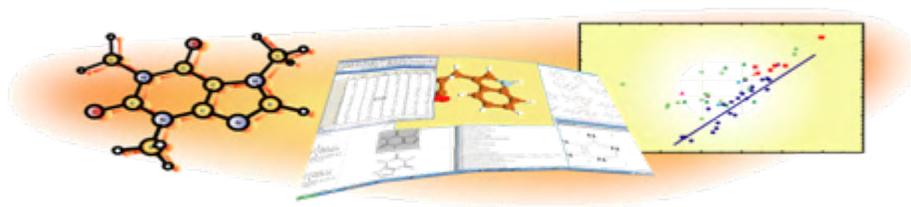


# A Tiered Approach for the Use of Non-Testing Methods in the Regulatory Assessment of Chemicals

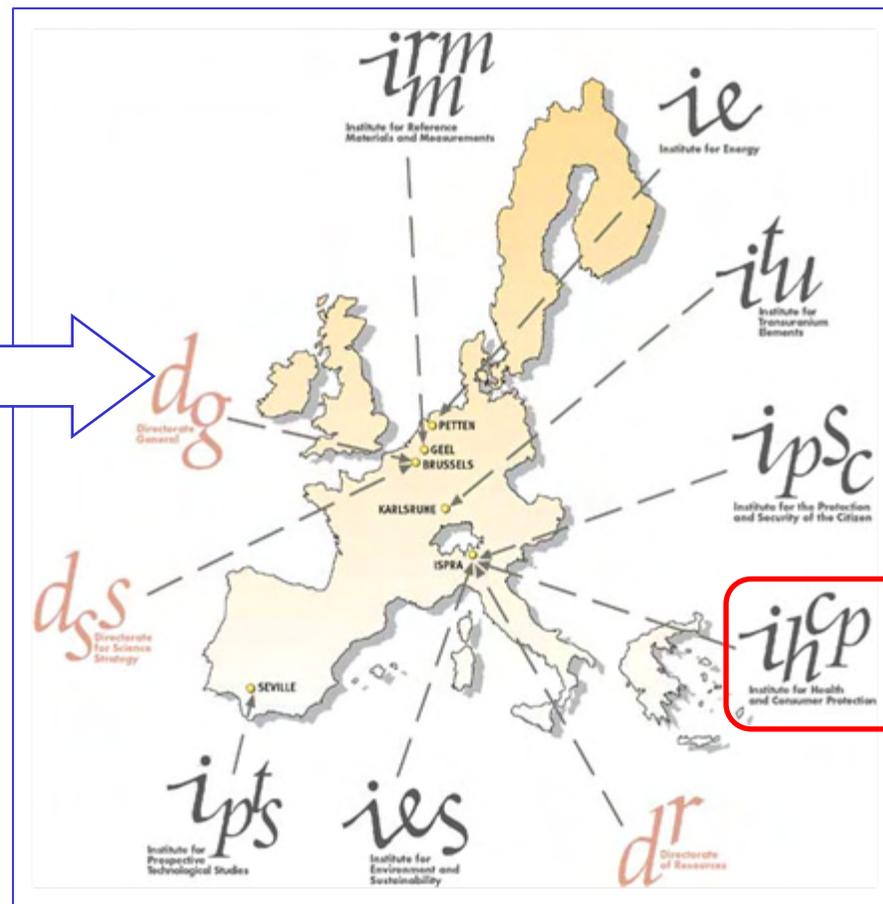
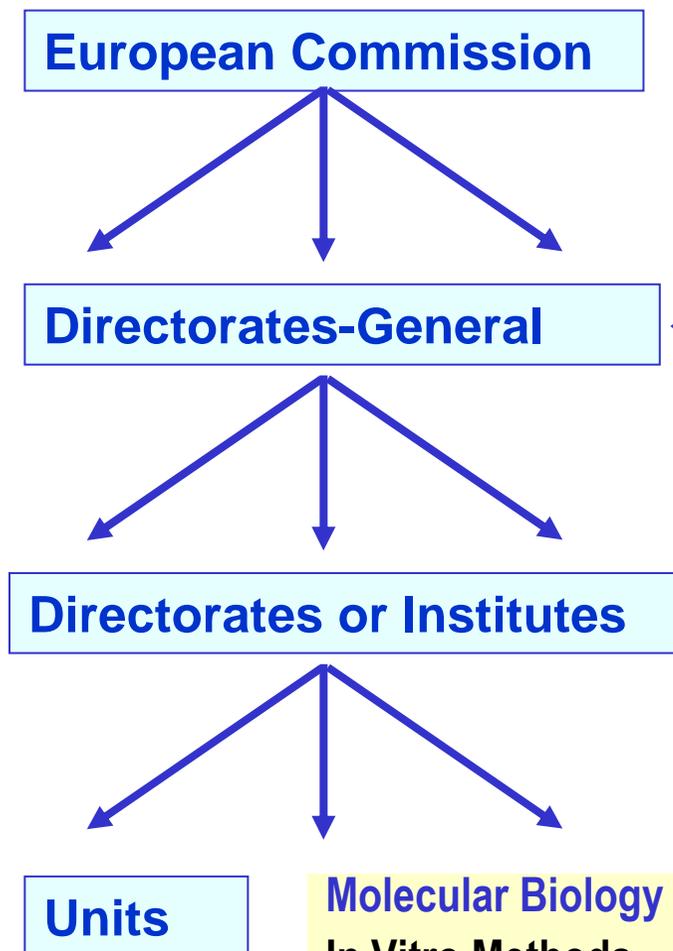
Andrew Worth  
Systems Toxicology Unit  
Institute for Health & Consumer Protection  
Joint Research Centre, European Commission

**US EPA Chemical Prioritisation meeting; 25 June 2009**

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



1. **Computational toxicology at the JRC**
2. **Role of non-testing methods in hazard and risk assessment**
3. **Assessing and documenting the adequacy of predictive methods**
4. **Non-testing strategy – a stepwise approach**
5. **Conclusions**



**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**

**Computational Toxicology:** development, assessment and application of computer-based assessment methods (e.g. QSARs and biophysical models) for chemicals

**High-throughput screening:** automation of robust and informative cell-based and biochemical assays, to generate high quality *in vitro* data for the targeted assessment of chemicals

**Metabolomics:** use of metabolomics and metabonomics to describe the metabolic status and biochemical events associated with a cellular or biological system, both in its steady-state and in its dynamic responses to environmental stressors such as chemicals

**Chemometrics and Biostatistics:** application of statistical approaches to support IHCP projects, e.g. experimental design and analysis of validation studies

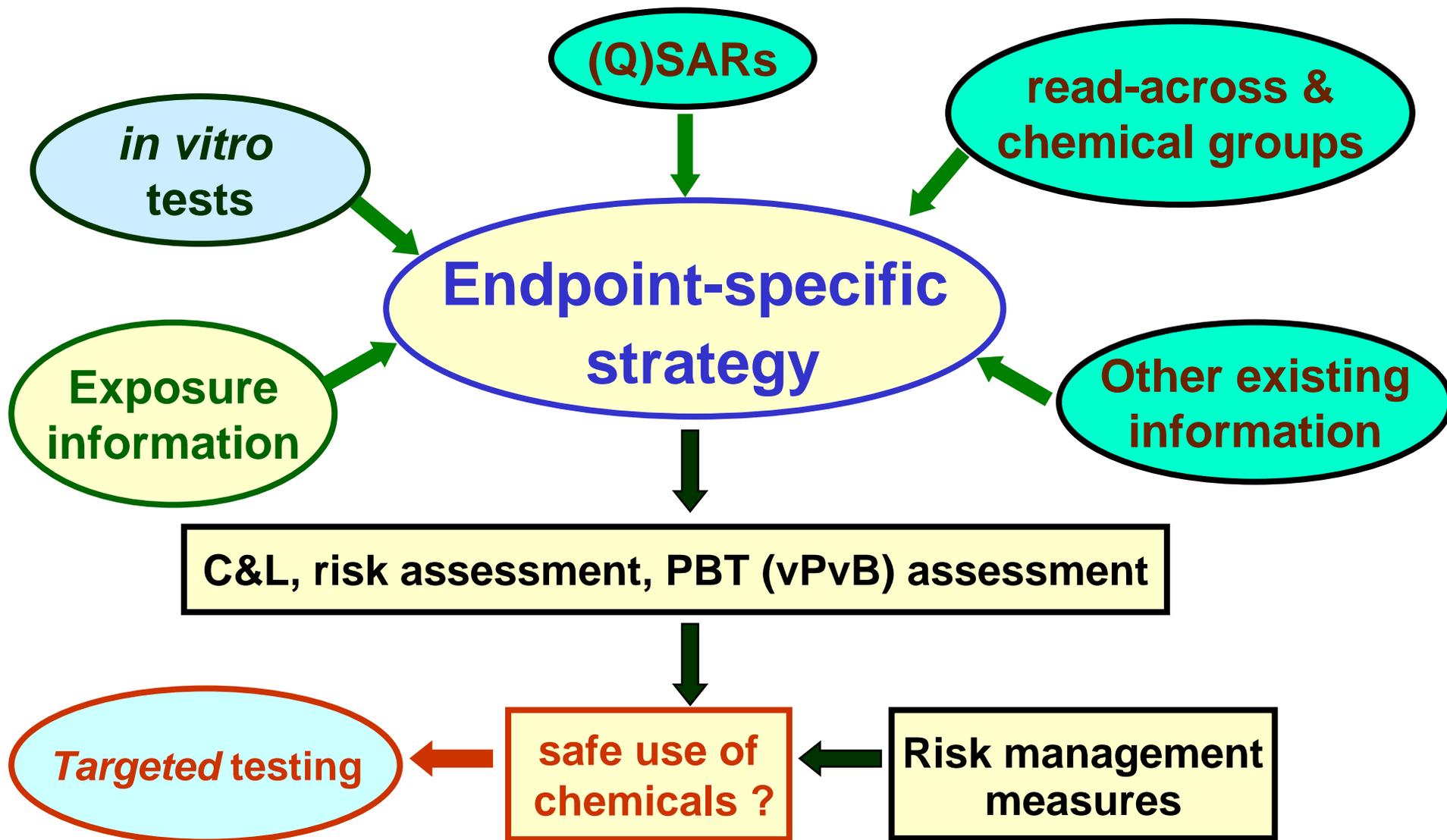
Information on chemical properties, fate and (eco)toxicological effects is used for various regulatory purposes: classification & labelling, risk assessment, PBT and vPvB assessment

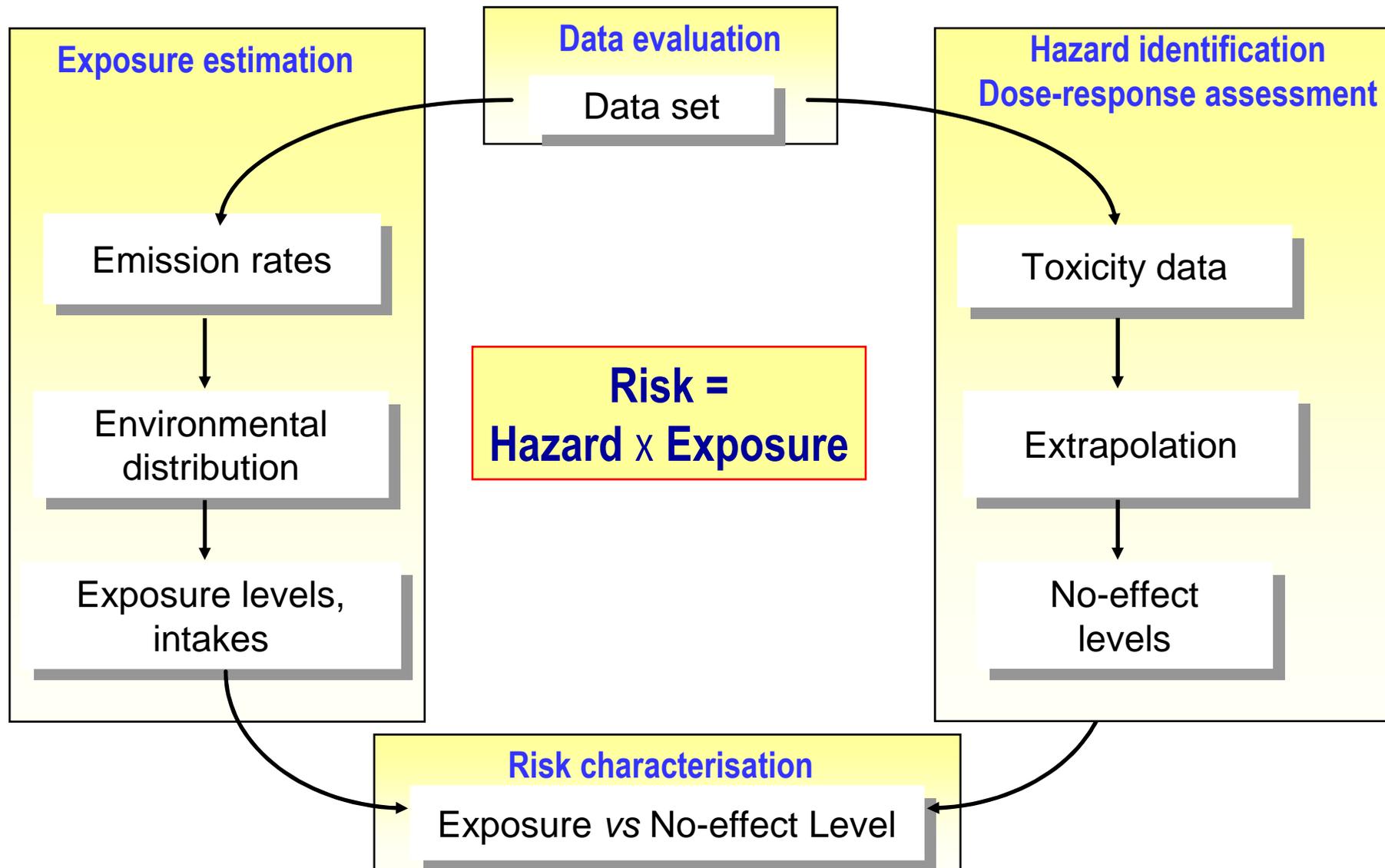
Information requirements are largely tonnage dependent, however ...

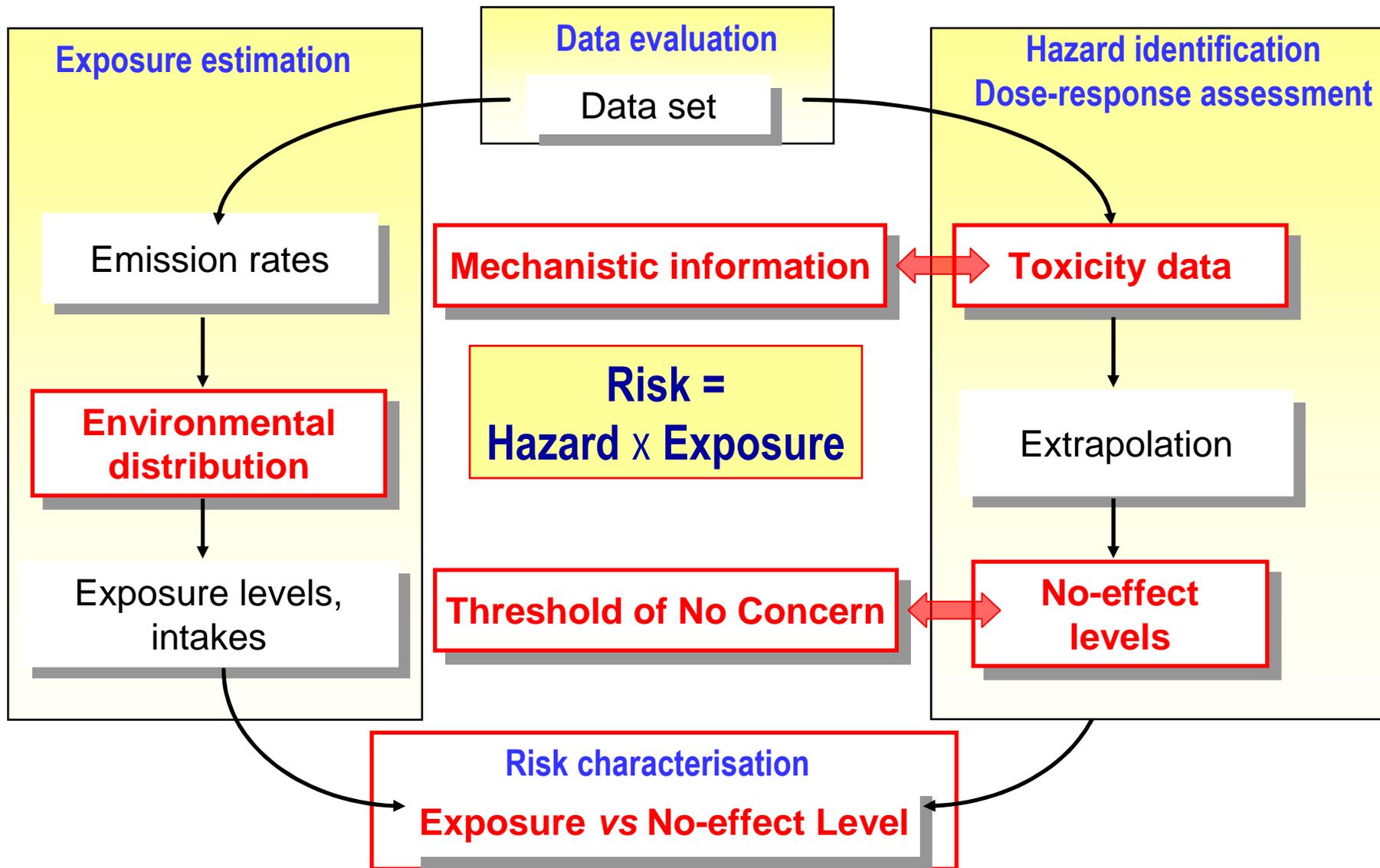
... (animal) testing can be reduced or avoided by “replacing traditional test data with predictions or equivalent data”

- *in silico* predictions (SARs, QSARs, expert systems, read-across)
- *in vitro* data
- Integrated Testing Strategies

**“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, REACH)**

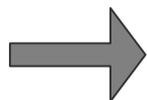




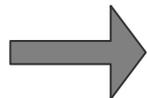


In order for a (Q)SAR result to be adequate for a given regulatory purpose, the following conditions must be fulfilled:

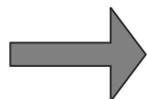
- the estimate should be generated by a valid (reliable) model
- the model should be applicable to the chemical of interest with the necessary level of reliability
- the model endpoint should be relevant for the regulatory purpose



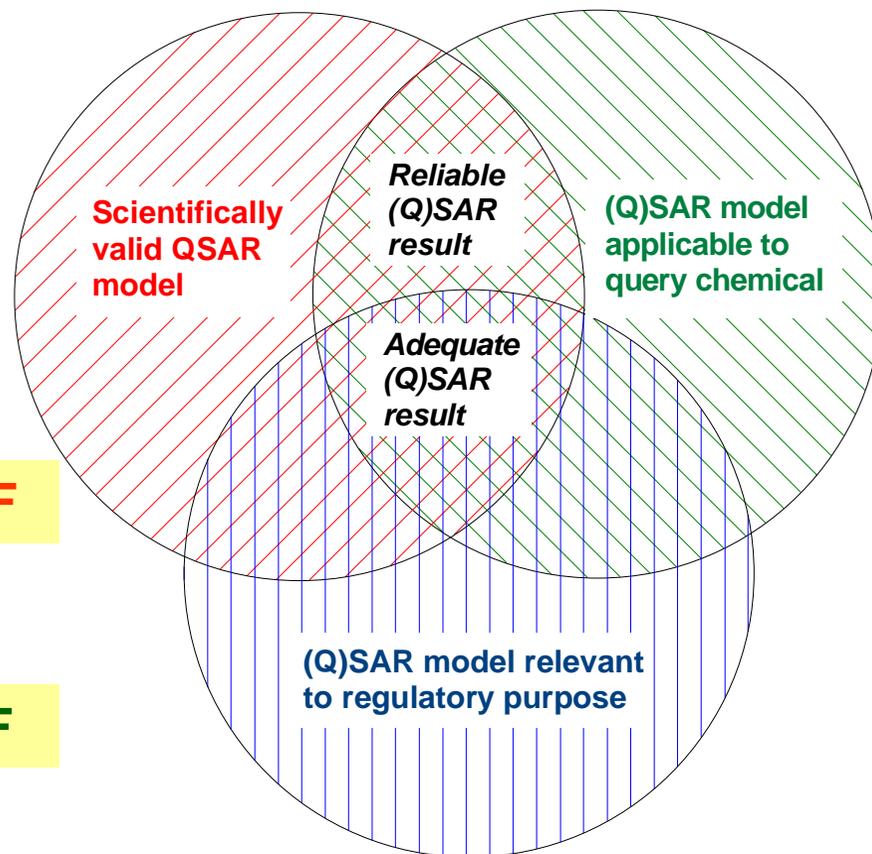
**QMRF**



**QPRF**



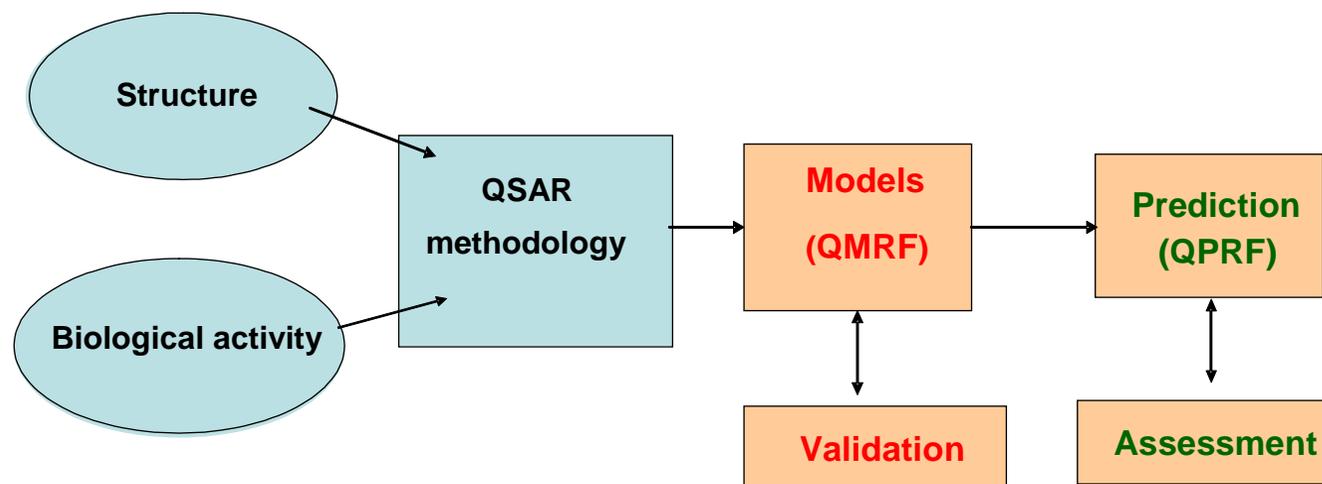
**QPRF**



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

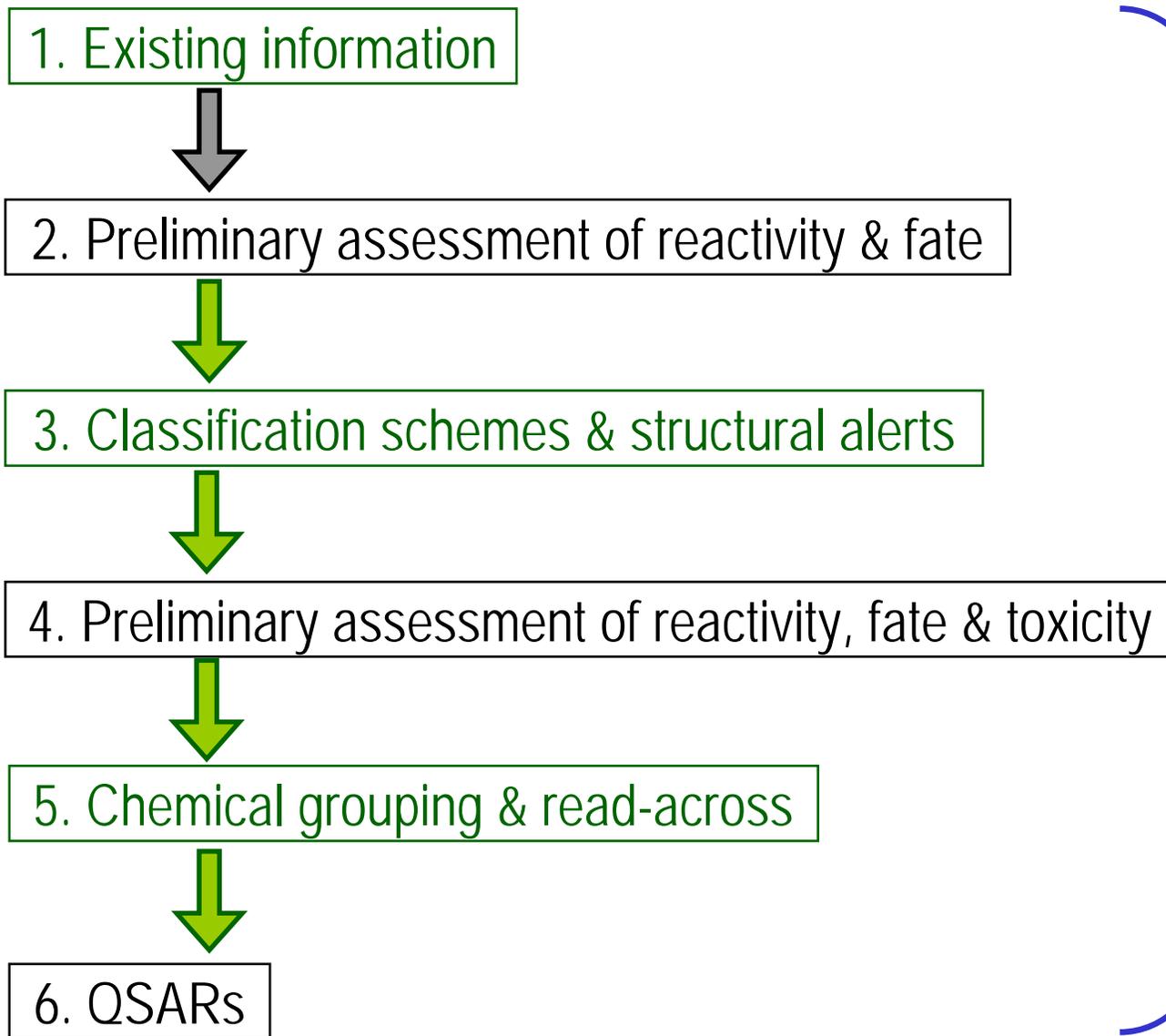
A (Q)SAR should be associated with the following information:

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)

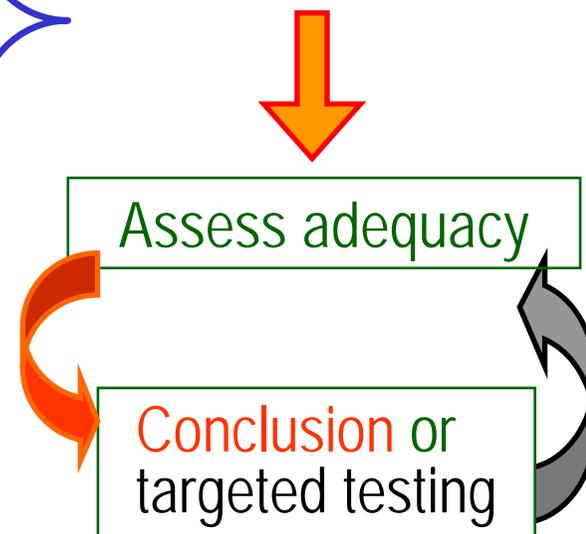
**QPRF** captures information on the substance and its prediction, and is intended to facilitate considerations of the adequacy of a prediction

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

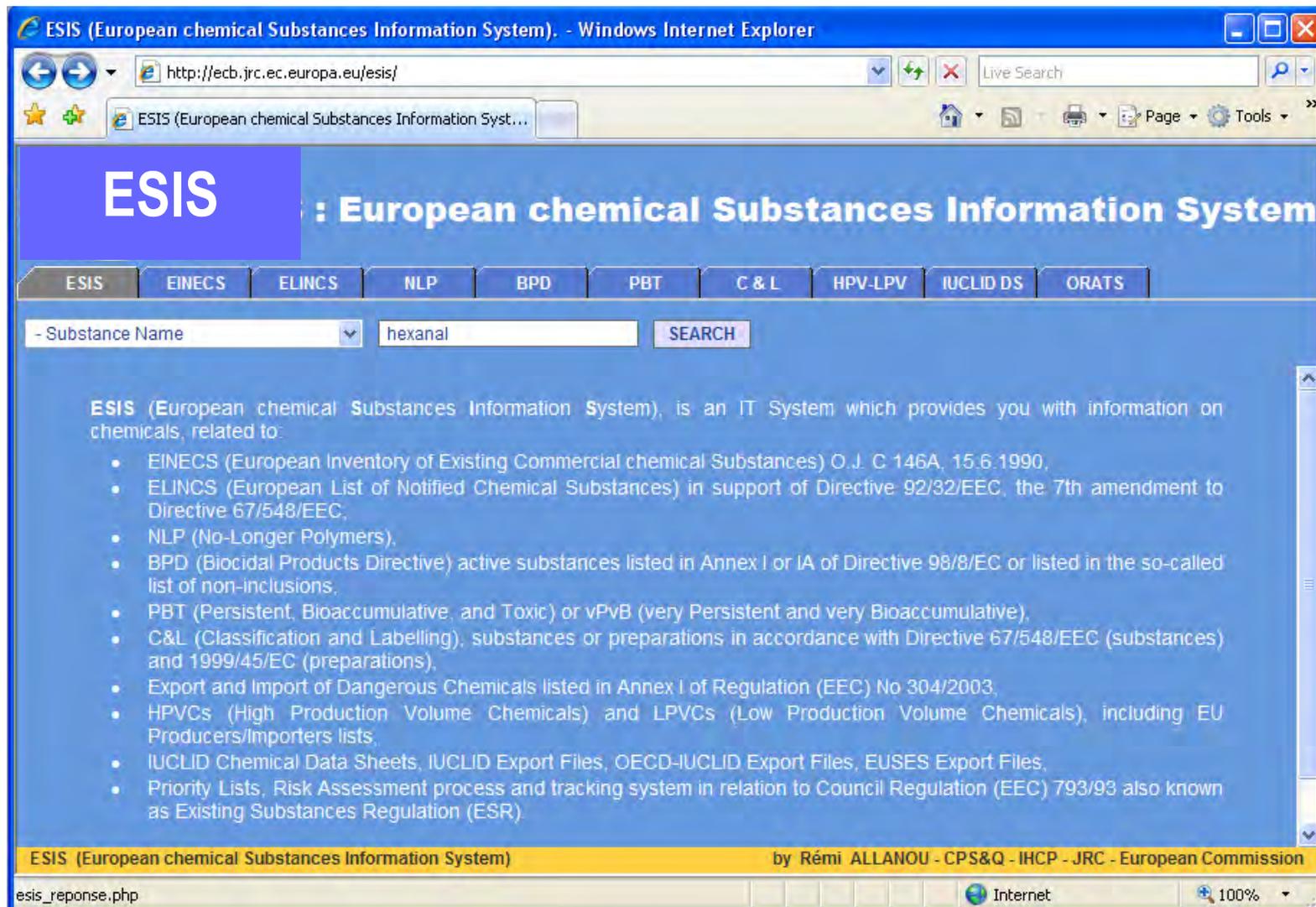


## Working Matrix

	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)



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

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

ESIS (European chemical Substances Information Syst...)

## 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/>

European Commission  
**Joint Research Centre**  
Institute for Health and Consumer Protection

European Commission JRC IHCP CPSQ ENDDA New Search

Home Search External Resources Qsar Models

Search

name:   
cas:   
smiles:

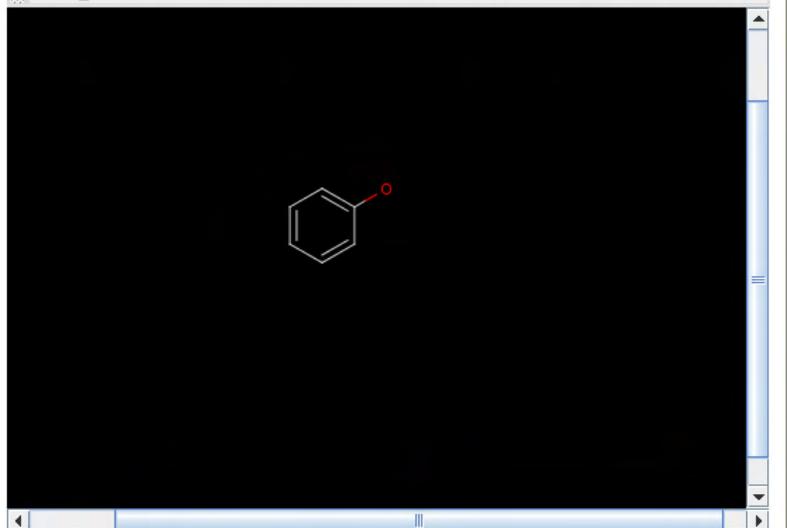
Search by Structure

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: Substructure

**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)
- Dinitroanilide
- Dioxins
- Diphenylpro
- Dithiocarba
- Furans
- Hexachloro
- Hydroxyben
- Linuron, diu
- Methoxychl
- Musk Fragre

**Moa**

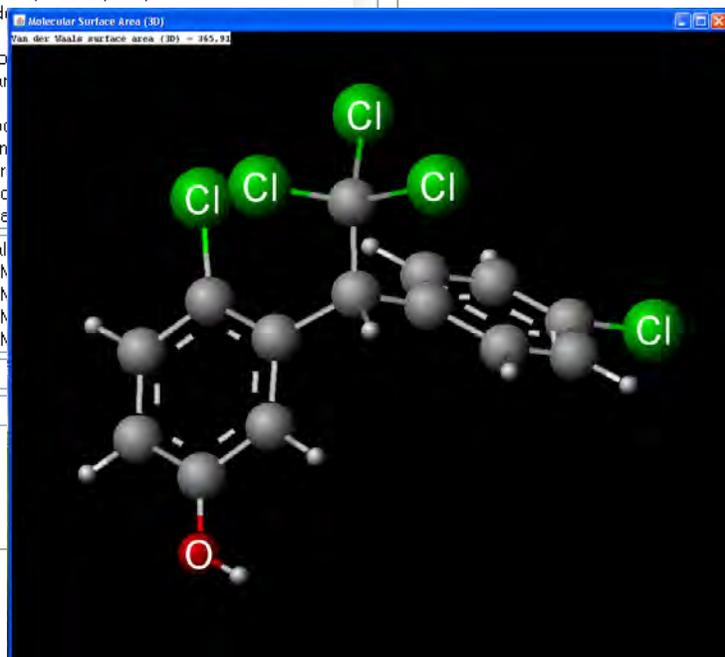
-all- (+)e (+)th +e +pi

mol. w. between   
logP between

Send Reset

Molecular Surface Area (3D)

Van der Waals surface area (3D) = 165.91



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 ...
  - *InSilico*First (MetaboGen and CRAFT)

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

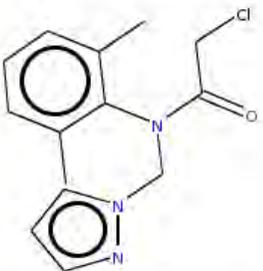
File Edit Chemical Compounds Toxic Hazard Method Help

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'**

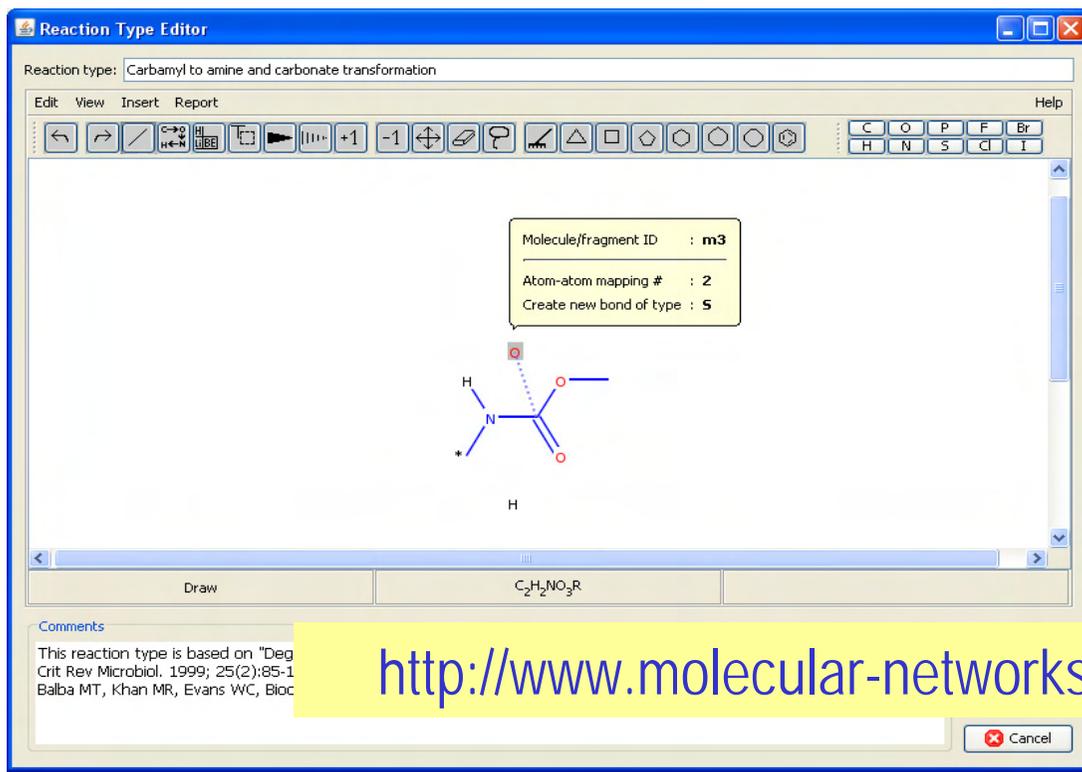


There are example molecules for each rule outcome. Select which one to display.

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

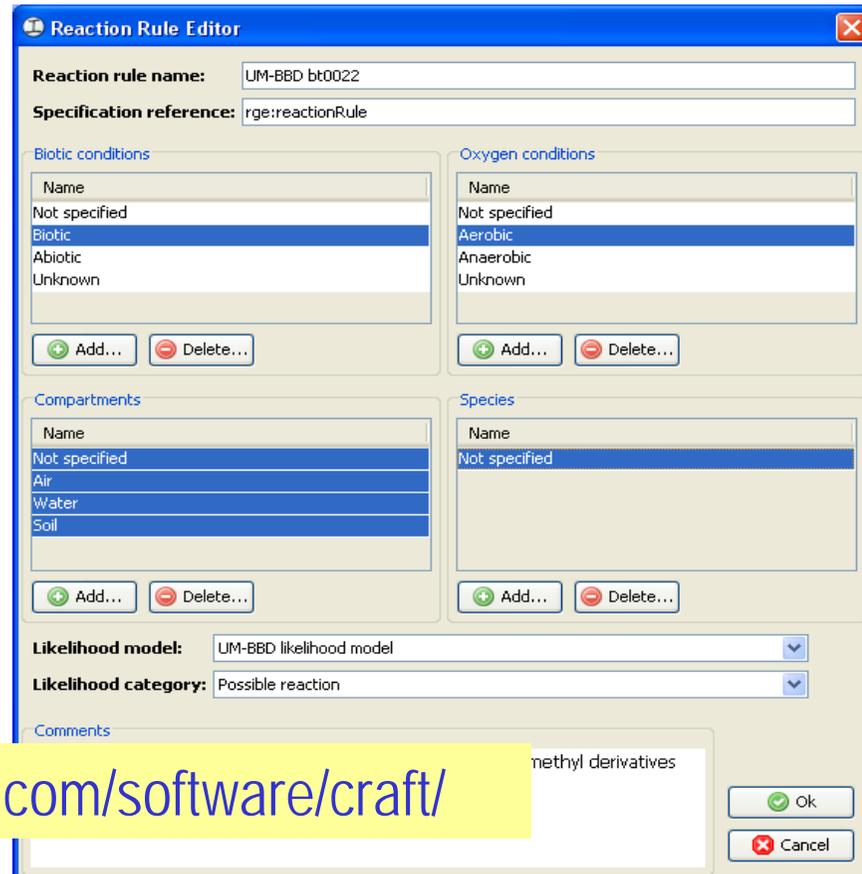


Reaction type: Carbamyl to amine and carbonate transformation

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

Draw C<sub>2</sub>H<sub>2</sub>NO<sub>3</sub>R

Comments  
This reaction type is based on "Deg Crit Rev Microbiol. 1999; 25(2):85-1 Balba MT, Khan MR, Evans WC, Bioc



Reaction rule name: UM-BBD bt0022  
Specification reference: rge:reactionRule

**Biotic conditions**

Name
Not specified
Biotic
Abiotic
Unknown

**Oxygen conditions**

Name
Not specified
Aerobic
Anaerobic
Unknown

**Compartments**

Name
Not specified
Air
Water
Soil

**Species**

Name
Not specified

Likelihood model: UM-BBD likelihood model  
Likelihood category: Possible reaction

Comments  
methyl derivatives

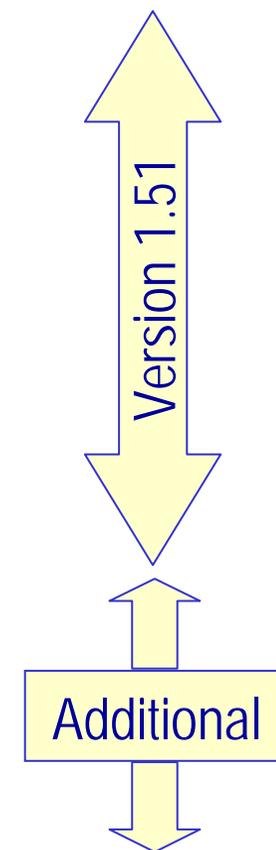
<http://www.molecular-networks.com/software/craft/>

- 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 ...
  - *Insilico*first consortium (Multicase Inc, Lhasa Ltd, Molecular Networks GmbH, Leadscope Inc)
- QSAR Model Databases (QMDBs)
  - JRC QSAR Model Database
  - OECD Toolbox

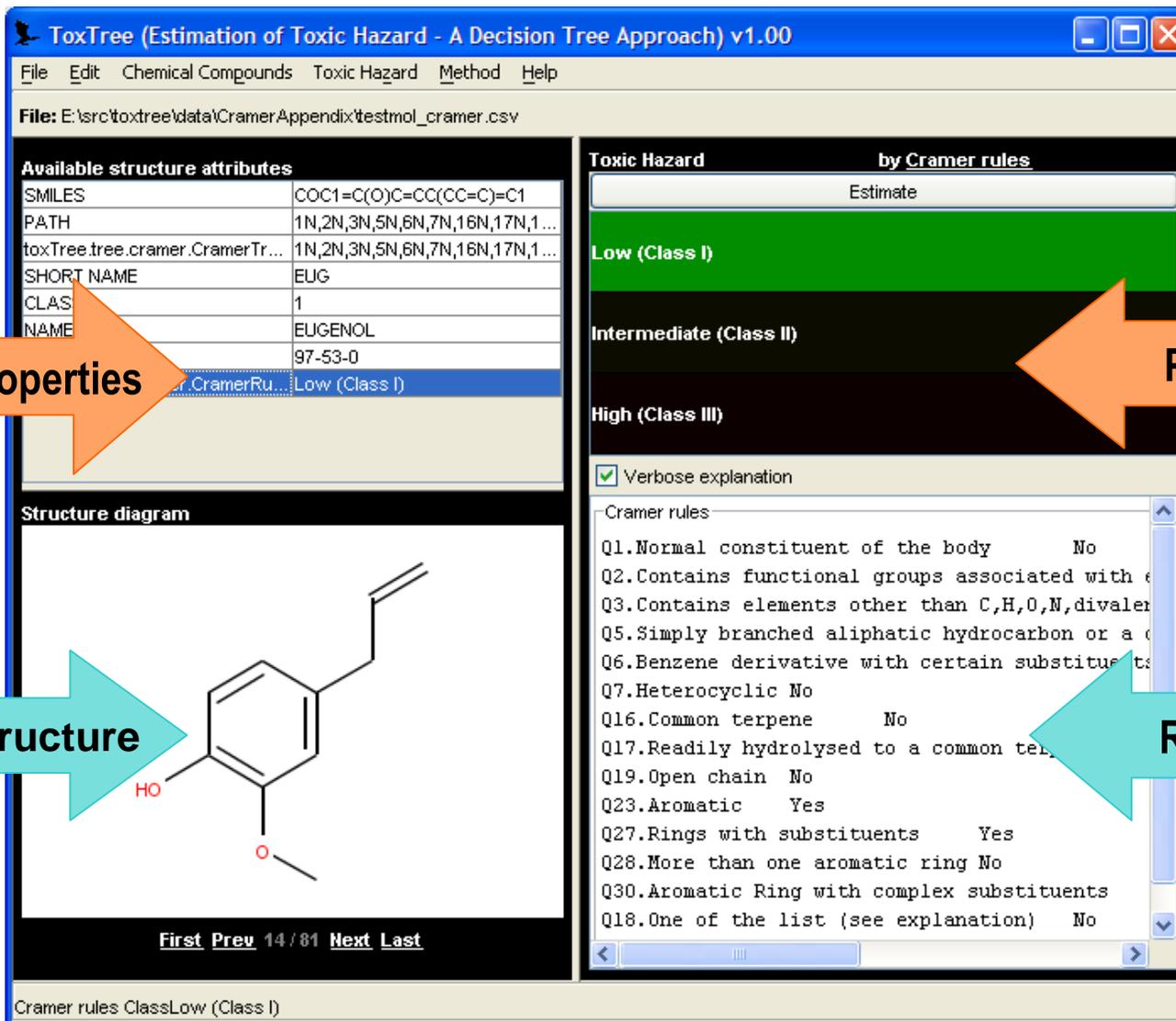
Toxtree is a flexible, user-friendly, open source application, which is able to estimate toxic hazard by applying decision tree approaches

Rulebases available:

- Oral systemic toxicity (Cramer scheme)
- Acute Fish Toxicity (Verhaar scheme)
- Skin irritation & corrosion potential (BfR rulebase)
- Eye irritation & corrosion potential (BfR rulebase)
- Mutagenicity & carcinogenicity (Benigni-Bossa rulebase)
- Mutagenicity & carcinogenicity (*In Vivo* Micronucleus rulebase)
- Biodegradation potential (START 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
CLAS	1
NAME	EUGENOL
	97-53-0
	Low (Class I)

**Structure diagram**

COC1=C(O)C=CC(CC=C)=C1

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

The Cramer classification scheme is probably the best known approach for structuring chemicals in order to estimate a **Threshold of Toxicological Concern**.

Chemicals are divided into **three structural classes** based on a decision tree. This comprises 33 structural rules and places evaluated compounds into one of three classes:

**Class I** substances are simple chemical structures with efficient modes of metabolism suggesting a **low order of oral toxicity**

**Class II** are of **intermediate toxicity**

**Class III** substances are those that permit no strong initial presumption of safety, or may even suggest **significant toxicity** or have reactive functional groups

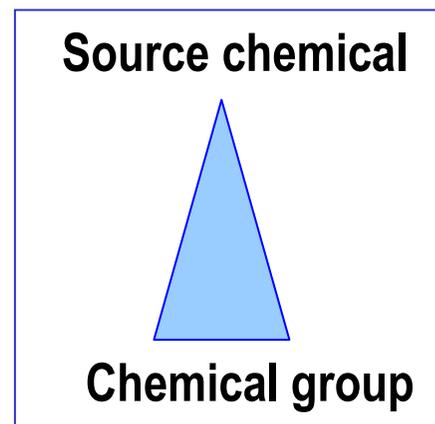
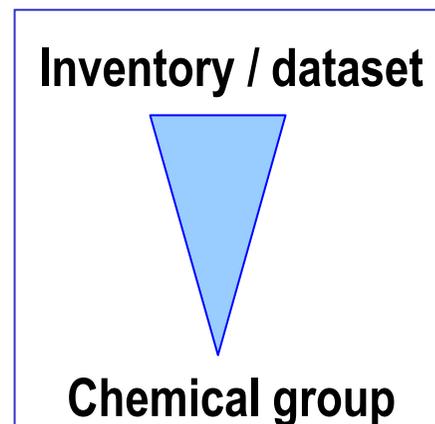
**Cramer GM, Ford RA & Hall RL (1978). Estimation of Toxic Hazard - A Decision Tree Approach. *J. Cosmet. Toxicol.*, Vol.16, pp. 255-276, Pergamon Press.**

**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.**



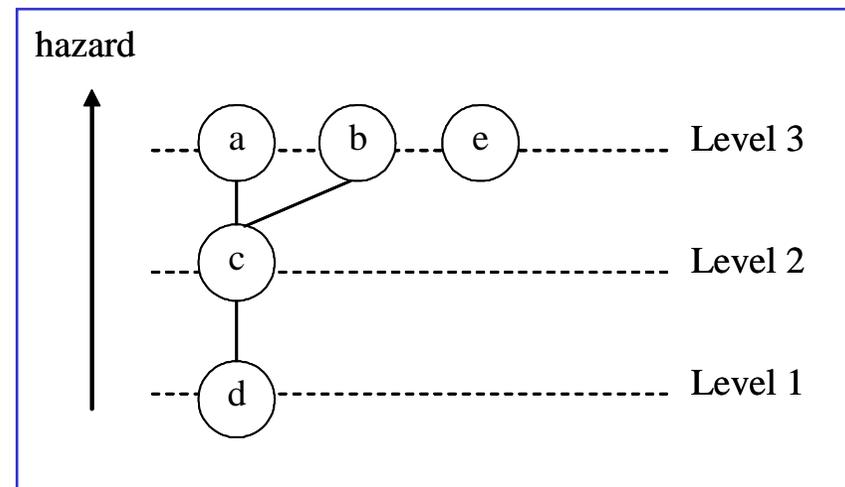
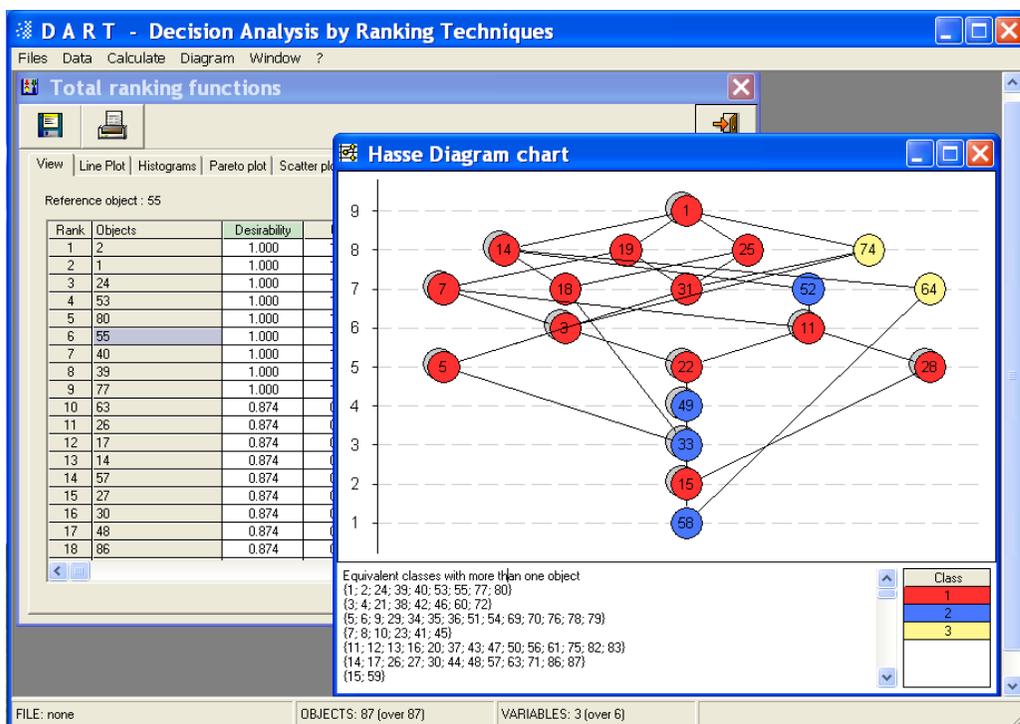
## Chemical read-across within analogue and category approaches

- 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)



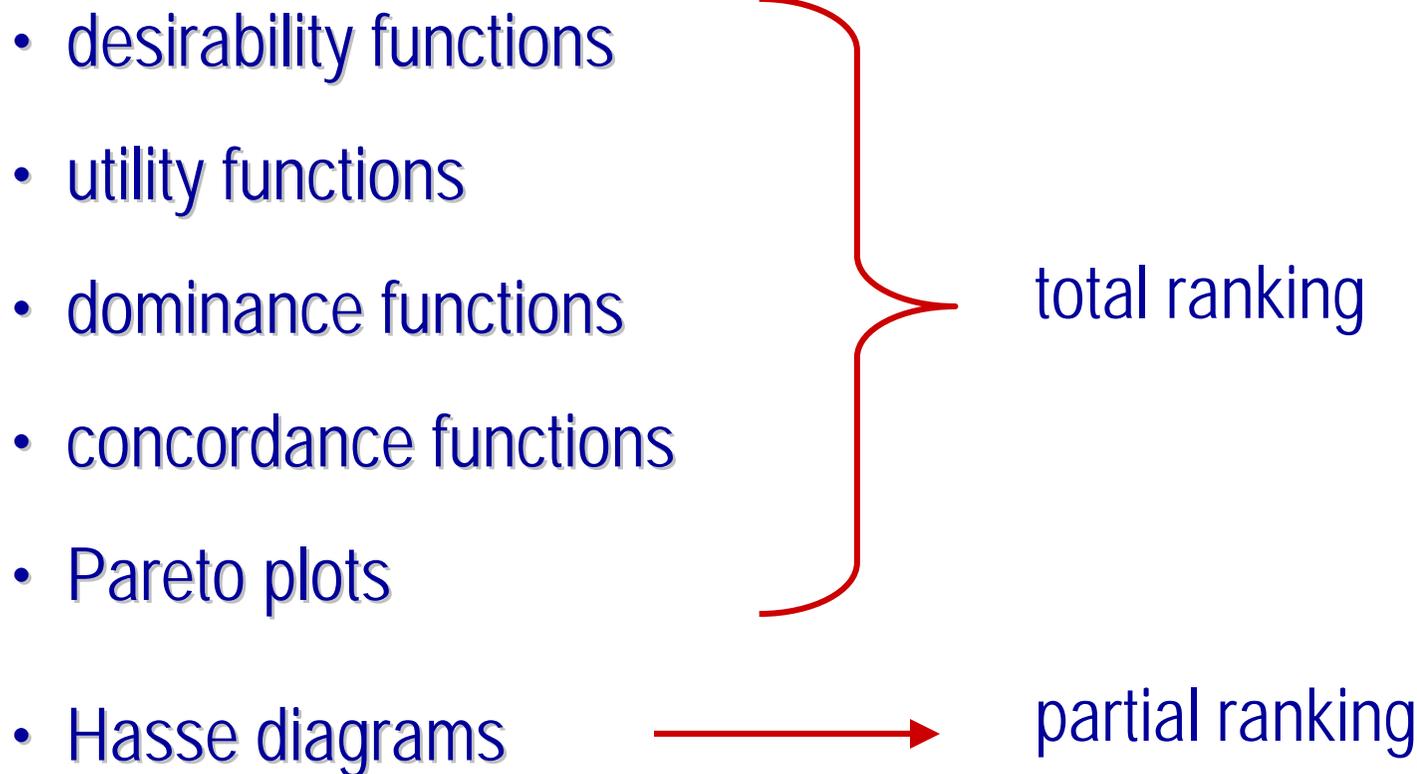
Worth A et al (2007). The Use of Computational Methods in the Grouping and Assessment of Chemicals - Preliminary Investigations. EUR 22941 EN

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



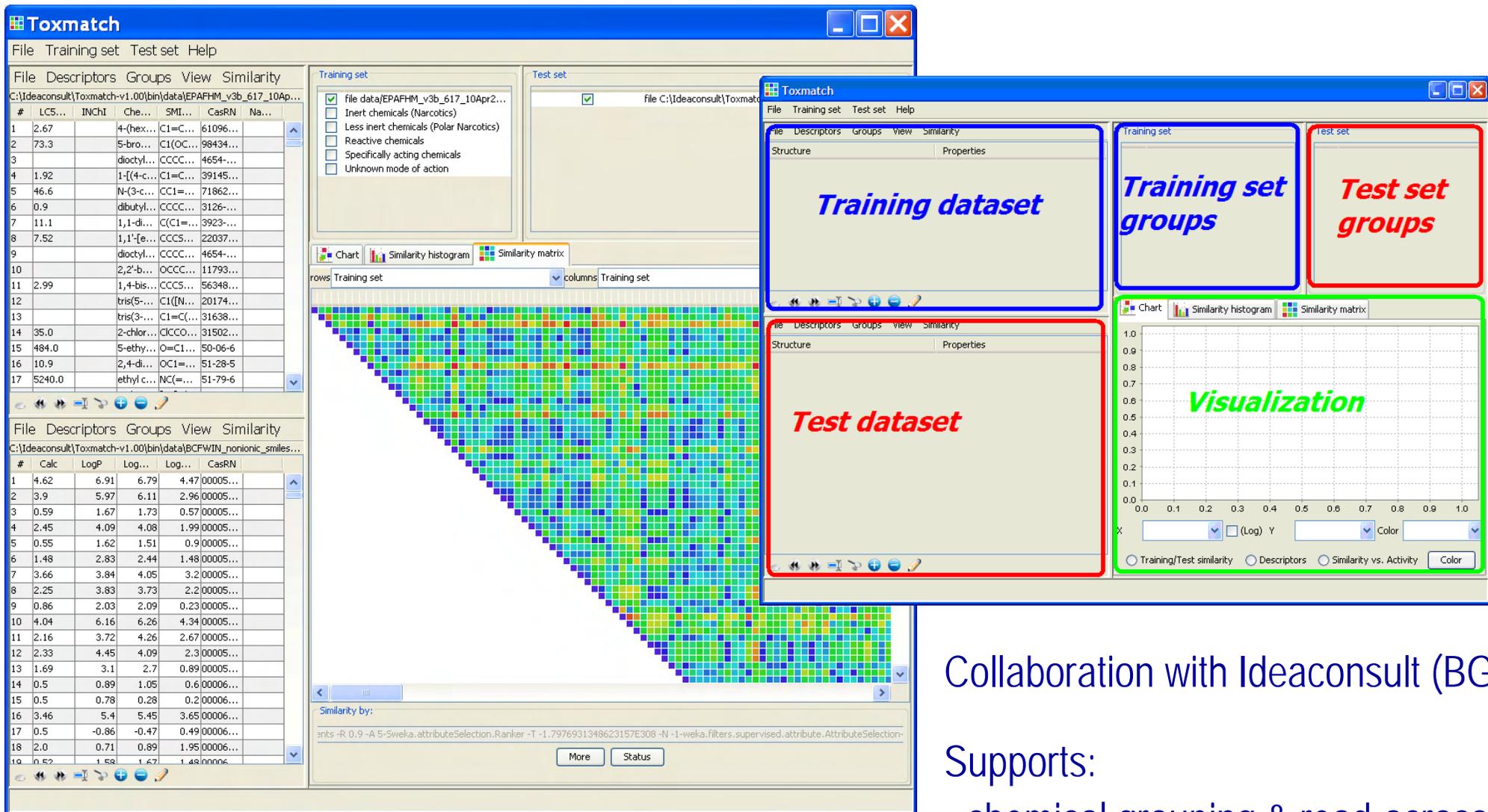
- collaboration with Talete srl (Italy)
- supports priority setting of chemicals

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.



Pavan M & Todeschini R (2008). Optimization: Multi-criteria Decision Making methods, in *Comprehensive Chemometrics*. B Walczak, RT Ferré & S Brown (Eds), in press. Elsevier.

Pavan M & Todeschini R, Eds (2008). *Scientific Data Ranking Methods: Theory and Applications*. 1st Edition. Elsevier, pp. 51-72.



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 single entry 'file C:\Ideaconsult\Toxmat'.
- Training dataset:** A panel labeled 'Training dataset' in blue text, outlined in blue.
- Test set groups:** A panel labeled 'Test set groups' in red text, outlined in red.
- Similarity matrix:** A large triangular matrix of colored squares (green, yellow, red, blue) representing similarity values between training and test set chemicals.
- Visualization:** A plot area labeled 'Visualization' in green text, outlined in green, showing a grid with axes from 0.0 to 1.0. Below the plot are options for 'X' and 'Y' axes (Log/Color) and radio buttons for 'Training/Test similarity', 'Descriptors', and 'Similarity vs. Activity'.
- Tables:** Two tables on the left side of the interface. The top table has columns: #, LC5..., INChI, Che..., SMI..., CasRN, Na... The bottom table has columns: #, Calc, LogP, Log..., Log..., CasRN.
- Similarity by:** A section at the bottom with a 'More' button and a 'Status' button.

Collaboration with Ideaconsult (BG)

Supports:

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

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

Toxmatch includes descriptor-based (distance-like) and fragment-based (correlation-like) similarity indices

- Euclidean distance
- Cosine similarity
- Hodgkin-Richards index
- Tanimoto distance on descriptors
- Tanimoto distance on fingerprints
- Hellinger distance on atom environments
- Maximum Common Structure similarity

$$D_{AB}(k, x) = [Z_{AA} + Z_{BB} - 2Z_{AB}]^{1/2}$$

$$C_{AB} = Z_{AB} [Z_{AA} Z_{BB}]^{-1/2}$$

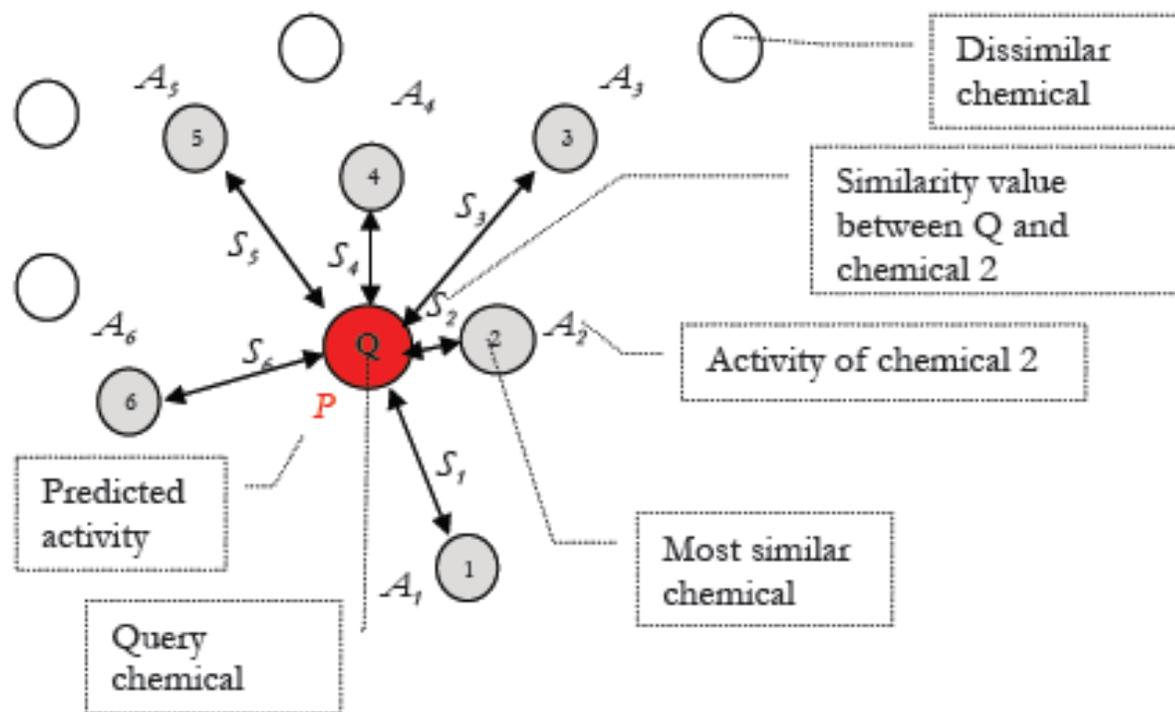
$$H_{AB} = 2Z_{AB} [Z_{AA} + Z_{BB}]^{-1}$$

$$T_{AB} = 2Z_{AB} [Z_{AA} + Z_{BB} - Z_{AB}]^{-1}$$

$$T_{AB} = \frac{N_{AB}}{N_A + N_B - N_{AB}}$$

**Gallegos Saliner A & Worth AP (2007). Development and Beta Testing of the Toxmatch Similarity Tool. JRC report EUR 22854 EN.**

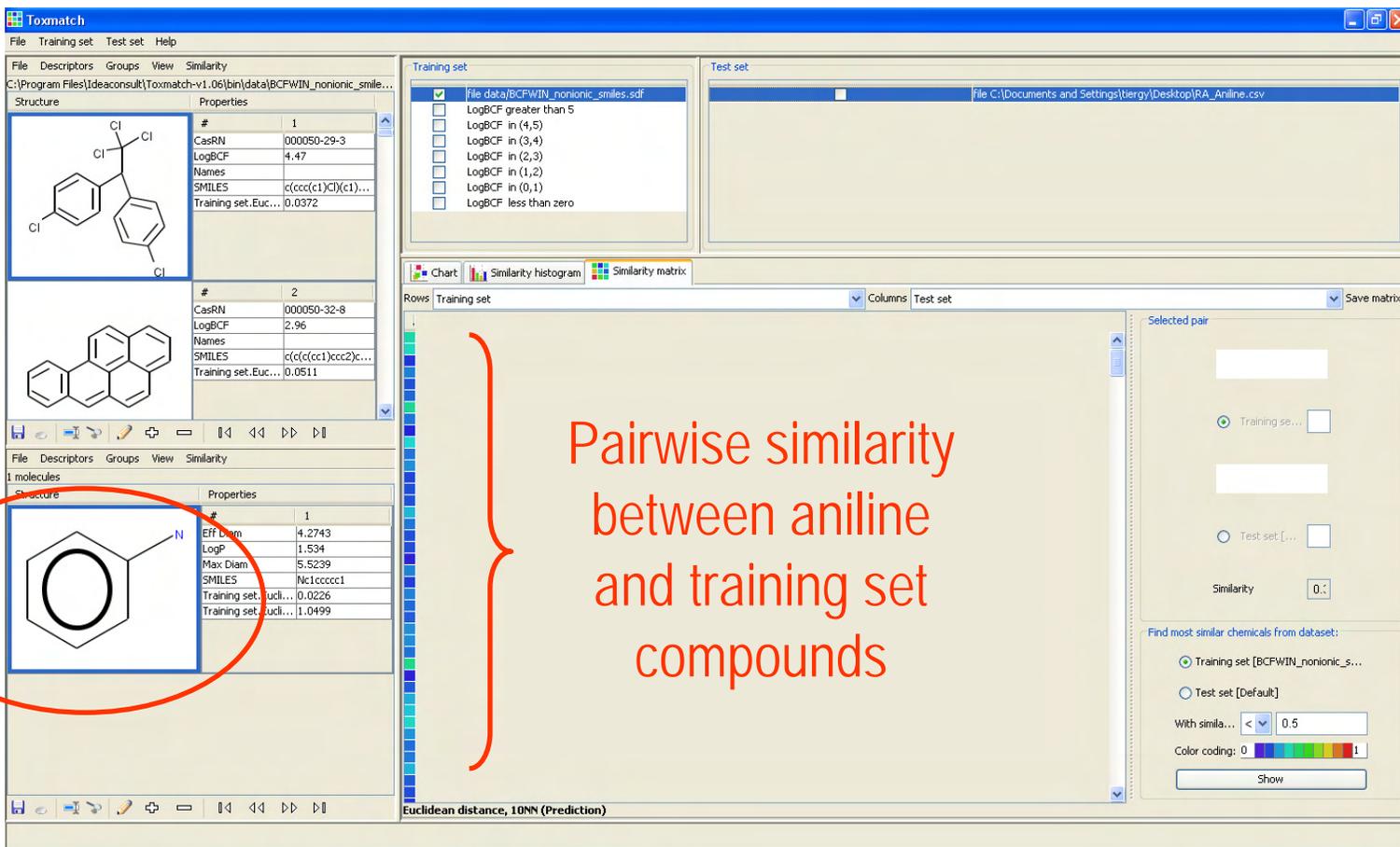
Many-to-one read-across of a quantitative property ( $k$  Nearest Neighbours)



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

Test set

File Descriptors Groups View Similarity

Structure Properties

#	1
CasRN	000050-29-3
LogBCF	4.47
Names	
SMILES	c(ccc1)C(c1)...
Training set.Euc...	0.0372

File Descriptors Groups View Similarity

#	2
CasRN	000050-32-8
LogBCF	2.96
Names	
SMILES	c(c(c(cc1)ccc2)...
Training set.Euc...	0.0511

File Descriptors Groups View Similarity

#	1
Eff Diam	4.2743
LogP	1.534
Max Diam	5.5239
SMILES	Nc1ccccc1
Training set. ucl...	0.0226
Training set. ucl...	1.0499

Rows: Training set Columns: Test set

Selected pair

Training set [BCFWIN\_nonionic\_s...

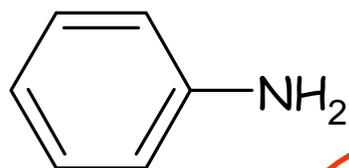
Test set [Default]

With simila... 0.5

Color coding: 0 1

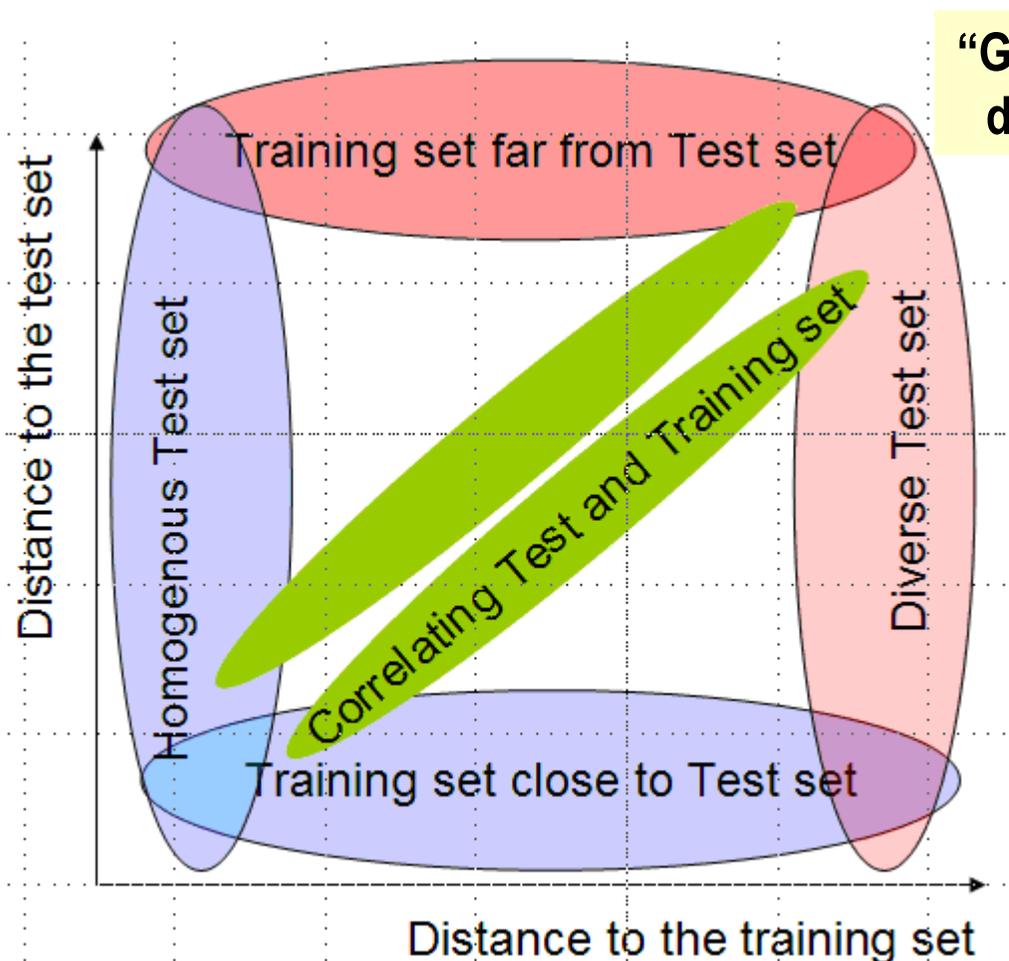
Show

Euclidean distance, 10NN (Prediction)

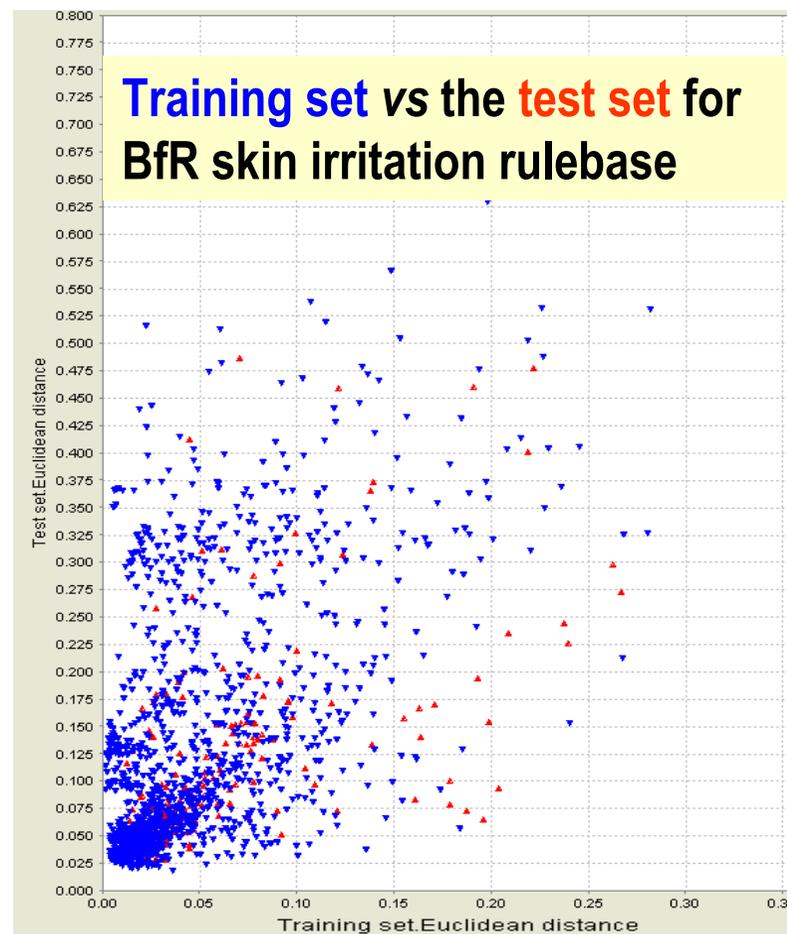


Aniline -  
test chemical

Pairwise similarity  
between aniline  
and training set  
compounds



“Gallegos” diagram

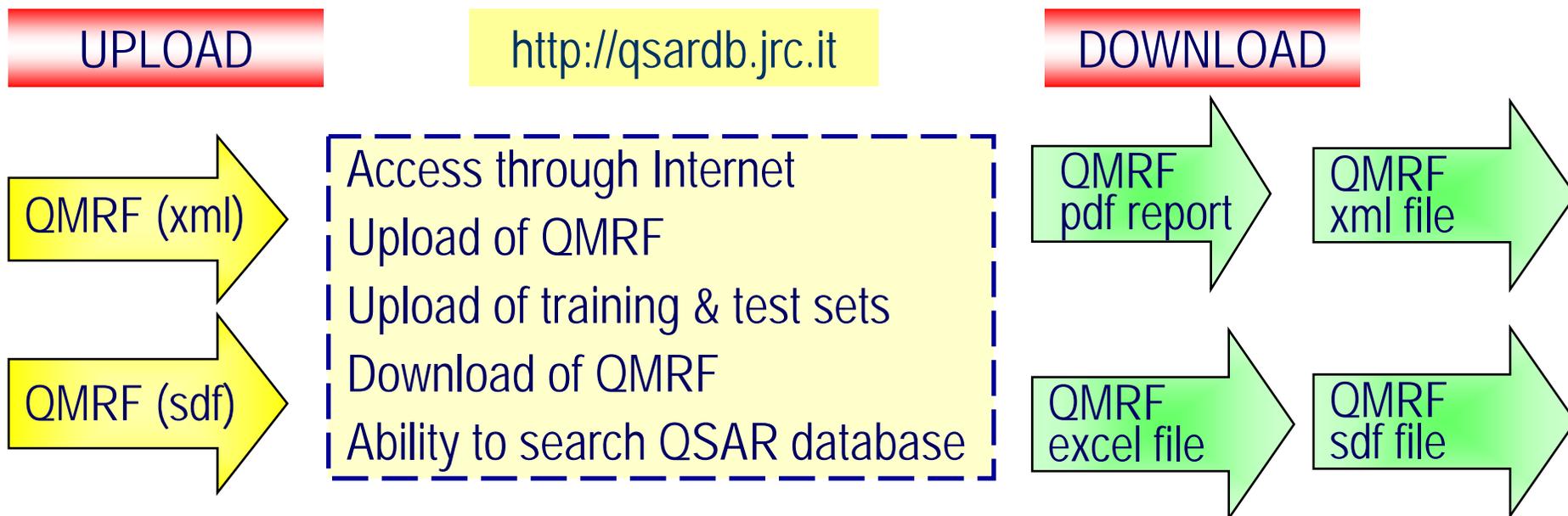


Gallegos Saliner A, Poater A, Jeliaskova N, Patlewicz G & Worth AP (2008). Toxmatch - A Chemical Classification and Activity Prediction Tool based on Similarity Measures. Regulatory Toxicology and Pharmacology 52, 77-84.

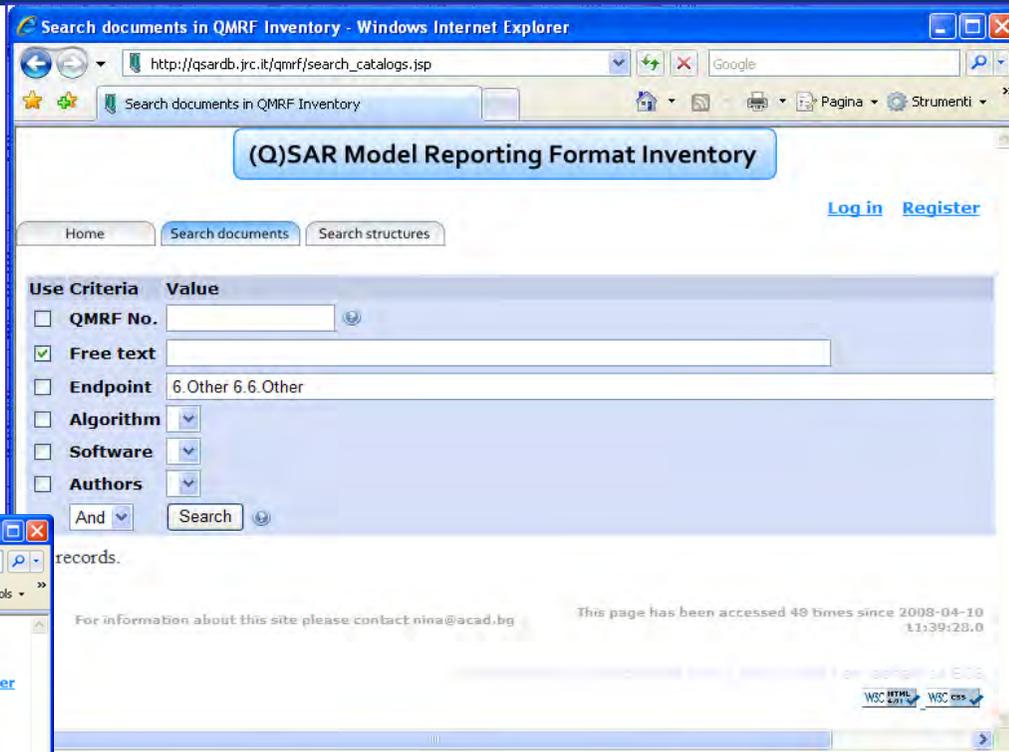
Need to identify and use **relevant, reliable and well documented** (Q)SARs

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



Search documents in QMRF Inventory - Windows Internet Explorer

http://qsar.db.jrc.it/qmrf/search\_catalogs.jsp

Search documents in QMRF Inventory

(Q)SAR Model Reporting Format Inventory

Log in Register

Home Search documents Search structures

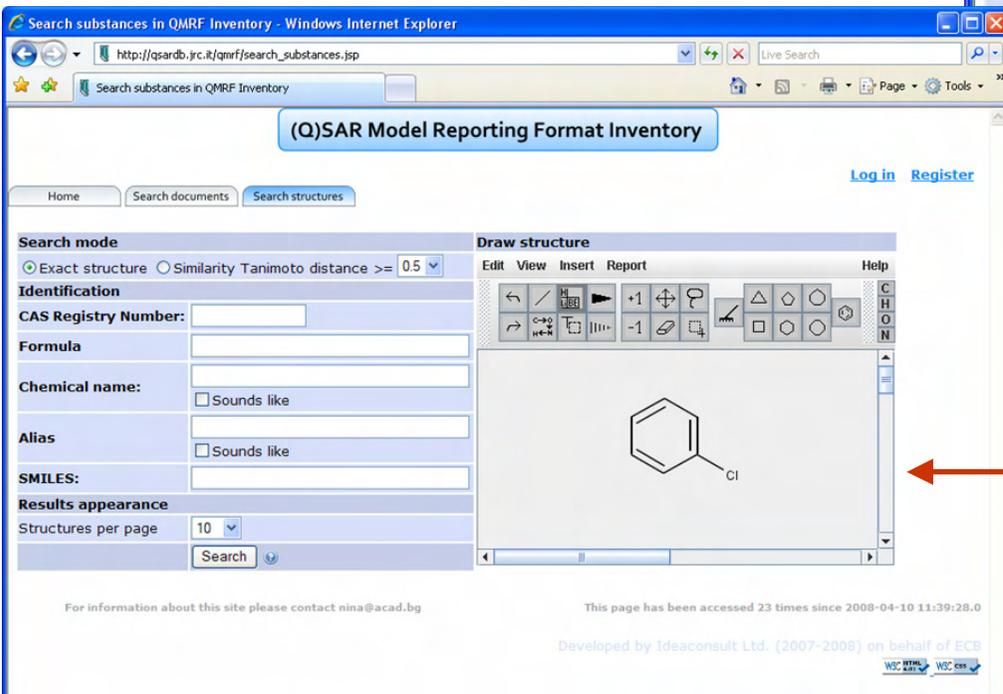
Use Criteria	Value
<input type="checkbox"/> QMRF No.	
<input checked="" type="checkbox"/> Free text	
<input type="checkbox"/> Endpoint	6.Other 6.6.Other
<input type="checkbox"/> Algorithm	
<input type="checkbox"/> Software	
<input type="checkbox"/> Authors	

And Search

records.

For information about this site please contact nina@acad.bg

This page has been accessed 49 times since 2008-04-10 11:39:28.0



Search substances in QMRF Inventory - Windows Internet Explorer

http://qsar.db.jrc.it/qmrf/search\_substances.jsp

Search substances in QMRF Inventory

(Q)SAR Model Reporting Format Inventory

Log in Register

Home Search documents Search structures

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

Search

Draw structure

Edit View Insert Report Help

Chemical structure: c1ccccc1Cl

For information about this site please contact nina@acad.bg

This page has been accessed 23 times since 2008-04-10 11:39:28.0

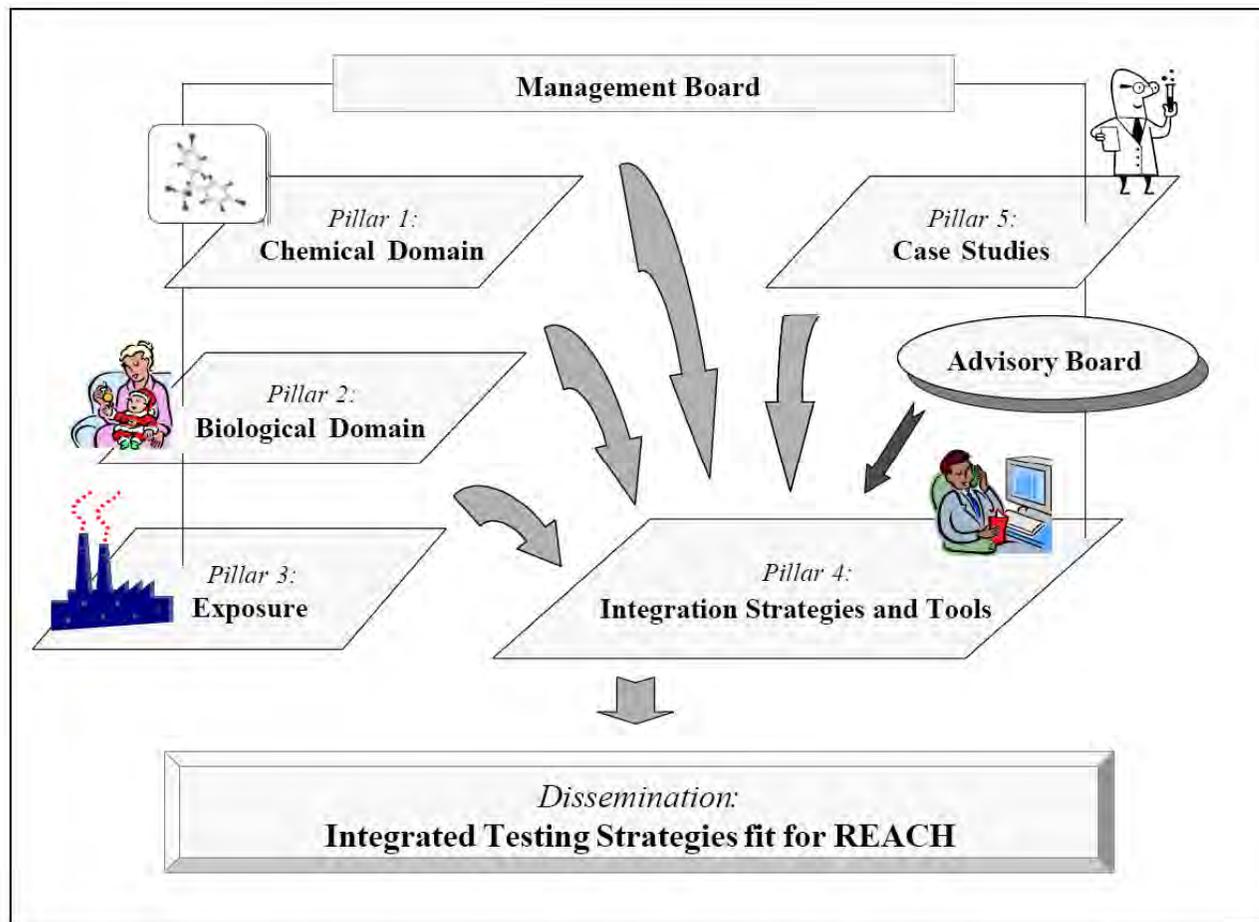
Developed by Ideaconult Ltd. (2007-2008) on behalf of ECB

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

<http://qsar.db.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)



<http://www.osiris-reach.eu/>

- To optimise the use of non-testing data, a conceptual framework is provided in the REACH guidance documentation
- An increasing number of models are being implemented in a range of software tools
- There is a need to incorporate mechanistic knowledge in the models (e.g. based on chemical reactivity and “omic” data)
- 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