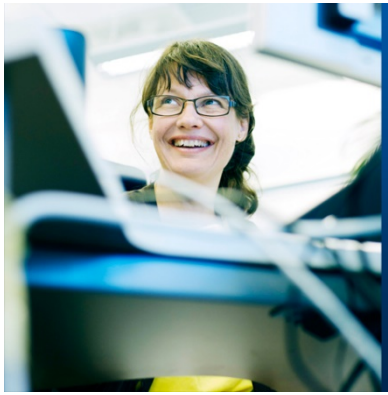

Webinar:
Questions and answers

ECHA organised a webinar on 3 June 2021 on **QSARs and their assessment under dossier evaluation.** It explained the scope and main requirements for the use of QSAR results as adaptations to standard information in REACH registrations. It also showed how ECHA evaluates the compliance of QSAR information.

This document compiles the questions and answers from the webinar. Minor editorial changes have been made to correct spelling mistakes and similar questions have been combined into one. The document will not be updated.

For the most up-to-date advice on the use of QSAR, contact us or refer to our support material.

Question	Answer
Does ECHA accept only predictions that fulfil all the requirements presented in the webinar? And what is the level of uncertainty related to the QSAR prediction that ECHA could still accept?	A QSAR result used as an adaptation to REACH standard information requirements has to meet the conditions listed in the REACH regulation, i.e. Annex XI 1.3. Ideally, this is the case when all points covered today are addressed. Any other case has to be duly justified and its acceptance is case dependent.
Do all fragments need to be present in the training set to have the prediction acceptable, or can we consider a lower reliability (with weight of evidence discussion) if one fragment of the molecule is not (well) represented in the training set?	Normally yes. However, in exceptional cases and with proper justification, the prediction can be accepted even if not fully within the fragment domain of the model.

For poorly soluble, high log Kow substances (e.g. log > 8) measurements for log Kow or log Koc are not possible anymore. The substances are out of AD. How would you recommend to proceed the presentation of QSARs in such a case? Would a WoE approach using at least two models be acceptable?	If there are no experimental data for these solubility and logKow ranges, also the QSAR models cannot make reliable prediction in these ranges, since the QSAR predictions in these cases are extrapolations. Especially when such data are used as input to predict other endpoints, the uncertainty is propagated. What can be indicated by QSAR predictions is that the logKow is very high, and solubility is very low. If it is indeed impossible to derive an experimental value, QSAR predictions (with the given uncertainty) may be an option to fill the data gap.
If a QSAR prediction meets all quality criteria, is it equally acceptable if it shows a concern (e.g., toxicity) as if it shows lack of concern?	Generally, positive and negative QSAR predictions are both acceptable under REACH. However, the acceptance does not only depend on the quality of the model/prediction, but also on the relevance for the endpoint and the other conditions presented today. For negative predictions, ECHA also considers the difference between lack of alerts and proof of lack of toxicity (i.e. the absence of alerts for an endpoint on its own is not sufficient to prove the lack of toxicity of a substance).
Is the response domain (i.e. value of the prediction) considered as a needed domain? In some extent, isn't it a repetition with the descriptor(s) domain?	ECHA looks at the four applicability domain layers presented in the slides: Descriptor domain, Structural domain, Mechanistic domain and Metabolic domain. Considerations on the response domain might be taken into account when considering the reliability and therefore adequacy of the prediction .
How is "similarity" defined for a AD check? Is there guidance on what is similar?	Similarity is based on chemical and mechanistic metrics. It is also endpoint dependent and case specific. At the moment, there is no official guidance on this. ECHA is part of an OECD working group to establish a QSAR assessment framework that, among others, could address the harmonisation of the definition of "similarity".
How is the ratio of training set number/number of descriptors >=5 derived? Does that apply only for descriptor domain or structural as well?	To calculate the Topliss ratio, you need to divide the number of substances included in the training set of the model and the number of descriptors used by the model. As indicated in the OECD guidance on QSAR: (https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=env/jm/mono(2007)2), the ratio must be equal or above 5 to consider the model scientifically valid. The Topliss ratio is not associated to the applicability domain of the model.
Can QSARs ever be used reliably enough for ECHA for UVCBs when the substance composition is by definition at least partially unknown and variable? How can validity of the QSAR be checked/confirmed?	For some type of UVCBs, it is possible to identify constituents considered to be representative for the substance. In such cases, QSAR predictions can be applied for data gap filling or at least be helpful to support the assessment and a possible further testing strategy.
When evaluating "simple" UVCBs such as esterification products of aliphatic alcohols and fatty acids, is it mandatory to clearly explain how the representative constituents have been selected? Or simply referring to what reported in the IUCLID composition section is enough?	Yes, please always justify the selection of representative structures of a UVCB used for QSAR predictions.
Again on UVCB. If I'm evaluating N repr. constituents with M models: One entry with all	Preferably individual (i.e. N x M) entries with a summary of the main outcome. If the number of individual entries is high (e.g. more than 10), you can consider "grouping" the results (e.g. by constituent, thus having N IUCLID endpoint

results and a "expert-based" evaluation of the AD compliance of the substance as a whole (based on the single results) OR NxM separate entries, with reliability evaluation of each single result?	study records (ESR), where each ESR includes the results for all models for that constituent)
Do I have to provide information on the validity of the model /QMRF for such well known models like those from the EPIsuite?	Not necessarily, if the documentation is publicly available, ECHA will try to access it. Since models (and related QMRF) are updated over time, it is still a good practice to attach the version of the QMRF of the exact version of the model used. Be aware that even well-known models may not be considered scientifically valid for the purpose of fulfilling REACH information requirements. (see examples in the presentation)
Does ECHA plan to publish a list of models for which the QMRFs are already available to ECHA? This would be of great help to the dossier submitters as we wouldn't have to submit identical documentation each time a model is used and would also simplify ECHA's work as there would be less documentation to process.	ECHA does not have a repository of QMRFs. Models also do change over time, hence ECHA prefers that the QMRFs are included in the registration dossiers. You can find QMRFs of a number of models in the QSAR Toolbox, or in the JRC QSAR model database: https://ec.europa.eu/jrc/en/scientific-tool/jrc-qsar-model-database. More recent software usually provide this QMRFs for their models. Please note that in addition to the QMRF, it is important also to provide documentation on the prediction itself (QPRF)
Are we moving to a future where a QSAR model could be OECD approved and be given a standard "OECD-approved model report" (in a similar way that assays are given OECD Test Guidelines), then perhaps REACH registrants would only need to submit a QSAR Prediction Report?	Currently there is no process to validate QSARs at OECD level. However, some QSARs are becoming part of OECD Guidelines, such as the Guideline for Defined Approaches for Skin Sensitisation, which combines two *in vitro* and one *in silico* method. In a way, the use of *in silico* methods as specified by the OECD Guideline (e.g. in combination with other information sources) for the particular purpose addressed by the Guideline can be seen as "approved".
Do we need a harmonised definition of an AD, or specific guidance, on how to define an AD? Will this help re development of new QSAR tools, or is it better to retain flexibility?	Yes, a harmonised definition of applicability domain would be beneficial for regulatory purposes. ECHA is part of an OECD working group that has just started to work on it. Please check OECD website for any public information available.
If a QSAR is used to fulfil a standard information requirement instead of an experimental study has then also a waiver to be provided that the (experimental) study is scientifically not necessary?	QSARs can be provided as adaptation to the standard information requirement according to REACH Annex XI, 1.3 (as a key study) or as part of a weight of evidence (REACH Annex XI, 1.2) . No additional waiver has then to be provided in the registration dossier.
Is the OECD Toolbox a QSAR Model? It is possible to generate a QMRPF using the OECD Toolbox?	The OECD QSAR Toolbox is not a model as such, but it is a platform that includes many functionalities. One functionality is to run external QSAR models implemented in the Toolbox or connected to it. You can also build your own local QSAR. For these models a QMRF can be accessed or generated. Note that we plan a Webinar on the QSAR Toolbox by the end of the year.

Is a trend analysis in the OECD toolbox a QSAR or a read across?	The Toolbox has this functionality as a way to fill data gaps. REACH registrants can decide if they want to justify this data gap filling as a QSAR or as Read-across (category approach), also taking into account the number of analogues.