ASCCT - American Society for Cellular and Computational Toxicology

Webinar

The New OECD (Q)SAR Assessment Framework: Details and Examples

Part1

Presenter:
Olga Tcheremenskaia
Department of Environment and Health, Istituto Superiore di Sanità (ISS), Italy

OECD (Q)SAR Assessment Framework (QAF):
Assessment of (Q)SAR models and examples

(Q)SAR Assessment Framework

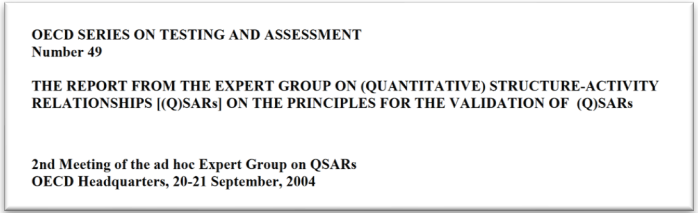
Table of contents

Principles for assessment of (Q)SAR models

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

Principles for QSAR model evaluation were established almost twenty years ago and extensively used so far by the scientific and regulatory communities:
https://one.oecd.org/document/env/jm/mono(2004)24/en/pdf

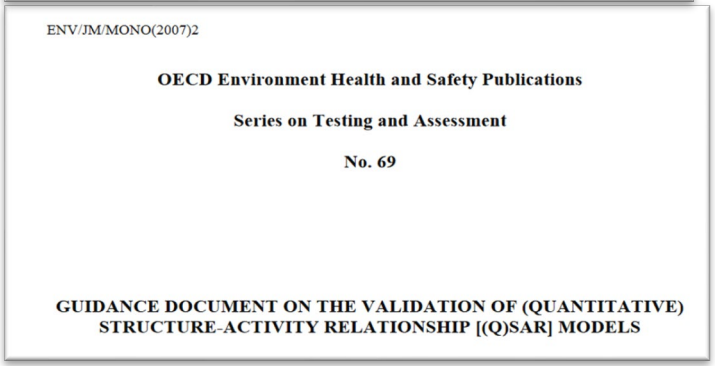
Guidance Document on the Validation of (Q)SAR Models was published in 2007 with the aim of providing guidance on how specific (Q)SAR models can be evaluated with respect to the OECD principles
https://one.oecd.org/document/env/jm/mono%282007%292/en/pdf



OECD SERIES ON TESTING AND ASSESSMENT
Number 49

THE REPORT FROM THE EXPERT GROUP ON (QUANTITATIVE) STRUCTURE-ACTIVITY RELATIONSHIPS [(Q)SARs] ON THE PRINCIPLES FOR THE VALIDATION OF (Q)SARs

2nd Meeting of the ad hoc Expert Group on QSARs
OECD Headquarters, 20-21 September, 2004



ENV/JM/MONO(2007)2

OECD Environment Health and Safety Publications

Series on Testing and Assessment

No. 69

GUIDANCE DOCUMENT ON THE VALIDATION OF (QUANTITATIVE) STRUCTURE-ACTIVITY RELATIONSHIP [(Q)SAR] MODELS

Assessment elements for (Q)SAR models in the guidance and in the checklist

Description of the algorithm and/or software *(AE 2.1 in the Model Checklist)*

The first element to be checked is the availability of a transparent description of the algorithm. The model equation, if applicable, including all descriptors and approach used for their selection, should be detailed. Furthermore, if applicable, a list of fragments/structure alerts (e.g., active, inactive, masks) and their description should be provided. The rationale that guided their identification could also be included. Calculated descriptors should be denoted with the software name and version used for their calculation. Furthermore, the version, developers' contact information and any available description of the software for the (Q)SAR model should also be provided. When an exact description of the algorithm is not publicly available (e.g., for commercial models), any available relevant information should still be assessed.

Inputs and other options *(AE 2.2 in the Model Checklist)*

Secondly, assessors should check if the documentation includes a description of inputs and settings of the model software. The allowed (or preferred) input formats for the chemical structure and its descriptors, including applicable pre-processing procedures (e.g., for salts and tautomers) should be documented. Further, customisable options/settings on the software should be reported and explained. Unless justified otherwise, the recommended input formats and options are expected to be the same as those used by model developers when developing the model and assessing its performance.

Model accessibility *(AE 2.3 in the Model Checklist)*

Finally, it should be checked if the model version under assessment is publicly accessible. A working link to access or download the model is expected in the QMRF documentation. When assessors have access to a different version of the model under assessment (e.g. a newer version), any differences in the outputs should be investigated.

OECD (Q)SAR Model Principle 2 is further considered in the Prediction and Result Checklists under the element "Reproducibility". Note that when the model is implemented in a software program that is accessible to the assessor, the reproducibility of the results should be possible even for cases when the description of the algorithm is not fully disclosed. Assessors may decide that this is acceptable for some regulatory purposes.

Principle	Assessment element	Practical advice	Examples
Unambiguous algorithm			
2.1	Description of the algorithm and/or software	An exact description of the algorithm might not be publicly available for commercial models. In such cases, any available relevant information should still be assessed. When the model is implemented in a computer program that is accessible to the assessor, the reproducibility of the results should be possible even for cases when the description of the algorithm is not fully disclosed, and assessors may decide that this is acceptable for some regulatory uses.	User manuals, publications, help files, such as EPISuite help file
2.2	Inputs and other options	The extent of this description depends on the complexity of the computer program. Simple programs with no customisable options require less explanations than programs that allow editing of the settings of the algorithm.	Instructions on the preparation of the input may include instructions how to pre-process salts and tautomers.
2.3	Model accessibility	When a different model version is available to the assessor, consider using it and compare the results.	"In vitro mutagenicity (Ames test) alerts" fragment-based model implemented in Toxtree 3.1.0 software available at https://toxtree.sourceforge.net/ has been used for generate a prediction.

Each principle is broken down to assessment elements (AEs)

The Guidance gives more details for each AE, the checklists - more practical advice and examples

Glossary of selected terms

- **Model checklist**: a separate document to facilitate the assessment of a (Q)SAR models according to QAF principles. **It includes a list of assessment elements** to consider, columns to record the outcome of the assessment, practical advice, and examples.
- **Assessment element (AE)**: a critical aspect to consider when assessing (Q)SAR models, predictions and overall results meet. AEs are associated with the OECD (Q)SAR principles for models and results.

- **(Q)SAR model**: a model that predicts the **property of a substance** using as input information on the structure,
- **Property**: a physicochemical, toxicological, ecotoxicological, or fate property; chemical reactivity or biological interaction. In this document, the term "property" is preferred to "endpoint" because of the different understanding of the meaning of the term endpoint depending on the audience.

Checklist for the regulatory assessment of (Q)SAR models

Annex C. (Q)SAR Model, Prediction and Result Checklists

The checklist (EXCEL version) is available at the following link:
https://www.oecd.org/chemicalsafety/testing/QAF-Checklist.xlsx

Model name and version:
Software name and version (if applicable):
Predicted property:
Intended purpose of use of the model:
QMRF availability:
Assessor name and date of the assessment:

Model 1

when more than one model is considered, add a comment here to identify to which model the checklist refers to (e.g. model name)

Principle	Assessment element	Outcome	Comments
Defined endpoint			
1.1	Clear scientific and regulatory purpose		
1.2	Transparency of the underlying experimental data		
1.3	Quality of the underlying experimental data		
Unambiguous algorithm			
2.1	Description of the algorithm and/or software		
2.2	Inputs and other options		
2.3	Model accessibility		
Defined domain of applicability			
3.1	Clear definition of the applicability domain and limitations of the model		
Appropriate measures of goodness-of-fit, robustness and predictivity			
4.1	Goodness-of-fit, robustness		
4.2	Predictivity		
Mechanistic interpretation			
5.1	Plausibility of the mechanistic interpretation		
Conclusion on the model		The conclusion is based on the outcome of the assessment elements as decided by i	
Comments			

A list of critical elements to which the assessor should assign a predefined value (**i.e., fulfilled, not fulfilled, not applicable/assessed, not documented**).

The analysis of each element supports the **overall decision** on whether the model is suitable for the intended regulatory purpose.

Model criteria and QMRF mapping

- Checklist provides details, practical advice, examples and mapping to the (Q)SAR model reporting format (QMRF) for each AE

Checklist for the regulatory assessment of (Q)SAR models
Details on the assessment elements

Principle	Assessment element	Objective	What to check and how	Practical advice	Examples	Mapping to the most relevant QMRF field(s)
Defined endpoint						
1.1	Clear scientific and regulatory purpose	The predicted endpoint is clearly defined in relation to a scientific and/or regulatory purpose.	The predicted endpoint is clearly defined andis consistent with the data used to build the model. For a clear scientific purpose: the predicted endpoint refers to physicochemical, biological or environmental effects o that can be measured and therefore modelled. For a clear regulatory purpose: the predicted endpoint refers to a specific regulatory requirement or test method or test guideline.	The description of the predicted endpoint should be as detailed as possible by including all elements that have been taken into account (e.g. the unit of measurement, timescale, observations such as growth, mortality, etc.).	Clear scientific (and regulatory) purpose: predicted endpoint = "Fish-short term toxicity (96 hours) as LC50 according to the OECD Test Guideline 203". Clear regulatory purpose: Predicted endpoint = "Classification for skin sensitisation according to GHS criteria".	3.2 Endpoint 3.3 Comment on endpoint 3.5. Dependent variable 3.6. Experimental protocol
1.2	Transparency of the underlying experimental data	The documentation is sufficient to independently assess the quality of the experimental data used to build the model for the next assessment element.	Check to what extent the following information is available : - Clear identification of the substances tested (name, structures, SMILES numerical identifiers, etc.); -A (primary) reference to the original studies; - Description of relevant experimental conditions that could affect the prediction (e.g. sex, species, temperature, exposure period, protocol, measurements units); -The original value in the case of data processing before modelling, information on data processing, unit or scale conversion; - Availability of the description of the data aggregation procedure and individual values for datasets where multiple data for the same substance are aggregated for modelling; - Information in the experimental data selection and curation procedure.	It is rare to have full details on each data point used to build the model, but a general description about the experimental data selection and curation procedure can be expected.	Example 1: The model documentation includes the list of substances part of the training set, the experimental values for the predicted property and details or reference for each data point. This assessment element is fulfilled. Example 2: The predicted endpoint is "Bacterial mutagenicity according to OECD TG 471", but the information on the underlying data does include information on the strains tested or presence of metabolic activation. This assessment element is not fulfilled.	3.1 Species 3.4 Endpoints units 3.5 Dependent variable 3.6 Experimental protocol 6.2 Available information for the training set 6.3 Data for each descriptor variable for the training set 6.4 Data for the dependent variable for the training set 6.5 Other information about the training set
1.3	Quality of the underlying experimental data	Ensure that the model is built on data of sufficient quality to obtain acceptable predictions.	- Assess the experimental data curation procedure; - Assess the quality of the data point individually, if possible;	Ideally data points should be evaluated individually. However, especially for large training sets, this may be not possible. In these cases, assessors can verify how the relevant experimental conditions that could affect the results of experimental studies (e.g., sex, species, temperature, exposure period, protocol) have been considered when selecting data to build the model. For models with large training sets, spot check some data points. In some cases, lower data quality can be compensated by large number of data points fitting the same trend.	The model documentation indicates that the predicted endpoint is fish long-term toxicity. The assessment of the data used to build the model shows that the duration of the exposure was not taken into account when selecting data to build the model. It is suspected that some of the data used to build the model refer to results from fish short-term toxicity studies. Outcome: This assessment element is not fulfilled and the model not considered valid for predicting fish long-term toxicity.	3.7 Endpoint data quality and variability 6.6 Pre-processing of data before modelling
Unambiguous algorithm						
2.1	Description of the algorithm and/or software	Ensure that it is clear how the prediction is obtained and that it can be reproduced by others	- Check if a sufficient description of all descriptors and of approach used for their selection and calculation is provided; - Check the availability of a transparent description of the algorithm and/or software, explaining how the predictions were produced. - For fragment/alert based models, the list of the fragments (active, inactive, masks, etc. as relevant) together with information of all substructures and identification of its substituents should be provided. - For equation based models, a description of the equation and all data/descriptors and approach used for their selection should be provided.	An exact description of the algorithm might not be publicly available for commercial models. In such cases, any available relevant information should still be assessed. When the model is implemented in a computer program that is accessible to the assessor, the reproducibility of the results should be possible even for cases when the description of the algorithm is not fully disclosed, and assessors may decide that this is acceptable for some regulatory uses.	User manuals, publications, help files, such as EPISuite help file	4.1 Type of model 4.2 Explicit algorithm 4.3 Descriptors in the model 4.4 Descriptor selection 4.5 Algorithm and descriptor generation 4.6 Software name and version for descriptor generation 4.7 Chemicals/Descriptors ratio 6.1 Availability of the training set
2.2	Inputs and other options	Allowed input formats, pre-processing procedure for the input structures and customisable options/settings are explained.	- Availability of instructions to prepare the input. - Availability of information on the editable options/settings (if any).	The extent of this description depends on the complexity of the computer program. Simple programs with no customisable options require less explanations than programs that allow editing of the settings of the algorithm.	Instructions on the preparation of the input may include instructions how to pre-process salts and tautomers.	1.3 Software coding the model 2.8 Availability of information about the model 6.6 Pre-processing of data before modelling
2.3	Model accessibility	Assess if the model or computer program is or can be available to the assessor.	- Availability of the same model and version described in the documentation	When a different model version is available to the assessor, consider using it and compare the results.	"In vitro mutagenicity (Ames test) alerts" fragment-based model implemented in Toxtree 3.1.0 software available at https://toxtree.sourceforge.net/ has been used for generate a prediction.	1.3 Software coding the model 2.5 Model developer(s) and contact details 2.6 Date of model development and/or publication 2.7 Reference(s) to main scientific papers and/or software package 2.8 Availability of information about the model

1. Defined endpoint

A (Q)SAR should be associated with a “defined endpoint”, where endpoint refers to any physicochemical, biological, or environmental property that can be measured and therefore modelled.

The intent of this principle is to ensure transparency in the endpoint being predicted by a given model, since an endpoint could be determined by different experimental protocols and under different experimental conditions.

The AEs to verify that the endpoint is clearly defined:

- Clear scientific and regulatory purposes
- Transparency of the underlying experimental data
- Quality of the underlying experimental data

Clear scientific and regulatory purposes (AE 1.1 in the Model Checklist)

Objective

The predicted endpoint is clearly defined in relation to a scientific and/or regulatory purpose.

What to check and how

- The predicted endpoint is clearly defined and is consistent with the data used to build the model.
- For a clear scientific purpose: the predicted endpoint refers to physicochemical, biological or environmental effects, can be measured and therefore modelled.
- For a clear regulatory purpose: the predicted endpoint refers to a specific regulatory requirement or test method or test guideline.

Clear scientific and regulatory purposes (AE 1.1 in the Model Checklist)

Example

Clear scientific (and regulatory) purpose:

Predicted endpoint = “Fish-short term toxicity (96 hours) as LC50 according to the OECD Test Guideline 203”

The AE is fulfilled.

Clear regulatory purpose:

Predicted endpoint = "Classification for skin sensitisation according to GHS criteria”

The AE is fulfilled.

Practical advice

The description of the predicted endpoint should be as detailed as possible by including all elements that have been considered (e.g., the unit of measurement, timescale, observations such as growth, mortality, etc).

Transparency of the underlying experimental data (AE 1.2 in the Model Checklist)

Objective

The documentation is sufficient to independently assess the quality of the experimental data used to build the model for the next assessment element.

What to check and how

Check to what extent the following information is available :

- Clear identification of the substances tested (name, structures, SMILES numerical identifiers, etc.)
- Reference to the original studies
- Description of relevant experimental conditions that could affect the prediction (e.g., sex, species, temperature, exposure period, protocol, measurements megument units)
- The original value in the case of data processing before modelling, information on data processing, unit or scale conversion
- Availability of the description of the data aggregation procedure and individual values for datasets where multiple data for the same substance are aggregated for modelling
- Information in the experimental data selection and curation procedure

Transparency of the underlying experimental data (AE 1.2 in the Model Checklist)

Examples

Example 1: The model documentation includes the list of substances part of the training set, the experimental values for the predicted property and details or reference for each data point. This assessment element is fulfilled.

Example 2: The predicted endpoint is "Bacterial mutagenicity according to OECD TG 471", but the information on the underlying data does include information on the strains tested or presence of metabolic activation. This assessment element is not fulfilled.

Practical advice

It is rare to have full details on each data point used to build the model, but a general description about the experimental data selection and curation procedure can be expected.

Quality of the underlying experimental data (AE 1.3 in the Model Checklist)

Objective

Ensure that the model is built on data of sufficient quality to obtain acceptable predictions.

What to check and how

- Assess the experimental data curation procedure
- Assess the quality of the data point individually, if possible

Quality of the underlying experimental data (AE 1.3 in the Model Checklist)

Example

The predicted endpoint is fish long-term toxicity.

The assessment of the data used to build the model shows that the duration of the exposure was not considered when selecting data to build the model.

It is suspected that some of the data used to build the model refer to results from fish short-term toxicity studies.

Outcome: This assessment element is not fulfilled, and the model not considered valid for predicting fish long-term toxicity.

Practical advice

- Ideally data points should be evaluated individually. However, especially for large training sets, this may be not possible. In these cases, assessors can verify how the relevant experimental conditions that could affect the results of experimental studies (e.g., sex, species, temperature, exposure period, protocol) have been considered when selecting data to build the model.
- For models with large training sets, spot check some data points.
- In some cases, lower data quality can be compensated by large number of data points fitting the same trend.

2. Unambiguous algorithm

A (Q)SAR model should be expressed in the form of an unambiguous algorithm (intended as unambiguous description of the algorithm). **The intent of this principle is to ensure transparency in the description of the model algorithm to allow an independent reproducibility of its predictions.**

The Model Checklist includes the following AEs to verify the principle of an unambiguous algorithm:

- **Description of the algorithm and/or software**
- **Inputs and other options**
- **Model accessibility**

Description of the algorithm and/or software (AE 2.1 in the Model Checklist)

Objective

Ensure that it is clear how the prediction is obtained and that it can be reproduced by others

What to check and how

- Check if a sufficient description of all descriptors and of approach used for their selection and calculation is provided;
- Check the availability of a transparent description of the algorithm and/or software, explaining how the predictions were produced.
- For fragment/alert-based models, the list of the fragments (active, inactive, masks, etc. as relevant) together with information of all substructures and identification of its substituents should be provided.
- For equation-based models, a description of the equation and all data/descriptors and approach used for their selection should be provided.

Description of the algorithm and/or software (AE 2.1 in the Model Checklist)

Example

Availability of user manuals, publications, help files, such as EPISuite help file

The AE is fulfilled.

Practical advice

- An exact description of the algorithm might not be publicly available for commercial models. In such cases, any available relevant information should still be assessed.
- When the model is implemented in a computer program that is accessible to the assessor, the reproducibility of the results should be possible to assess even for cases when the description of the algorithm is not fully disclosed, and assessors may decide that this is acceptable for some regulatory uses.

Inputs and other options (AE 2.2 in the Model Checklist)

Objective

Allowed input formats, pre-processing procedure for the input structures and customisable options/settings are explained.

What to check and how

- Availability of instructions to prepare the input.
- Availability of information on the editable options/settings (if any).

Example

Instructions on the preparation of the input include instructions how to pre-process salts. AE is fulfilled

Practical advice

The acceptable level of details depends on the complexity of the computer program. Simple programs with no customisable options require less explanations than programs that allow editing of the settings of the algorithm.

Model accessibility (AE 2.3 in the Model Checklist)

Objective

Assess if the model or computer program is or can be available to the assessor.

What to check and how

-Availability of the same model and version described in the documentation

Example

"In vitro mutagenicity (Ames test) alerts" fragment-based model implemented in Toxtree 3.1.0 software available at https://toxtree.sourceforge.net/ has been used for generate a prediction.
The AE si fulfilled

Practical advice

When a different model version is available to the assessor, consider using it and compare the results.

3. A defined domain of applicability

The AD of a (Q)SAR model, as described in the Guidance (OECD, 2007), is the response and chemical structure space in which the model makes predictions with a given reliability.

Elaborating on the AD definition given above, the **AD should therefore consider the parametric, structural, mechanistic, metabolic and response space of the model.**

The QAF does not prescribe a specific way to define the AD of a model because multiple valid methodologies can be used but focuses on practical aspects of the assessment within the QAF.

The Model Checklist includes one AE related to the applicability domain:

- **Clear definition of the applicability domain and limitations of the model**

Clear definition of the applicability domain and limitations of the model (AE 3.1 in the Model Checklist)

Objective

Ensure that the AD definition is sufficiently detailed to allow the assessment of how a given substance relates to the AD of the model (is the substance within the AD of the model?)

What to check and how

- Check that the AD definition has sufficient details to decide if a substance is within AD

Example

The prediction report obtained using a model includes in the information on the applicability of the model to the input substance explaining how the assessment is done fulfil this criteria.

Practical advice

- Many modern models automatically assess if a substance falls within their applicability domain.
- Some models include global and local definitions of applicability domain. At this stage, the assessment might be limited to the definition of the global domain, since many aspects associated to local domain are also assessed in the prediction checklist

4. Appropriate measures of goodness-of-fit, robustness and predictivity

A (Q)SAR should be associated with "appropriate measures of goodness-of–fit, robustness and predictivity."

This principle expresses the need to provide information on the goodness-of-fit and robustness of a model (as determined by internal validation) and the predictivity of a model (as determined by external validation).

The performance should be measured within the applicability domain defined by its developers. The Guidance Document (OECD, 2007) can be consulted for further scientific aspects concerning Principle 4.

The Model Checklist includes the following AEs to verify the appropriateness of measures of goodness-of- fit, robustness and predictivity of the model:

- **Goodness-of-fit, robustness**
- **Predictivity**

Goodness-of-fit, robustness (AEs 4.1 in the Model Checklist)
Predictivity (AEs 4.2 in the Model Checklist)

Objective

Measures of performance for goodness-of-fit and robustness are provided and considered adequate.
Measures of performance for predictivity are provided and considered adequate.

What to check and how

Check the available information on the statistical method(s) used for internal/external validation of the model :

- For models predicting continuous endpoints, availability of at least basic statistics such as r2 value and standard error;
- For models predicting categorical endpoints, availability of at least basic statistics such as accuracy, sensitivity and specificity;
- If the regulatory context sets some reference values, compare the performance of the model to the reference values.
- An indication whether cross-validation or resampling was performed, if yes, by which method.

Goodness-of-fit, robustness (AEs 4.1 in the Model Checklist)
Predictivity (AEs 4.2 in the Model Checklist)

Example

For a model predicting categorical endpoints, the information on accuracy, sensitivity and specificity on the training set and on the external set is provided and considered good enough for the intended regulatory purpose. The AE is fulfilled.

Practical advice

- These measures estimate the general performance of the model. When assessing individual results, local performance assessed within the prediction checklist may be more important.
- In some cases, models lack measures of predictivity (i.e., external validation). Authorities responsible for the assessment should clarify if this is acceptable for their applications.
- If the external validation has been performed using data selected according to different criteria than the training set (e.g., at a different time or by different authors), assess the quality of these data.

5. Mechanistic interpretation

A (Q)SAR “should be associated with a mechanistic interpretation, if possible”.

Assessors may require that the model documentation includes considerations on how the rationale behind a (Q)SAR model is consistent with the knowledge related to the predicted property (such as known Adverse Outcome Pathways, AOPs, relevant for the predicted property), namely a mechanistic interpretation. Toxicokinetic considerations are also part of the mechanistic interpretation, if relevant for the property of interest.

The Model Checklist includes the following AE related to mechanistic interpretation:

- **Plausibility of the mechanistic interpretation**

Plausibility of the mechanistic interpretation (AE 5.1 in the Model Checklist)

Objective

To assess if the provided mechanistic interpretation is scientifically sound.

What to check and how

- Scientific plausibility of the proposed mechanistic interpretation (e.g., reference to scientific literature), when available.
- Check if a sufficient explanation and interpretation of the descriptors that is consistent with a known mechanism of (biological) action are provided.
- Check at what stage of modelling the mechanistic basis of the model was determined is provided.
- If relevant, an explanation and interpretation of the molecular events that underlie the properties of molecules containing the substructure should be provided.
- Consider that a mechanistic interpretation is optional in the OECD document on model validity ("if possible")

Plausibility of the mechanistic interpretation (AE 5.1 in the Model Checklist)

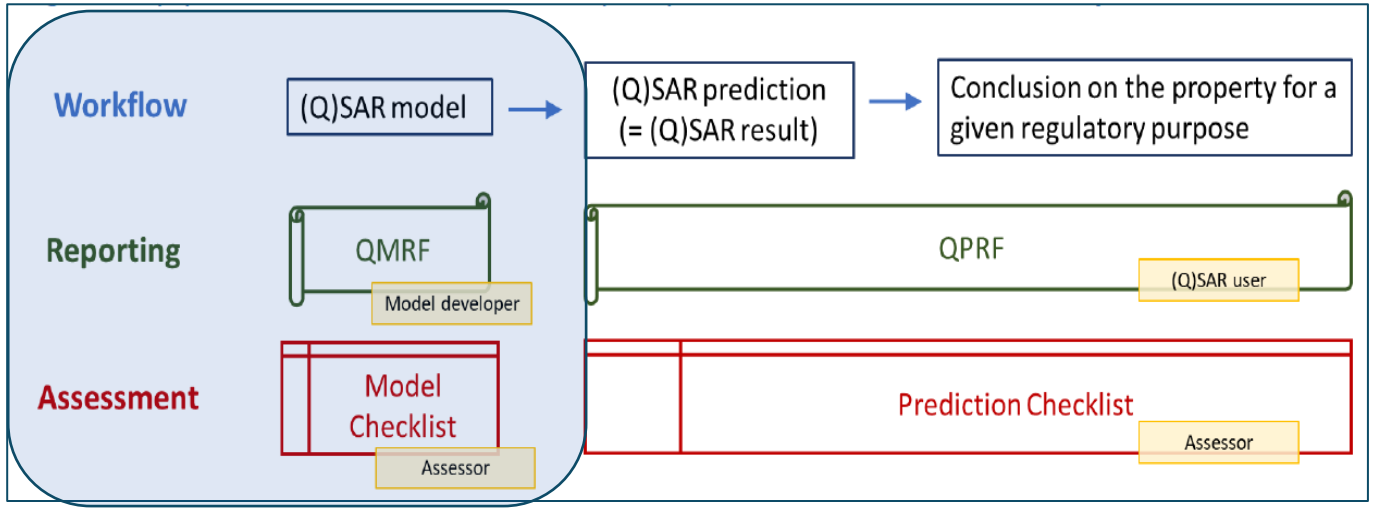
Example

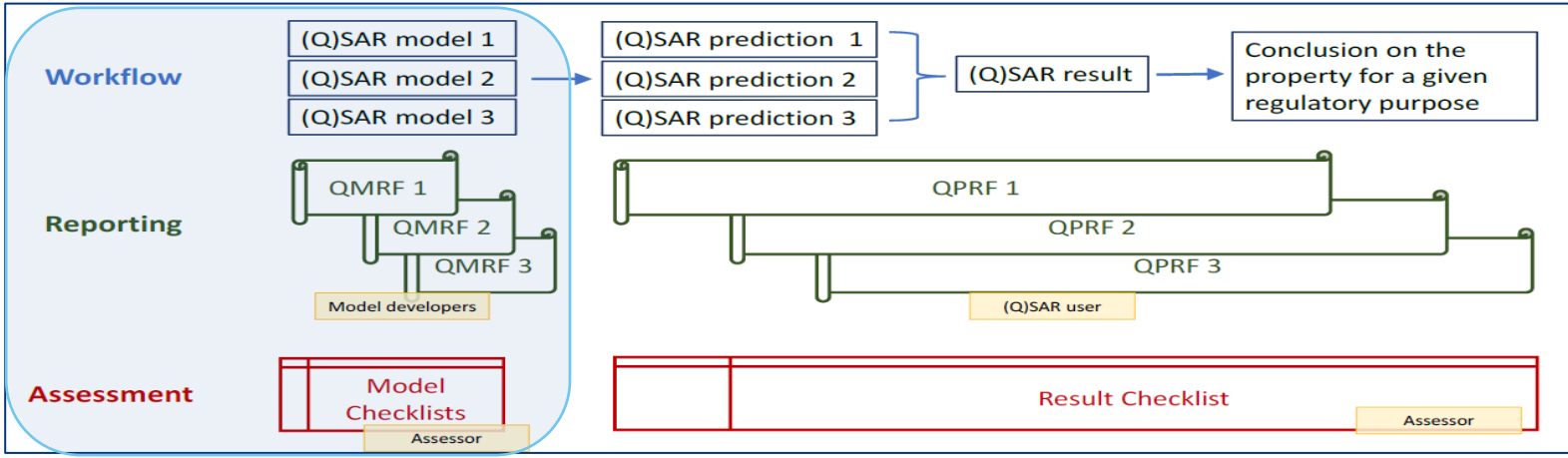
The documentation of a model predicting skin sensitisation based on structural-alerts includes an explanation on how the structural-alerts are supposed to bind to proteins causing skin sensitization The AS is fulfilled.

Practical advice

- For endpoints for which the mechanisms are known, the availability of a mechanistic interpretation facilitates the regulatory acceptance.
- Literature references that support the (purported) mechanistic basis can be used to support of the hypothesis.
- An indication whether the mechanistic basis of the model was determined a priori (i.e., before modelling, by ensuring that the initial set of training structures and/or descriptors were selected to fit a pre-defined mechanism of action) or a posteriori (i.e., after the modelling, by interpretation of the final set of training structures and/or descriptors) is also useful.

Model Checklist in the QAF workflow for assessing predictions and results based on multiple predictions





Final remarks on the (Q)SAR model checklist

- The assessment of a model is specific for the regulatory purpose
  - It should be repeated when assessing the use of same model for a different purpose
  - If the regulatory purpose is the same, assessors do not need to repeat the evaluation of the model for each prediction

- The model checklist can be used to verify that a QMRF contains all necessary information
  - Models developers could use it when preparing the model documentation

- Assessment of individual predictions may not be feasible when running prediction of a large number of substances, e.g., for screening of databases
  - In this case, assessors may need to rely solely on the assessment of the model/model checklist

Conclusions

- The compilation of the Model Checklist is the first step in the assessment of predictions and results from multiple predictions. When a model is considered not acceptable, then the assessment could be concluded without further considering predictions and results.
- Our expectation is that the application of the QAF for model assessment will improve the clarity and transparency of the models' evaluations.
- The evaluation (even partial) of each principle will guide the assessors in assessment of the model regarding its suitability for the specific regulatory purpose.
- Furthermore, completion of the model checklist serves to verify that the documentation accompanying the model contains all the information necessary to carry out the assessment of each prediction
- The Model Checklist can be used as a standalone tool when e.g., (Q)SARs are used for screening databases without the possibility to assess predictions individually, or to keep a separate record for the assessment of a model that could be reused in future
- For models' developers: best practices for model documentation to facilitated regulatory acceptance. A mapping between the Model Checklist and the QMRF can serve as feedback to model developers for further improvement of their models and related documentation.

Thank you very much!

Coordination group of the project from ISS

Cecilia Bossa cecilia.bossa@iss.it
Chiara Laura Battistelli chiara.battistelli@iss.it
Olga Tcheremenskaia olga.tcheremenskaia@iss.it

ISTITUTO SUPERIORE DI SANITÀ

DEPARTMENT
ENVIRONMENT AND HEALTH

ECHA co-lead: Andrea Gissi
OECD Secretariat: Patience Browne, Tomoko Aoyagi
QAF expert group