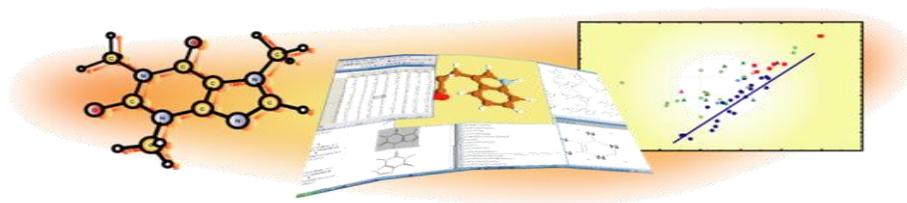


Computational Tools and Guidance developed by the JRC

Klaus Daginnus
Institute for Health & Consumer Protection
Joint Research Centre, European Commission

CAESAR Final Workshop
Milan, 11 March 2009

<http://ecb.jrc.ec.europa.eu/qsar/>



1. **Computational toxicology at the JRC**
2. **Role of computational methods in risk assessment**
3. **Filling data gaps: read-across and (Q)SARs**
4. **Documenting the results of read-across and (Q)SARs**
5. **Non-testing strategy – a stepwise approach**
6. **Conclusions**

European Commission

Directorates-General

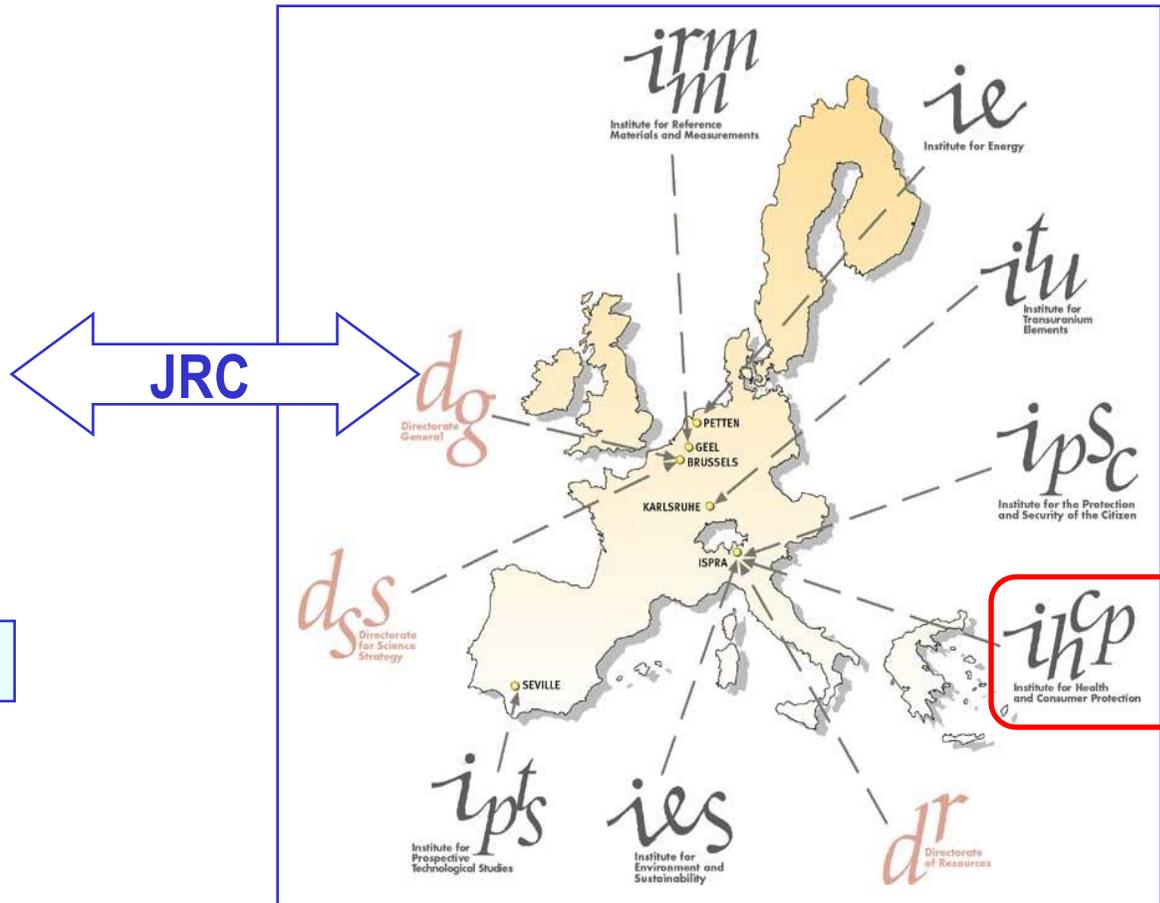
Directorates or Institutes

Units

Molecular Biology
In Vitro Methods
Nanobiosciences
Chemical Assessment & Testing
Systems Toxicology

Genetically Modified Organisms
Alternative Methods & ECVAM
Nanotechnology
Health and Environment
Consumer Products & Nutrition

Policy Areas



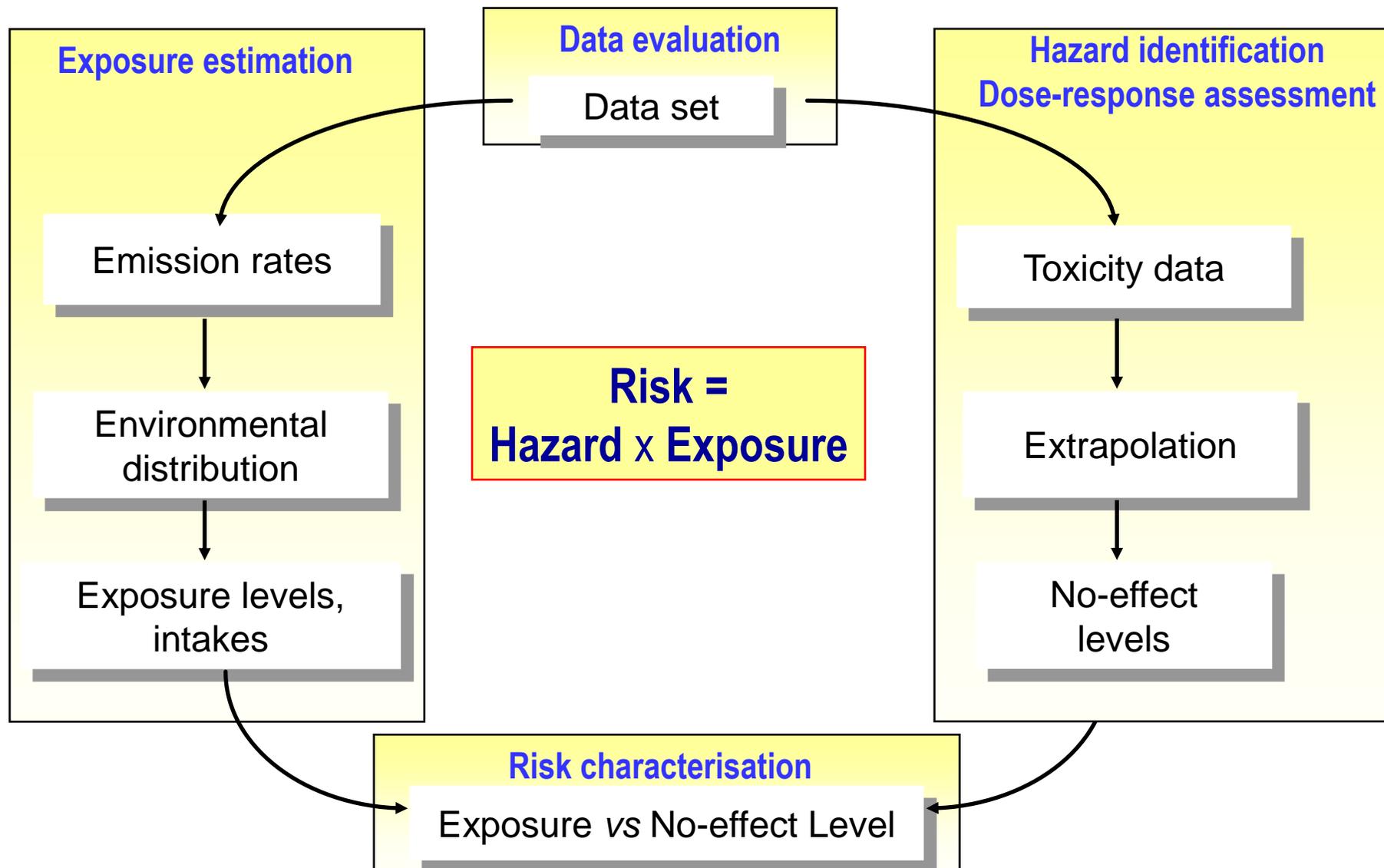
Overall aim: to promote the development, assessment, acceptance and implementation of computer-based methods potentially suitable for the regulatory assessment of chemicals

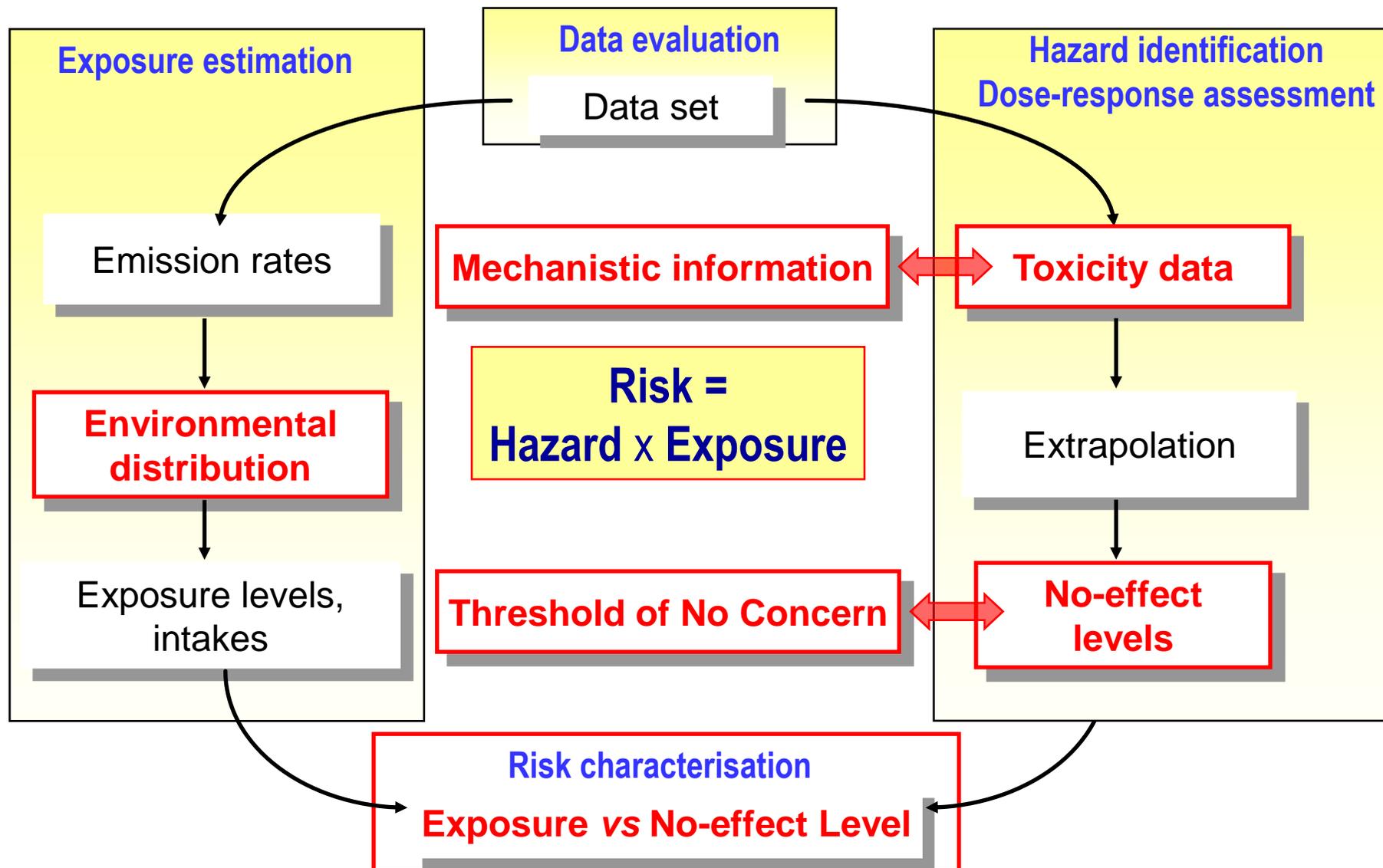
Applications: REACH, Water Framework Directive, Cosmetics Directive, Biocides Directive, Plant Protection Products Directive

Main approaches: SAR, QSAR, molecular modelling, ranking

Computational methods provide information for use in hazard and risk assessment → “non-testing” or alternative methods

<http://ecb.jrc.ec.europa.eu/qsar/>



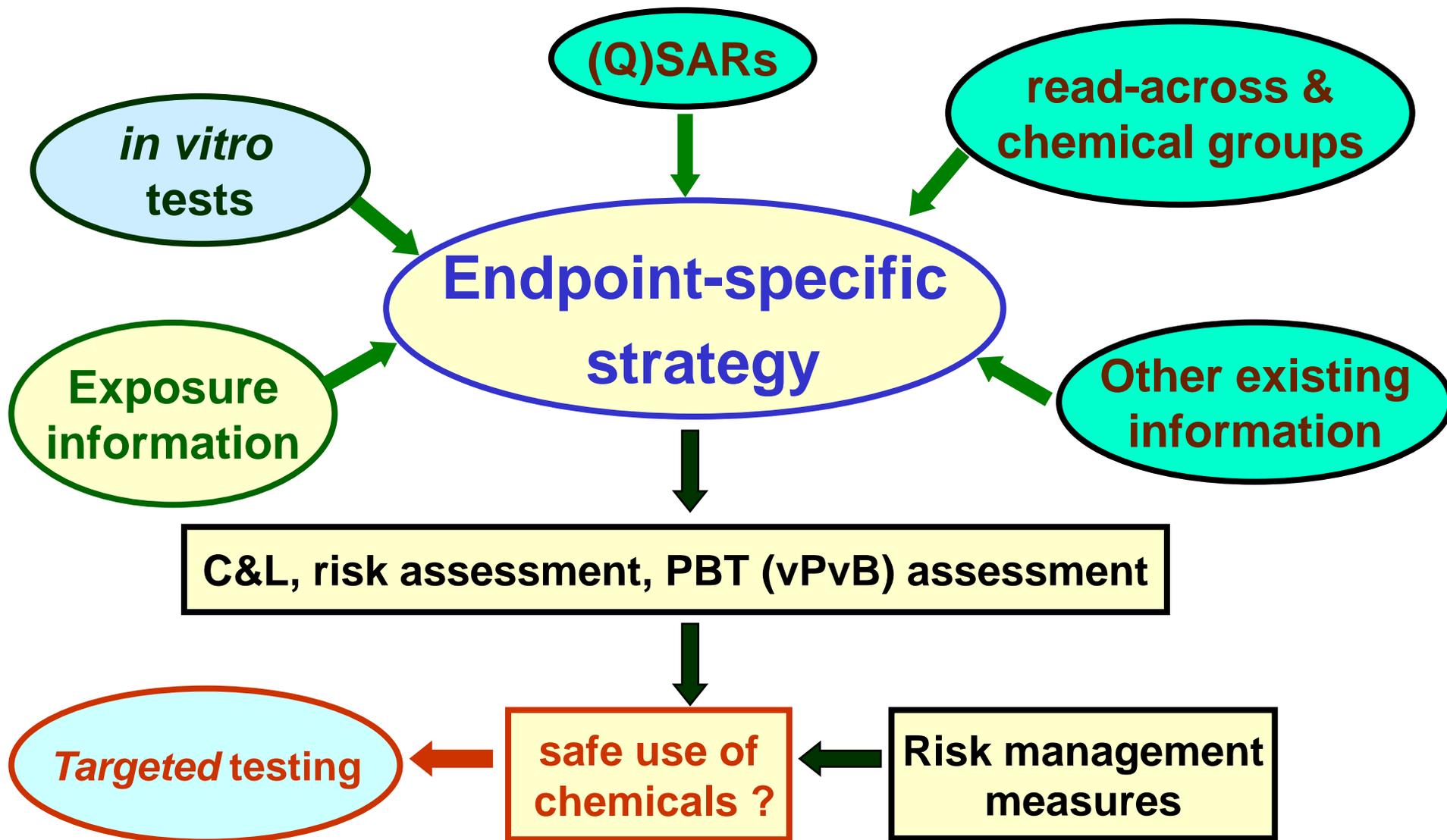


Information requirements are largely tonnage dependent, however ...

“Information on intrinsic properties of substances may be generated by means other than tests, provided that the conditions set out in Annex XI are met” (Article 13)

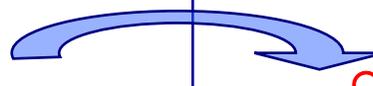
(Animal) testing can be reduced or avoided by “replacing traditional test data with predictions or equivalent data”

... however a number of conditions apply

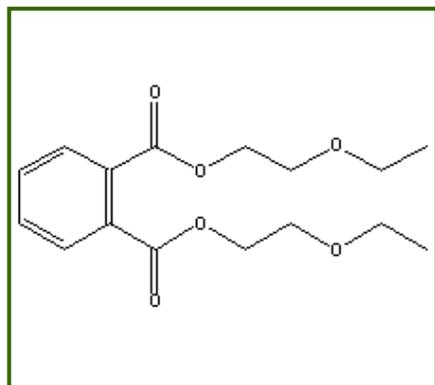


Known information on the property of a substance (source chemical) is used to make a prediction of the same property for another substance (target chemical) that is considered “similar”

	Source chemical	Target chemical
Property	●	○



**1,2-Benzenedicarboxylic acid,
bis(2-ethoxyethyl) ester**

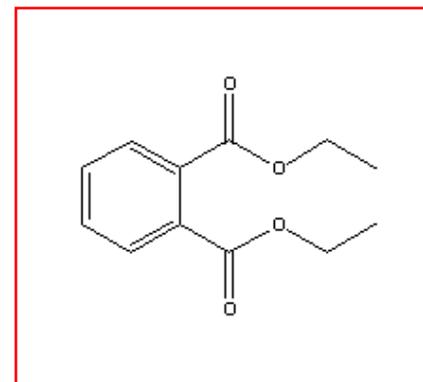


Known to be harmful: $1 < \log LC50 < 2$

Acute fish toxicity?



diethyl phthalate



Predicted to be harmful

The **analogue approach** refers to the grouping of chemicals and application of read-across for a single endpoint based on a relatively small number of analogues

one-to-one

	Substance 1	Substance 2
Property	●	○



● reliable data point

○ missing data point

many-to-one

	Substance 1	Substance 2	Substance 3
Property	●	○	●



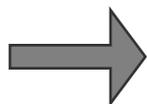
	Chemical 1	Chemical 2	Chemical 3	Chemical 4	
Property 1	●	○	●	○	read-across
Property 2	●	○	●	●	
Property 3	○	●	●	○	
Property 4	●	●	●	●	Trend analysis
Activity 1	○	○	○	○	
Activity 2	●	●	●	●	
Activity 3	○	○	○	○	
Activity 4	○	○	○	○	

● reliable data point
○ missing data point

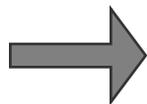
The **category approach** refers to a wider approach, based on more analogues, multiple endpoints, and in which trends are also apparent

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

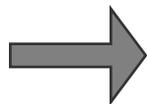
- the estimate should be generated by a valid (reliable) model
- the model should be applicable to the chemical of interest with the necessary level of reliability
- the model endpoint should be relevant for the regulatory purpose



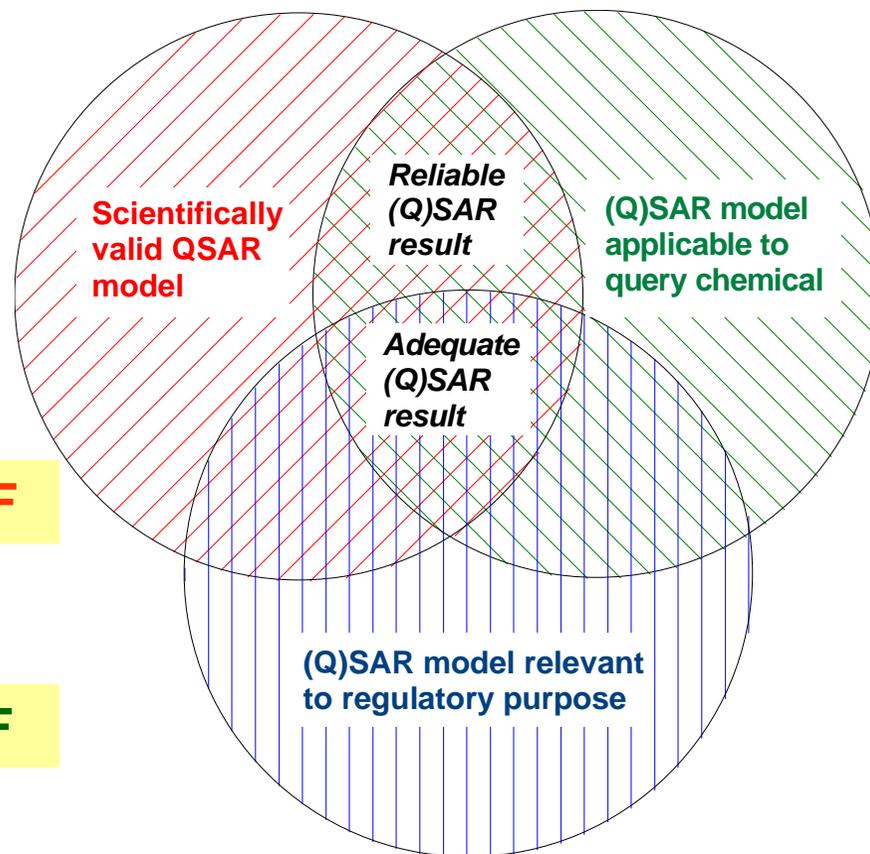
QMRF



QPRF



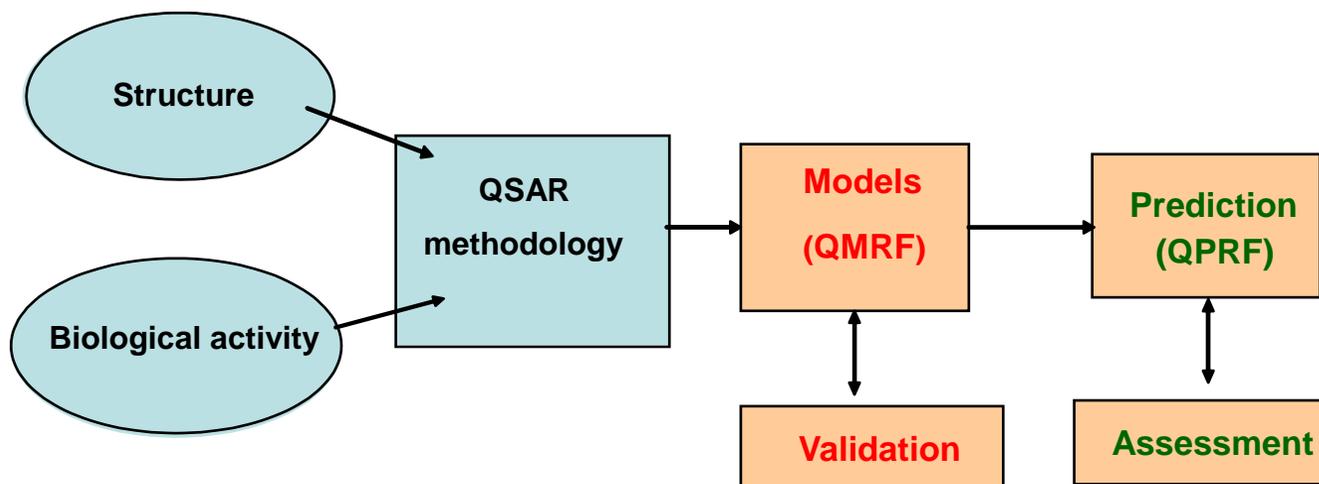
QPRF



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

A (Q)SAR Model Reporting Format (QMRF) is a robust summary of a (Q)SAR model, which reports key information on the model according to the OECD validation principles

A (Q)SAR Prediction Reporting Format (QPRF) is a description and assessment of the prediction made by given model for a given chemical



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

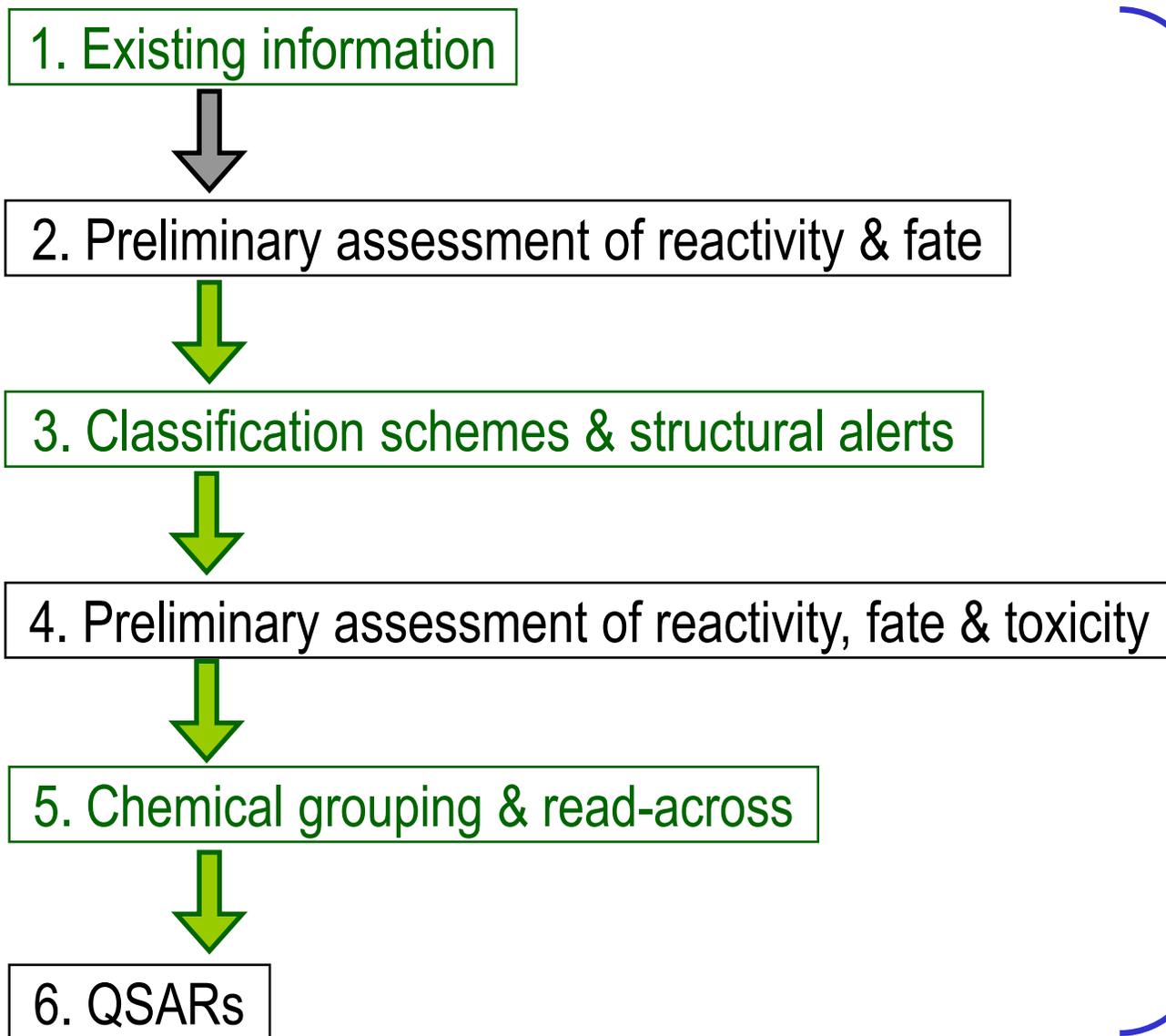
A (Q)SAR should be associated with the following information:

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)

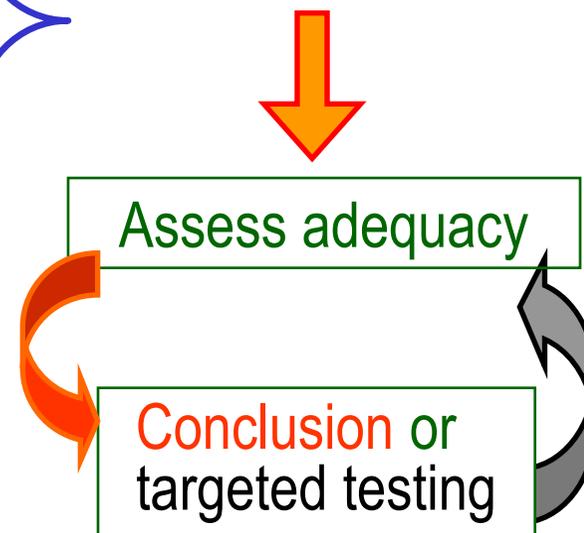
QPRF captures information on the substance and its prediction, and is intended to facilitate considerations of the adequacy of a prediction

1. Substance information
 2. General (administrative) information on QPRF
 3. Information on prediction (endpoint, algorithm, applicability domain, uncertainty, mechanism)
 4. Adequacy (optional, legislation-specific, and 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

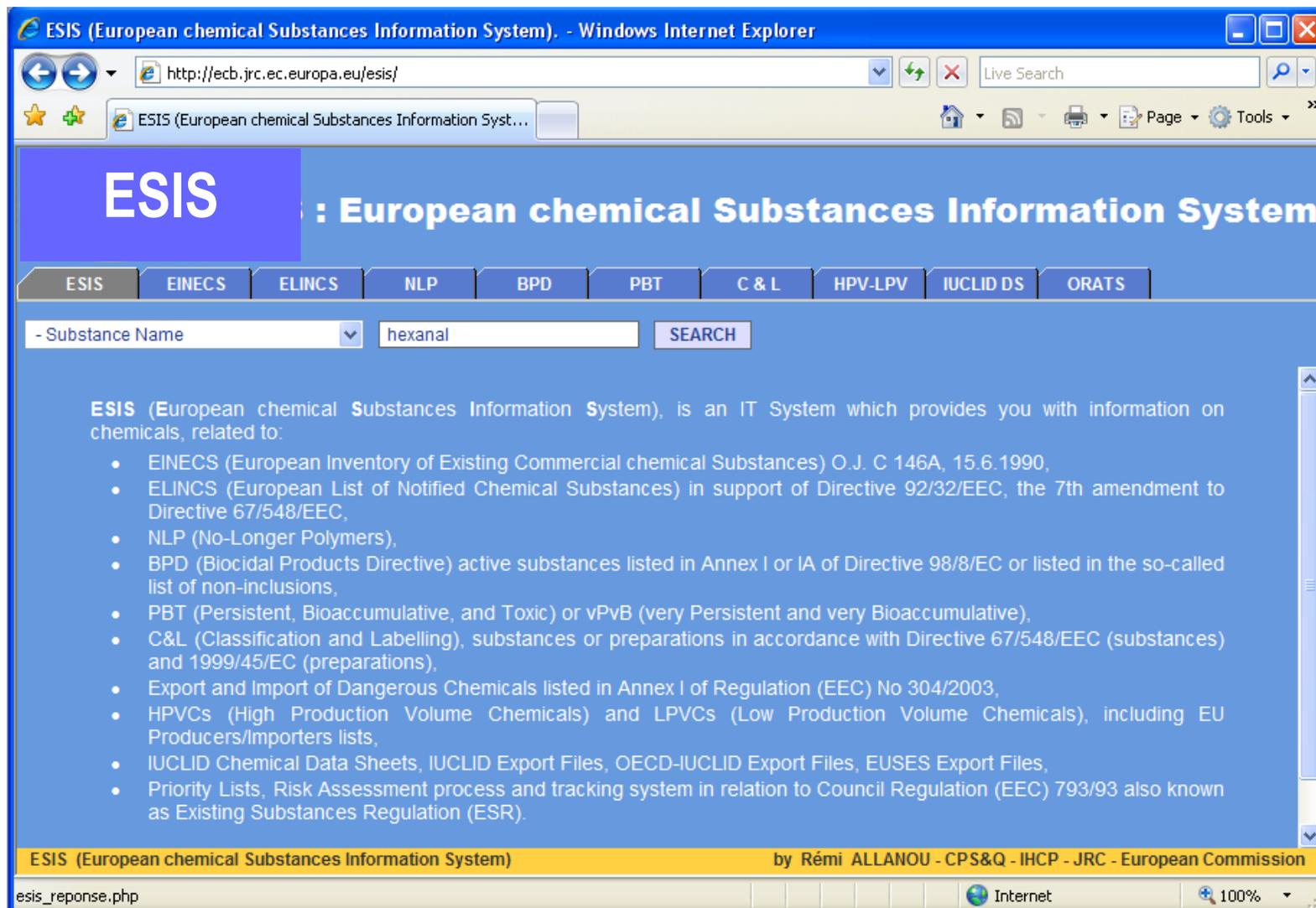


Working Matrix

	A	B	C
Chemical			
Metabolite 1			
Metabolite 2			



- Chemical composition (components, purity/impurity profile)
- Structure generation and verification
- Key chemical features (functional groups, protonation states, isomers)
- Experimental data: physicochemical properties, (eco)toxicity, fate
 - Freely-accessible web resources (ESIS, ChemSpider, PubChem, AMBIT2)
 - Databases in freely-available software tools (OECD Toolbox)
 - Commercial databases (Vitic, ...)
- Estimated data: pre-generated QSAR or read-across estimates
 - Freely-accessible web resources (ChemSpider, Danish QSAR database)
 - Chemical category databases (OECD Toolbox)



The screenshot shows a Windows Internet Explorer browser window displaying the ESIS website. The address bar shows the URL <http://ecb.jrc.ec.europa.eu/esis/>. The page title is "ESIS (European chemical Substances Information System)". The main content area has a blue background with the text "ESIS : European chemical Substances Information System". Below this is a navigation menu with buttons for "ESIS", "EINECS", "ELINCS", "NLP", "BPD", "PBT", "C & L", "HPV-LPV", "IUCLID DS", and "ORATS". A search bar is present with a dropdown menu set to "- Substance Name" and the text "hexanal" entered. A "SEARCH" button is next to the search bar. Below the search bar, there is a paragraph describing ESIS and a bulleted list of the types of information it provides. The footer of the page contains the text "ESIS (European chemical Substances Information System) by Rémi ALLANOU - CPS&Q - IHCP - JRC - European Commission". The browser's status bar at the bottom shows "Internet" and "100%" zoom.

ESIS (European chemical Substances Information System) - Windows Internet Explorer

<http://ecb.jrc.ec.europa.eu/esis/>

ESIS : European chemical Substances Information System

ESIS EINECS ELINCS NLP BPD PBT C & L HPV-LPV IUCLID DS ORATS

- Substance Name

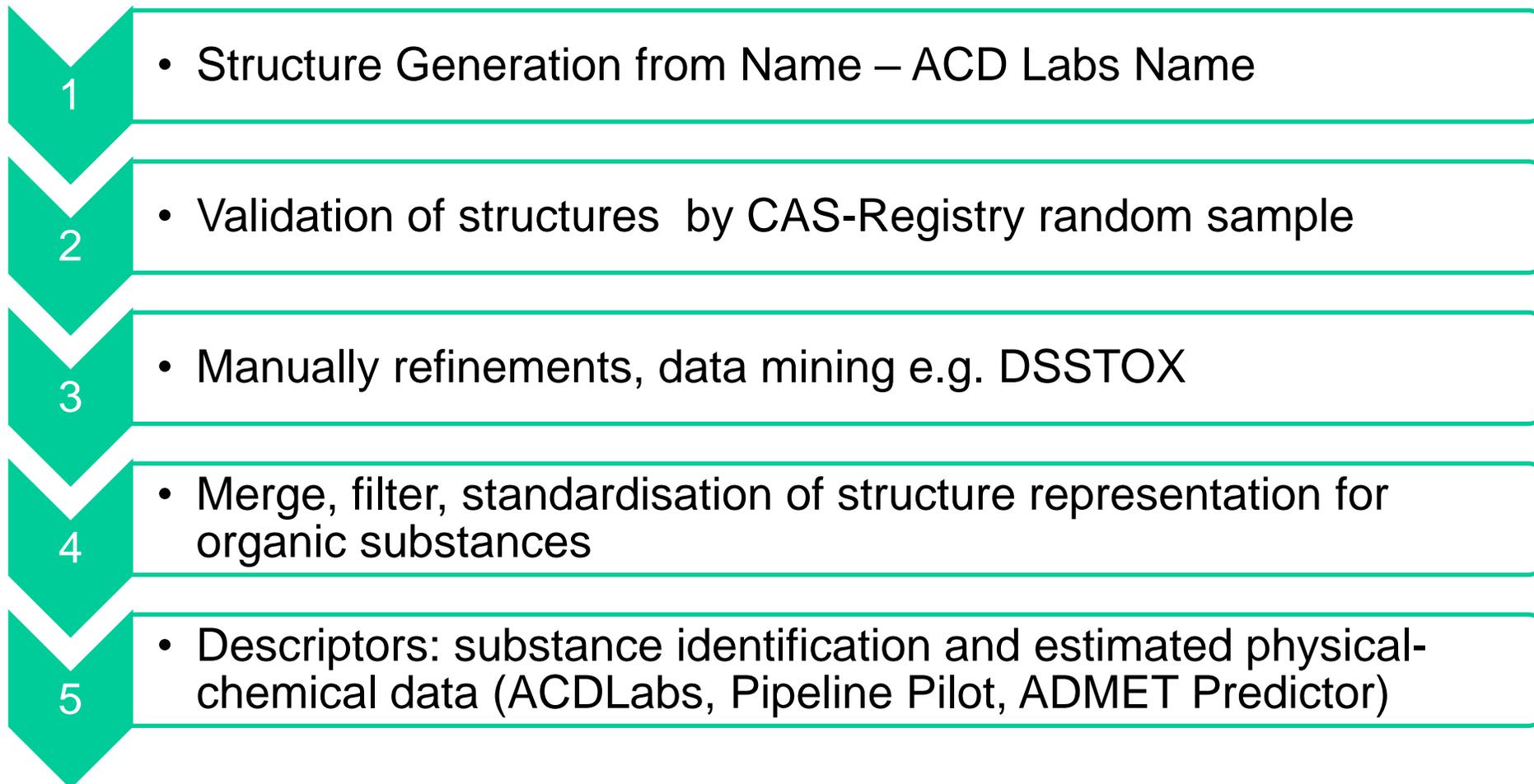
ESIS (European chemical Substances Information System), is an IT System which provides you with information on chemicals, related to:

- EINECS (European Inventory of Existing Commercial chemical Substances) O.J. C 146A, 15.6.1990,
- ELINCS (European List of Notified Chemical Substances) in support of Directive 92/32/EEC, the 7th amendment to Directive 67/548/EEC,
- NLP (No-Longer Polymers),
- BPD (Biocidal Products Directive) active substances listed in Annex I or IA of Directive 98/8/EC or listed in the so-called list of non-inclusions,
- PBT (Persistent, Bioaccumulative, and Toxic) or vPvB (very Persistent and very Bioaccumulative),
- C&L (Classification and Labelling), substances or preparations in accordance with Directive 67/548/EEC (substances) and 1999/45/EC (preparations),
- Export and Import of Dangerous Chemicals listed in Annex I of Regulation (EEC) No 304/2003,
- HPVCs (High Production Volume Chemicals) and LPVCs (Low Production Volume Chemicals), including EU Producers/Importers lists,
- IUCLID Chemical Data Sheets, IUCLID Export Files, OECD-IUCLID Export Files, EUSES Export Files,
- Priority Lists, Risk Assessment process and tracking system in relation to Council Regulation (EEC) 793/93 also known as Existing Substances Regulation (ESR).

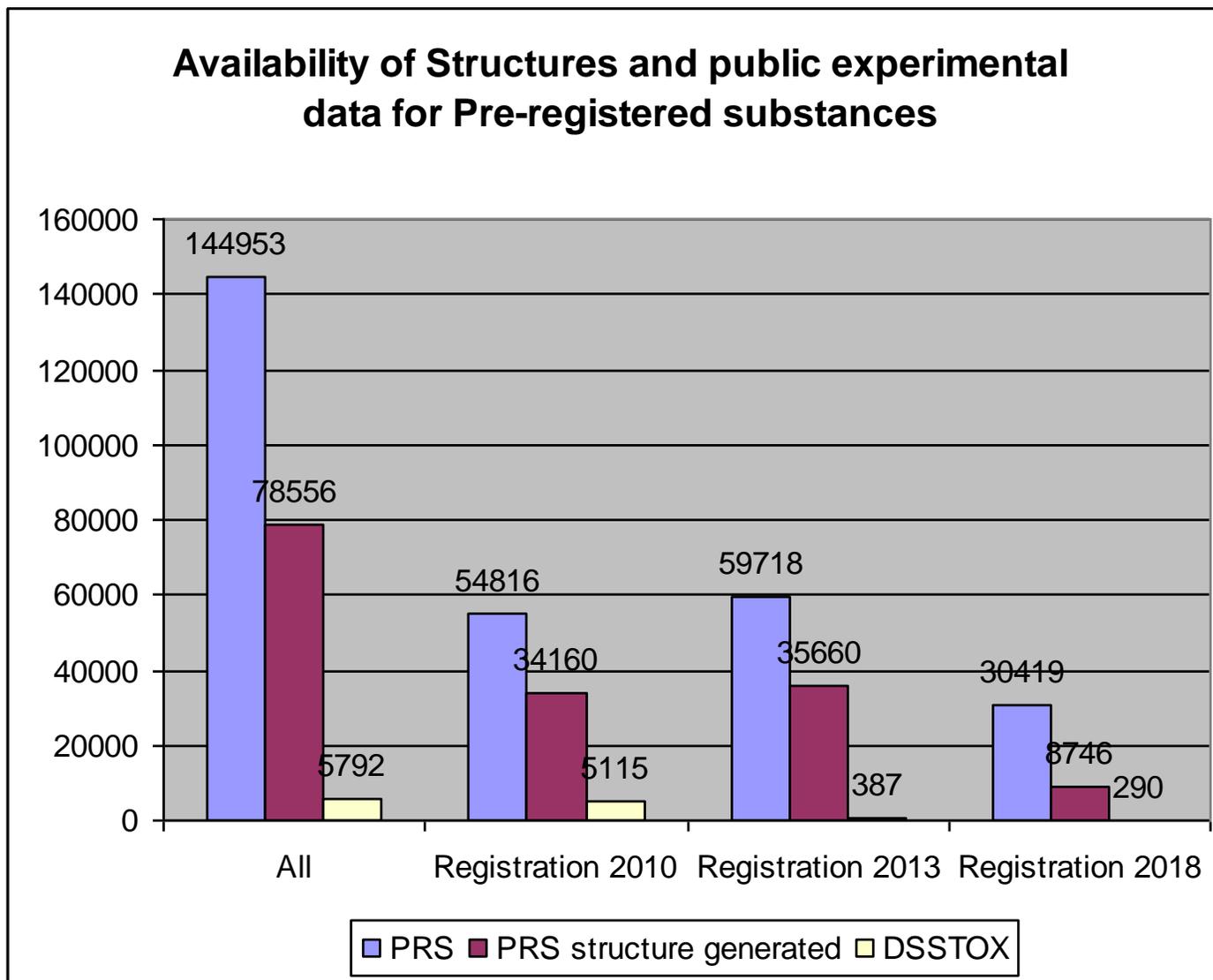
ESIS (European chemical Substances Information System) by Rémi ALLANOU - CPS&Q - IHCP - JRC - European Commission

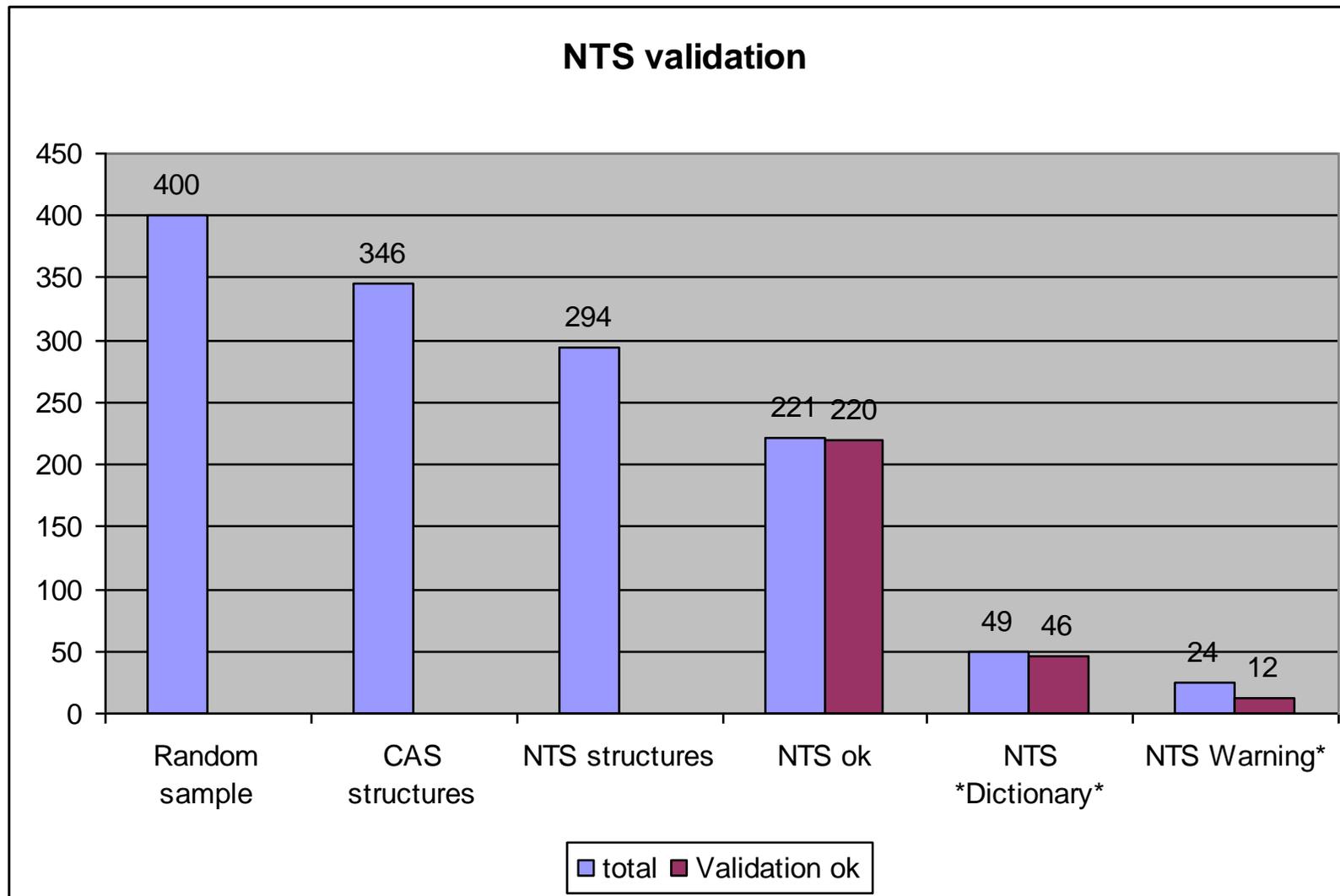
esis_reponse.php Internet 100%

<http://ecb.jrc.ec.europa.eu/esis/>

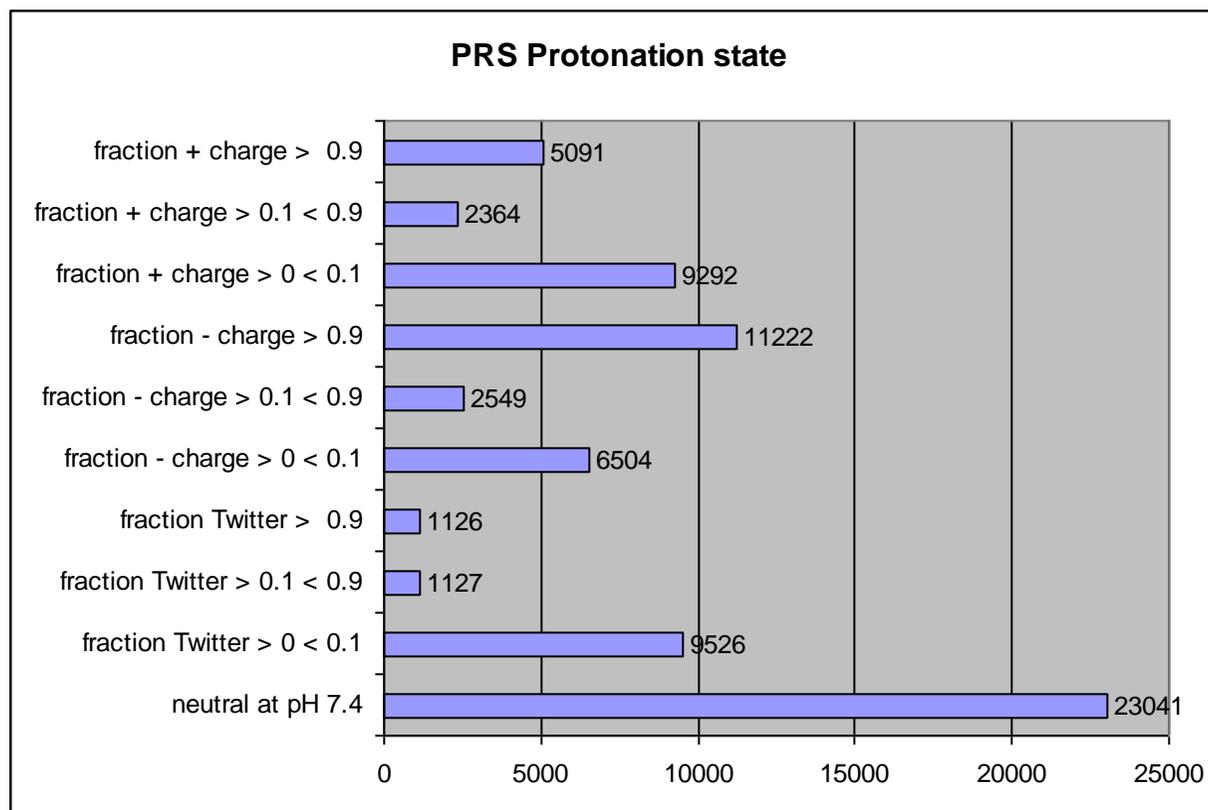


http://ecb.jrc.ec.europa.eu/qsar/information-sources/ec_inventory/





logP, pKa, logD, Water Solubility by ADMET Predictor, see benchmarking:
 Calculation of molecular lipophilicity: State-of-the-art and comparison of logP
 methods on more than 96,000 compounds, Mannhold, R.; Poda, G.I.; Ostermann,
 C.; Tetko, I.V. *Journal of Pharmaceutical Sciences* **2009**, 98(3), 861-893.



Search

name:

cas:

smiles:

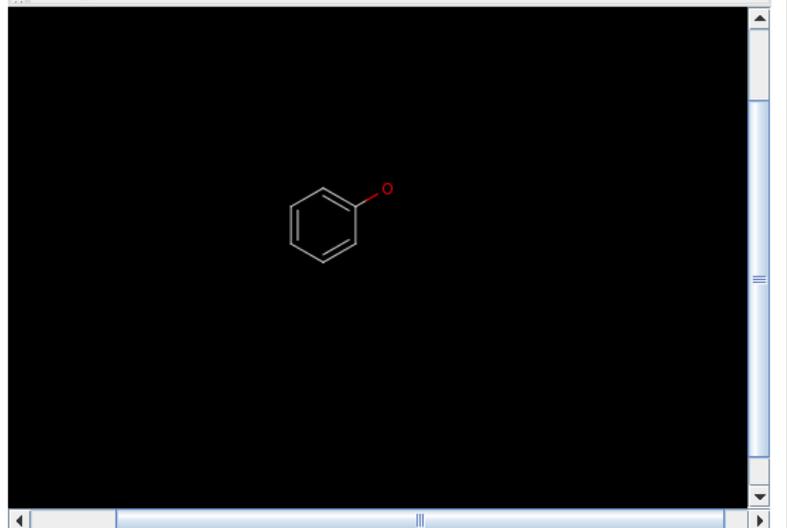
Search by Structure

File Edit View Insert Tools Help

H C N O F React Select Erase Paste Undo Redo Zoom

- + P S Cl →

More Br I



Type of Search:

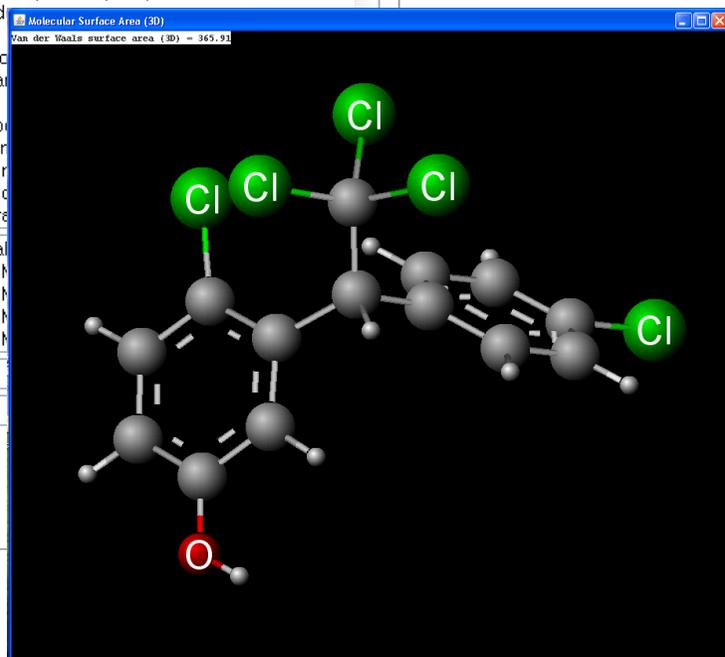
Chemical Group

- all-
- (organo) metals
- Alkylbenzenes and styrenes
- Alkylphenols and derivatives
- Benzamidazoles
- Biphenyls
- Bisphenols
- Carbamates
- Chlorinated cyclodienes and camphenes
- Chlorinated paraffins (CPs)
- Chlorophenols and benzenes
- Chlorophenoxy compounds
- DDT, derivatives and metabolites
- Dicarboximides
- Diesel exhaust particle (DEP)
- Dinitroanilid
- Dioxins
- Diphenylpro
- Dithiocarba
- Furans
- Hexachloro
- Hydroxyber
- Linuron, diu
- Methoxychl
- Musk Fragre

Moa

mol. w. between

logP between



web-accessible database under development

- Prediction of abiotic / biotic reactivity to identify reactive potential and possible transformation products / metabolites
- Freely-available software
 - CRAFT (Chemical Reactivity & Fate Tool)
 - START (Structural Alerts in Toxtree)
 - OECD Toolbox
- Commercial software and databases
 - CATABOL, TIMES, Meteor, Mexalert, MetabolExpert ...
 - MetaPath, SciFinder, MDL Reaction Database ...

Toxtree (Estimation of Toxic Hazard - A Decision Tree Approach) v1.50

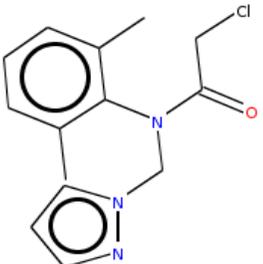
File Edit Chemical Compounds Toxic Hazard Method Help

File: H:\User\Alexei\Projects\ECB sandbox\ECB1\sample.smi

Available structure attributes

Biodegradability	Persistent chemical
Biodegradability#explanation	1N,2N,3N,4N,5N,6N,7N,8...

Structure diagram



First Prev 1 / 1 Next Last

Completed.

Toxic Hazard by Biodegradability

Estimate

Easily biodegraded chemical

Persistent chemical

Unknown biodegradability

Verbose explanation

Biodegradability

Q1. Terminal tert-butyl branch No

Q2. Epoxides No

Q3. FusedAliphaticNonBranchedRings No

Q4. At least two terminal isopropyl groups for a non-cyclic chemical No

Q5. Aliphatic cyclic chemicals with no branches No

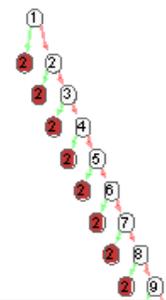
Q6. One or more halogen substitutions on a branched molecule No

Q7. Unbranched, non-cyclic chemicals with two halogen substitutions No

Q8. More than two hydroxy substituents on aromatic ring No

Q9. Two or more rings Yes Class Persistent chemical

Biodegradability



Decision node: Q1. Terminal tert-butyl branch

If 'NO' go to: Q.2

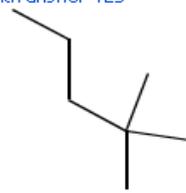
If 'YES' assign: Persistent chemical

Rule ID	Rule title
1	Terminal tert-butyl branch

Rule explanation

Returns true if the query contains substructures specified by SMARTS patterns.

Example with answer 'YES'

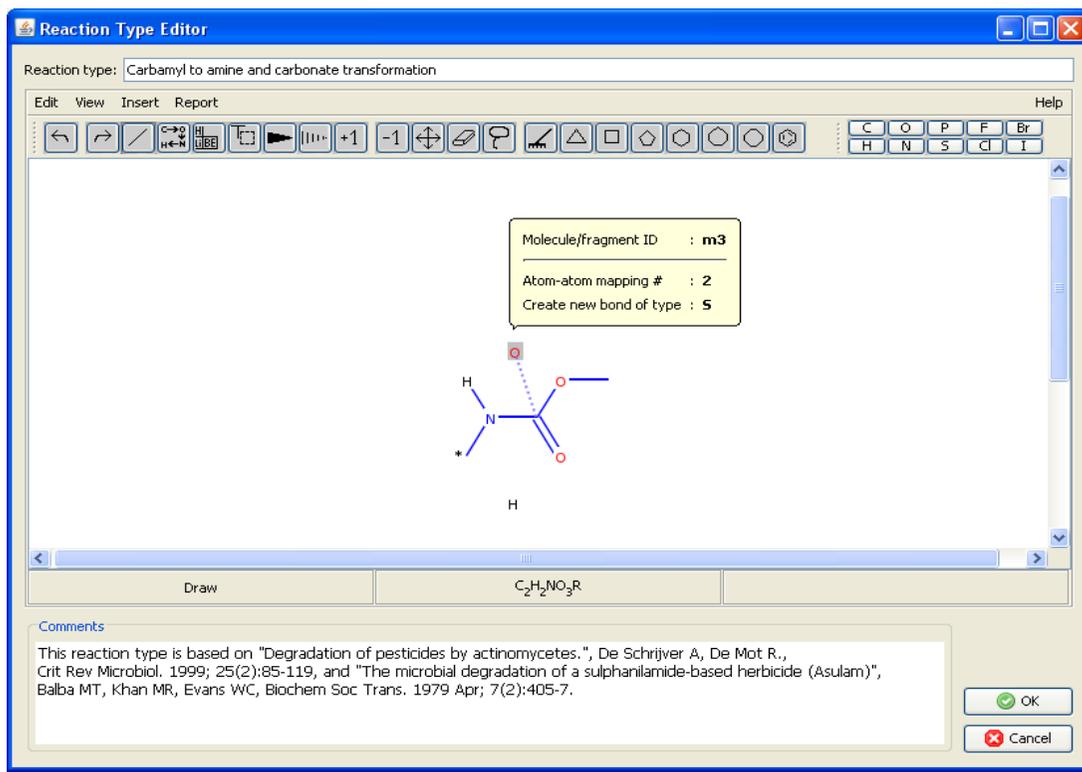


There are example molecules for each rule outcome. Select which one to display.

Yes branch No branch

- Collaboration with Molecular Networks (Germany)
- Toxtree plug-in
- Estimates biodegradation potential

- Collaboration with Molecular Networks (Germany)
- Generates & visualises reactions, ranks transformation products
- Initial emphasis on abiotic processes & microbial biodegradation
- Data model based on AMBIT technology
- User can modify knowledge base and rulebase



Reaction Type Editor

Reaction type: Carbamyl to amine and carbonate transformation

Edit View Insert Report Help

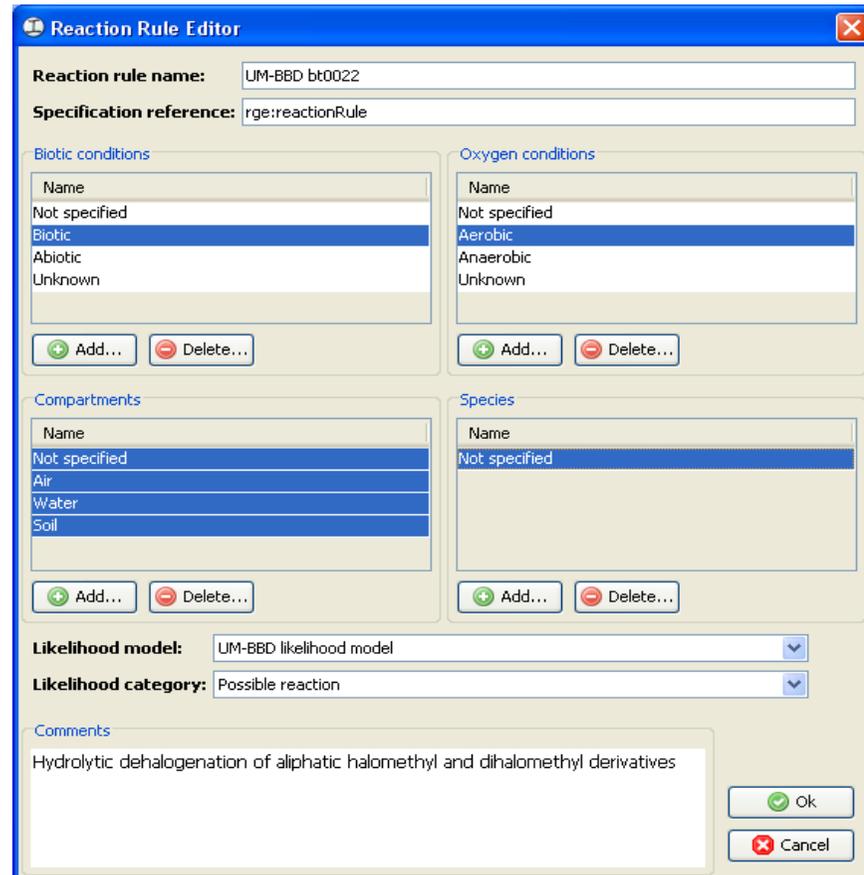
Molecule/fragment ID : m3
Atom-atom mapping # : 2
Create new bond of type : 5

C2H2NO3R

Comments

This reaction type is based on "Degradation of pesticides by actinomycetes.", De Schrijver A, De Mot R., Crit Rev Microbiol. 1999; 25(2):85-119, and "The microbial degradation of a sulphanilamide-based herbicide (Asulam)", Balba MT, Khan MR, Evans WC, Biochem Soc Trans. 1979 Apr; 7(2):405-7.

OK Cancel



Reaction Rule Editor

Reaction rule name: UM-BBD bt0022

Specification reference: rge:reactionRule

Biotic conditions

Name
Not specified
Biotic
Abiotic
Unknown

+ Add... - Delete...

Oxygen conditions

Name
Not specified
Aerobic
Anaerobic
Unknown

+ Add... - Delete...

Compartments

Name
Not specified
Air
Water
Soil

+ Add... - Delete...

Species

Name
Not specified

+ Add... - Delete...

Likelihood model: UM-BBD likelihood model

Likelihood category: Possible reaction

Comments

Hydrolytic dehalogenation of aliphatic halomethyl and dihalomethyl derivatives

OK Cancel

- Models and rulebases for mode-of-action classification, hazard identification, hazard classification and potency prediction
- Freely-available software
 - Episuite, Toxtree, AMBIT2, OECD Toolbox ...
 - OpenTox framework (<http://www.opentox.org>)
- Commercial software
 - DEREK, MultiCASE, HazardExpert, ToxAlert, ToxBoxes ...
 - *Insilicofirst* consortium (Multicase Inc, Lhasa Ltd, Molecular Networks GmbH, Leadscope Inc)
- QSAR Model Databases (QMDBs)
 - JRC QSAR Model Database
 - OECD Toolbox

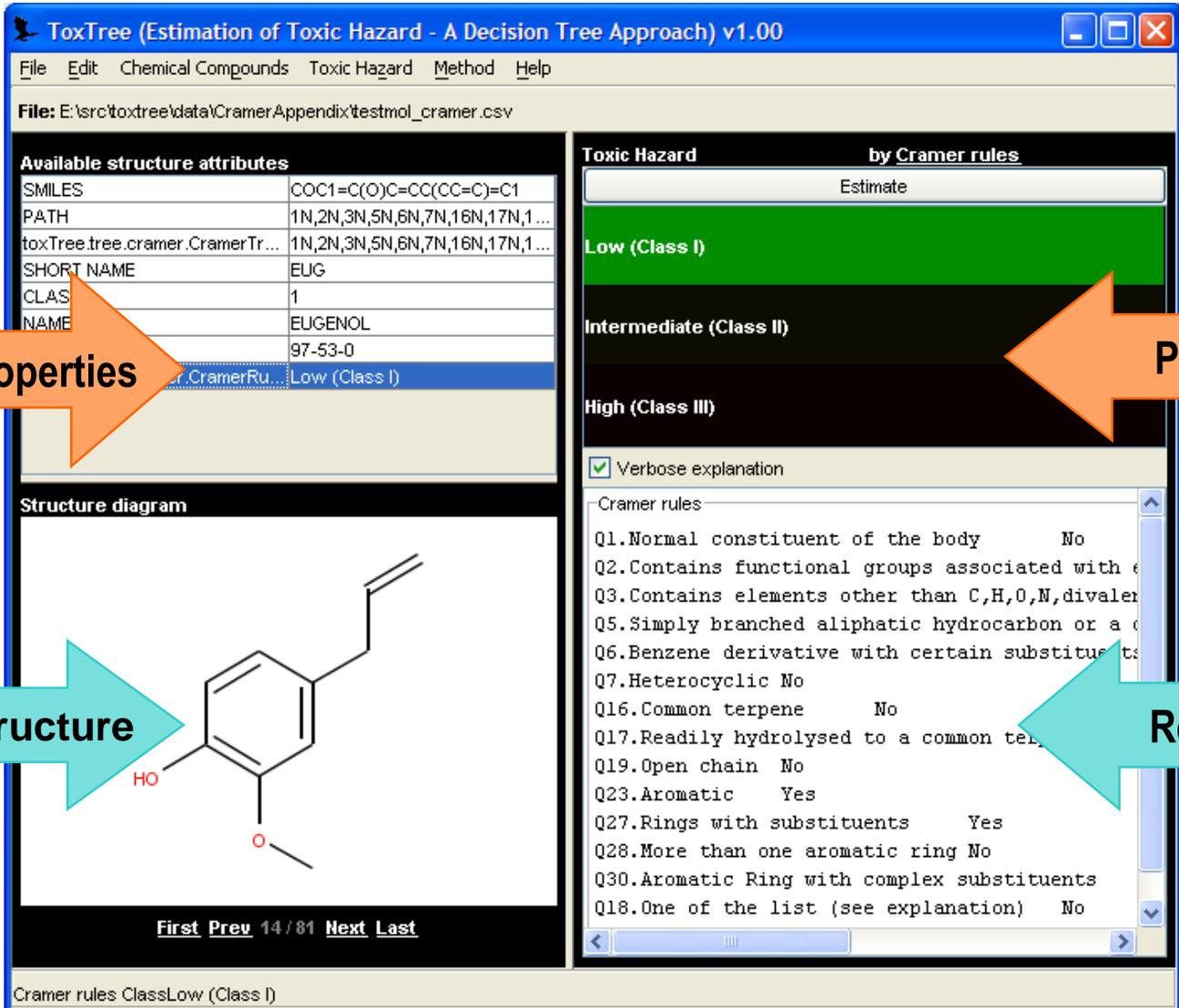
Toxtree is a flexible, user-friendly, open source application, which is able to classify chemicals into modes of action and estimate toxic hazard by applying decision tree approaches

Collaboration with Ideaconult (BG)

Rulebases in version 1.51 (June 2008):

- Acute Fish Toxicity (Verhaar scheme)
- Oral systemic toxicity (Cramer scheme)
- Skin irritation & corrosion potential (BfR rulebase)
- Eye irritation & corrosion potential (BfR rulebase)
- Mutagenicity & carcinogenicity (Benigni-Bossa rulebase)

<http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/>



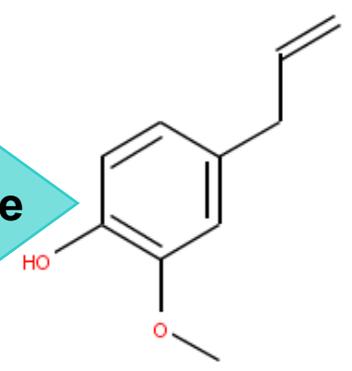
ToxTree (Estimation of Toxic Hazard - A Decision Tree Approach) v1.00

File Edit Chemical Compounds Toxic Hazard Method Help

File: E:\src\toxtree\data\Cramer.Appendix\testmol_cramer.csv

Available structure attributes	
SMILES	COC1=C(O)C=CC(CC=C)=C1
PATH	1N,2N,3N,5N,6N,7N,16N,17N,1...
toxTree.tree.cramer.CramerTr...	1N,2N,3N,5N,6N,7N,16N,17N,1...
SHORT NAME	EUG
CLASS	1
NAME	EUGENOL
	97-53-0
	Low (Class I)

Structure diagram



First Prev 14 / 81 Next Last

Cramer rules ClassLow (Class I)

Toxic Hazard by Cramer rules	
Estimate	
Low (Class I)	
Intermediate (Class II)	
High (Class III)	

Verbose explanation

Cramer rules

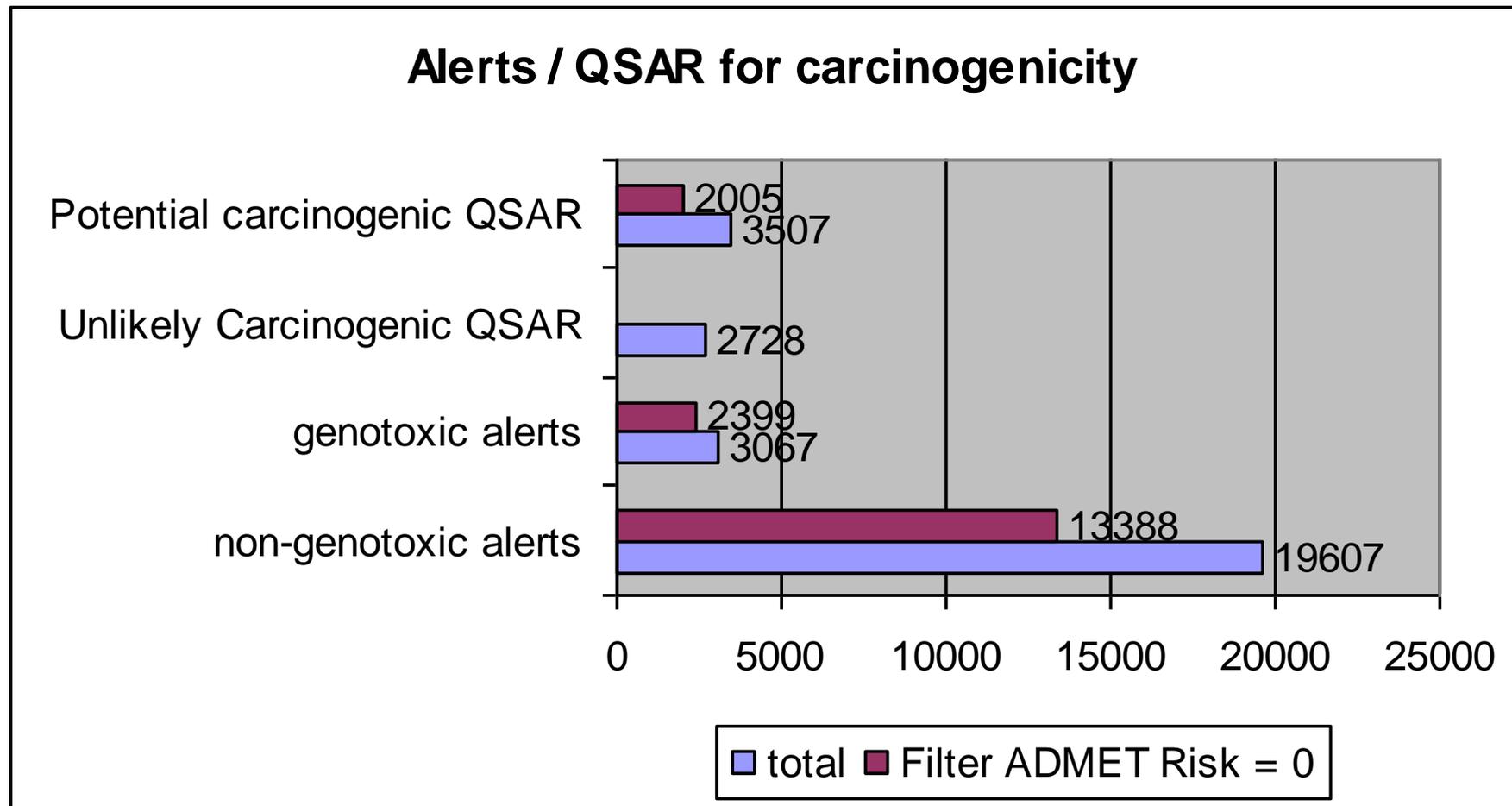
- Q1. Normal constituent of the body No
- Q2. Contains functional groups associated with e
- Q3. Contains elements other than C,H,O,N,divalen
- Q5. Simply branched aliphatic hydrocarbon or a c
- Q6. Benzene derivative with certain substituents
- Q7. Heterocyclic No
- Q16. Common terpene No
- Q17. Readily hydrolysed to a common tea
- Q19. Open chain No
- Q23. Aromatic Yes
- Q27. Rings with substituents Yes
- Q28. More than one aromatic ring No
- Q30. Aromatic Ring with complex substituents
- Q18. One of the list (see explanation) No

Compound properties

Prediction

Compound structure

Reasoning



ADMET Risk estimates from ADMET Predictor

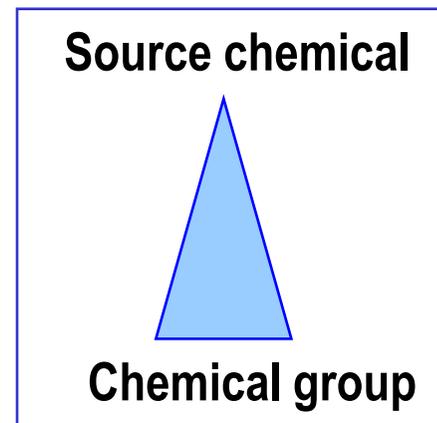
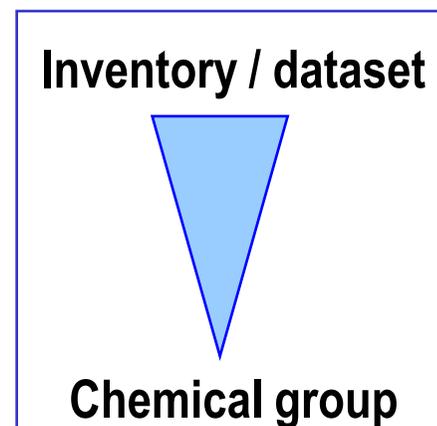
The implementation of the Cramer classification scheme in the Toxtree software was evaluated to evaluate its concordance and highlight potential software modifications

The results were promising with an overall good concordance between the reported classifications and those generated by Toxtree

Improvements for Toxtree were proposed. Notable of these is a necessity to update the lists of common food components and normal body constituents as these accounted for the majority of false classifications observed.

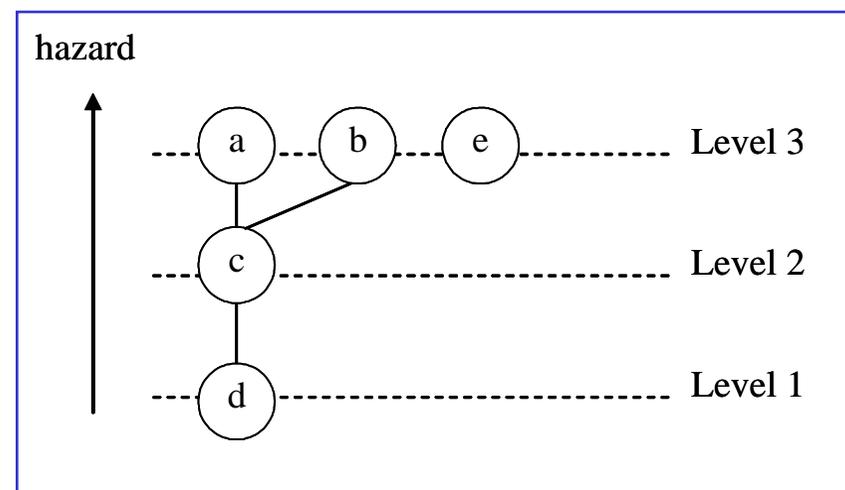
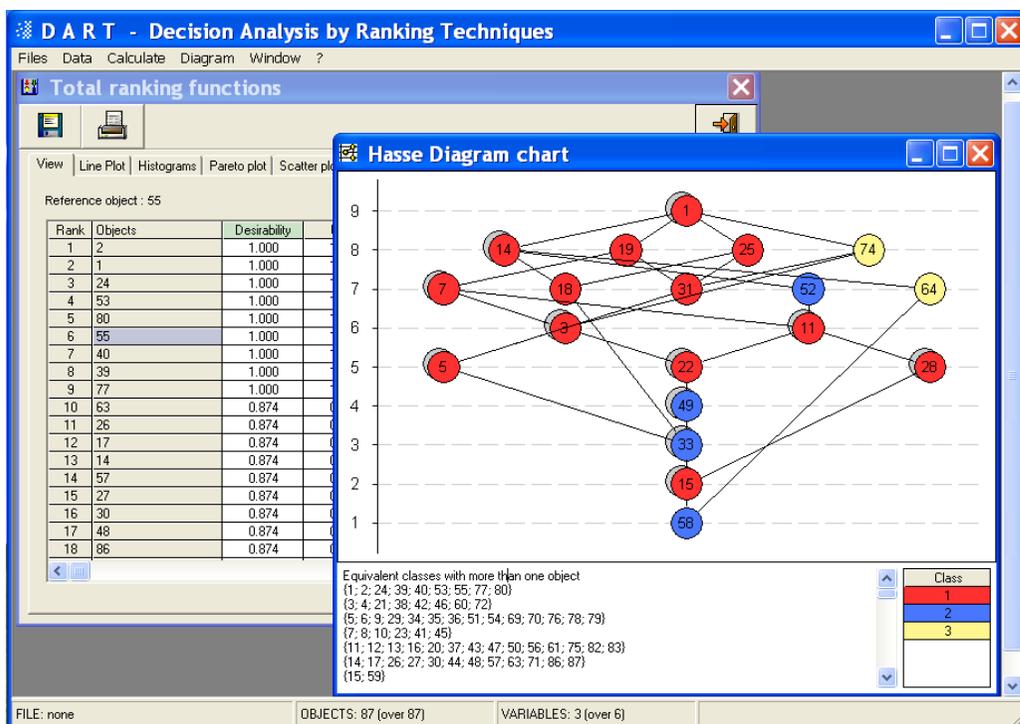
Patlewicz G, Jeliaskova N, Safford RJ, Worth AP & Aleksiev B (2008). An evaluation of the implementation of the Cramer classification scheme in the Toxtree software. *SAR and QSAR in Environmental Research* 19, 495-524.

- Chemical read-across within analogue and category approaches
- Biological read-across (between endpoints or species)
- Chemical grouping by a **top-down approach**
 - Supervised and unsupervised statistical methods
 - Ranking methods (DART)
- Chemical grouping by a **bottom-up approach**
 - Freely available tools with analogue-searching capability (Toxmatch, AMBIT2, AIM, PubChem, OECD Toolbox)



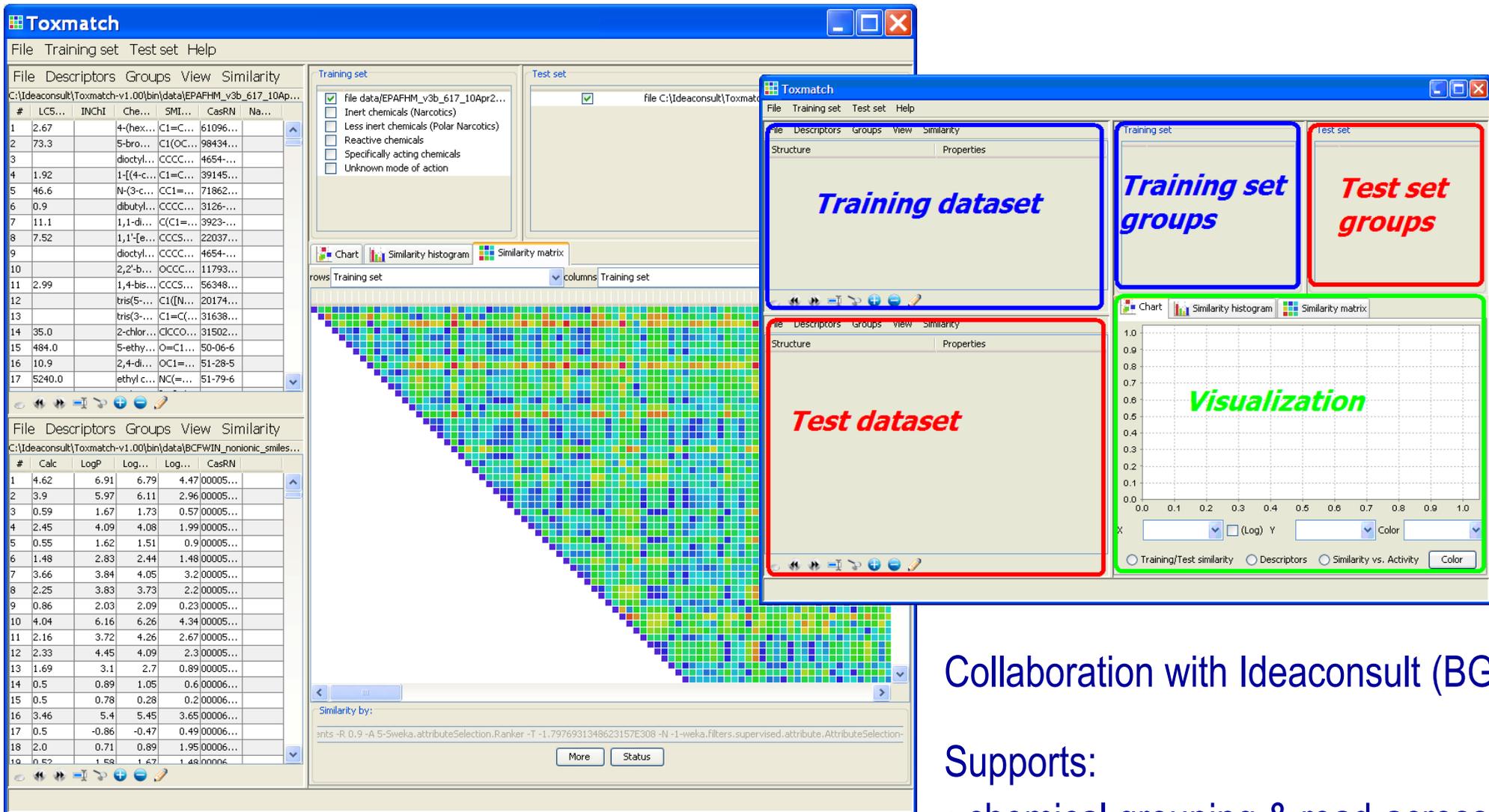
Worth A et al (2007). The Use of Computational Methods in the Grouping and Assessment of Chemicals - Preliminary Investigations. EUR 22941 EN

DART (Decision Analysis by Ranking Techniques) is a flexible, user-friendly, open source application, which is able to rank and group chemicals according to properties of concern



- collaboration with Talete srl (Italy)
- supports priority setting of chemicals

Pavan M & Worth AP (2008). A set of case studies to illustrate the applicability of DART (Decision Analysis by Ranking Techniques) in the ranking of chemicals. EUR 23481 EN.



The screenshot displays the Toxmatch software interface. The main window is titled 'Toxmatch' and contains several panels:

- Training set:** A list of chemical groups with checkboxes, including 'file data/EPAFHM_v3b_617_10Apr...', 'Inert chemicals (Narcotics)', 'Less inert chemicals (Polar Narcotics)', 'Reactive chemicals', 'Specifically acting chemicals', and 'Unknown mode of action'.
- Test set:** A list of chemical groups, currently showing 'file C:\Ideaconsult\Toxm...'
- Similarity matrix:** A large triangular heatmap showing similarity scores between training and test set chemicals. The x and y axes are labeled 'rows Training set' and 'columns Training set'.
- Visualization:** A chart area with a grid and axes ranging from 0.0 to 1.0. It includes a legend for 'Training/Test similarity', 'Descriptors', and 'Similarity vs. Activity'.
- Tables:** Two tables on the left side of the interface. The top table lists chemical groups with columns for '#', 'LC5...', 'INCHI', 'Che...', 'SMI...', 'CasRN', and 'Na...'. The bottom table lists chemical groups with columns for '#', 'Calc', 'LogP', 'Log...', 'Log...', and 'CasRN'.

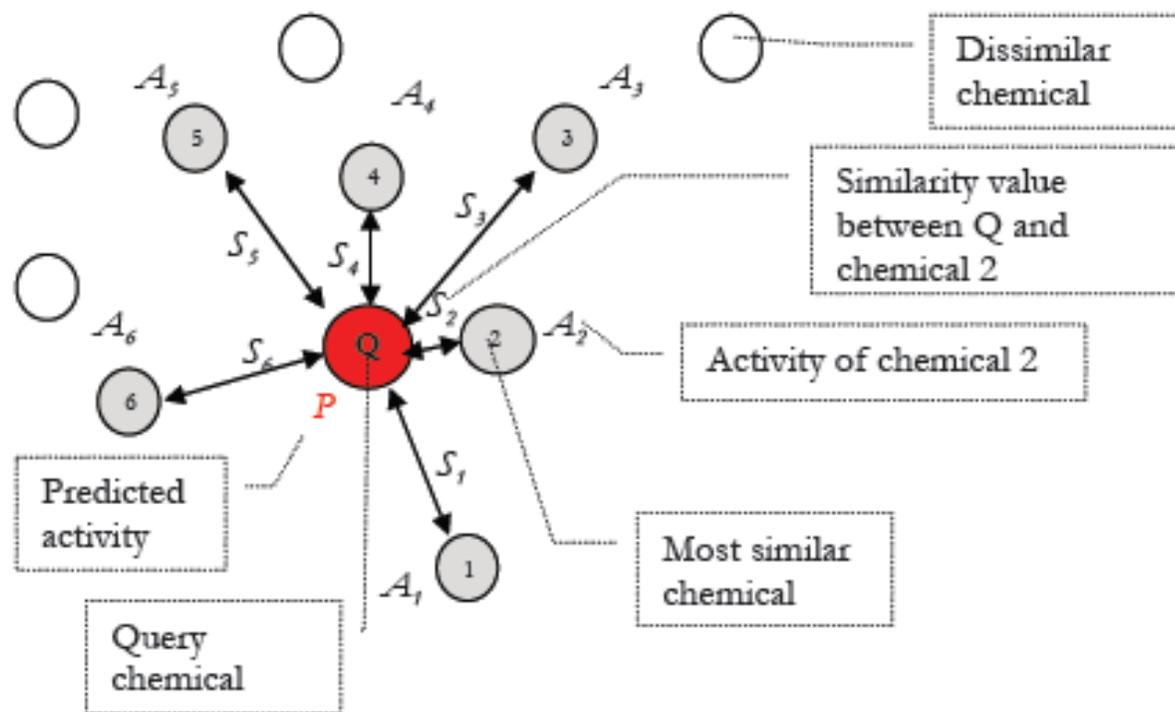
Collaboration with Ideaconsult (BG)

Supports:

- chemical grouping & read-across
- comparison of training & test sets

<http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/>

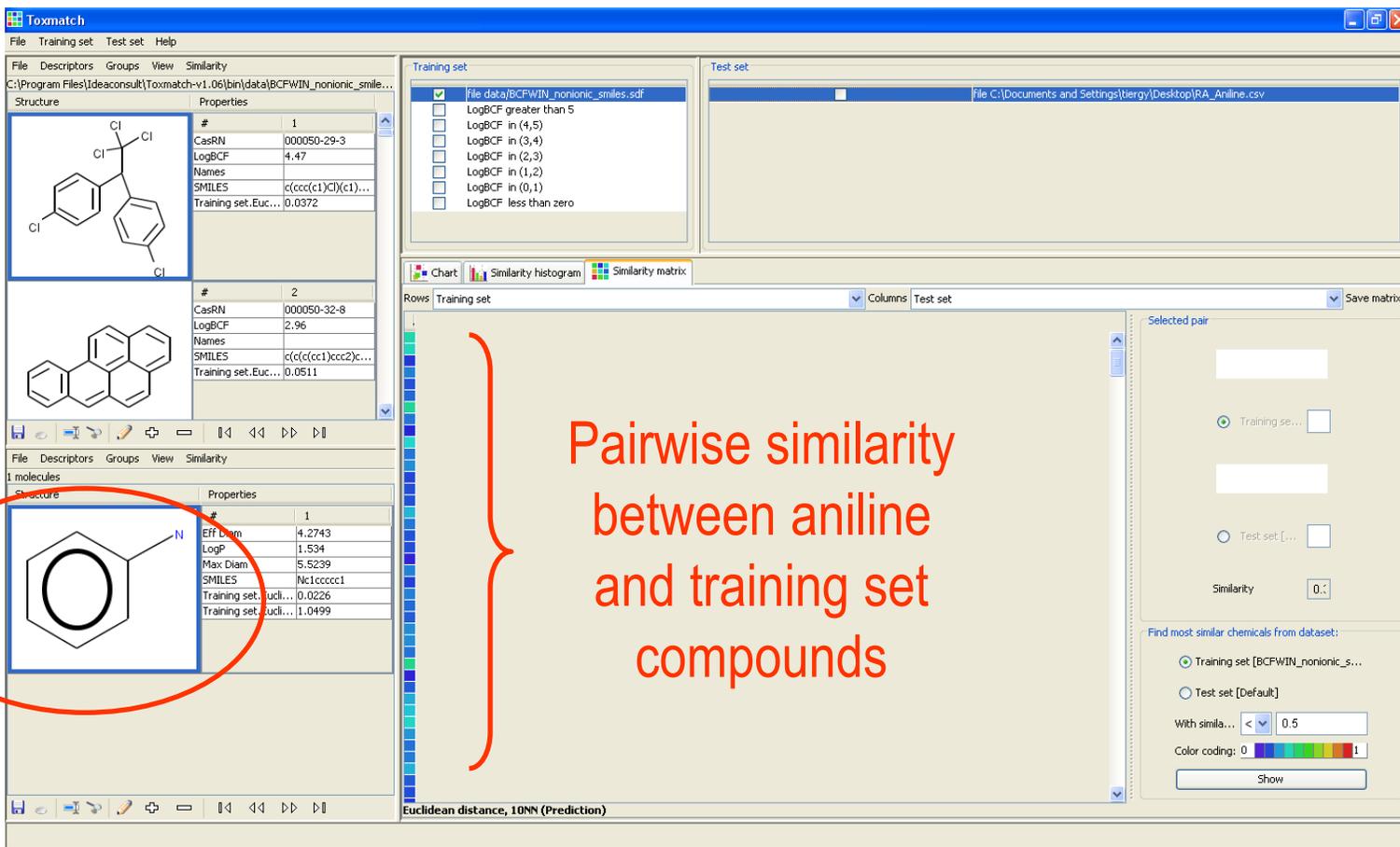
Many-to-one read-across of a quantitative property (k Nearest Neighbours)



Patlewicz G, Jeliaskova N, Gallegos Saliner A & Worth AP (2008). Toxmatch – A new software tool to aid in the development and evaluation of chemically similar groups. *SAR and QSAR in Environmental Research* 19, 397-412.

- BCF of aniline predicted on basis of effective diameter, maximum diameter and LogP
- Predicted LogBCF = 1.05
- Experimental LogBCF = 0.78 (Hazardous Substances Databank)

Training set of
610 chemicals



Training set

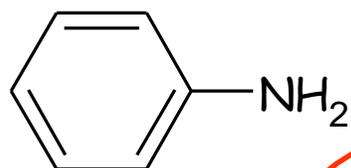
#	CasRN	LogBCF	Names	SMILES	Training set.Euc...
1	000050-29-3	4.47		<chem>c(ccc1)C(c1)...</chem>	0.0372
2	000050-32-8	2.96		<chem>c(c(c(cc1)ccc2)...</chem>	0.0511

Test set

#	CasRN	LogBCF	Names	SMILES	Training set.Euc...
1				<chem>Nc1ccccc1</chem>	1.0499

Pairwise similarity between aniline and training set compounds

Euclidean distance, 10NN (Prediction)

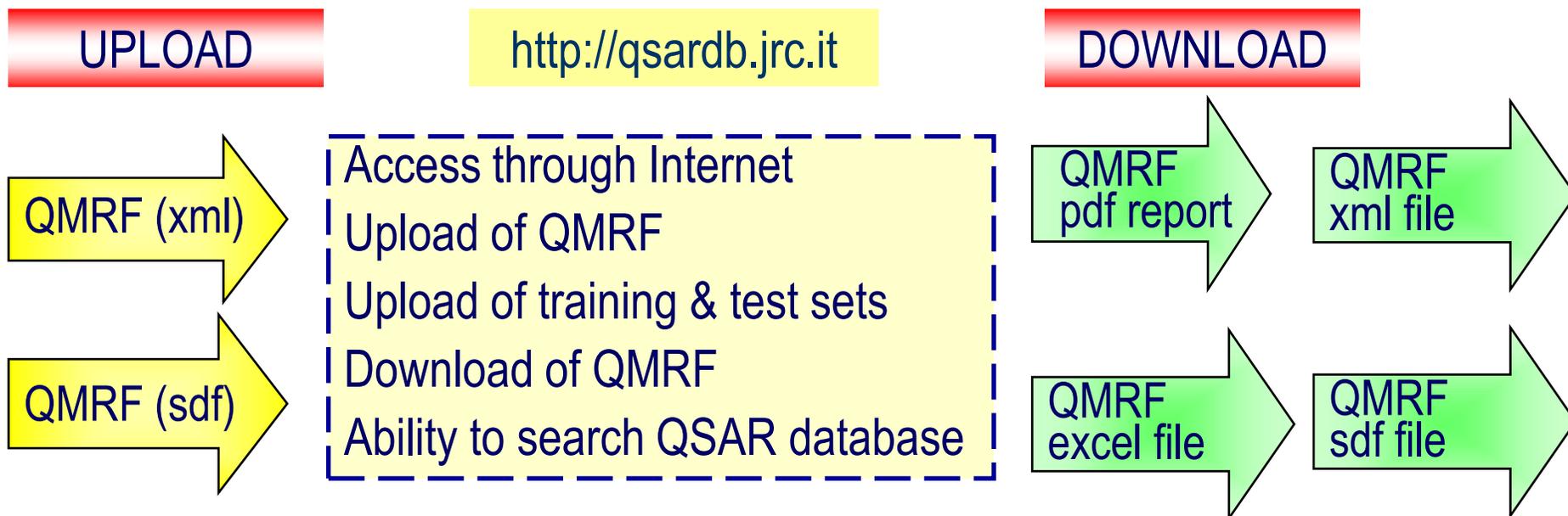


Aniline -
test chemical

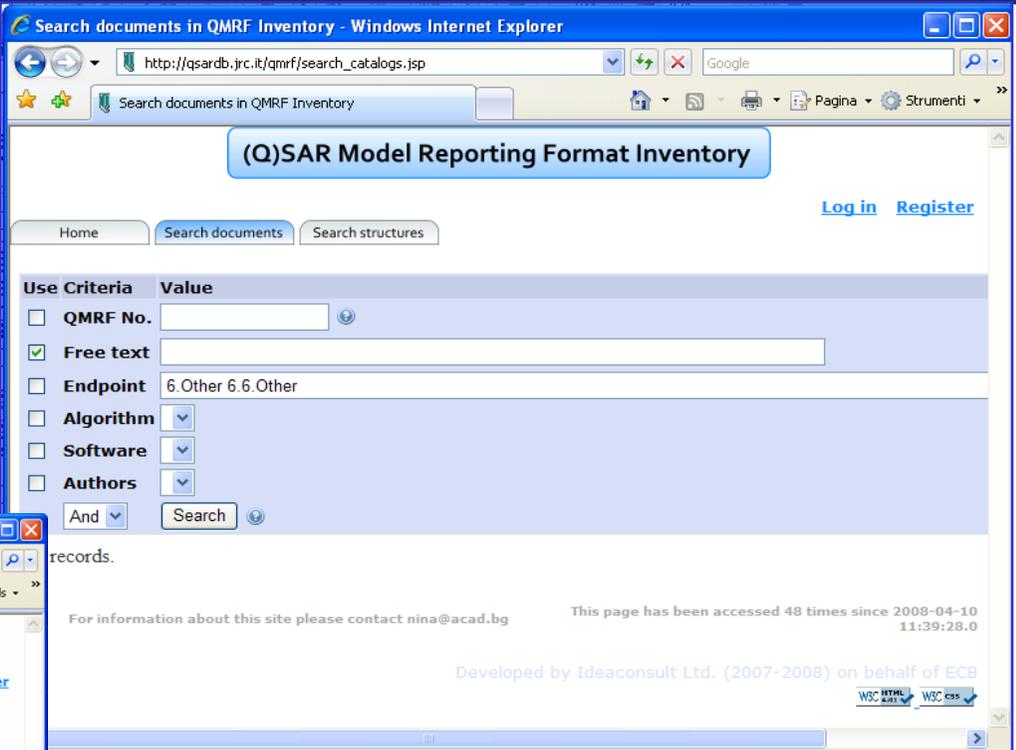
Need to identify and use **relevant, reliable and well documented** (Q)SARs

The **JRC QSAR Model Database** is a searchable inventory of peer-reviewed information on (Q)SAR models

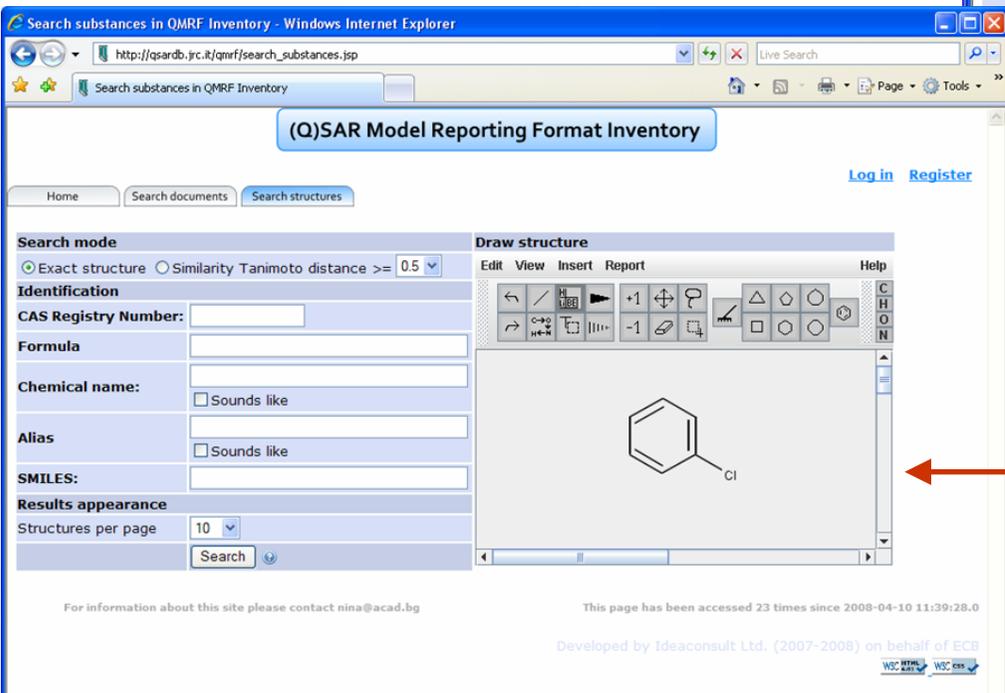
Developers and users of (Q)SAR models can submit information on (Q)SARs by using the **(Q)SAR Model Reporting Format (QMRF)**



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



Use Criteria	Value
<input type="checkbox"/> QMRF No.	<input type="text"/>
<input checked="" type="checkbox"/> Free text	<input type="text"/>
<input type="checkbox"/> Endpoint	6.Other 6.6.Other
<input type="checkbox"/> Algorithm	<input type="text"/>
<input type="checkbox"/> Software	<input type="text"/>
<input type="checkbox"/> Authors	<input type="text"/>



Search mode
 Exact structure Similarity Tanimoto distance >= 0.5

Identification
 CAS Registry Number:
 Formula:
 Chemical name: Sounds like
 Alias: Sounds like
 SMILES:

Results appearance
 Structures per page: 10

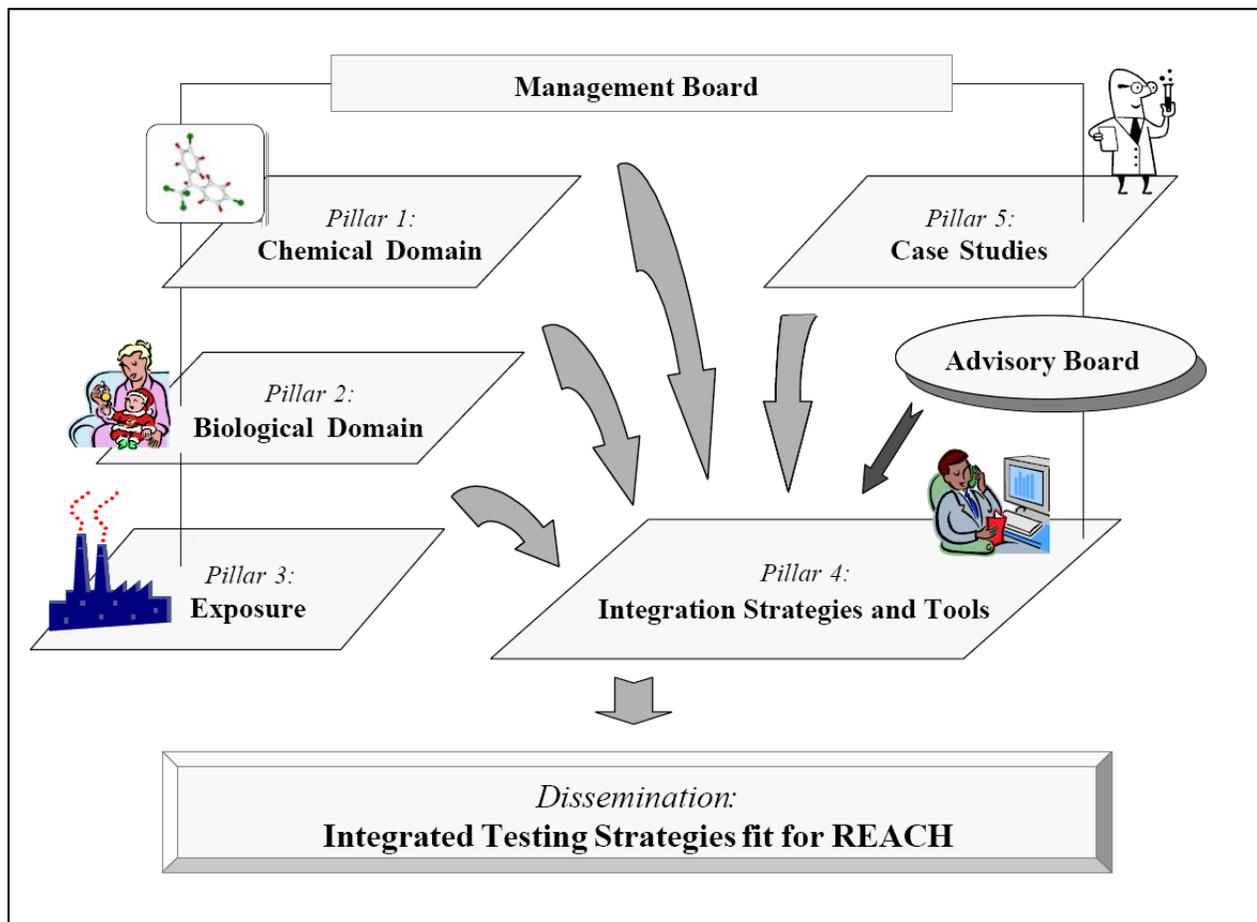
Draw structure
 Edit View Insert Report Help
c1ccccc1Cl

- CAS No.
- Formula
- Chemical name
- Alias
- SMILES

<http://qsardb.jrc.it>

- Need to assess the toxicological significance of pesticide active metabolites and degradation products (not tested under *Directive 91/414/EEC*)
- Three projects funded by EFSA (2009-2010)
 - Applicability of QSAR analysis in assessing metabolite toxicity
 - Applicability of the TTC concept in assessing metabolite toxicity
 - Impact of metabolism on toxicological properties
- Next steps by EFSA
 - Opinion of the PPR panel (2010-2011)
 - Guidance document on pesticide residue definition for dietary risk assessment (2011-2012)

Optimized Strategies for Risk Assessment of Industrial Chemicals through Integration of Non-Test and Test Information (OSIRIS)



- To optimise the use of non-testing data, a conceptual framework is provided in the REACH guidance documentation
- There is a need to incorporate mechanistic knowledge in the models (e.g. based on chemical reactivity and “omic” data)
- An increasing number of models are being implemented in a range of software tools
- There is a need to facilitate the use of multiple tools by developing automated workflows
- Further guidance is needed on how to assess the adequacy of non-testing and alternative test data by weight-of-evidence approaches