The use and assessment of QSAR predictions under REACH

Advanced Toxicology Course: Computational Toxicology

Virtual event organised by the German Toxicology Society
01 March 2021

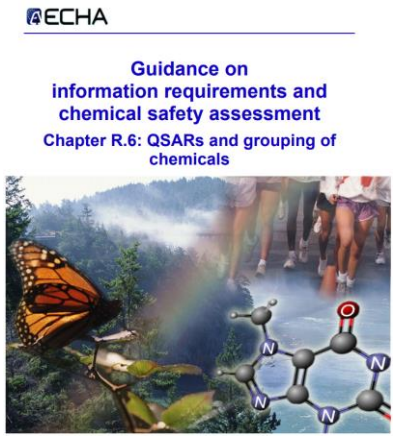
Andrea Gissi, PhD
Computational assessment unit
European Chemicals Agency

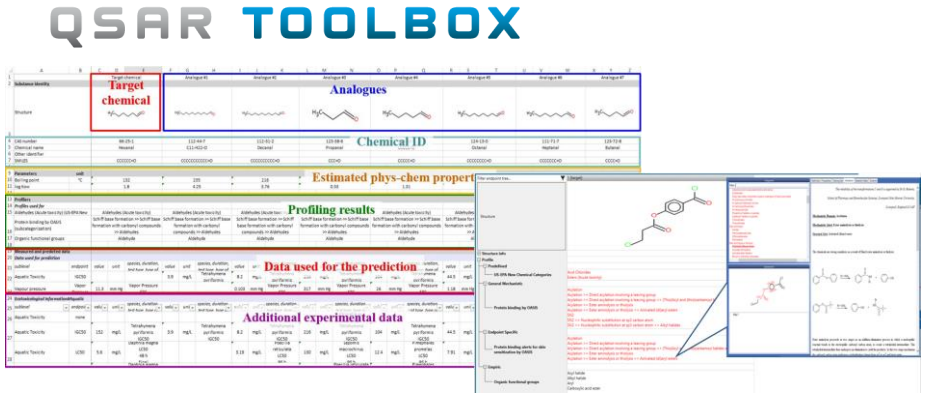
The views expressed in this presentation are solely those of the authors and the content of the presentation does not represent the views or position of the European Chemicals Agency

Outline of the presentation

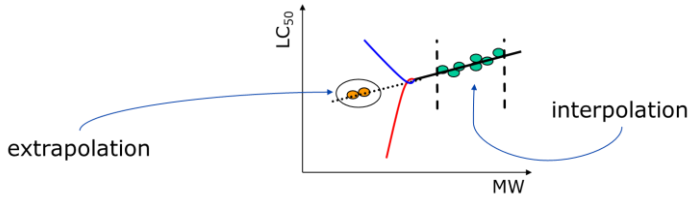
1. Promotion of the correct use of non-animal methods

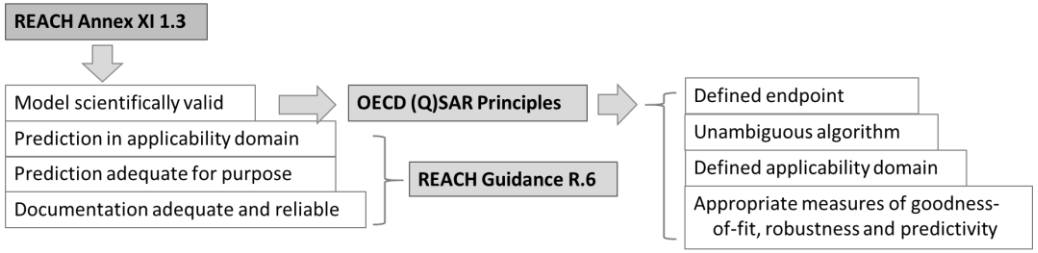
APCRA
ACCELERATING THE PACE OF CHEMICAL RISK ASSESSMENT





2. ECHA's assessment of QSAR predictions under REACH





REACH – main processes and actors

Data sharing
Registration
Self-classification

Industry gathers information and is responsible for risk management

Member States

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

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

Member States

Authorisation
Restriction
Harmonised C&L

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

ECHA's promotion of the correct use of non-animal methods for hazard assessment = first part of the presentation

ECHA assesses QSARs submitted by Industry under Dossier evaluation = second part of the presentation

Promotion of the correct use of non-animal methods

Minimisation of unnecessary animal testing

Testing on vertebrate animals for the purpose of REACH as last resort:

1. Avoid repetition of studies with data sharing and joint submission

2. Adapt standard information requirements (Annex XI):

   – Use of existing data

   – Weight of evidence

   – (Quantitative) Structure-activity relationships

   - Study is technically not possible

   – *In vitro* methods

   – **Grouping and read-across**

   – Exposure considerations

3. The use of adequate testing strategies can also support the reduction of animal tests

Promotion of correct use of non-animal methods

ECHA's contributions:

- **OECD QSAR Toolbox** - Co-ownership and co-development
- Participation to **OECD activities** – development of new OECD test guidelines that include non-animal methods, drafting of guidance documents
- Participation to **APCRA** - an international project on accelerating the pace of chemical risk assessment (APCRA) through the use of new approach methodologies (NAMs)
- **Dissemination of data** - making non-confidential data from REACH registrations more and more available for further developments of non-animal methods
- **Preparation of REACH Guidance and other documents** – e.g. Report on the use of alternatives to testing on animals for the REACH regulation (art 117.3 report), read-across assessment framework, practical guide on the use and reporting of QSARs, etc.

What is the QSAR Toolbox?

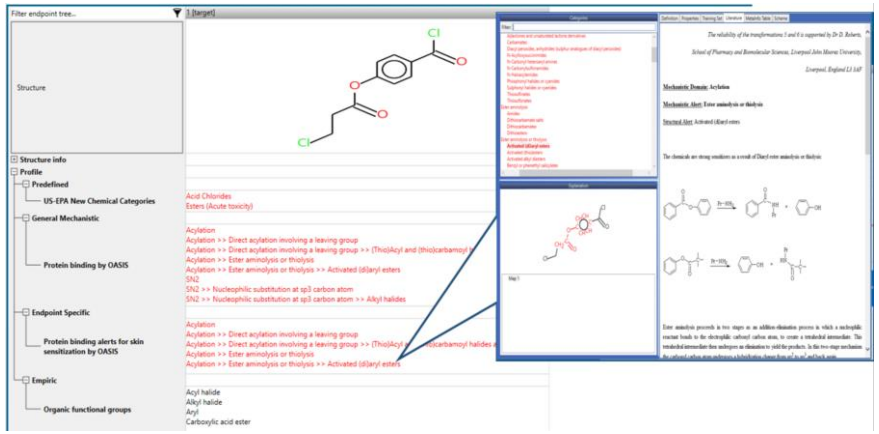
- The Toolbox is a free software application that supports reproducible and transparent chemical hazard assessment. ECHA and OECD co-own and co-develop it
- Freely available software. Official website: www.qsartoolbox.org
- It offers functionalities for retrieving experimental data, simulating metabolism and profiling properties of chemicals.
- These information and tools can be used to find structurally and mechanistically defined analogues and chemical categories, which can serve as sources for read-across and trend analysis for data gap filling.

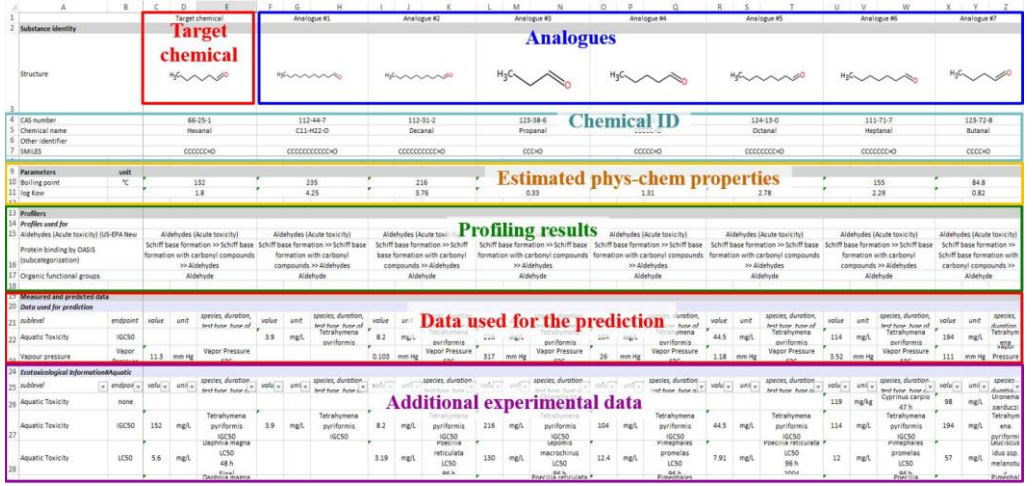
Use of QSAR Toolbox under REACH

- Widely used under REACH: thousands of registrations include Toolbox results
- Toolbox mainly used for read-across predictions or as supporting information to prove mechanistic similarity
- A read-across with Toolbox can be accepted only if the prediction report is complemented with manually compiled read-across justification according to RAAF principles (unfortunately, not often the case)
- **<u>Most of the predictions provided by the QSAR Toolbox are read-across from analogues, not QSAR results -> evaluated according to read-across rules (RAAF)</u>**

Typical workflow with the QSAR Toolbox

- Search for existing experimental data on the input substance
- Profile the substance to predict mode of action
- Find structural analogues with the same predicted mode of actions and experimental data on the desired endpoint
- Use data from analogues as source for read-across, trend-analysis or QSAR predictions
- Generate prediction and category reports, and data matrix





Participation to other OECD activities

ECHA has nominated experts for several OECD groups related to the development and promotion of correct use of alternatives. Some examples include:

1. For establishing a QSAR assessment framework: to develop a concrete framework for the assessment of QSAR predictions

2. For establishing Good Computational Modelling Practices: to develop criteria equivalent to GLP for computational methods

3. Defined approaches for skin sensitisation (DASS) expert group: for developing an OECD Test Guideline for skin sensitisation that includes a combination of in-vitro and in-silico methods (including the QSAR Toolbox) to predict skin sensitisation potential.

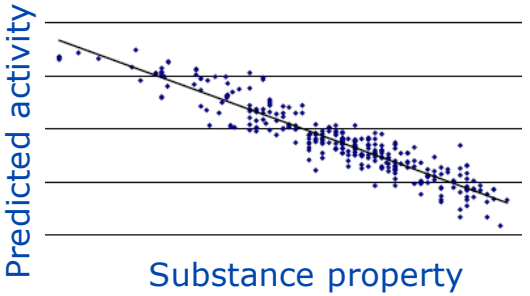
ECHA’s assessment of QSAR predictions under REACH

(Q)SARs in a nutshell

(Q)SARs are models that **relate chemical structure and activity of substances**. They can be used to predict the unknown activity of (new) substances.

QSAR

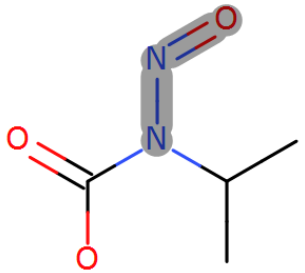
- Use an algorithm or equation to derive a **quantitative** prediction
- The prediction is a function of other properties of the substance ("molecular descriptors")



Quantitative relationship between logKow and fish LC_{50} for neutral organic compounds

SAR

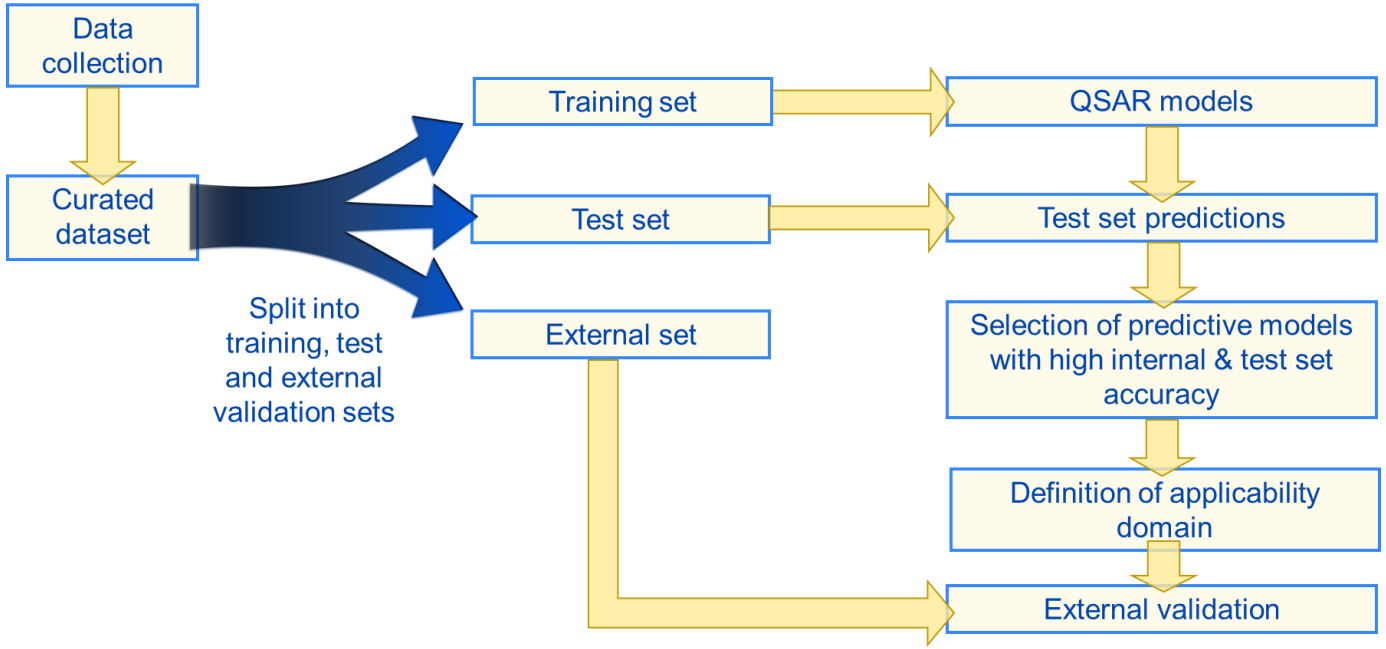
- Outcome **qualitative** (yes/no) or semi-quantitative (e.g. weak/strong)
- Often based on structural alerts, may not include a mathematical function



Positive prediction due to a structural alert:

Compounds with an N-nitroso group may be mutagenic or carcinogenic

Training and test sets



Curated dataset: high quality structures and experimental data used to build the model

Training set: used to derive the (Q)SAR algorithm

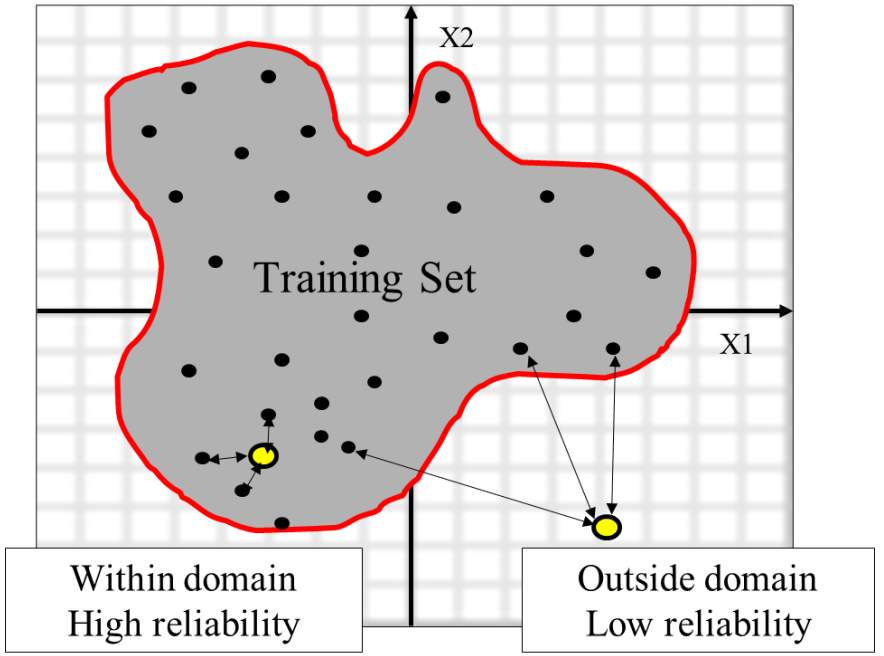
Test set: used to verify the ability of the model to predict "new" substances

High quality data = better model

Applicability domain (I)

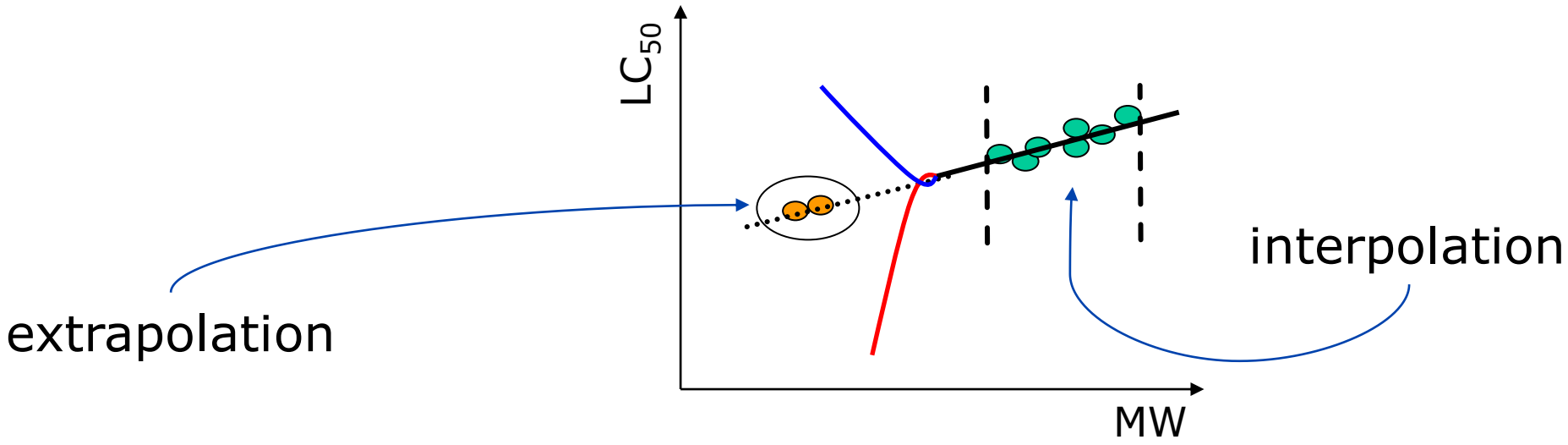
- The **applicability domain (AD)** of a QSAR model is the physico-chemical, structural or biological space, knowledge or information on which the **training set** of the model has been developed, and for which it is applicable to make predictions for new compounds.

- Lack of single generally accepted methodology for determining the applicability domain.



Applicability domain (II)

- The purpose of AD is to define for which chemicals the model can be reliably applied. In general, this is the case for interpolation rather than for extrapolation.
- E.g. the model is trained with structures with a molecular weight ranging from 16 to 350, and substances with a MW outside these boundaries should be considered out of its domain.



Acceptance of non-animal methods under REACH

Depending on the complexity of the endpoint, different level of acceptance:

1. For some "simpler" endpoints, *in vitro* approaches are the standard information requirements

2. For other "middle level of complexity" endpoints, the use is possible as **adaptations**, if properly justified*

3. For the "most complex" endpoints, QSAR results are not sufficient as stand-alone information to cover the requirements

* In principle, any endpoint can be adapted, as long as results provide equivalent information as compared to the standard test

Non-animal methods as Adaptations (I)

REACH allows the use of **"adaptations"** to standard information.

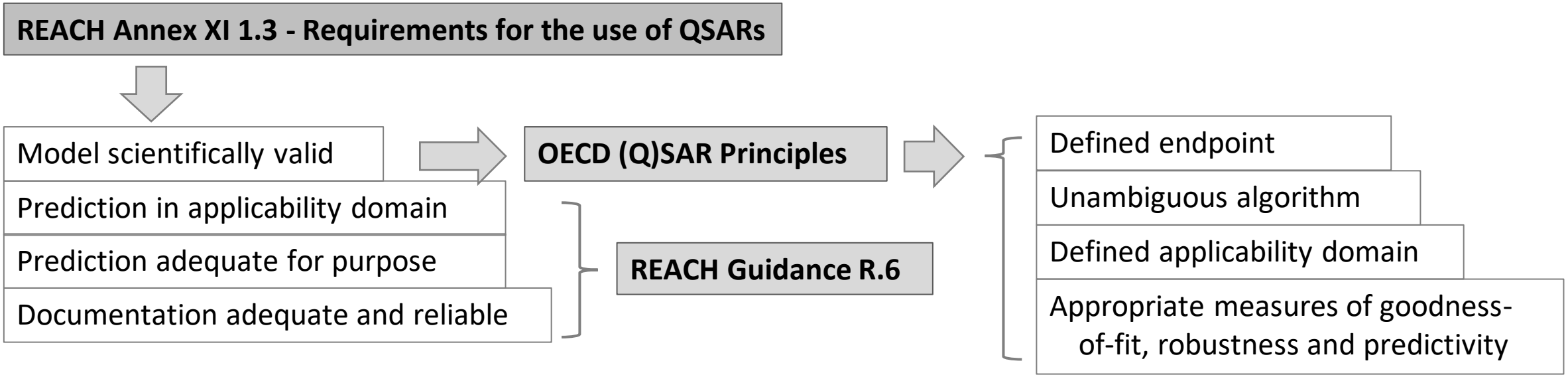
Two different types of adaptations in REACH:

- 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 indicates the types of adaptations possible and the requirements for their validity

Starting point for ECHA when assessing a QSAR prediction

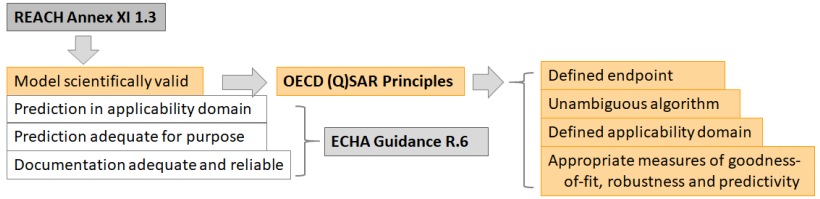


OECD principles: OECD ENV/JM/MONO(2007)2

http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=env/jm/mono(2007)2

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

https://echa.europa.eu/documents/10162/13632/information_requirements_r6_en.pdf



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 model scientifically valid? (OECD principles)

Defined endpoint – most common shortcoming related to model validity

Typical problems relate to the data used for the training set of the model:

- not sufficiently documented,
- inhomogeneous test data mixed together,
- mismatch between the effects measured by the OECD test guidelines and the QSAR model. For complex endpoints, usually the effects measured and reported in the experiment are many more than those reported by the QSAR prediction.

Example of endpoint not well defined for REACH

A QSAR model for developmental toxicity provides results based on the table shown here.

Two possible results:

- Non developmental toxicant
- Developmental toxicant

These results cannot be used to adapt REACH developmental toxicity information requirements because the results are not comparable with the outcome of standard tests (e.g. OECD TG 414).

Can be used as supporting information, e.g. in addition to read-across

FDA classes	Definition	[redacted] Binary class
Category A	Negative human studies	Non developmental toxicant
Category B	Negative animal studies & No human studies executed OR Positive animal studies & Negative human studies	
Category C	Postive animal studies & No human studies executed OR No studies at all	Developmental toxicant
Category D	Postive human studies	
Category X	Animal OR human studies show abnormalities AND/OR Evidence of foetal risk based on human experience	

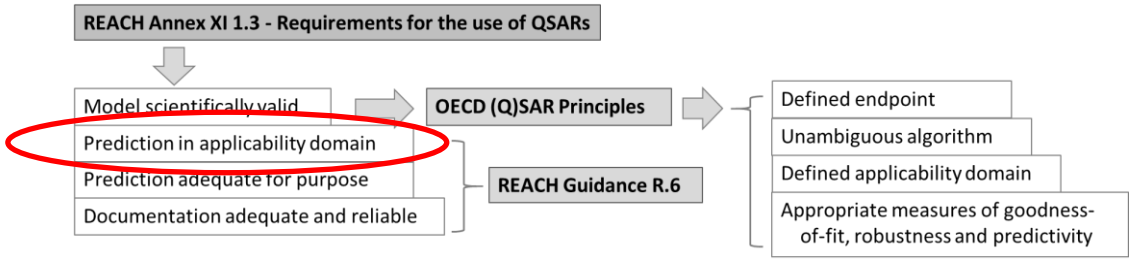
Acceptable model ≠ Acceptable prediction

A scientifically valid model is a necessary but not sufficient condition for a prediction to be accepted.

Predictions need to fulfil additional requirements!

OECD QSAR principles only cover the scientific validity of the model.

Prediction within domain?

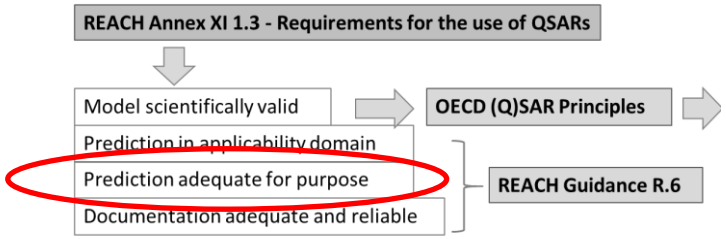


The model developers definition of applicability domain is the starting point for the assessment.

In addition, the following aspects are checked:

- Descriptor domain
- Structural domain
- Mechanistic domain
- (Metabolic domain)

Prediction adequate for purpose?



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

Substance identity -> correct structure(s) must be predicted.

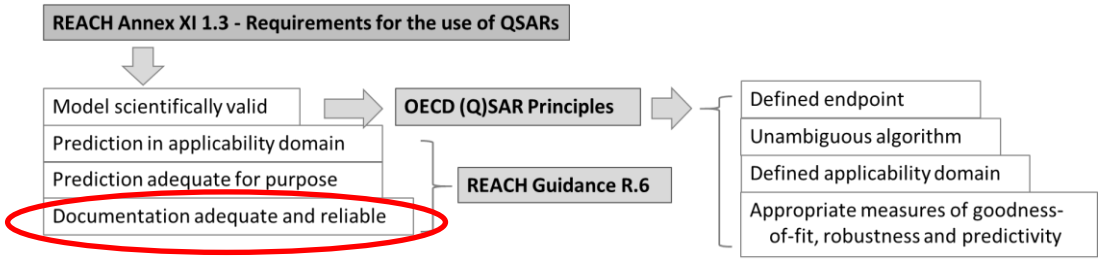
This is not trivial in case of multi-constituents or substances with Unknown or Variable composition, Complex reaction product or Biological origin (UVCB).

Reliability of the prediction

In addition to being within the applicability domain of the model, the reliability of a prediction can be assessed by considering the following aspects:

- reliability of input parameters (e.g. predicted vs experimental descriptor values)
- 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 (same or different endpoints)

Adequate documentation?



Model (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.

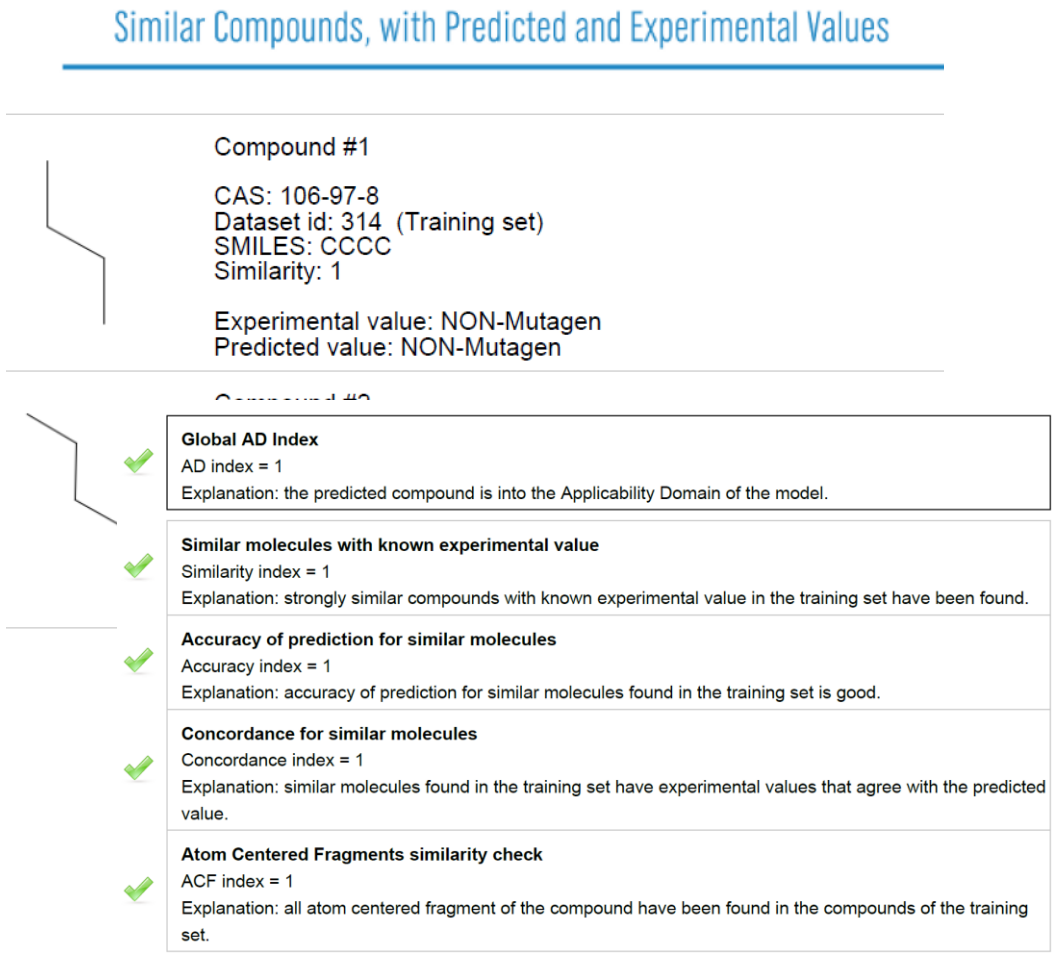
Prediction (QPRF) must include information on:

- 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.**

***in bold the most common shortcomings**

Example of domain and reliability assessment automatically provided by a software

- There are software that make automatic assessments or applicability domain and reliability of their predictions
- Results need to be critically investigated; no automatic acceptance based on software results



(Personal) reflections and conclusions

Current status on acceptance of QSARs under REACH (1/3)

For **ecotoxicological and environmental endpoints**, valid QSAR models for REACH are available for aquatic toxicity (especially short-term) for different trophic levels and bioaccumulation in fish:

- Enough experimental data are available for building robust models,
- The measured effects have good correlation with "simple" physico-chemical properties, especially when no specific mode of actions are expected,
- The results reported from experimental studies include one or few effects (e.g., mortality as LC50 for fish short-term toxicity) that are also predicted by the QSAR models.

If a valid model is used and the prediction fulfils the other Annex XI requirements and it is properly justified, then the results are accepted.

Current status on acceptance of QSARs under REACH (2/3)

For **human health endpoints**, the acceptance of QSAR results is hindered mainly by the complexity of the endpoint and effects measured by the experimental studies.

QSAR models provide simplified results, which are not adequate for replacement of the standard tests for **high tier endpoints**, e.g.:

- a QSAR model may predict a NOAEL for repeat 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)

QSAR results should be used only as supporting information in these cases.

Current status on acceptance of QSARs under REACH (3/3)

Valid QSAR models for some **lower tier human health endpoints** are available. A crucial requirement is that the predicted endpoint and results match those provided by the standard test:

- For bacterial reverse mutation (Ames) test if the model explicitly considers all the 5 strains and metabolic activation as required by the OECD TG 471;
- For skin sensitization if the model provides results that allow skin sensitization classification based on GHS and CLP criteria.

If a valid model is used and the prediction fulfils the other Annex XI requirements and properly justified, then the results are accepted.

Thinking about future…

- QSAR models and New Approach Methodologies (NAMs) in general keep being developed;
- NAMs produce results related to in-vivo endpoints required by REACH, however they do not measure the same effects of the animal methods;
- REACH objective is the protection of human health and environment. Acceptance of alternatives should be granted only if **equivalent level of protection** is expected with their use. How can this be achieved and measured? The question is still open…
- ECHA is involved in APCRA projects that compare point of departure estimates based on NAMs and in-vivo results for hazard assessment to investigate if this is an option

Take home messages

- ECHA is leading or involved in many activities for the promotion of the use of alternative methods by companies and authorities

- ECHA assesses the compliance of QSAR results according to Guidance R6 requirements. It is challenging to successfully use QSARs to adapt (high tier) human health endpoints due to limitations of the existing models

- Keeping in mind that the objective is to ensure the protection of human health and environment, it is still an open question how in future the new methods can be used in the regulatory framework for ensuring safe use and/or for screening and prioritisation of substances

Thank you!

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