Assessment of the validity of QSAR results under dossier evaluation

Webinar

3 June 2021

Andrea Gissi
Doris Hirmann
Mounir Bouhifd

Computational assessment unit
European Chemicals Agency

The views expressed in this presentation are solely those of the authors and the content of the presentation does not represent the views or position of the European Chemicals Agency

The presenters (from ECHA's computational assessment unit)

Andrea Gissi

- QSAR expert for human health and environmental endpoints
- OECD QSAR Toolbox
- QSAR related OECD activities

Mounir Bouhifd

- QSAR expert for human health and environmental endpoints
- New Approach Methodologies (NAM)
- Omics and High throughput data

Doris Hirmann

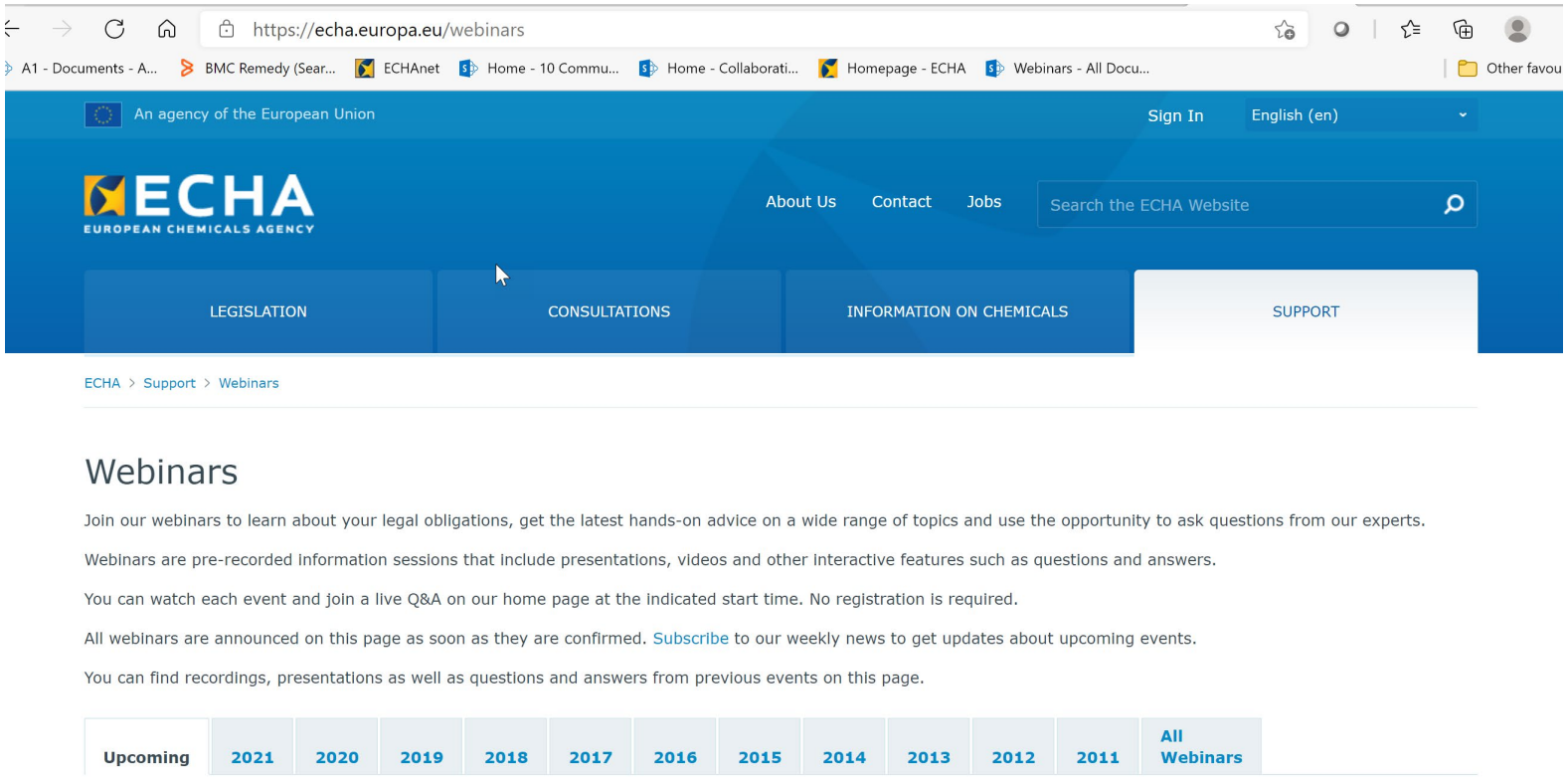
- QSAR expert for environmental endpoints
- New Approach Methodologies (NAM) environment
- Assessment of Persistent, Bioaccumulative and Toxic (PBT) properties

Questions

- Join Q&A at: slido.com
  Event code: **#qsars21**
- Send questions from **11:00 to 13:00 Helsinki time**
- Only questions within scope
- Question not answered?
  Contact us: echa.europa.eu/contact

Material available

- Video recording, presentations and Q&A: echa.europa.eu/webinars



What you can expect from today

Introduction – REACH and ECHA

REACH requirements for the use of QSAR

Most common issues with the compliance of QSAR results

QSARs as adaptations in REACH – coverage of information requirements

Conclusions

OECD QSAR Toolbox (TB)

- ECHA develops the TB, however this software will be discussed in a separate webinar planned for the end of the year

- In REACH registrations, TB is mainly used for read-across studies

  -> evaluated according to the read across assessment framework (RAAF)

- In cases of QSAR predictions from TB, the QSAR assessment discussed today applies

Introduction

REACH and ECHA

REACH (EC 1907/2006): EU Regulation on **R**egistration, **E**valuation, **A**uthorisation (restriction) of **C**hemicals

- Under REACH, industry is responsible for the safe use of the substances they produce or import.

- **Hazard information is the starting point for chemical safety assessment**. A minimum set of information (SIRs: standard information requirements) must be submitted by industry in form of REACH registration dossiers in IUCLID format.

- SIRs depend on the production volume of the chemical: higher volume = higher requirements

ECHA: **E**uropean **Ch**emicals **A**gency

- Established by REACH (2007), now the Agency is responsible for the implementation of a number of chemical-related regulations.

- One core process under REACH is **dossier evaluation**, where ECHA checks that the information submitted by industry for their substances is compliant with information requirements.

- Another core process is **substance evaluation**, where ECHA and EU Member States use the information in the dossiers to identify substances of very high concern for which risk management measures are needed (e.g. authorisation, restriction).

REACH – main processes and actors

Data sharing
Registration
Self-classification

Industry gathers information and is responsible for risk management

Evaluation

- Dossier evaluation
- Substance evaluation

ECHA and Member States Competent Authorities (MSCAs) control and request for further info

Authorisation
Restriction
Harmonised C&L

Commission, with support of ECHA and MSCAs, applies community wide risk management measures

Industry may include QSARs in REACH dossiers as adaptations to standard information requirements.

ECHA assesses these QSAR results under Dossier evaluation.

Promotion of alternatives to testing on animals

- Testing on vertebrate animals for the purpose of REACH as last resort.
- Promotion of alternative methods by ECHA:
  - Contribution to OECD activities (e.g. OECD QSAR Toolbox)
  - Participation to the international project on accelerating the pace of chemical risk assessment (APCRA)
  - Making non-confidential data from REACH registrations more and more available
  - Preparation of Guidance and other supporting documents, such as the report on the use of alternatives to testing on animals for the REACH regulation art 117.3 report)

QSARs as adaptations

- Standard test can be "adapted", if properly justified.

Two types of **adaptations** in REACH:

- Specific Annexes VII-X column 2 adaptations
- Annex XI general adaptations
  - Testing not scientifically necessary
  - Testing scientifically not possible
  - Exposure-based adaptation (i.e. no exposure is demonstrated)

1. Use of existing data
2. Weight of evidence (WoE)
3. **Qualitative or quantitative structure-activity relationship ((Q)SAR)**
4. *In vitro* methods
5. Grouping of substances and read-across approach

- REACH Annex XI also indicates the requirements for their validity

REACH requirements for the use of QSAR

QSARs under dossier evaluation

For endpoints that are “adapted” with QSAR studies, ECHA checks that the study meets the requirements indicated in **Annex XI 1.3**

1. Scientific valid **model**
2. **Substance** within applicability domain
3. **Results** adequate for the purpose
4. Adequate and reliable **documentation of the applied method**

Valid model: OECD principles for (Q)SARs

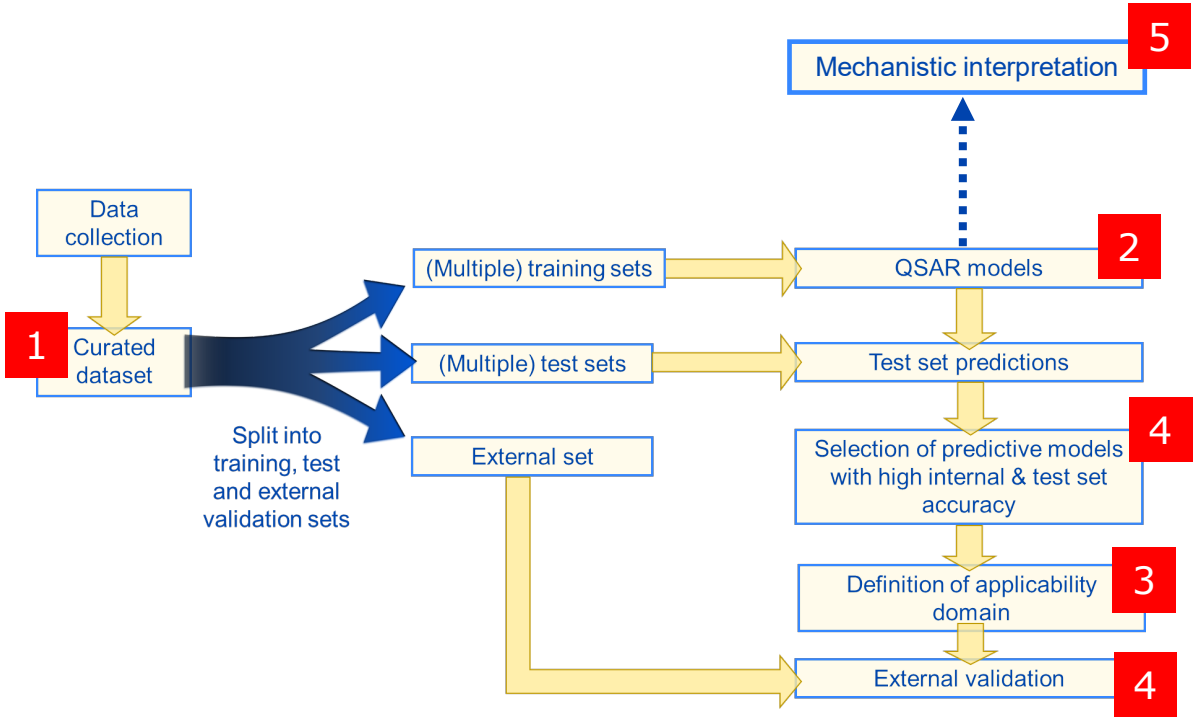
OECD principles for the validation, for regulatory purposes, of QSAR models:

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

NB! The OECD principles only cover the first REACH condition for acceptable QSAR results, i.e., **scientific validity of the model**.

QSAR modelling steps and principles for validity

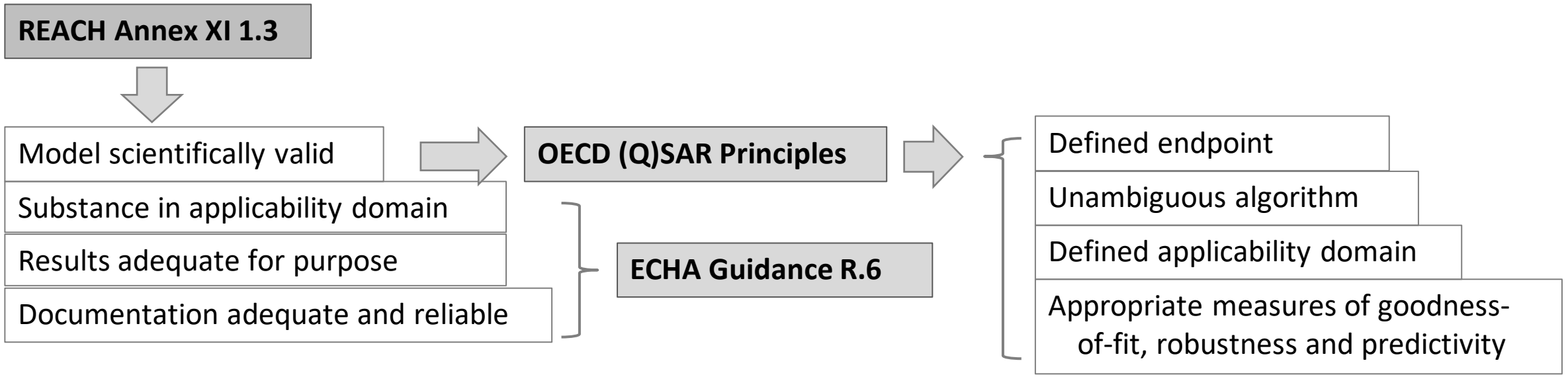
1. Defined endpoint
2. Unambiguous algorithm
3. Defined domain of applicability
4. Appropriate measures of goodness of fit, robustness and predictivity
5. Mechanistic interpretation, if possible



Valid model vs acceptable prediction

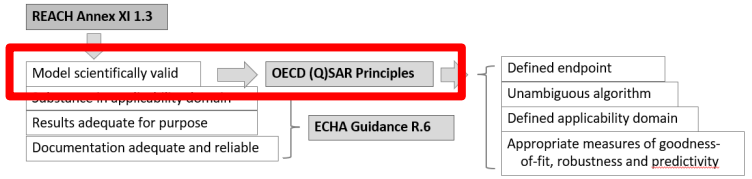
- The use of a scientifically valid model is required but not sufficient to obtain a valid prediction
- Criteria for the assessment of the validity of QSAR predictions are indicated in REACH Guidance R6 (no agreed principles at OECD level, but an OECD project has just started)
- In principle a **model** needs to be assessed only once (for a specific purpose), while each **prediction** needs to be assessed individually

(Q)SAR assessment overview



OECD ENV/JM/MONO(2007)2

Guidance on information requirements and chemical safety assessment Chapter R.6: QSARs and grouping of chemicals



Is the model scientifically valid? (OECD principles)

1. **Defined endpoint** -> Check the data used to build the model (i.e. training set)

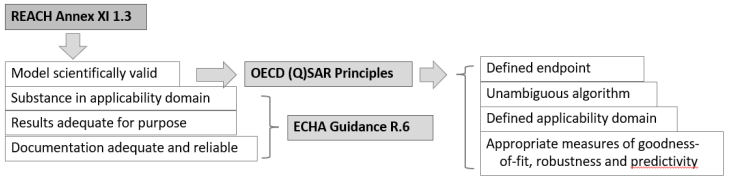
2. **Unambiguous algorithm** -> Check that the prediction is reproducible (same input and settings = same output)

3. **Defined domain of applicability** -> Check that the applicability domain is defined

4. **Appropriate measures of goodness of fit, robustness and predictivity** -> check the availability of measures of performances

5. **Mechanistic interpretation, if possible** -> Not formally checked since it is an optional requirement

https://www.oecd.org/chemicalsafety/risk-assessment/37849783.pdf
ENV/JM/MONO(2004)24: Report from the Expert Group on Validation of (Q)SARs
ENV/JM/MONO(2007)2: OECD Guidance on the Validation of (Q)SAR Models



Is the prediction valid? (R6 Guidance)

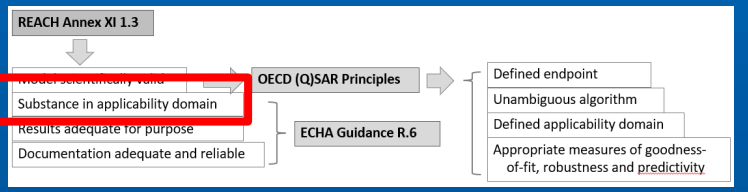
- **Scientifically valid model** -> Check OECD QSAR principles listed in the previous slide

- **Substance within domain** -> Check domain as defined by model developers + parametric, structural, mechanistic and metabolic domain, as relevant

- **Results adequate for purpose** -> Check the input structure and the reliability of the prediction

- **Documentation ->** Check QPRF and QMRF, or equivalent content

Applicability domain



Definition of AD

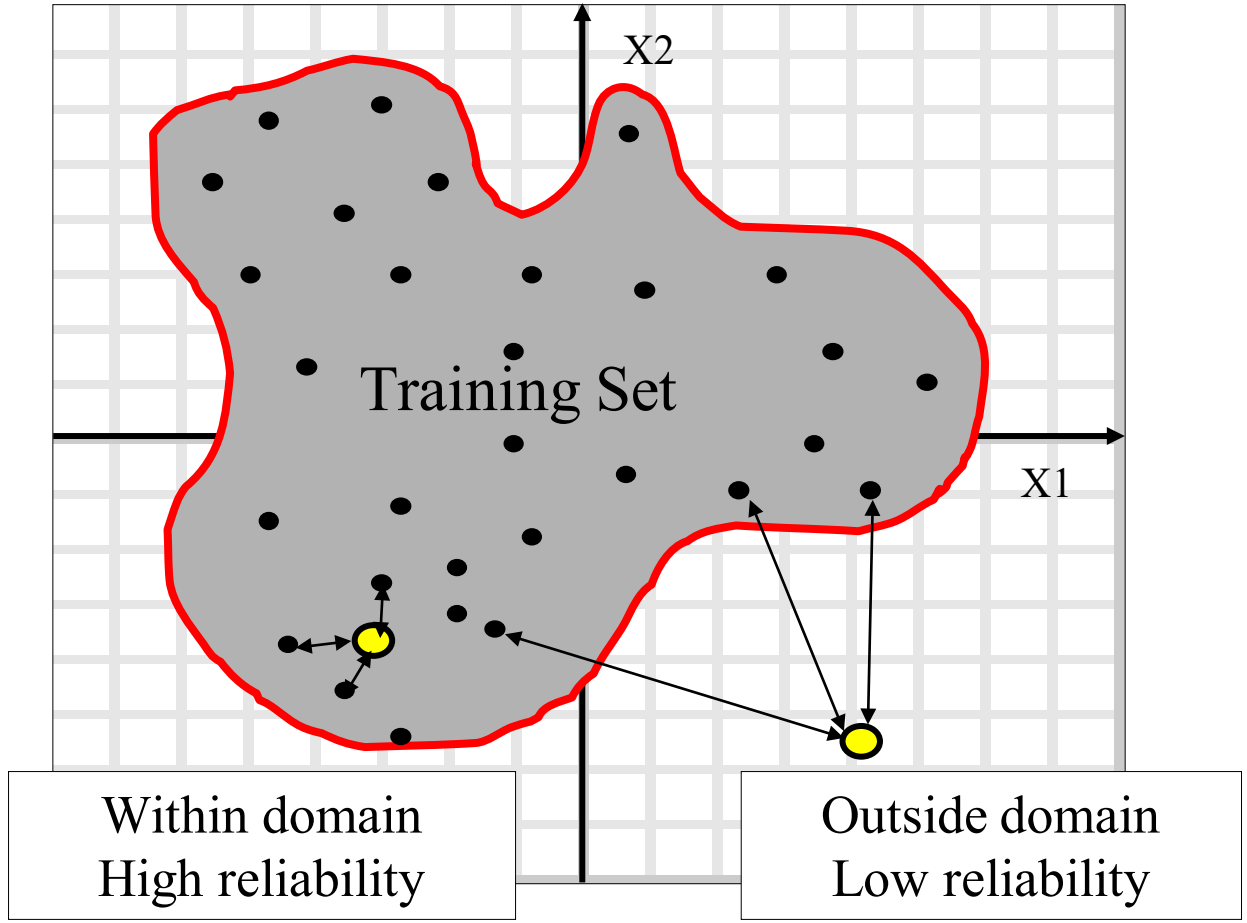
*The applicability domain (AD) of a QSAR model is the physico-chemical, structural or biological space, knowledge or information on which the **training set** of the model has been developed, and for which it is applicable to make predictions for new compounds.*

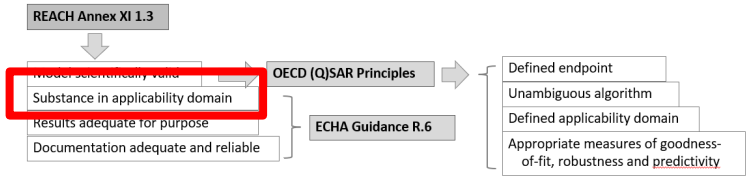
- Lack of agreed definition and methodology at regulatory but also scientific level
- Each model developer applies/suggests its own method

Determination of the applicability domain (AD)

- The structures in the training set of the model and their properties can be plotted to define an interpolation space called "applicability domain" of the model

- Different properties can be considered:
  1. Descriptors
  2. Structural fragments
  3. Mechanistic of action
  4. Metabolism





Substance within domain?

The model developers definition of applicability domain is the starting point for ECHA's assessment.

In addition, ECHA considers the following aspects, as relevant:

- Descriptor domain
- Structural domain
- Mechanistic domain
- Metabolic domain

Applicability Domain “layers”

Descriptor domain
Based on the descriptor values used by the model for the structures in the Training Set (TS)

Structural domain
Based on the fragments present in the structures in the TS

Mechanistic domain
Based on the different mechanisms of effect/toxicity covered by the structures in the TS

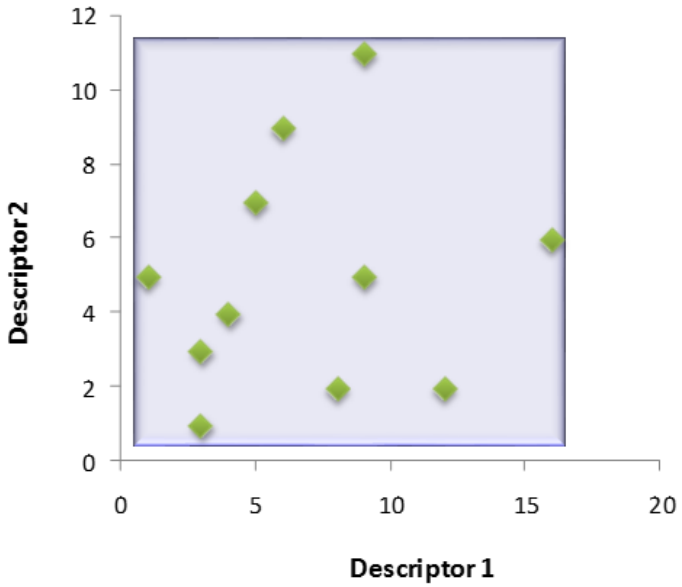
Metabolic domain
Applicable only if metabolism is relevant for the endpoint. Based on metabolic considerations for the structures in the TS

Descriptor domain – Range method

Uses the range of the chemical descriptors for the chemicals in the training set

A new chemical with descriptors out of the range is considered out of AD

EPISuite suggests to combine this approach (for MW and logP) with the fragment based approach



Training Set (527 Compounds):

Molecular Weight:
Minimum MW: 68.08 (Furan)
Maximum MW: 991.80 Ionic: (2,7-Naphthalenedisulfonic acid, 4-amino-5-hy
bis[[4-[[2-(sulfooxy)ethyl]sulfonyl]phenyl]azo]-, tetrasodium salt)
Maximum MW: 959.17 Non-Ionic: (Benzene, 1,1 -oxybis[2,3,4,5,6-pentabro
Average MW: 244.00

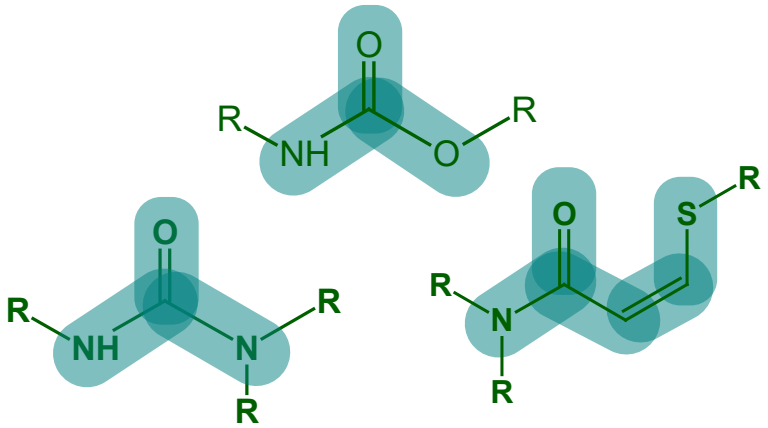
Log Kow:
Minimum LogKow: -6.50 Ionic: (2,7-Naphthalenedisulfonic acid, 4-amino-5-h
Minimum LogKow: -1.37 Non-Ionic: (1,3,5-Triazine-2,4,6-triamine)
Maximum LogKow: 11.26 (Benzenamine, ar-octyl-N-(octylphenyl)-)

Structural domain

The list of fragments (and eventually their maximum number of occurrences per molecule) is determined for the training set

Only new molecules with the fragments represented in the training set are considered within domain

This approach is suggested in EPISuite together with the descriptor space



Correction Factor	BCFBAF New Coef	BCFWIN Old Coef	No. Compounds containing Factor in Training set	Maximum Number of each Factor in any Individual compound
Ketone (aromatic connection)	-0.5851	-0.84	12	1
Phosphate ester	-0.8250	-0.78	14	1
Multi-halogenated biphenyl/PAH	0.5860	0.62	13	1
Aromatic ring-CH-OH	-0.2556	-0.65	5	1
Aromatic sym-triazine ring	-0.5169	-0.32	2	1
Tert-Butyl ortho-phenol type	-0.2220	-0.45	8	1
Phenanthrene ring	0.6609	0.48	2	1
Cyclopropyl-C(=O)-O- ester	-1.2591	-1.65	3	1
Alkyl chains (8+ CH2 groups)log Kow >4&<7.0	-1.3743	-1.00	2	1
Alkyl chains (8+ CH2 groups)log Kow 7-10	-0.5965	-1.50	2	1
Disulfide (-S-S-) ** NEW Descriptor	-1.3404	---	2	1
Multi-halogenated phenol **NOT Used BCFBAF	---	-0.40		

Mechanistic domain

Does the target chemical act according to the same mode or mechanism of action as other chemicals for which the model is applicable?

The model can reliably predict substances that act according to the same mode of actions covered by the substances in the training set.

Example:

Many BCF models calculate the BCF potential based only on the logKow of the substance assuming accumulation in the lipid tissues of the fish. Other mechanisms are not considered.

Metabolic domain

Does the chemical of interest undergo transformation or metabolism, and how does this affect the prediction for the parent compound?

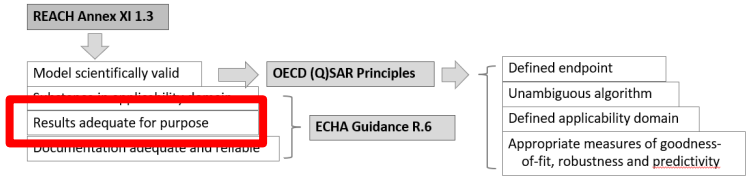
Metabolism can increase or reduce the toxicity of the substance.

Example:

Many skin sensitisers cause their effect after biotic or abiotic transformation (pre- and pro-haptens).

Some QSAR models simulate skin metabolism and autooxidation and then predict the skin sensitisation potential of the metabolites too. The applicability domain of the simulators must also be taken into account.

Back to the REACH requirements for QSARs



Adequate results?

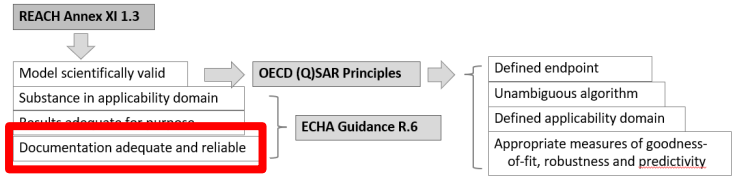
(Purpose relevant for ECHA -> mainly adaptation of REACH information requirements)

Input structure

Choosing the correct input structure(s) is not trivial in case of multi-constituents or substances with Unknown or Variable composition, Complex reaction product or Biological origin (UVCB).

Reliability of the prediction

- reliability of input parameters
- presence of analogues in the training/test sets and the accuracy of their predictions
- consistency of the prediction with other information available for the substance



Adequate documentation?

QSAR Model Reporting Format (QMRF) must include information on:

- the predicted endpoint, including information on experimental protocol and data quality for the data used to develop the model,
- an unambiguous definition of the algorithm, the descriptor(s) of the model and its applicability domain,
- an estimate of the goodness-of-fit and of the predictivity of the model, including information on training set and validation statistics.

QSAR Prediction Reporting Format (QPRF) must include information on:

- the model prediction(s), including the endpoint,
- a precise identification of the substance modelled,
- the relationship between the modelled substance and the defined applicability domain,
- the identities of close analogues, including considerations on how predicted and experimental data for analogues support the prediction.

Most common issues with the compliance of QSAR results

Most common shortcomings

The most common shortcomings found by ECHA in QSAR studies in REACH dossiers relate to:

1. **Scientific validity of the model**:
    - 1.1 Definition of the endpoint
    - 1.2 Measures of performance
2. **Substance within applicability domain**:
    - 2.1 As defined by the model developers
    - 2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
    - 3.1 Input structure(s)
    - 3.2 Reliability of the prediction
4. **Documentation:**
    - 4.1 Model
    - 4.2 Prediction

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation:**
   4.1 Model
   4.2 Prediction

1.1 Definition of the endpoint

Models predict the endpoint and effects according to the experimental data they are trained on.

Homogeneity of the training set data: If the model is built on inhomogeneous data and effects, it is considered not well-defined from a regulatory point of view for the purpose of meeting the information requirements.

Effects covered by training set data: all effects measured by the standard test should be taken into account by the model. If the effects are different or do not include all the effects measured by the defined test protocol; then the predictions from the model are not acceptable to meet the information requirements.

Examples of issues:

- Model for long term toxicity to fish mixing tests of different durations, including 14 days studies
- Ames model lacking 5th strain
- Model for repeated dose toxicity predicting only the NOAEL

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation:**
   4.1 Model
   4.2 Prediction

1.1 Endpoint not well-defined

Real case:

3. Prediction

3.1 Endpoint (OECD Principle 1)

a. **Endpoint**: *Developmental toxicity dysmorphogenesis, functional toxicity, mortality, growth toxicity, general effects relating to fertility and reproductive capability.*

b. **Dependent variable:** *CASE units (active=30-80, inactive =10-19, marginal 20-29)*

Test animals

Species

other: rat, mouse, rabbit, human

Issues related to the definition of the endpoint:

1. The predicted value is expressed in CASE units, and then converted to qualitative values (ACTIVE/INACTIVE). NOAEL and specific effects are not mentioned.
2. The information from “species” and “endpoint” indicates that the training set considers data from different experimental protocols. The underlying data are not provided.

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation:**
   4.1 Model
   4.2 Prediction

1.2 Measures of performance

- **Performance not measured**

Measures of performances of the model are needed to estimate the reliability of a prediction. Without such information, predictions cannot be accepted.

Example of issue:

- Profilers from Toolbox should not be used directly as predictions for the purpose of fulfilling standard information requirements. They are grouping tools for the identification of analogues and measures of their performance to predict apical endpoints are usually not available.

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation:**
   4.1 Model
   4.2 Prediction

1.2 Measures of performance

- **Too small training set**

A model built on a too small number of substances is not considered **robust**.

Minimum requirement: Topliss ratio ≥ 5 (the number of substances in training set / number of descriptors)

Example:

ECOSAR class specific models (one descriptor, logKow) based on training set with less than 5 substances are considered not robust.

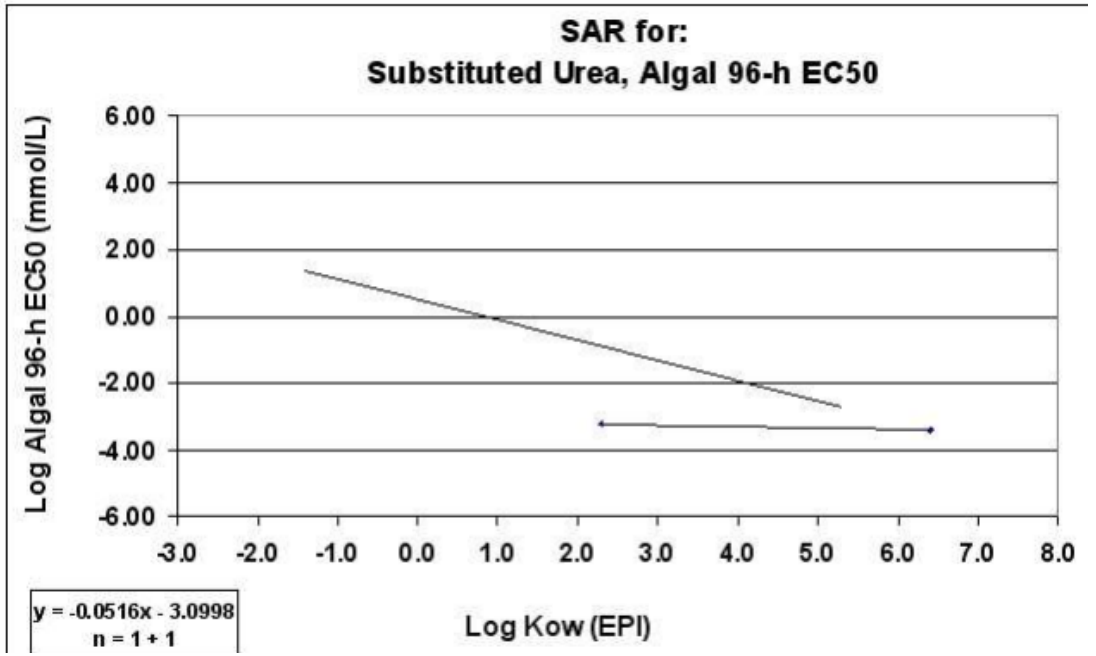
1. **Scientific validity of the model**:
   - 1.1 Definition of the endpoint
   - 1.2 Measures of performance
2. **Substance within applicability domain**:
   - 2.1 As defined by the model developers
   - 2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   - 3.1 Input structure(s)
   - 3.2 Reliability of the prediction
4. **Documentation:**
   - 4.1 Model
   - 4.2 Prediction

1.2 Model not robust

Issue:

The Algae EC50 model in ECOSAR for the class "Substituted Urea" from 2008 used in the dossier is built on only one data point (CBI) and the logKow limit.

This is not a scientifically valid model.



DATA:

CAS No.	Chemical Name	M.W.	log Kow (CLogP)	log Kow (EPI)	log Kow (M)	Algal 96-h EC50 (mg/L)	Log Algal 96-h EC50 (mmol/L)	Reference (Meas. Kow)	Reference (Algal 96-h EC
CBI	CBI	215	2.3	2.3	2.3	0.13	-3.22		8(e)-
	Kow Limit		6.4	6.4			-3.43	NO Cutoff	NO SAR
SAR Data Not Included in Regression Equation:									
Data Not Included in SAR:									
						* Indicates no effects at saturation			

1. **Scientific validity of the model**:
   - 1.1 Definition of the endpoint
   - 1.2 Measures of performance
2. **Substance within applicability domain**:
   - 2.1 As defined by the model developers
   - 2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   - 3.1 Input structure(s)
   - 3.2 Reliability of the prediction
4. **Documentation:**
   - 4.1 Model
   - 4.2 Prediction

2.1 AD as defined by model developers

Automatic assessment: Many "modern" software automatically assess the applicability of the model to the target substance.

Manual assessment: the definition of the applicability domain is often included in the documentation of the model.

When the substance falls outside the applicability domain as defined by the model developers, the prediction is not accepted unless a good justification is provided.

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation**:
   4.1 Model
   4.2 Prediction

Example of domain and reliability assessment automatically provided by a software

- There are software that make automatic assessments of applicability domain and reliability of their predictions

- Results need to be critically investigated; no automatic acceptance based on software results

Similar Compounds, with Predicted and Experimental Values

Compound #1

CAS: 106-97-8
Dataset id: 314 (Training set)
SMILES: CCCC
Similarity: 1

Experimental value: NON-Mutagen
Predicted value: NON-Mutagen

Global AD Index
AD index = 1
Explanation: the predicted compound is into the Applicability Domain of the model.

Similar molecules with known experimental value
Similarity index = 1
Explanation: strongly similar compounds with known experimental value in the training set have been found.

Accuracy of prediction for similar molecules
Accuracy index = 1
Explanation: accuracy of prediction for similar molecules found in the training set is good.

Concordance for similar molecules
Concordance index = 1
Explanation: similar molecules found in the training set have experimental values that agree with the predicted value.

Atom Centered Fragments similarity check
ACF index = 1
Explanation: all atom centered fragment of the compound have been found in the compounds of the training set.

2.2 As defined by REACH Guidance R6

- **Descriptor, structural, mechanistic and metabolic domains**

When model developers define the applicability domain of their model, they may not always consider all aspects needed for a regulatory acceptable prediction. These aspects are described in ECHA's Guidance R6, and include descriptor, structural, mechanistic and metabolic domains.

For this reason, a model may indicate a structure to be within its domain but ECHA may overrule the automatic assessment of the model.

2.2 Substance out of AD

1. **Scientific validity of the model**:
   - 1.1 Definition of the endpoint
   - 1.2 Measures of performance
2. **Substance within applicability domain**:
   - 2.1 As defined by the model developers
   - 2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   - 3.1 Input structure(s)
   - 3.2 Reliability of the prediction
4. **Documentation:**
   - 4.1 Model
   - 4.2 Prediction

Issue:

logKow of the constituents of the substance > 8

upper limit of the logKow range covered by the training set of the model is ~7

The model developers define the applicability domain as the range of molecular weight covered by the model.

In addition, we checked the descriptor domain and concluded that the substance is out of descriptor domain for the logKow based KOC model

CAS	Name	LogKoc	MCI	LogKow	KocEst (MCI)	Residual	KocEst (Kow)	Residual	Source Koc
000050-29-3	P,P'-DDT	5.31	8.8760	6.91	5.23	-0.08	6.00	0.69	Schuurmann,G et al (2006)
000050-32-8	Benzo[a]pyrene	5.95	9.9158	6.13	5.77	-0.18	5.32	-0.63	SRC (1991); Meylan et al (1992)
000053-70-3	1,2,5,6-DIBENZANTHRACENE	6.22	10.8990	6.75	6.28	0.06	5.68	-0.54	Schuurmann,G et al (2006)
000056-23-5	Tetrachloromethane	1.85	2.0000	2.83	1.64	-0.21	2.46	0.61	SRC (1991); Meylan et al (1992)
000056-49-5	3-METHYLCHOLANTHRENE	6.10	10.3265	6.42	5.98	-0.12	5.57	-0.53	Schuurmann,G et al (2006)
000056-55-3	BENZ(A)ANTHRACENE	5.30	8.9158	5.76	5.25	-0.05	5.00	-0.30	Schuurmann,G et al (2006)
000057-97-6	7,12-DIMETHYLBENZ(A)ANTHRACENE	5.37	9.7709	5.80	5.69	0.32	5.03	-0.34	Schuurmann,G et al (2006)
000058-89-9	GAMMA-HEXACHLOROCYCLOHEXANE	3.04	5.4641	3.72	3.45	0.41	3.59	0.55	Schuurmann,G et al (2006)
000067-66-3	Trichloromethane	1.60	1.7321	1.97	1.50	-0.10	1.71	0.11	SRC (1991); Meylan et al (1992)
000071-43-2	Benzene	1.75	3.0000	2.13	2.16	0.41	1.85	0.10	SRC (1991); Meylan et al (1992)
000071-55-6	1,1,1-Trichloroethane	1.95	2.0000	2.49	1.64	-0.31	2.16	0.21	SRC (1991); Meylan et al (1992)
000072-54-8	P,P'-DDD	4.70	8.5754	6.02	5.07	0.37	5.22	0.52	SRC (1991); Meylan et al (1992)
000072-55-9	P,P'-DDE	4.82	8.5754	6.51	5.07	0.25	5.65	0.83	Schuurmann,G et al (2006)
000074-83-9	BROMOMETHANE	1.34	1.0000	1.19	1.12	-0.22	1.03	-0.31	Schuurmann,G et al (2006)
000075-09-2	Dichloromethane	1.44	1.4142	1.25	1.34	-0.10	1.08	-0.36	SRC (1991); Meylan et al (1992)
000078-87-5	1,2-Dichloropropane	1.67	2.2701	1.98	1.78	0.11	1.72	0.05	SRC (1991); Meylan et al (1992)
000079-00-5	1,1,2-Trichloroethane	1.89	2.2701	1.89	1.78	-0.11	1.64	-0.25	SRC (1991); Meylan et al (1992)
000079-01-6	Trichloroethylene	2.00	2.2701	2.42	1.78	-0.22	2.10	0.10	SRC (1991); Meylan et al (1992)
000079-	1,1,2,2,-Tetrachloroethane	1.90	2.6427	2.39	1.98	0.08	2.07	0.17	SRC (1991); Meylan

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation**:
   4.1 Model
   4.2 Prediction

3.1 Input structure(s)

The whole composition of the substance needs to be taken into account by:

a. Predicting individually each constituent, or
b. Selecting one or more representative structures (with justification).

Examples:

a. A substance has 3 constituents, quite different from each other. A single representative structure cannot be selected, and the registrant should predict all 3 constituents individually.

b. UVCB with variable alkyl branching. The registrant should use one or more structures supposed to cover the worst case and justify the selection.

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation:**
   4.1 Model
   4.2 Prediction

3.1 Input structures not consistent

Issue:

The Substance is a UVCB.

Two input structures are used and described as the "main components"

A link between the structures and the composition of the substance **is not given**.

Constituent	Concentration range
Const 1	30-80 %
Const 2	0-20 %
Const 3	0-10 %
Const 4	0-10 %
Const 5	0-10 %
Const 6	0-10 %
Const 7	0-5 %
Const 8	0-5 %

Tip: Do not forget to take into account impurities, which may drive the concern for substances as e.g.:

- persistent, bioaccumulative and toxic (PBT) or very persistent and very bioaccumulative (vPvB); or
- carcinogenic, mutagenic or toxic to reproduction (CMR)

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation:**
   4.1 Model
   4.2 Prediction

3.2 Reliability of the prediction

- Accuracy of predictions for similar substances
- Reliability of input parameters
- Consistency with information from other endpoints

3.2 Accuracy of predictions for similar substances

A model is expected to predict similar substances with comparable accuracy. For confirming the reliability of the predictions:

1. Similar substances with experimental data for the target endpoint are needed, and
2. The model must predict well these substances

If one of the two conditions is not met, then the reliability of the prediction may not be confirmed.

The definition of similarity depends on the endpoint and has some flexibility. Complete lack of similar substances with data (in the training set) may also rise concerns of the applicability of the model to the target substance.

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation**:
   4.1 Model
   4.2 [illegible]

3.2 The prediction is not adequate due to low reliability

- **Reliability of the input parameters**

The input parameters must be reliable. Most of the models produce a prediction based on the input structure and its properties. The model can generate reliable predictions if the input parameters are reliable. The properties can be predicted by the model itself after input of the structure or manually inserted by the user.

Example: if a reliable experimental logKow is available and differs significantly from the calculated one, the calculated logKow is considered unreliable and should not be used for BCF predictions

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation:**
   4.1 Model
   4.2 Prediction

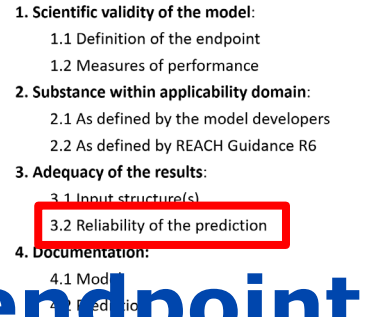
3.2 Reliability of the input parameters

Information in the dossier:

- Prediction: BCF ~ 1000 L/kg (not B) using manual input value of logKow ~9, substance is within applicability domain and prediction well documented.
- An experimental value of LogKow 8 ± 1 is provided in the dossier.
- The registrant used logKow 9 as input to the model.

logKow – used as input	BCF prediction
8 (experimental value)	2500
9 (experimental value - upper limit)	**1000**
7 (experimental value - lower limit)	4000

- This leads to a BCF predicted to be relatively low, while the prediction with the measured logKow indicates much higher BCF values.
- Taking into account the uncertainty of the measured logKow, the predicted BCF of the main constituent is either very low or above the B threshold of 2000.
- Therefore, the prediction is not adequate for the purpose of risk assessment and PBT/vPvB assessment.



3.2 Consistency of the information for other endpoints

The information from the prediction must be consistent with the reliable information available for other related endpoints, which is usually not taken into account by QSAR models. If this is not the case, then the prediction is considered unreliable

Example: reliable experimental short-term fish LC50 is available. A QSAR prediction for fish long-term toxicity reports a NOEC higher than the short-term LC50.

3.2 Consistency with information for other endpoints

Predicted value fish long-term toxicity: ChV = 80 mg/L

The following experimental data are available:

- **LC50 short term fish = 0.01 mg/L**
- EC50/LC50 for freshwater invertebrates 0.2mg/L
- EC10, LC10 or NOEC for freshwater invertebrates 0.05mg/L
- EC50 for freshwater algae 0.3 mg/L

Among others, the prediction would be rejected due to inconsistency with information for other endpoints.

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation**:
   4.1 Model
   4.2 Prediction

4. Model documentation (QMRF)

The following information is expected:

- **Predicted endpoint**, including information on the experimental protocol and data quality for the data used to develop the model
- Unambiguous definition of the **algorithm** of the model, including descriptors and applicability domain
- Estimate of **performance**: the goodness-of-fit and predictivity of the model, including information on training set and validation statistics

Ideally it should be provided in form of a QMRF, but could also be included in the endpoint study record in the IUCLID dossier or simply be publicly available on the internet.

If this information is missing, the scientific validity of the model cannot be established.

1. **Scientific validity of the model**:
   1.1 Definition of the endpoint
   1.2 Measures of performance
2. **Substance within applicability domain**:
   2.1 As defined by the model developers
   2.2 As defined by REACH Guidance R6
3. **Adequacy of the results**:
   3.1 Input structure(s)
   3.2 Reliability of the prediction
4. **Documentation:**
   4.1 Model
   4.2 Prediction

4. Prediction documentation (QPRF)

The following information is expected:

- The model **prediction**(s), including the endpoint
- Precise identification of the **substance modelled**
- Relationship between the substance and **applicability domain**
- **Identity of close analogues**, including considerations on how predicted and experimental data for analogues support the prediction

Ideally it should be provided in form of a QPRF, but could also be included in the endpoint study record.

If this information is missing, the validity of the prediction cannot be established.

QSARs as adaptations in REACH - coverage of information requirements

Observations on the current status

QSARs as adaptations in REACH - coverage of information requirements

Not all REACH requirements can be successfully adapted by stand-alone QSARs due to the limitations of the currently available QSAR models.

One of the limiting factor is the complexity of high tier endpoints.

In general, QSARs work best when:

- a high number of good quality experimental data is available
- the mechanism of action/toxicity is well understood
- the results of adapted experimental studies consider a limited number of effects that can be reasonably predicted by QSAR models

QSARs for physico-chemical and environmental requirements

The use of QSARs as adaptation is possible for many **physicochemical**, **ecotoxicological and fate requirements:**

- experimental data of sufficient number and quality are available for building robust models;
- mode of actions are well understood;
- the measured effects have good correlation with molecular descriptors (such as lipophilicity and aquatic toxicity in absence of specific mode of actions);
- the results reported from experimental studies include one or few effects (e.g., mortality as LC_{50} for fish short-term toxicity) that match those predicted by the QSAR models.

QSARs for high tier human health endpoints

The use of QSARs for **high tier human health requirements** should be limited to supporting information due to the complexity of the endpoints and effects measured by the experimental studies:

- a QSAR model may predict a NOAEL for repeated dose toxicity without providing information on target organs or other parameters that may be relevant to trigger further studies (e.g., specific organ toxicity or endocrine related effects) or classification;

- details on the underlying data used to build the QSAR model are often neither homogeneous nor of sufficient quality (due to the limited number of available data, model developers need to lower the quality standards for the data they use) or not reported at all.

QSARs for low tier human health endpoints

The use of QSARs as adaptation is possible for some **lower tier human health endpoints** for the same reasons mentioned for environmental endpoints.

A crucial requirement is that the predicted endpoint and results match those provided by the standard experimental test:

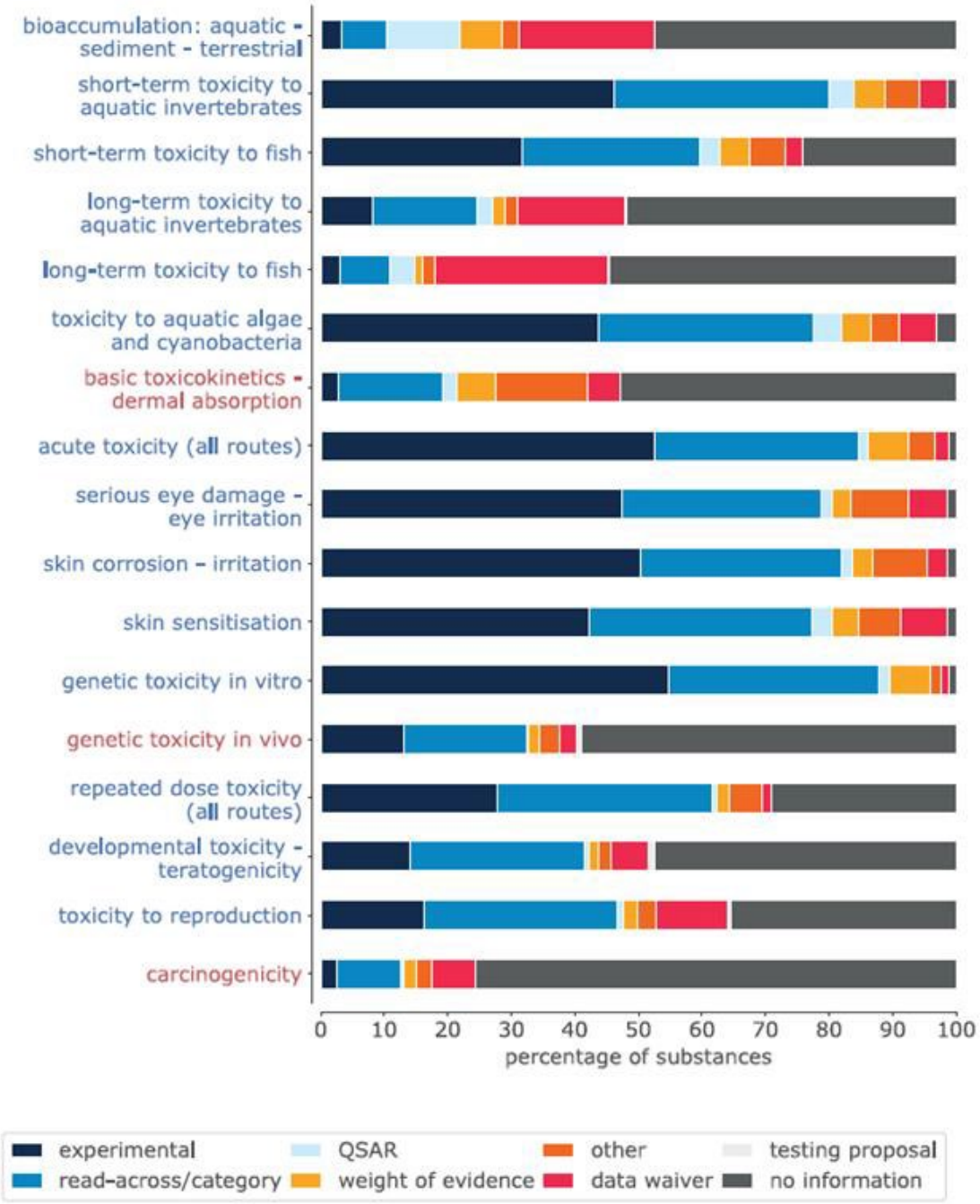
- For bacterial reverse mutation (Ames) test, the QSAR model must explicitly consider all the 5 strains and metabolic activation as required by the OECD TG 471;

- For skin sensitisation, the model must provide results that allow skin sensitisation classification based on GHS and CLP criteria.

Frequency of use of QSARs for different REACH endpoints

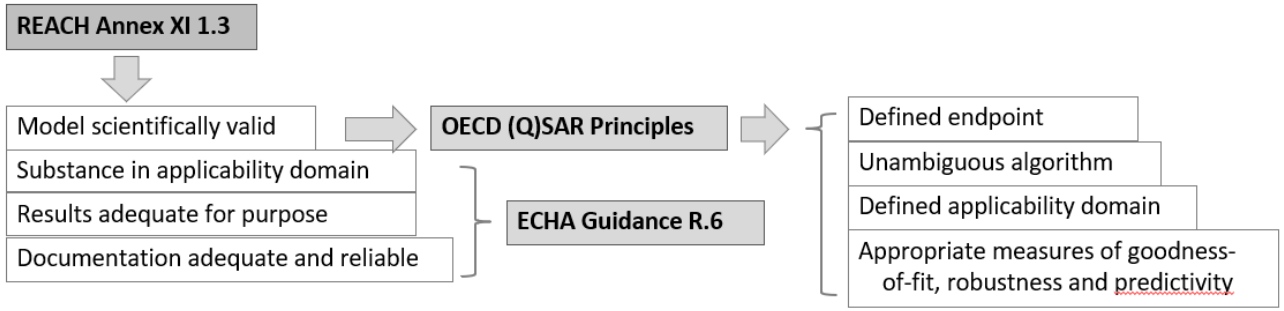
Source: Report on the use of alternatives to testing on animals for the REACH regulation (art 117.3 report)

- QSARs represent ~2.5% of the submitted information for the endpoints considered in the report
- For some endpoints, QSARs were used more (bioaccumulation, aquatic toxicity, low tier human health)



Take home messages

Conclusions



- Sufficient information must be provided to allow ECHA to assess each of the criteria in REACH Annex XI 1.3 – document both the model and the prediction;
- The validity of each individual prediction must be justified, the use of a valid model is not sufficient;
- The availability of valid QSAR models depends on the requirement addressed.

Thank you!

Doris.hirmann@echa.europa.eu
Mounir.Bouhifd@echa.europa.eu
Andrea.gissi@echa.europa.eu

Subscribe to our news at
echa.europa.eu/subscribe

Follow us on Twitter
@EU_ECHA

Follow us on Facebook
Facebook.com/EUECHA