.	Photosensitisation	
		QMRF 4.18.b.	Receptor binding and gene expression (specify receptor)	
		QMRF 4.14.	Repeated dose toxicity	
			OECD 452	Chronic Toxicity Studies
			EC B.30.	Chronic Toxicity Test
			OECD 453	Combined Chronic Toxicity/Carcinogenicity Studies
			EC B.33.	Combined Chronic Toxicity/Carcinogenicity Test
			OECD 422	Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test
			EC B.38.	Delayed Neurotoxicity of Organophosphorus Substances 28 Days Repeated Dose Study
			EC B. 9.	Repeated Dose (28 Days) Toxicity (Dermal)
			EC B. 7.	Repeated Dose (28 Days) Toxicity (Oral)
			EC B. 8.	Repeated Dose (28 Days) Toxicity Inhalation
			EC B.27.	Repeated Dose 90-day Oral Toxicity Study in Non-Rodents
			OECD 409	Repeated Dose 90-day Oral Toxicity Study in Non-Rodents
			EC B.26.	Repeated Dose 90-day Oral Toxicity Study in Rodents
			OECD 408	Repeated Dose 90-day Oral Toxicity Study in Rodents
			OECD 410	Repeated Dose Dermal Toxicity: 21/28 Day
			OECD 412	Repeated Dose Inhalation Toxicity: 28/14-Day
			OECD 407	Repeated Dose Oral Toxicity-Rodent 28/14-Days
			OECD 419	Subchronic Delayed Neurotoxicity of Organophosphorus Substances: 28-Day
			EC B.28.	Subchronic Dermal Toxicity Study: 90-Day Repeated Dermal Dose Study Using Rodent Species
			OECD 411	Subchronic Dermal Toxicity: 90-Day
			EC B.29.	Subchronic Inhalation Toxicity Study: 90-Day Repeated Inhalation Dose Study Using Rodent Species
			OECD 413	Subchronic Inhalation Toxicity: 90-Day
		QMRF 4. 7.	Respiratory sensitisation	
		QMRF 4. 4.	Skin irritation /corrosion	

Endpoint classification 7

QMRF 4.		QMRF 4. 6.	Skin sensitisation	
			OECD 429	LLNA
			EC B. 6.	Skin Sensitisation
			EC B. 6.	Skin Sensitisation
			OECD 406	Skin Sensitisation
			OECD 406	Skin Sensitisation
			EC B.42.	Skin Sensitisation: Local Lymph Node Assay
QMRF 5.	**Toxicokinetics**	EC B.36.	Toxicokinetics	
		EC B.36.	Toxicokinetics	
		OECD 417	Toxicokinetics	
		OECD 417	Toxicokinetics	
		QMRF 5. 4.	Toxicokinetics.Blood-brain barrier penetration	
		QMRF 5. 7.	Toxicokinetics.Blood-lung barrier penetration	
		QMRF 5. 6.	Toxicokinetics.Blood-testis barrier penetration	
		QMRF 5.10.	Toxicokinetics.DNA-binding	
		QMRF 5. 3.	Toxicokinetics.Gastrointestinal absorption	
		QMRF 5. 8.	Toxicokinetics.Metabolism (including metabolic clearance)	
		QMRF 5. 2.	Toxicokinetics.Ocular membrane penetration	
		QMRF 5. 5.	Toxicokinetics.Placental barrier penetration	
		QMRF 5. 9.	Toxicokinetics.Protein-binding	
		QMRF 5. 1.	Toxicokinetics.Skin penetration	
QMRF 6.	**Other**	OECD 5XX	Crop Field Trial Test Guideline	
		OECD 508	Magnitude of Pesticide Residues in Processed Commodities	
		OECD 507	Nature of Pesticide Residues in Processed Commodities High Temperature Hydrolysis	
		QMRF 6. 6.	Other	
		OECD 505	Residues in Livestock	
		OECD 504	Residues in Rotational Crops (limited Field Studies)	
		OECD 506	Stability of Pesticide Residues in Stored Commodities	

Creating a QSAR Model Report

Please, try to fill in the fields of the QMRF for the model of interest. If the field is not pertinent with the model you are describing, or if you cannot provide the requested information, please answer "no information available". **The set of information that you provide will be used to facilitate regulatory considerations of (Q)SARs.** For this purpose, the structure of the QMRF is devised to reflect as much as possible the OECD principles for the validation, for regulatory purposes, of (Q)SAR models. You are invited to consult the OECD "Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship Models" that can aid you in filling in a number of fields of the QMRF.

Step-by Step Guidance on the compilation of QSAR Models for the purpose of publication via the JRC QSAR Model database is provided from page 28 onwards.

The QSAR Model Reporting Format (QMRF) is summarised hereafter covering the following sections:

1. QSAR identifier

1.1 QSAR identifier (title): Provide a short and indicative title for the model including relevant keyword. Some possible keywords are: endpoint modelled (as specified in field 3.2, recommended), name of the model, name of the modeller, and name of the software coding the model. *Examples: "BIOWIN for Biodegradation"; "TOPKAT Developmental Toxicity Potential Aliphatic Model".*
1.2 Other related models: If appropriate, identify any model that is related to the model described in the present QMRF. *Example: "TOPKAT Developmental Toxicity Potential Heteroaromatic Model and TOPKAT Developmental Toxicity Potential Carboaromatic Model" (these two models are related to the primary model "TOPKAT Developmental Toxicity Potential Aliphatic Model").*
1.3 Software coding the model: If appropriate, specify the name and the version of the software that implements the model. *Examples: "BIOWIN v. 4.2 (EPI Suite)"; "TOPKAT v. 6.2".*

2. General information

2.1 Date of QMRF: Report the date of QMRF drafting (day/month/year). *Example: "5 November 2006".*
2.2 QMRF author(s) and contact details: Indicate the name and the contact details of the author(s) of the QMRF (first version of the QMRF).
2.3 Date of QMRF update(s): Indicate the date (day/month/year) of any update of the QMRF. The QMRF can be updated for a number of reasons such as additions of new information (e.g. addition of new validation studies in section 7) and corrections of information.
2.4 QMRF update(s): Indicate the name and the contact details of the author(s) of the updates QMRF (see field 2.3) and list which sections and fields have been modified.

2.5 Model developer(s) and contact details: Indicate the name of model developer(s)/author(s), and the corresponding contact details; possibly report the contact details of the corresponding author.

2.6 Date of model development and/or publication: Report the year of release/publication of the model described in the current QMRF.

2.7 Reference(s) to main scientific papers and/or software package: List the main bibliographic references (if any) to original paper(s) explaining the model development and/or software implementation. Any other reference such as references to original experimental data and related models can be reported in field 9.2 "Bibliography".

2.8 Availability of information about the model: Indicate whether the model is proprietary or non-proprietary and specify (if possible) what kind of information about the model cannot be disclosed or are not available (e.g., training and external validation sets, source code, and algorithm). ***Example: "The model is non-proprietary but the training and test sets are not available"; "The model is proprietary and the algorithm and the data sets are confidential".***

2.9 Availability of another QMRF for exactly the same model: Indicate if you are aware or suspect that another QMRF is available for the current model you are describing. If possible, identify this other QMRF.

3. Defining the endpoint – OECD Principle 1

PRINCIPLE 1: "A DEFINED ENDPOINT". ENDPOINT refers to any physicochemical, biological, or environmental effect that can be measured and therefore modelled. The intent of PRINCIPLE 1 (a (Q)SAR should be associated with a defined endpoint) is to ensure clarity in the endpoint being predicted by a given model, since a given endpoint could be determined by different experimental protocols and under different experimental conditions. It is therefore important to identify the experimental system that is being modelled by the Q)SAR.

3.1 Species: Indicate the species for the endpoint being modelled.

3.2 Endpoint: Choose the endpoint (physicochemical, biological, or environmental effect) from the pre-defined classification. If the pre-defined classification does not include the endpoint of interest, select "Other" and report the endpoint in the subsequent field 3.3.

3.3 Comment on the endpoint: Include in this field any other information to define the endpoint being modelled. Specify the endpoint further if relevant, e.g. according to test organism such as species, strain, sex, age or life stage; according to test duration and protocol; according to the detailed nature of endpoint etc. You can also define here the endpoint of interest in case this is not listed in the pre-defined classification (see field 3.2) or you can add information about a second endpoint modelled by the same model. ***Example: Nitrate radical degradation rate constant: kNO3.***

3.4 Endpoint units: Specify the units of the endpoint measured.

3.5 Dependent variable: Specify the relationship between the dependent variable being modelled and the endpoint measured since the two quantities may be different. ***Example: For modelling purposes all rate constants (i.e. Nitrate radical degradation rate constant***

kNO3) were transformed to logarithmic units and multiplied by −1 to obtain positive values. The dependent variable is: −log(kNO3).

3.6 Experimental protocol: Make any useful reference to a specific experimental protocol (or protocols) followed in the collection and evaluation of the experimental data sets.

3.7 Endpoint data quality and variability: Provide available information about the test data selection and evaluation and include a description of the data quality used to develop the model. This includes provision of information about the variability of the test data, i.e. repeatability (variability over time) and reproducibility (variability between laboratories) and sources of error (confounding factors which may influence testing results).

4. Defining the algorithm – OECD Principle 2

PRINCIPLE 2: "AN UNAMBIGUOUS ALGORITHM". The (Q)SAR estimate of an endpoint is the result of applying an ALGORITHM to a set of structural parameters which describe the chemical structure. The intent of PRINCIPLE 2 (a (Q)SAR should be associated with a unambiguous algorithm) is to ensure transparency in the model algorithm that generates predictions of an endpoint from information on chemical structure and/or physicochemical properties. In this context, algorithm refers to any mathematical equation, decision rule or output approach.

4.1 Type of model: Describe the type of model (e.g., SAR, QSAR, Expert System, Neural Network, etc.).

4.2 Explicit algorithm: Report the algorithm (only the algorithm) for generating predictions from the descriptors; more text information about the algorithm can be reported in the following fields of this section or as supporting information (see field 9.3). If the algorithm is too long and complicated and thus cannot be reported here, include in this field a reference to a paper or a document where the algorithm is described in detail. This material can be attached as supporting information.

4.3 Descriptors in the model: Identify the number and the name or identifier of the descriptors included in the model. In this context, descriptors refers to e.g. physicochemical parameters, structural fragments etc

4.4 Descriptor selection: Indicate the number and the type (name) of descriptors /decision rules initially screened, and explain the method used to select the descriptors and develop the model from them.

4.5 Algorithm and descriptor generation: Explain the approach used to derive the algorithm and the method (approach) used to generate each descriptor.

4.6 Software name and version for descriptor generation: Specify the name and the version of the software used to generate the descriptors. If relevant, report the specific settings chosen in the software to generate a descriptor.

4.7 Chemicals/ Descriptors ratio: Report the following ratio: number of chemicals (chemicals from the training set) to number of descriptors , if applicable (if not, explain why).

5. Defining the applicability domain – OECD Principle 3

PRINCIPLE 3: "A DEFINED DOMAIN OF APPLICABILITY". APPLICABILITY DOMAIN refers to the response and chemical structure space in which the model makes predictions with a given reliability. Ideally the applicability domain should express the structural, physicochemical and response space of the model. The CHEMICAL STRUCTURE (x variable) space can be expressed by information on physicochemical properties and/or structural fragments. The RESPONSE (y variable) can be any physicochemical, biological or environmental effect that is being predicted. According to PRINCIPLE 3 a (Q)SAR should be associated with a defined domain of applicability. Section 5 can be repeated (e.g., 5.a, 5.b, 5.c, etc) as many time as necessary if more than one method has been used to assess the applicability domain.

5.1 Description of the applicability domain of the model: Describe the response and chemical structure and/or descriptor space in which the model makes predictions with a given reliability. Discuss if relevant whether: a) fixed or probabilistic boundaries define the applicability domain; b) structural features, a descriptor or a response space defines the applicability domain; c) in the case of SAR, there exists a description of the limits on its applicability (inclusion and/or exclusion rules regarding the chemical classes to which the substructure is applicable); d) in the case of SAR, there exist rules describing the modularity effects of the substructure's molecular environment; e) in the case of QSAR, there exist inclusion and/or exclusion rules that define the descriptor variable ranges for which the QSAR is applicable; f) in the case of QSAR, there exist inclusion and/or exclusion rules that define the response variable ranges for which the QSAR is applicable; g) there exists a (graphical) expression of how the descriptor values of the chemicals in the training set are distributed in relation to the endpoint values predicted by the model.

5.2 Method used to assess the applicability domain: Describe the method used to assess the applicability domain of the model.

5.3 Software name and version for applicability domain assessment: Specify the name and the version of the software used to apply the applicability domain method, where applicable. If relevant, report the specific settings chosen in the software to apply the method.

5.4 Limits of applicability: Describe for example the inclusion and/or exclusion rules (fixed or probabilistic boundaries, structural features, descriptor space, response space) that define the applicability domain.

6. Defining goodness-of-fit and robustness – OECD Principle 4

PRINCIPLE 4: "APPROPRIATE MEASURES OF GOODNESS-OF-FIT, ROBUSTENESS AND PREDICTIVITY". PRINCIPLE 4 expresses the need to perform validation to establish the performance of the model. GOODNESS-OF-FIT and ROBUSTNESS refer to the internal model performance.

6.1 Availability of the training set: Indicate whether the training set is somehow available (e.g., published in a paper, embedded in the software implementing the model, stored in a database) and appended to the current QMRF as supporting information (field 9.3). If it is not available, explain why. ***Example: "It is available and attached" "It is available but not attached"; "It is not available because the data set is proprietary"; "The data set could not be retrieved".***

6.2 Available information for the training set: Indicate whether the following information for the training set is reported as supporting information (see field 9.3): a) Chemical names (common names and/or IUPAC names); b) CAS numbers; c) SMILES; d) InChI codes; e) MOL files; f) Structural formula; g) Any other structural information.

6.3 Data for each descriptor variable for the training set: Indicate whether the descriptor values of the training set are available and are attached as supporting information (see field 9.3).

6.4 Data for the dependent variable (response) for the training set: Indicate whether dependent variable values of the training set are available and attached as supporting information (see field 9.3).

6.5 Other information about the training set: Indicate any other relevant information about the training set (e.g. number and type of compounds in the training set (e.g. for models predicting positive and negative results the number of positives and the number of negatives in the training set)).

6.6 Pre-processing of data before modelling: Indicate whether raw data have been rocessed before modelling (e.g. averaging of replicate values); if yes, report whether both raw data and processed data are given.

6.7 Statistics for goodness-of-fit: Report here goodness-of-fit statistics (r^2, r^2 adjusted, standard error, sensitivity, specificity, false negatives, false positives, predictive values etc).

6.8 Robustness – Statistics obtained by leave-one-out cross-validation: Report here the corresponding statistics.

6.9 Robustness – Statistics obtained by leave-many-out cross-validation: Report here the corresponding statistics, the strategy for splitting the data set (e.g. random, stratified), the percentage of left out compounds and the number of cross-validations.

6.10 Robustness – Statistics obtained by Y-scrambling: Report here the corresponding statistics and the number of iterations.

6.11 Robustness – Statistics obtained by bootstrap: Report here the corresponding statistics and the number of iterations.

6.12 Robustness – Statistics obtained by other methods: Report here the corresponding statistics.

7. Defining predictivity – OECD Principle 4

PRINCIPLE 4: "APPROPRIATE MEASURES OF GOODNESS-OF-FIT, ROBUSTENESS AND PREDICTIVITY". PRINCIPLE 4 expresses the need to perform validation to establish the performance of the model. PREDICTIVITY refers to the external model validation. Section 7 can be repeated (e.g., 7.a, 7.b, 7.c, etc) as many time as necessary if more validation studies needs to be reported in the QMRF.

7.1 Availability of the external validation set: Indicate whether an external validation set is available and appended to the current QMRF as supporting information (field 9.3). If it is not available, explain why.

7.2 Available information for the external validation set: Indicate whether the following information for the external validation set is reported as supporting information (see field 9.3): a) Chemical names (common names and/or IUPAC names); b) CAS numbers; c) SMILES; d) InChI codes; e) MOL files; f) Structural formula; g) Any other structural information.

7.3 Data for each descriptor variable for the external validation set: Indicate whether descriptor values of the external validation set are somehow available and attached as supporting information (see field 9.3).

7.4 Data for the dependent variable for the external validation set: Indicate whether dependent variable values of the external validation set are somehow available and attached as supporting information (see field 9.3).

7.5 Other information about the external validation set: Indicate any other relevant information about the validation set. ***Example: "External validation set with 56 compounds appended".***

7.6 Experimental design of test set: Indicate any experimental design for getting the test set (e.g. by randomly setting aside chemicals before modelling, by literature search after modelling, by prospective experimental testing after modelling, etc.).

7.7 Predictivity – Statistics obtained by external validation: Report here the corresponding statistics. In the case of classification models, include false positive and negative rates.

7.8 Predictivity – Assessment of the external validation set: Discuss whether the external validation set is sufficiently large and representative of the applicability domain. Describe for example the descriptor and response range or space for the validation test set as compared with that for the training set. Here the descriptor values of the chemicals predicted by the model (training set) should be compared with the descriptor value range of the test set. In addition the distribution of the response values of the chemicals in the training set should be compared to the distribution of the response values of the test set.

7.9 Comments on the external validation of the model: Add any other useful comments about the external validation procedure.

8. Providing a mechanistic interpretation – OECD Principle 5

PRINCIPLE 5: "A MECHANISTIC INTERPRETATION, IF POSSIBLE". According to PRINCIPLE 5, a (Q)SAR should be associated with a mechanistic interpretation, if possible.

8.1 Mechanistic basis of the model: Provide information on the mechanistic basis of the model (if possible). In the case of SAR, you may want to describe (if possible) the molecular features that underlie the properties of the molecules containing the substructure (e.g. a description of how sub-structural features could act as nucleophiles or electrophiles, or form part or all of a receptor-binding region). In the case of QSAR, you may give (if possible) a physicochemical interpretation of the descriptors used (consistent with a known mechanism of biological action). If it is not possible to provide a mechanistic interpretation, try to explain why.

8.2 A priori or a posteriori mechanistic interpretation: Indicate whether the mechanistic basis of the model was determined a priori (i.e. before modelling, by ensuring that the initial set of training structures and/or descriptors were selected to fit pre-defined mechanism of action) or a posteriori (i.e. after modelling, by interpretation of the final set of training structures and or descriptors).

8.3 Other information about the mechanistic interpretation: Report any other useful information about the (purported) mechanistic interpretation described in the previous fields (8.1 and 8.2) such as any reference supporting the mechanistic basis.

9. Miscellaneous information

9.1 Comments: Add here other relevant and useful comments (e.g. other related models, known applications of the model) that may facilitate regulatory considerations on the model described. Include if relevant experience obtained by use of model prediction for various types of regulatory decisions (incl. references as appropriate).

9.2 Bibliography: Report useful references other than those directly associated with the model development (references describing the model development are reported in <u>field 2.5</u>).

9.3 Supporting information: Indicate whether supporting information is attached (e.g. external documents) to this QMRF and specify its content and possibly its utility.

10. Summary for the JRC QSAR Model Database (compiled by JRC)

The summary section is specific for the JRC QSAR Model Database. If the model is submitted to JRC for inclusion in the JRC Database of QSAR models, then this summary is compiled by JRC after QMRF submission. **The QMRF author does not have to fill in any of the fields of the summary section.**

10.1 QMRF number: A unique number (numeric identifier) is assigned to any QMRF that is published in the JRC QSAR Model Database. The number encodes the following information: Q YEAR-ENDPOINT-No *Example: Q11-417-002 refers to a QMRF published in 2011, for the endpoint 4.17. It is the second QMRF published in 2011.* The number is unique for any QMRF uploaded and stored in the JRC QSAR Model Database.

10.2 Publication date: The date (day/month/year) of publication in the JRC Database is reported here.

10.3 Keywords: Any relevant keywords associated with the present QMRF are reported here.

10.4 Comments: Any comments that are relevant for the publication of the QMRF in the JRC Database (e.g., comments about updates and about supporting information) are reported here.

Supporting Information – QMRF 2.0.0

Templates will be provided for submitting information about the training and test sets. Storage of searchable information about the training and the test sets in the database of the JRC Database will be possible if the submitter uses specific file formats (Excel file or preferably SDF files) with predefined fields (most important in bold):

- o   Chemical Name (IUPAC)
- o   Chemical Name (Not IUPAC)
- o   CAS Number
- o   SMILES
- o   InChᵢ

MOL *(file name is reported for Excel files; if it is an SDF file, coordinate can be simply included in it).*

- o   Structural Formula
- o   Dependent Variable

Descriptor1 Value *(the name of the descriptors should be specified by the user)*

- o   Descriptor 2 Value
- o   Descriptor 3 Value
- o   Descriptor 4 Value
- o   Descriptor X Value

The template for the test set is identical to the one for the training set. Any other supporting information has to be provided in PDF format.

QMRF Numeric Identifier

QMRF numbering convention is set up as follows: **Q YEAR-ENDPOINT-No**

Example: Q11-417-002 refers to a QMRF published in 2011, for the endpoint 4.17. It is the second QMRF published in 2011.

Q	YEAR	ENDPOINT	NO
Q	11	4.17	002

1. **Q** is a prefix always used to introduce the QMRF numeric identifier
2. **Year**: This number (e.g. 11) identifies the year of publication in the JRC QSAR Model Database.
3. **Endpoint**: The Endpoint number (e.g. 4.17) indicates the endpoint selected for the model in the course of the report preparation.
4. **Number:** This number shows the number of reports released until the day of publishing (e.g. 002). The number always consists of three digits, in which the ten and hundred positions are filled with zeros, if not available.

The combination of the information reported in these four fields should always result in a unique numeric identifier.

Standard Output (editor v.1.3)

QMRFs can be printed in a readable output format (HTML, PDF or XLS).
The implemented standard output is shown on the following pages.

QMRF identifier (JRC Inventory):	
QMRF Title:	
Printing Date:	

1. QSAR identifier

1.1. QSAR identifier (title)

1.2. Other related models

1.3. Software coding the model

2. General information

2.1. Date of QMRF

2.2. QMRF author(s) and contact details

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

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

2.8. Availability of information about the model

2.9. Availability of another QMRF for exactly the same model

3. Defining the endpoint - OECD Principle 1: A QSAR should be associated with a defined endpoint) is to ensure clarity in the endpoint being predicted by a given model, since a given endpoint could be determined by different experimental protocols and under different experimental conditions. It is therefore important to identify the experimental system that is being modelled by the QSAR.

3.1. Species

3.2. Endpoint

3.3. Comment on endpoint

3.4. Endpoint units

3.5. Dependent variable

3.6. Experimental protocol

3.7. Endpoint data quality and variability

4. Defining the algorithm - OECD Principle 2: A QSAR should be associated with a unambiguous algorithm) is to ensure transparency in the model algorithm that generates predictions of an endpoint from information on chemical structure and/or physicochemical properties. In this context, algorithm refers to any mathematical equation, decision rule or output from a formalised modelling approach.

4.1. Type of model

4.2. Explicit algorithm

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

5. Defining the applicability domain - OECD Principle 3: A QSAR should be associated with a defined domain of applicability. Section 5 can be repeated (e.g., 5.a, 5.b, 5.c, etc) as many time as necessary if more than one method has been used to assess the applicability domain.

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: Expresses the need to perform validation to establish the performance of the model. GOODNESS-OF-FIT and ROBUSTNESS refer to the internal model performance.

6.1. Availability of the training set

6.2. Available information for the training set

6.3. Data for each descriptor variable for the training set

6.4. Data for the dependent variable for the training set

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

8. Providing a mechanistic interpretation - OECD Principle 5: "A MECHANISTIC INTERPRETATION, IF POSSIBLE". According to this principle, a QSAR should be associated with a mechanistic interpretation, if possible.

9. Miscellaneous information

Step by Step

1 **Go to** EU Science Hub:
https://ec.europa.eu/jrc/en/

2 **Select** Research

3 **Select** Science area: *Health and Consumer Protection*
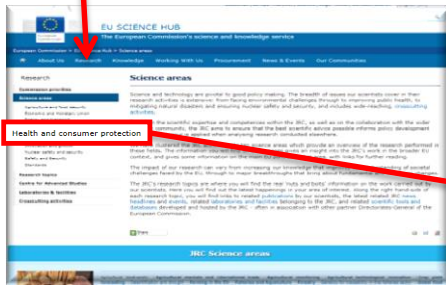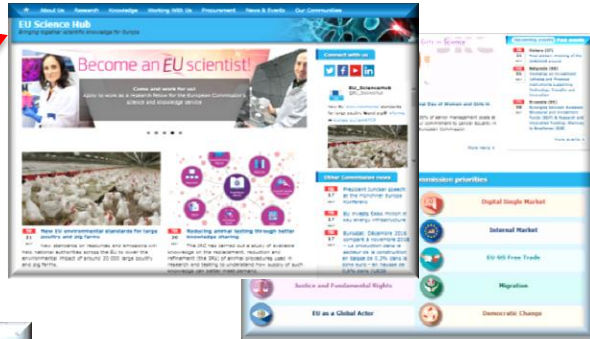
4 **Select** Alternatives to Animal testing and safety assessment of chemicals

5 **Select** Databases: *JRC QSAR Model Data base*
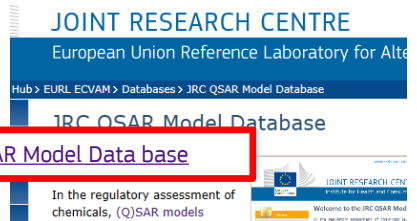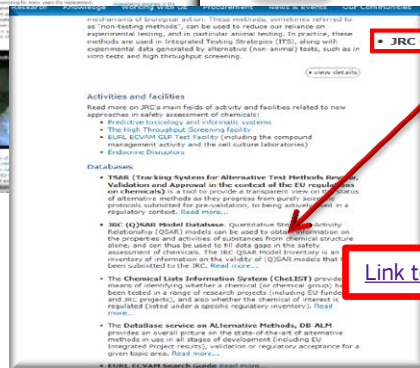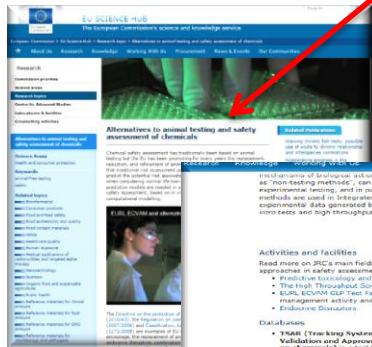
Welcome to the JRC QSAR Model Database

register

⊗

log in

Get QMRF Editor | Submit QMRF | User support | Legal Notice | Cookies | Contact | Search   English (en)

JOINT RESEARCH CENTRE
The European Commission's science and knowledge service

European Commission > EU Science Hub > EURL ECVAM > QSARDB > Home

Access
Get QMRF Editor
OECD Principles
User support

Statistics
Datasets   398
Chemical structures 57422

Welcome to the JRC QSAR Model Database

In the regulatory assessment of chemicals (e.g. under REACH), QSAR 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 QSARs and to use them to derive estimates and/or have access to their precalculated estimates. To help meet these needs, we are developing a database of QSAR models (i.e. an inventory of information on the models). The JRC QSAR Model Database is freely accessible through this web site.

The **QSAR Model Reporting Format (QMRF)** is a harmonised template for summarising and reporting key information on QSAR models, including the results of any validation studies. The information is structured according to the OECD QSAR Validation principles.

The **QSAR Prediction Reporting Format (QPRF)** is a harmonised template for summarising and reporting substance-specific predictions generated by QSAR models.

All substances, available in the QMRF Database, can be searched by exact or similar structure, or by a substructure.

Please send us your models to have them included in the database: JRC-COMPUTOX@ec.europa.eu.

Last update: 31/03/2017 | Get QMRF Editor | Submit QMRF | User support | Legal Notice | Cookies | Contact | Search | Top

6

AUTHORS:
Download the QMRF Editor
Submit a report

Functional mailbox:
JRC-COMPUTOX@ec.europa.eu

7

ALL USERS:
Read about the OECD Principles
Learn more about QSARs and find supporting documents
Access the Data base

Access
Get QMRF Editor
OECD Principles
User support
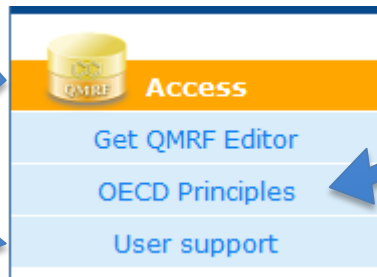
8

Start searching the Data base

Home
QMRF document search
Structures search
Endpoints

Get QMRF Editor | Submit QMRF | User support | Legal Notice | Cookies | Contact | Search   English (en)

JOINT RESEARCH CENTRE
The European Commission's science and knowledge service

European Commission > EU Science Hub > EURL ECVAM > QSARDB > QMRF document search

Home
QMRF document search
Structures search
Endpoints

Display 10 ▾ QMRF documents.   Search:

QMRF Number	Title	Endpoint	Last updated	Download
Q15-66-0018	QSARINS model for Global Half-Life Index	6.6.Other	Jun 12 2015	
Q15-26Partitioncoefficient-0017	QSARINS model for log Koc	2.6.Partitioncoefficient.Organic carbon-so	Jun 12 2015	
Q15-66-0016	QSARINS model for PBT Index	6.6.Other	Jun 12 2015	
Q15-32-0015	QSARINS model for (benzo-)triazole toxicity in Pseudokirchneriella subcapitata	3.2.Short-term toxicity to algae (inhibi)	Jun 12 2015	
Q15-41-0014	QSARINS model for inhalation toxicity of polyfluorinated compounds in mouse	4.1.Acute Inhalation toxicity	Jun 12 2015	
	QSARINS model for aquatic toxicity of			

QMRF document search

○ Title ?
● Free text ?
○ Free text (boolean) ?
○ Endpoint ?
○ Author ?
○ QMRF number ?

Max results 100

Search

9 **Decide** what to see and to do

JOINT RESEARCH CENTRE

The European Commission's science and knowledge service

European Commission › EU Science Hub › EURL ECVAM › QSARDB › QMRF document search

Home

QMRF document search

Structures search

Endpoints

QMRF document search

○ Title ?
● Free text ?
○ Free text (boolean) ?
○ Endpoint ?
○ Author ?
● QMRF number ?

Max results 100

Search

Display 10 ▼ QMRF documents.   Search:

QMRF number	Title	Endpoint	Last updated	Download
Q15-...8	QSARINS model for Global Half-Life Index	6.6.Other	Jun 12 2015	
Q15-26PartitionCoefficient-0017	QSARINS model for log Koc	2.6.PartitionCoefficient.Organic carbon-so	Jun 12 2015	
Q15-66-0016	QSARINS model for PBT Index	6.6.Other	Jun 12 2015	
Q15-32-0015	QSARINS model for (benzo-)triazole toxicity in Pseudokirchneriella subcapitata	3.2.Short-term toxicity to algae (inhibiti	Jun 12 2015	
Q15-41-0014	QSARINS model for inhalation toxicity of polyfluorinated compounds in mouse	4.1.Acute Inhalation toxicity	Jun 12 2015	
Q15-33-0013	QSARINS model for aquatic toxicity of organic chemicals in Pimephales promelas (Fathead minnow)	3.3.Acute toxicity to fish (lethality)	Jun 12 2015	
Q15-33-0012	QSARINS model for prediction of (Benzo-)Triazoles toxicity in Oncorhynchus mykiss	3.3.Acute toxicity to fish (lethality)	Jun 11 2015	
Q15-31-0011	QSARINS model for (benzo-)triazole toxicity in Daphnia magna	3.1.Short-term toxicity to Daphnia (immobi	Jun 11 2015	
Q15-35-0007	Quantitative Structure Activity Relationship for CHemical Ecotoxicity (QSARCHE) – chronic fish toxicity by	3.5.Long-term toxicity to fish (egg/sac fr	Mar 05 2015	
		...rm toxicity to fish	Mar 05 2015	

Download as ⊘⊕

chose here the number of documents to display

chose here the allowed maximum number of results for your search

Opening _qmrf_protocol.csv

You have chosen to open:

🗎 _qmrf_protocol.csv

which is: Microsoft Excel Comma Separated Values File (19.2 kB)
from: http://ict-srv09.jrc.it

What should Firefox do with this file?

○ Open with [Microsoft Excel (default) ▼]
● Save File

☐ Do this automatically for files like this from now on.

OK   Cancel

Download as 📊

File | Home | Insert | Page Layout | Formulas | Data | Review | View | Developer

A2   http://ict-srv09.jrc.it/qmrf/protocol/Q15-66-0018

qmrf_protocol.csv [Read-Only]

http://ict-srv09.jrc.it/qmrf/protocol/Q15-66-0018

	A	B	C	D	E	F	G	H	I
1	URI	Identifier	Title	Published	Submitted	Updated	Owner	Owner URI	
2	66-0018	Q15-66-00	QSARINS	published	Jun 12 2015	Jun 12 201	editor	http://ict-srv09.jrc.it/qmrf/user/U75	
3	http://ict-	Q15-26Par	QSARINS	published	Jun 12 201	Jun 12 201	editor	http://ict-srv09.jrc.it/qmrf/user/U75	
4	http://ict-	Q15-66-00	QSARINS	published	Jun 12 201	Jun 12 201	editor	http://ict-srv09.jrc.it/qmrf/user/U75	
5	http://ict-	Q15-32-00	QSARINS	published	Jun 12 201	Jun 12 201	editor	http://ict-srv09.jrc.it/qmrf/user/U75	
6	http://ict-	Q15-41-00	QSARINS	published	Jun 12 201	Jun 12 201	editor	http://ict-srv09.jrc.it/qmrf/user/U75	
7	http://ict-	Q15-33-00	QSARINS	published	Jun 12 201	Jun 12 201	editor	http://ict-srv09.jrc.it/qmrf/user/U75	
8	http://ict-	Q15-33-00	QSARINS	published	Jun 11 201	Jun 11 201	editor	http://ict-srv09.jrc.it/qmrf/user/U75	
9	http://ict-	Q15-31-00	QSARINS	published	Jun 11 201	Jun 11 201	editor	http://ict-srv09.jrc.it/qmrf/user/U75	
10	http://ict-	Q15-35-00	Quantitati	published	Mar 05 20:	Mar 05 20:	editor	http://ict-srv09.jrc.it/qmrf/user/U75	
11	http://ict-	Q15-35-00	Quantitati	published	Mar 05 20:	Mar 05 20:	editor	http://ict-srv09.jrc.it/qmrf/user/U75	
12	http://ict-	Q15-66-00	iSafeRat&	published	Jun 11 201	Jun 11 201	editor	http://ict-srv09.jrc.it/qmrf/user/U75	

qmrf_protocol

Edit   100%

Download as 📊

http://ict-srv09.jrc.it/qmrf/protocol/**Q15-66-0018**

copy here the link to the report you are interested in for further use

Search options

10

Document search

• Title: Enter partial or full QMRF document title.
• Free text: Enter a phrase in free text. There are no special operators.
• Free text (Boolean): Enter search string in implied Boolean logic. By default the words are combined with "or". More rules: + the word must be present; – the word must not be present; ( ) group expressions
• Endpoint: Enter endpoint name
• Author: Enter QMRF author name
• QMRF number: Enter QMRF number

11

Structures search

• Auto (exact structure or search by identifier: CAS, Name, EINECS, SMILES or InChI)
• Similarity (enter SMILES or draw structure)
• Substructure (enter or draw a SMARTS query)

Search by Endpoint

12

Endpoints catalog:
347 endpoints

Source:

- ENDPOINTS according to harmonised classification and labelling (CLH)
- EU Test Method numbers
- OECD Test Guideline numbers
- Technical Guidance Document

Number of reports per endpoint

Structures search
Endpoints

			Show
Code	Name		
QMRF 1.	Physical Chemical Properties		
QMRF 2.	Environmental fate parameters		
QMRF 3.	Ecotoxic effects		
QMRF 4.	Human Health Effects		
QMRF 5.	Toxicokinetics		
QMRF 6.	Other		

ing 1 to 6 of 6 entries

Show 25 entries   Search:

	Code	Name	#
	EC B.36.	Toxicokinetics	
	EC B.36.	Toxicokinetics	
	OECD 417	Toxicokinetics	
	OECD 417	Toxicokinetics	
	QMRF 5. 4.	Toxicokinetics.Blood-brain barrier penetration	1
	QMRF 5. 7.	Toxicokinetics.Blood-lung barrier penetration	
	QMRF 5. 6.	Toxicokinetics.Blood-testis barrier penetration	
	QMRF 5.10.	Toxicokinetics.DNA-binding	
	QMRF 5. 3.	Toxicokinetics.Gastrointestinal absorption	
	QMRF 5. 8.	Toxicokinetics.Metabolism (including metabolic clearance)	
	QMRF 5. 2.	Toxicokinetics.Ocular membrane penetration	
	QMRF 5. 5.	Toxicokinetics.Placental barrier penetration	
	QMRF 5. 9.	Toxicokinetics.Protein-binding	1
	QMRF 5. 1.	Toxicokinetics.Skin penetration	

Showing 1 to 14 of 14 entries

select to see the report(s) for this endpoint

QMRF 5. 4. Toxicokinetics.Blood-brain barrier penetration			1
QMRF number	Title	Last updated	Download
Q13-54-0007	QSAR for blood-brain barrier partitioning	2013-06-21	
			Download as

13

Search Document by Author

Home
QMRF document search
Structures search
Endpoints

1. **Select** Document search
2. **Type** name of Author
3. **Tick off** Author
4. **Chose** SEARCH

QMRF document search

Klaus

○ Title ?
○ Free text ?
○ Free text (boolean) ?
○ Endpoint ?
● Author ?
○ QMRF number ?

Max results 100

Search

Display 10 ▾ QMRF documents.

QMRF Number ⇕	Title ⇕
🔍 Q13-33-0041	Nonlinear QSAR: artificial neural ne... for acute aquatic toxicity

Showing 1 QMRF documents (1 to 1)

Get QMRF Editor | Submit QMRF | User support | Legal Notice | Cookies | Contact | search   English (en) ▾

JOINT RESEARCH CENTRE
The European Commission's science and knowledge service

European Commission > EU Science Hub > EURL ECVAM > QSARDB > Endpoints

Home
QMRF document search
Structures search
Endpoints

Q13-33-0041   Nonlinear QSAR: artificial neural network for acute aquatic toxicity   2013-06-27

Please wait while QMRF document chapters are loading...

Download as 📊

5. **Download as** PDF, EXCEL, WORD, XML file

Q13-33-0041   Nonlinear QSAR: artificial neural network for acute aquatic toxicity   2013-06-27

General | Endpoint | Algorithm | App. domain | Robustness | Predictivity | Interpretation | Bibliography | Summary | Attachments

1.QSAR identifier

1.1.QSAR identifier (title) Nonlinear QSAR: artificial neural network for acute aquatic toxicity
1.2.Other related models
1.3.Software coding the model

Name: ADMET Predictor™ 3.0
Description: Software for estimating certain ADMET (Absorption, Distribution, Metabolism, Elimination, and ... of a drug-like chemical from its molecular structure; 1998-2008; Simulations Plus Inc ...ons Plus, Inc. 42505 10th Street West Lancaster, CA 93534-7059 USA, Phone: ...ternational). Toll free: 888.266.9294 (in the U.S. & Canada), Fax: +1.661.723.5524...

14 **STEP by STEP** the Chapters 1 – 4

QMRF editor

Chapter 1 & 2 QSAR Identifier & General information

Chapter 3 Endpoint

Chapter 4 Algorithm

STEP by STEP the Chapters 5 – 7

QMRF editor

Chapter 5 Applicability Domain

Chapter 6 Robustness
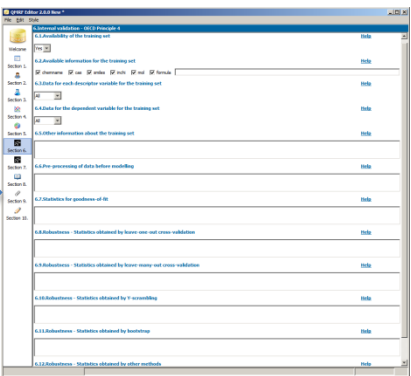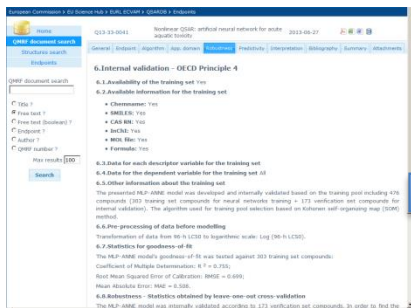
Chapter 7 Predictivity

14

STEP by STEP the Chapters 8 – 10

QMRF editor
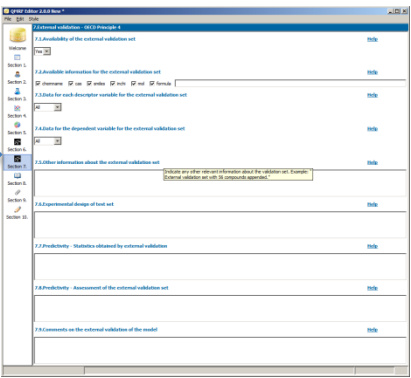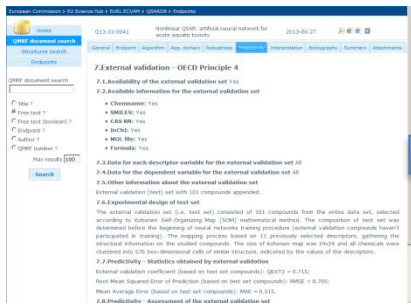
Chapter 8 Interpretation
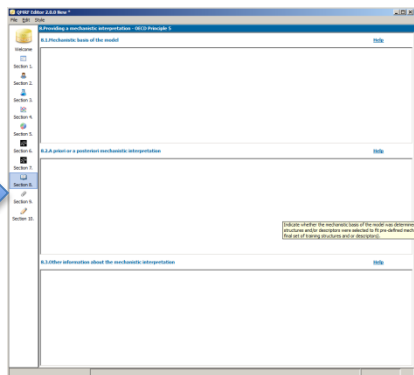
Chapter 9 Bibliography
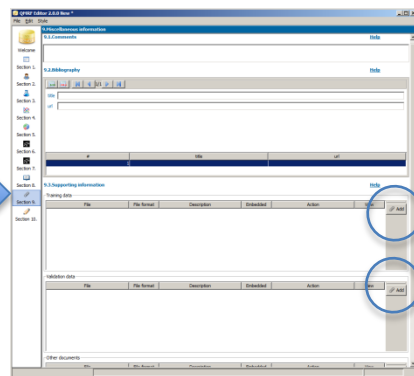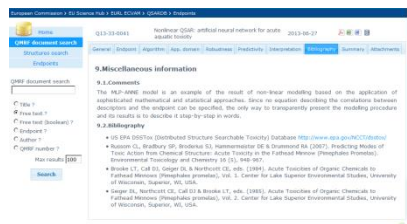
Add attachments

Chapter 10 Summary

Publication date

QMRF number

Keywords

Comments
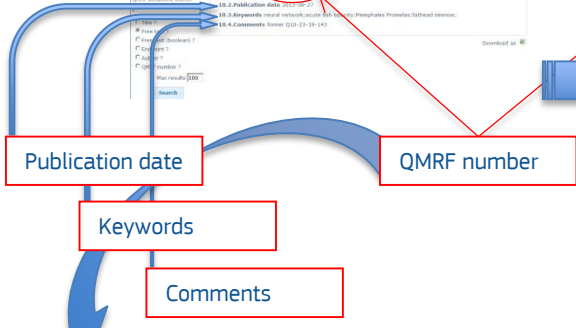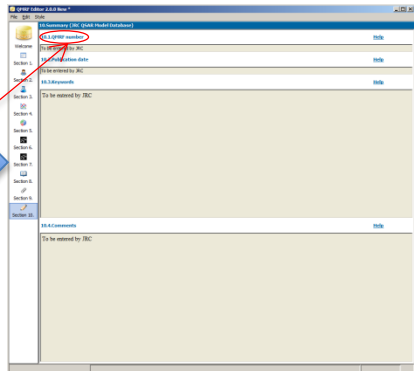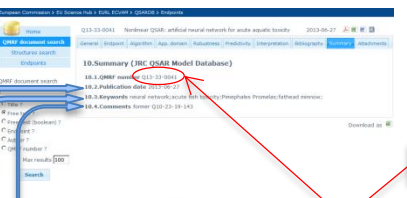
Q-YEAR-ENDPOINT-No, e.g. Q 13-33-0041 refers to a QMRF published in 2013, for the endpoint 3.3, and is the 41st QMRF to be published in 2013

Training data set

Validation data set

preferably in **sdf** format: **s**tructure-**d**ata **f**ile:Instructions on how to create a sdf file can be found on our website

Test data set

Snn files: STATISTICA **n**eural **n**etworks file

Any **other supporting document** (pdf, doc, docx, xls, xlsx)

Browse structures

Download file

Provides all information about the substances listed in the data set

All molecules listed in the training set

Show QMRF

All published reports with the selected molecule mentioned

Download a list of all published reports with the selected molecule mentioned

15   Structure search

auto (exact structure or by identifier):

Caffeine
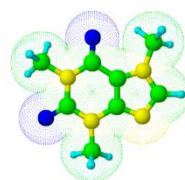
Insert caffeine
Select auto
Chose search

Similarity (enter SMILES or draw structure):

Similarity 0,6

Select similarity 0,6
Chose search

Substructure (enter or draw a SMARTS query)

Draw (sub)structure

SMILES

CN(C)C(=O)N(C)C

Select substructure
Select draw substructure
Chose submit
Select Search

15

URI uniform resource identifier
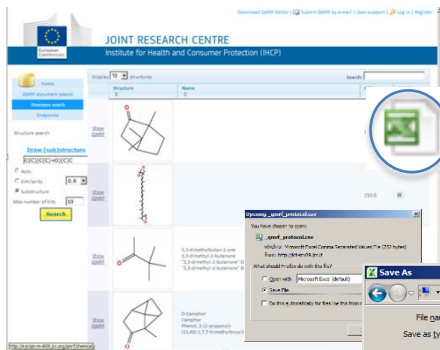
Download a list of all published reports with the selected molecule mentioned By default CSV: comma or character separated value

Save as excel work book

http://ict-srv09.jrc.it/qmrf/protocol/Q15-410-0008
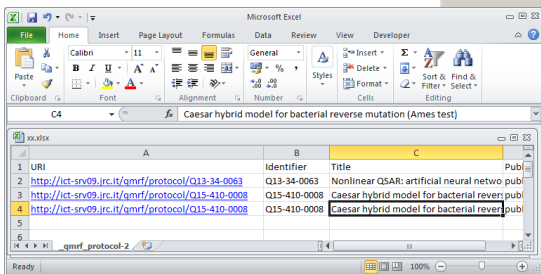
Copy the link and use it link in your blog, newsletter, article, newspaper

http://ict-srv09.jrc.it/qmrf/protocol/Q15-410-0008

Structure Data File (SDF) format

File Format Description

Structure Data Format (SDF) is a chemical file format to represent multiple chemical structure records and associated data fields. SDF was developed and published by Molecular Design Limited (MDL) and became the most widely used standard for importing and exporting information on chemicals. A chemical data file created in the Structure Data File (SDF) format is saved in plain text and contains chemical structure records. Molecular Design Limited was renamed to MDL Information Systems and then later was acquired by Symyx Technologies, the organization that now maintains the SDF format.

Below is a sample chemical record in SDF format:

```
 -CPSS- 0804941117

13 14 0 0 0 0 0 0 0 0
   0.8400  -0.1600   0.0000 N  0 0   0 0 0 0 0
   1.4800   0.4300   0.0000 N  0 0   0 0 0 0 0
   0.0900   0.2700   0.0000 N  0 0   0 0 0 0 0
   1.1100   1.2100   0.0000 C  0 0   0 0 0 0 0
   0.2700   1.1200   0.0000 C  0 0   0 0 0 0 0
   0.8400  -1.0300   0.0000 C  0 0   0 0 0 0 0
   1.5300   1.9900   0.0000 C  0 0   0 0 0 0 0
   1.0700   2.7400   0.0000 Cl 0 0   0 0 0 0 0
   1.5900  -1.4600   0.0000 C  0 0   0 0 0 0 0
   0.0800  -1.4600   0.0000 C  0 0   0 0 0 0 0
   1.5900  -2.3300   0.0000 C  0 0   0 0 0 0 0
   0.0700  -2.3200   0.0000 C  0 0   0 0 0 0 0
   0.8400  -2.7600   0.0000 C  0 0   0 0 0 0 0
 2 1 1 0 2 0 0
 3 1 1 0 2 0 0
 4 2 2 0 2 0 0
 5 3 2 0 2 0 0
 6 1 1 0 2 0 0
 7 4 1 0 2 0 0
 8 7 1 0 2 0 0
 9 6 1 0 1 0 0
10 6 2 0 1 0 0
11 9 2 0 1 0 0
12 10 1 0 1 0 0
13 12 2 0 1 0 0
 4 5 1 0 2 0 0
13 11 1 0 1 0 0
> <Sample Ref.>
OC101-12

> <Melting Point>
41.00 - 43.00

> <B1 Record No.>
304

> <ID>
304

$$$$
```

Short and clear explanation of SDF format

- The first three lines can contain general information about the molecule (e.g. substance name, version number, software used).
- The overall number of the atoms and of the bonds is stated in the fourth line in this case 7 atoms and 7 bonds.
- The following lines contain the x-, y-, and z-coordinates and the atom types of each atom in the molecule.
- At the end, the bonds between the atoms are described. The use of delocalised bond types can lead to misunderstandings. This bond type is not recommended and should not be used.


```
FLUORBENZOL
PHARMAINFORMATIC

7  7  0
    0.2162   -0.2524   -0.2302 C    0  0  0  0
   -0.7923   -0.9471   -0.9108 C    0  0  0  0
   -0.1225    0.7317    0.6962 C    0  0  0  0
    1.5136   -0.5123   -0.4863 F    0  0  0  0
   -2.1374   -0.6774   -0.6446 C    0  0  0  0
   -1.4689    1.0206    0.9734 C    0  0  0  0
   -2.4751    0.3067    0.2913 C    0  0  0  0
 1  2  1  0      0  0
 1  3  2  0      0  0
 1  4  1  0      0  0
 2  5  2  0      0  0
 3  6  1  0      0  0
 5  7  1  0      0  0
 6  7  2  0      0  0
M   END
> <IDENTIFIER NAME>
FLUORBENZOL

$$$$
Next molecule...
```
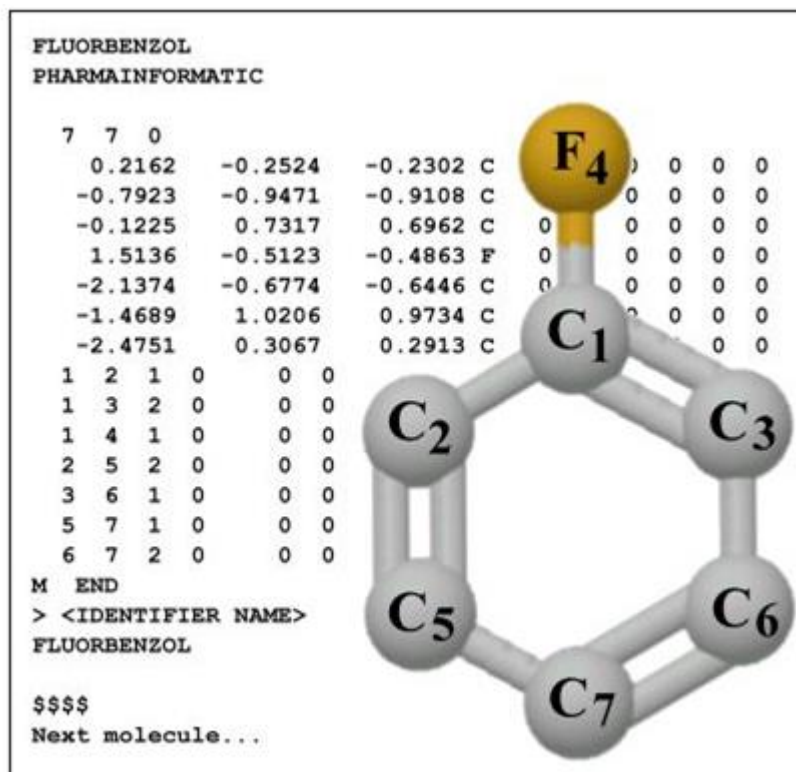
Resources on SDF

1. ChemFileBrowser - A win32 free sotfware for chemistry - ChemFileBrowser is a win32 free software for chemistry designed to visualize and works with SDFile (MDL® format) to exchange and analyse information associated with chemical structure. It includes descriptors calculation like TPSA, molecular weight, HBd,...
2. Chemical file format - Wikipedia, the free encyclopedia List of commonly used chemical MIME file formats including SDF....
3. Chemtool development page Chemtool is a small program for drawing chemical structures on Linux and Unix systems using the GTK toolkit under X11. A short and possibly outdated description of the available functions is available here. Chemtool relies on transfig by Brian Smith for postscript printing and exporting files in PicTeX and EPS formats. Its companion program, XFig, is recommended for enhancing the output of chemtool, and for creation of 2D diagrams and schematics in general. Both are included with most distributions of Linux, and are available through a number of websites including,

www.xfig.org. If you want to import chemtool drawings into word processing programs other than LaTeX you will probably want to add a preview bitmap to them, as neither StarOffice/OpenOffice nor that software from Redmond seem to be able to display postscript inserts on screen without them. For this purpose, using either ps2epsi, which comes with ghostscript, or epstool, a part of gsview is recommended. Since chemtool-1.6, this option is supported directly (through the equivalent function offered by recent versions of transfig).

4. Main Page - Open Babel Open Babel is a project designed to pick up where Babel left off, as a cross-platform program and library designed to interconvert between many file formats used in molecular modeling, computational chemistry, and many related areas. Features includes: A...

Open Babel: The Open Source Chemistry Toolbox

Open Babel is a chemical toolbox designed to speak the many languages of chemical data. It's an open, collaborative project allowing anyone to search, convert, analyze, or store data from molecular modeling, chemistry, solid-state materials, biochemistry, or related areas.

– Ready-to-use programs, and complete programmer's toolkit
– Read, write and convert over 110 chemical file formats
– Filter and search molecular files using SMARTS and other methods
– Supports molecular modeling, cheminformatics, bioinformatics
– Organic chemistry, inorganic chemistry, solid-state materials, nuclear chemistry

5. MN.CONVERT: Conversion of chemical file formats (SDF, MOL, MOL2,, CONVERT recognizes about 40 formats either by analyzing of the file's content or by using the file's extension (e.g. .mol, .smi, .sdf...), or the input format can also be specified....

6. http://www.hyleos.net/ Chemfile Browser hemFileBrowser is a win32 free software which was designed to visualize and work with SDFile (MDL® format). A format which is used by chemists to exchange and store compounds as well as associated data.

Some of the features mentioned are:

– ability to navigate forward and backward through an SDF
– introduce SDI file (SDF File index) for direct mapping
– adding and editing field names
– the option to export selected compounds
– exporting SDF with selected fields
– renaming structures with a given field value
– export the data as *.csv file
– copy to clipboard (compatible with IsisDraw, ChemDraw, ViewerPro and others)
– bookmark compound manager to create an SDFile from a selection
– splitting and merging of SDF
– chemical descriptors : TPSA, Hydrogen Bond donor and acceptor number, molecular weight

Figures

114 QMRFs in July 2017

Legend:
- 1. Physical Chemical Properties
- 2. Environmental fate parameters
- 3. Ecotoxic effects
- 4. Human Health Effects
- 5. Toxicokinetics
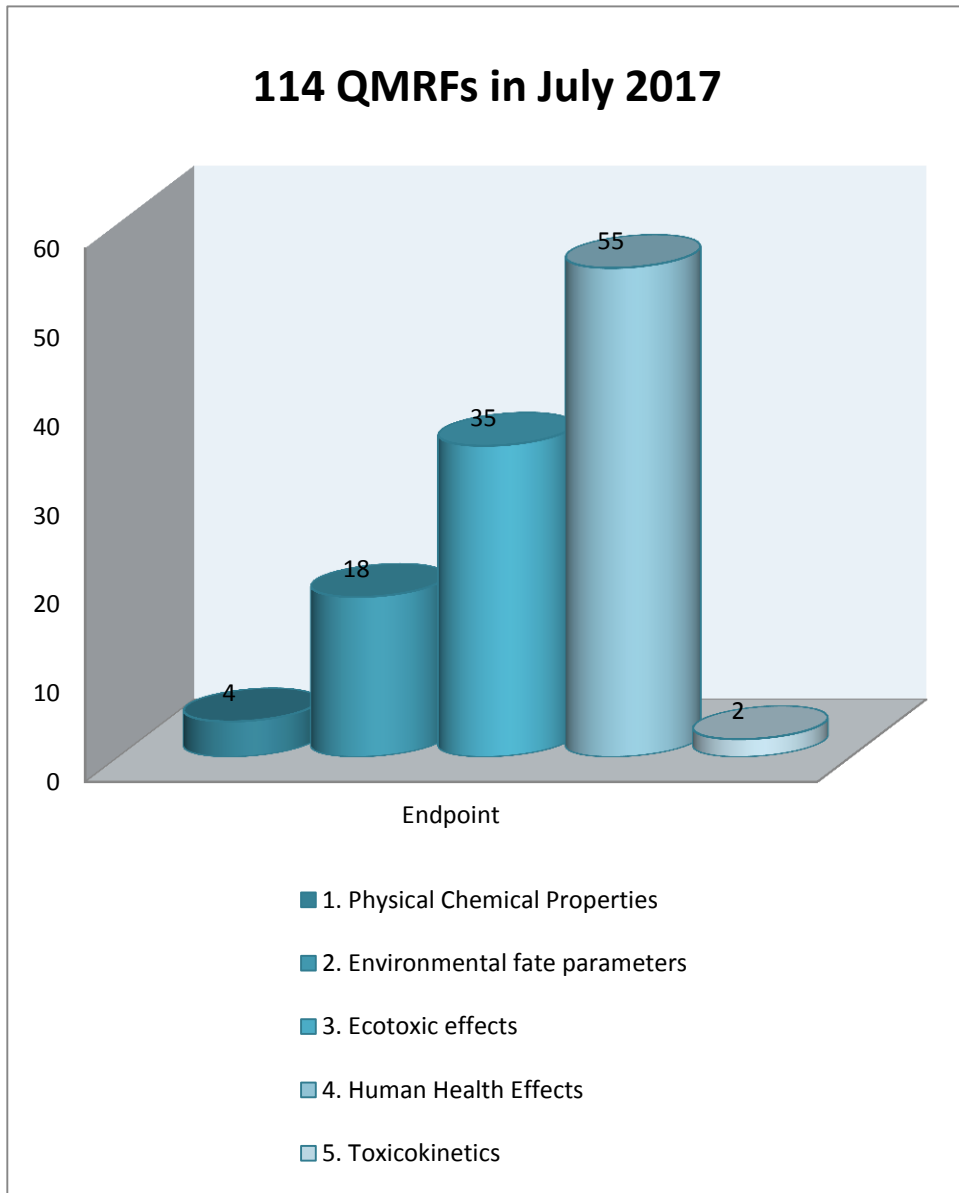
Glossary

ASCII

stands for **American Standard Code for Information Interchange**. It is a character encoding standard (the Internet Assigned Numbers Authority (IANA) prefers the name US-ASCII[2]). ASCII codes represent text in computers, telecommunications equipment, and other devices. Most modern character-encoding schemes are based on ASCII, although they support many additional characters

CAS Registry Number

also referred to as CASRN or CAS Number, is a unique numerical identifier assigned by Chemical Abstracts Service (CAS) to every chemical substance described in the open scientific literature (currently including those described from at least 1957 through the present), including organic and inorganic compounds, minerals, isotopes, alloys and non struc
urable materials (UVCBs, of unknown, variable composition, or biological origin)

DB-ALM

EURL ECVAM DataBase service on ALternative Methods to animal experimentation; access at:
http://ecvam-dbalm.jrc.ec.europa.eu

ECHA

The mission of the **European Chemicals Agency** is to:
 - Manage all REACH and CLP tasks by carrying out or co-coordinating the necessary activities
 - Ensure a consistent implementation at Community level
 - Provide Member States and the European institutions with the best possible scientific advice on questions related to the safety and the socio-economic aspects of the use of chemicals. *Source:* ECHA web site
 The European Chemicals Agency (ECHA; /ˈɛkə/ EK-ə)[citation needed] is an agency of the European Union which manages the technical, scientific and administrative aspects of the implementation of the European Union regulation called Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). ECHA is the driving force among regulatory authorities in implementing the EU's chemicals legislation. ECHA helps companies to comply with the legislation, advances the safe use of chemicals, provides information on chemicals and addresses chemicals of concern. It is located in Helsinki, Finland. *Source: WIKIPEDIA*

EURL

European Union Reference Laboratory: In the context of the EU strategy aimed at improving animal health and establishing the single market for live animals and animal products, a network of European Union and National reference laboratories dealing with major animal diseases has been gradually set up.
The Council and the Commission have designated European Union reference laboratories (EURLs) with scientific and technical expertise within the areas of animal health, public health and zootechnics in a number of legal acts. These legal acts contain provisions that specify the functions and duties of each designated EURL. The designation of EURL should contribute to a high quality and uniformity of analytical results. Source: *Europa, Food Safety web site*
The European Commission Joint Research Center currently hosts seven EURLs in support of EU Member States' National Reference Laboratories (NRLs) in the respective fields; three of these EURLs (in yellow below) are managed by the Institute for Health and Consumer Protection (IHCP):
- EURL GMFF (European Union Reference Laboratory for GMOs in Food and Feed)
- EURL for feed additives
- **EURL ECVAM** (European Union Reference Laboratory and European Centre for validation of Alternative Methods)

- EURL FCM (European Union Reference Laboratory for Food Contact Materials)
- EURL for heavy metals in feed and food
- EURL for mycotoxins in food and feed
- EURL for polycyclic aromatic hydrocarbons *Source: WIKIPEDIA*

IHCP Institute for Health and Consumer Protection (IHCP)

IHCP's mission is to provide scientific and technical support to the EU policies for the protection of the interests and health of European citizens in the areas of food, consumer products, chemicals and public health. As of 1.7.2016 its duties transferred to the **JRC Directorate F - Health, Consumers and Reference Materials**

OECD

The Organisation for Economic Co-operation and Development (OECD) (French: Organisation de coopération et de développement économiques, OCDE) is an intergovernmental economic organisation of 35 countries, founded in 1961 to stimulate economic progress and world trade. It is a forum of countries describing themselves as committed to democracy and the market economy, providing a platform to compare policy experiences, seeking answers to common problems, identify good practices and coordinate domestic and international policies of its members. *Source: WIKIPEDIA*

The **OECD (Organisation for Economic Co-operation and Development)** is an intergovernmental organisation in which representatives of 30 industrialised countries in North America, Europe and the Pacific, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and subsidiary groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and subsidiary groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions. *Source: OECD website*

QSARs

Structure-activity relationships and **quantitative structure-activity relationships**, collectively referred to as **(Q)SARs**, are simplified mathematical representations of complex chemical-biological interactions that can be used to predict the physicochemical and biological properties of molecules. They can take various forms of various complexity and either be qualitative or quantitative.
A **structure-activity relationship (SAR)** usually represents an association between a chemical substructure and the potential of a chemical containing the substructure to exhibit a certain biological effect.
A **quantitative structure-activity relationship (QSAR)** quantitatives relates the properties of a chemical (encoded in its chemical structure) to a physical property or to a biological effect (e.g. a toxicological endpoint).
Source: DB-ALM website
(Q)SARs are methods for estimating properties of a chemical from its molecular structure and have the potential to provide information on hazards of chemicals, while reducing time, monetary cost and animal testing currently needed.
To facilitate practical application of (Q)SAR approaches in regulatory contexts by governments and industry and to improve their regulatory acceptance, the OECD (Q)SAR project has developed various outcomes such as the principles for the validation of (Q)SAR models, guidance documents as well as the (Q)SAR Application Toolbox. The OECD (Q)SAR Project is carried out with the financial assistance of the EU.
Source: OECD

Animal tests can be avoided if the hazardous properties of a substance can be predicted using computer models. The **[(Q)SAR] [(quantitative) structure–activity relationship]** approach seeks to predict the intrinsic properties of chemicals by using various databases and theoretical models, instead of conducting tests. Based on knowledge of chemical structure, QSAR quantitatively relates characteristics of the chemical to a measure of a particular activity. QSAR should be distinguished from SAR, which makes qualitative conclusions about the presence or absence of a property of a substance, based on a structural feature of the substance. *Source: ECHA*

QMRFs

The **Q**SAR **M**odel **R**eporting **F**ormat (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.

REACH

REACH is the Regulation for **R**egistration, **E**valuation, **A**uthorisation and Restriction of **Ch**emicals. It entered into force on 1st June 2007 to streamline and improve the former legislative framework on chemicals of the European Union (EU). REACH places greater responsibility on industry to manage the risks that chemicals may pose to the health and the environment. In principle REACH applies to all chemicals: not only chemicals used in industrial processes but also in our day-to-day life, for example in cleaning products, paints as well as in articles such as clothes, furniture and electrical appliances. REACH makes industry bear most responsibilities to manage the risks posed by chemicals and provide appropriate safety information to their users. In parallel, it foresees that the European Union can take additional measures on highly dangerous substances, where there is a need for complementing action at EU level. REACH also creates the European Chemicals Agency (ECHA) with a central coordination and implementation role in the overall process. *Source: ECHA web site*

SAR

A **Structure–Activity Relationship** (SAR) usually represents an association between a chemical substructure and the potential of a chemical containing the substructure to exhibit a certain biological effect.Variants: structure-activity relationship

STU

Systems Toxicology Unit: As of 1.7.2016 its duties transferred to the Chemicals Safety and Alternative Methods unit F.3)

sdf – extension

SDF is one of a family of chemical-data file formats developed by MDL; it is intended especially for structural information. "SDF" stands for **structure-data file**, and SDF files actually wrap the molfile (MDL Molfile) format. Multiple compounds are delimited by lines consisting of four dollar signs ($$$$). A feature of the SDF format is its ability to include associated data.

snn – extension

The snn file extension **STATISTICA neural networks file** is associated with the STATISTICA an analytics solution developed by StatSoft, Inc. and now owned by Dell. The snn file stores neural networks data used by STATISTICA version 7.

SMARTS

The SMARTS language is designed to describe substructure patterns in molecules, and to overcome the limitations of simple substructure matching.
SMARTS query filter is a powerful and flexible tool to perform complex structural queries impossible to describe with a simple Substructure search. For example, if you are seeking for molecules containing a phenol substituted with any halogen atoms at para position, you can easily define the corresponding SMARTS pattern, and do the search.

SMILES

The **simplified molecular-input line-entry system (SMILES)** is a specification in form of a line notation for describing the structure of chemical species using short ASCII strings. SMILES strings can be imported by most molecule editors for conversion back into two-dimensional drawings or three-dimensional models of the molecules.

Tanimoto coefficient

A Tanimoto coefficient approaching 1 means two chemical structures are very similar. The lower the Tanimoto coefficient the more dissimilar two molecules are.

UVCBs,

Chemical Substances of **U**nknown or **V**ariable Composition, **C**omplex Reaction Products and **B**iological Materials (**UVCB Substance**)

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: http://europea.eu/contact

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: http://europa.eu/contact

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: http://europa.eu

EU publications

You can download or order free and priced EU publications from EU Bookshop at: http://bookshop.europa.eu. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact).

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub – Joint Research Centre

Joint Research Centre

EU Science Hub