Welcome

Webinar: New developments and regulatory applications of the OECD QSAR Toolbox

28 April 2022

Andrea Gissi
Scientific Officer
European Chemicals Agency

What you can expect today

- Recent and future developments of the QSAR Toolbox, including work supported by U.S. EPA
- Functionalities of the software with practical demonstrations
- How to use the QSAR Toolbox for regulatory applications, focussing on read-across for REACH

Live Q&A

- Join Q&A at: **slido.com**
  Event code: **#qsartoolbox22**
- Or with the QR code
- Send questions between
  14:00 - 16:30 (EEST, GMT +3)
- Only questions within scope
- Question not answered?
  Contact us: echa.europa.eu/contact

With you today

Andrea Gissi
Project Manager
QSAR Toolbox

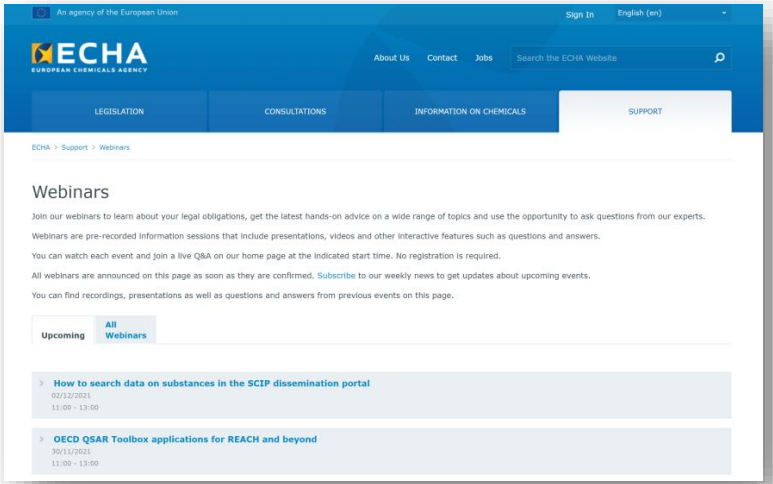
Doris Hirmann
Co-chair
PBT Expert Group

Andrea Richarz
Scientific Officer
Read-across

Tomasz Sobanski
Regulatory Officer
New approach methodologies

Material available

- Video recording
- Presentations
- Q&A (soon after the event)
- Supplementary video



echa.europa.eu/webinars

Programme

Time	Topic	Speaker
14:00	Welcome	Andrea Gissi (ECHA)
14:05	More than QSAR predictions	Patience Browne (OECD) Doris Hirmann (ECHA)
14:45	What is new in Toolbox v4.5?	Darina Yordanova (LMC) Stanislav Temelkov (LMC)
15:15	QSAR Toolbox at the U.S. EPA	Martin B. Philips, Brianne Raccor (US EPA)
15:40	Link between QSAR Toolbox and IUCLID	Andrea Gissi (ECHA)
16:00	QSAR Toolbox in support of read-across for REACH	Andrea Richarz (ECHA)
16:20	Live Q&A panel	ECHA experts
16:45	Conclusions	Tiago Pedrosa (ECHA)

Resources

- QSAR Toolbox website: software download, tutorials, manuals
- ECHA webinar: OECD QSAR Toolbox applications for REACH and beyond
- QSAR Toolbox on the OECD website

More than a QSAR tool

Webinar: New developments and practical regulatory applications of the QSAR Toolbox

28 April 2022

Doris Hirmann
Senior scientific officer
European Chemicals Agency

Different interfaces in QSAR Toolbox

- **Classical user interface**: Main interface. Includes all functionalities but needs training to use. Only runs on Windows
- **Simplified user interface**: Easy to use but only includes simple functionalities. Gradually be replaced by web client. Only runs on Windows
- **Web client**: Latest interface. Easy to use, more functionalities introduced at each release. Runs on all operating systems*

**server must be installed on a Windows machine*

Demo

- Toolbox desktop client
  - Simplified user interface (brief introduction, full demo available in supplementary video)
  - Classic user interface, from simpler functionalities to queries and searching for analogues (similar in structure and mechanism) taking into account metabolism
- Toolbox web client (demo later)

Thank you!

doris.hirmann@echa.europa.eu

Link between QSAR Toolbox and IUCLID

Webinar: New developments and practical regulatory applications of the QSAR Toolbox

28 April 2022

Andrea Gissi
Scientific officer
European Chemicals Agency

QSAR Toolbox and IUCLID

Database for chemicals and their properties implementing OECD harmonised templates

Structure and endpoint data
(Import functionality in Toolbox)

Search results
(Provided by Toolbox as links to IUCLID documents)

Software application to analyse and predict data on chemicals

QSAR TOOLBOX

Toolbox workflows

Knowledge, predictions

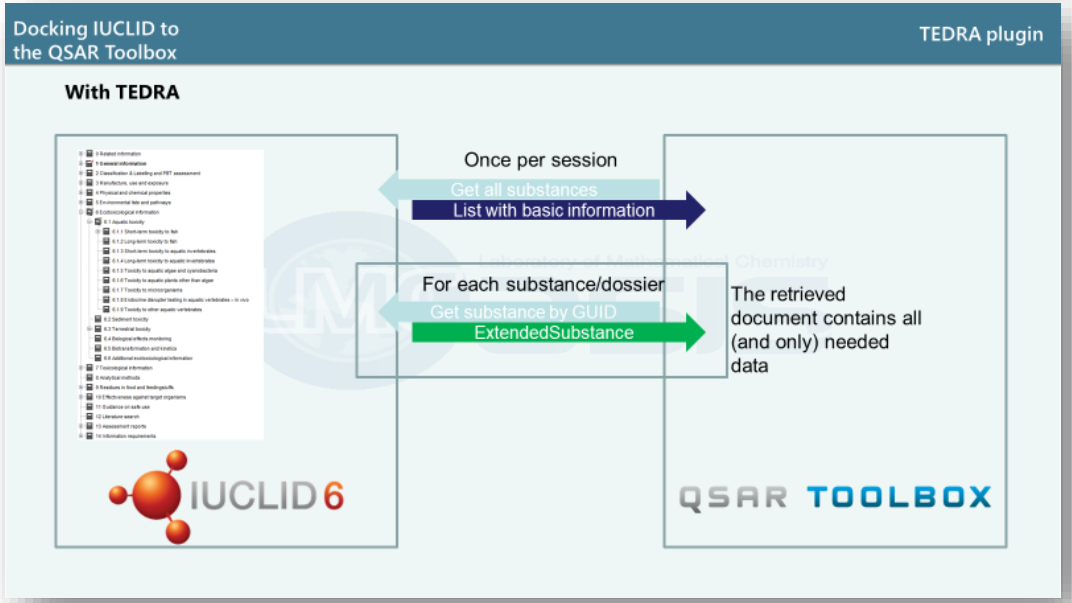
Endpoint study records
(manually compiled for use under REACH)

Download and instructions

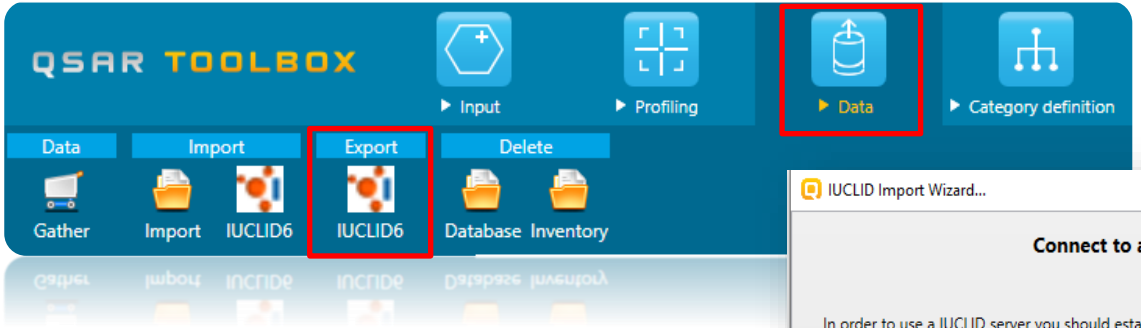
qsartoolbox.org/download/#tedra

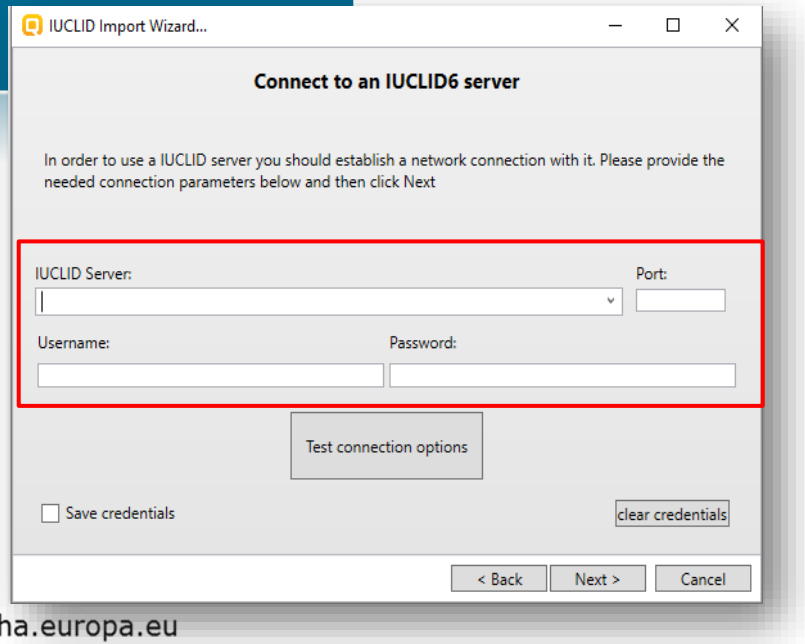
Toolbox plug-in for IUCLID to import data:

- TEDRA - **T**oolbox **E**xtended **D**ata **R**etrieval and **A**ggregation tool
- File downloaded and copied in the IUCLID server installation folder



Connect to a IUCLID database

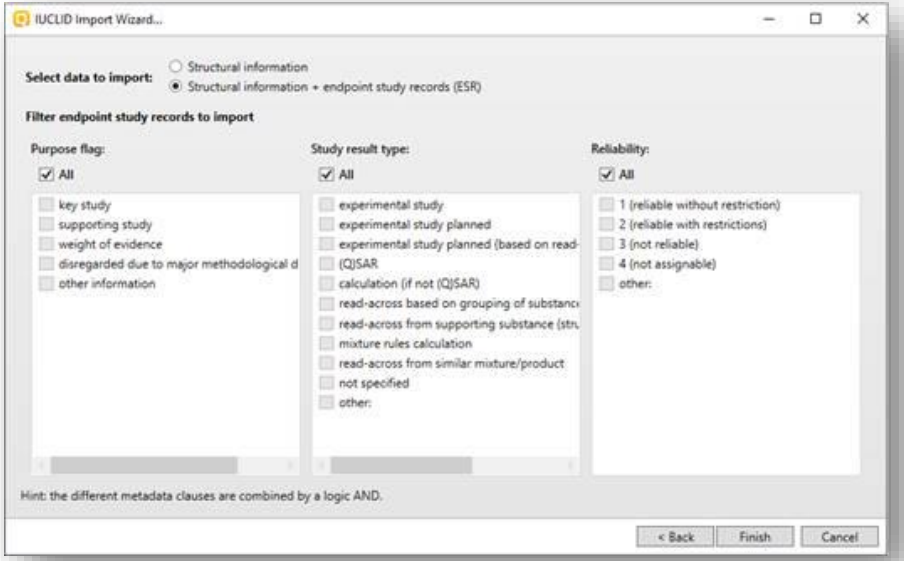




- From the data module in Toolbox
- Use your IUCLID credentials

Select endpoint study data to import

- Import structural information only or also endpoint study records (ESRs)
- ESRs can be filtered to limit import to information relevant for Toolbox (e.g., only experimental data)
- After first import, incremental updates possible to optimise time



ECHA REACH examples

- Examples and demonstration refer to ECHA REACH database
- ECHA REACH: only IUCLID data pre-imported in Toolbox
- Same functionalities available for user imported IUCLID data (except link to ECHA public chemicals database)

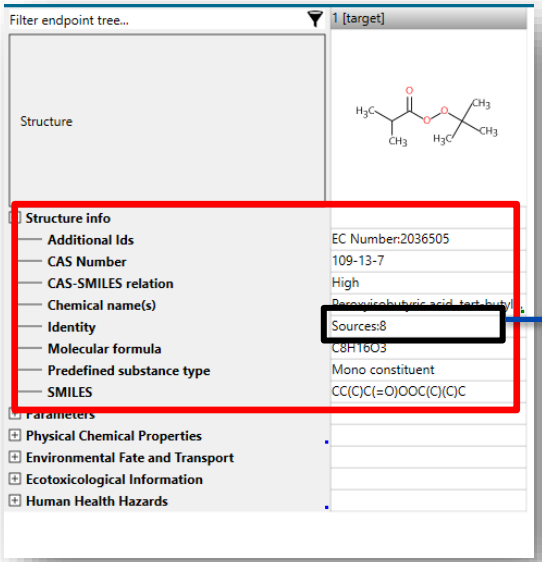
About ECHA REACH…

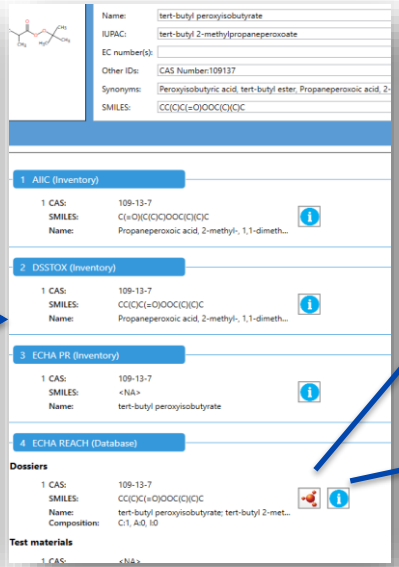
- Includes published data from REACH registrations
- QSAR Toolbox only imports structural and endpoint data. Structural information comes from IUCLID Section 1.1 (Identification), Section 1.2 (Composition) and from test materials reported in each endpoint study record of REACH dossiers
- **Compositions** include constituents, impurities and additives. Could be multiple compositions for each registered substance
- **Impurities** published only if relevant for classification
- **Test materials** in endpoint study records can be different from registered substance

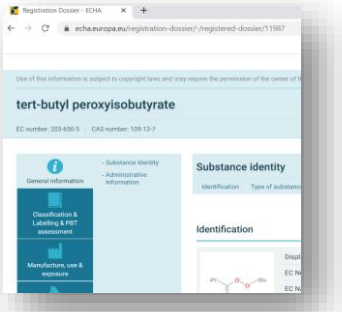
IUCLID substance identity

Expanded chemical structure

Sources (i.e. databases and/or inventories) where chemical is present)

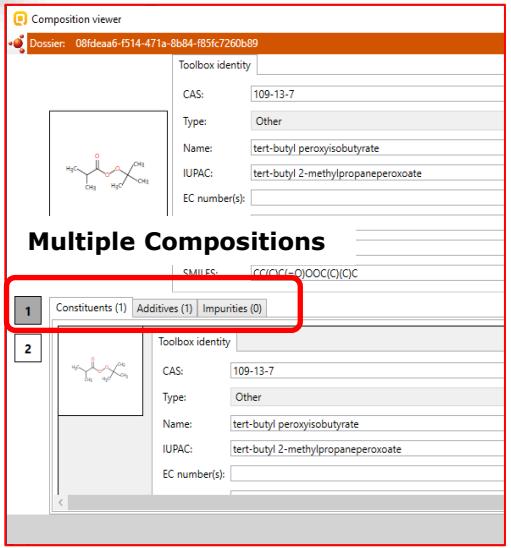




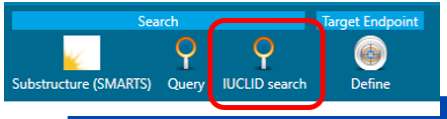


Link to ECHA public chemicals database

Composition details



Toolbox: IUCLID searches



Import IUCLID data in Toolbox

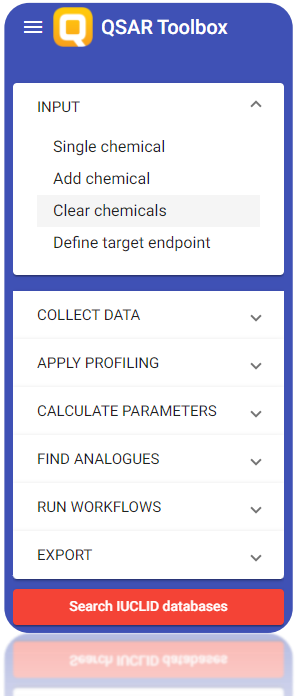
- Using Toolbox plug-in for IUCLID
- Not needed for ECHA REACH data, that is pre-imported in Toolbox

Search structure with Toolbox

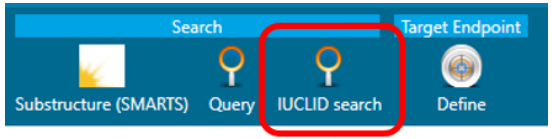
- Using IUCLID search functionality in Toolbox

Visualise search results in IUCLID

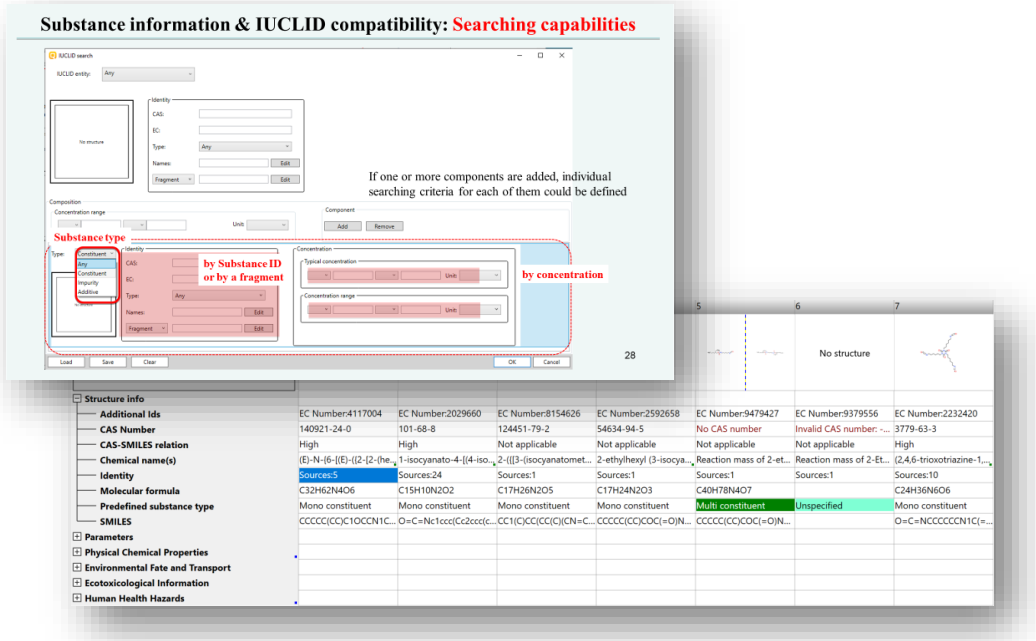
- Search results include link to IUCLID documents
- For ECHA REACH, IUCLID links are replaced by links to public information on REACH registrations



Desktop client

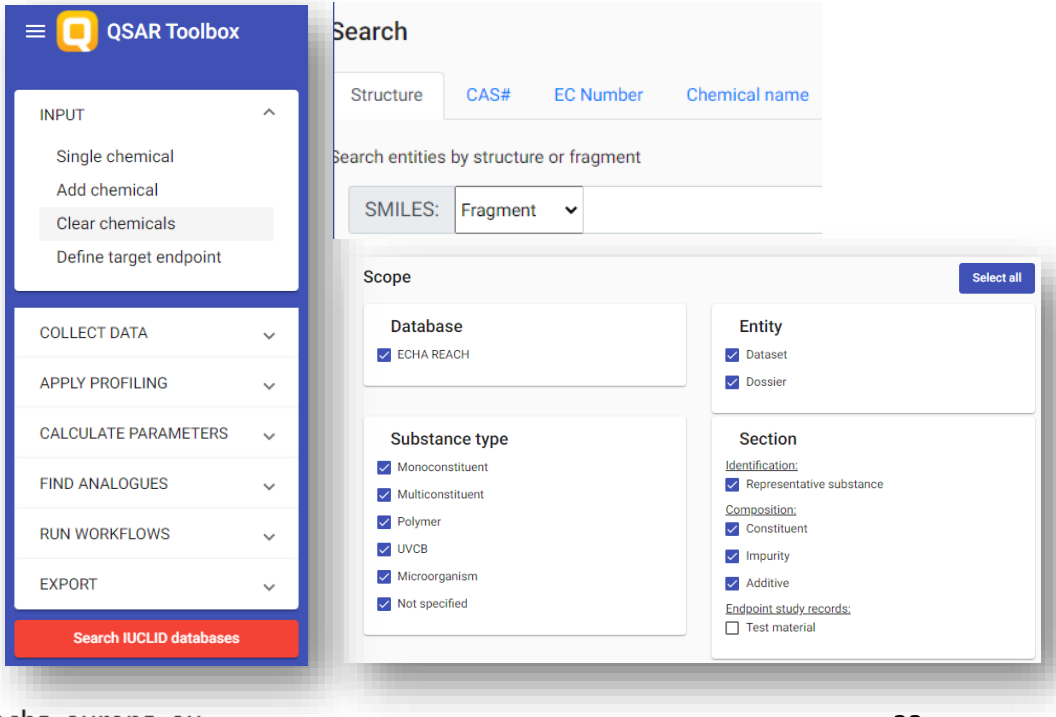


- From input module, click IUCLID search
- Desktop client search includes advanced functionalities, e.g. definition of concentration ranges and more than one criterion at once



Web client

- Select “Search IUCLID database” in web client
- Search for substance using SMILES, CAS#, EC#, or name
- Search scope can be defined:
  - Database
  - Entity (dossier or substance)
  - Substance type
  - Section
- Search results include link to IUCLID documents



Thank you!

andrea.gissi@echa.europa.eu

QSAR Toolbox: supporting read-across for REACH

Webinar: New developments and practical regulatory applications of the QSAR Toolbox

28 April 2022

Andrea Richarz
Scientific officer
European Chemicals Agency

Content

- Uses of the Toolbox in REACH submissions
- Read-across under REACH: requirements
- QSAR Toolbox to support read-across
  - Limitations and must haves
- Conclusions

Uses of the Toolbox in REACH submissions

Uses of the QSAR Toolbox

Read-across

- To identify analogues and source data ✓
- To support read-across hypothesis ✓
- Stand-alone predictions using trend analysis and read-across results without additional justification ✗

QSAR

- Predictions from external QSAR models ✓
- Use profiler results as QSAR predictions ✗

Defined approaches

- Results for DASS ITS v2 OECD Guideline ✓

Predictions from external QSAR models

- QSAR Toolbox includes external QSAR models e.g. EPISuite models
  - Results generated using these models are assessed against criteria for acceptable (Q)SAR results: REACH Annex XI, Section 1.3

More

- Webinar: QSARs and their assessment under dossier evaluation
- Practical guide: How to use and report (Q)SARs

Profiler results used as QSAR predictions

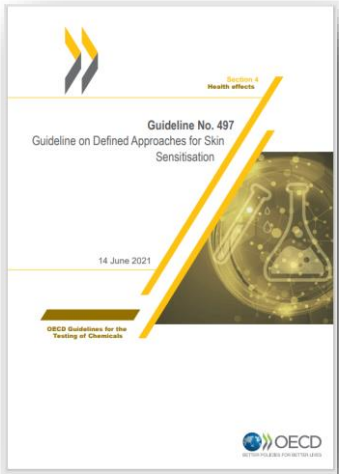
✗ Do not use profiler results (structural alerts) as (Q)SAR predictions to adapt REACH standard information requirements

- Profilers should be used for identification of analogues
- May not fulfil criteria for valid (Q)SAR models such as defined applicability domain, or appropriate measures of performances

e.g. lack of mutagenicity alerts cannot be used to conclude lack of mutagenicity potential of a substance; but can contribute to find suitable analogues with data to build read-across case

Defined approaches for skin sensitisation: Automated workflow

- Defined approaches for skin sensitisation (DASS) OECD Guideline No. 497: integrated testing strategy (ITS) combines in vitro with in silico results
- ITS v2 uses QSAR Toolbox results: skin sensitisation hazard prediction with automated workflow
  - Includes identification of analogues with skin sensitisation data
- Within applicability domain, DASS have the same weight as a murine local lymph node assay LLNA



OECD Guideline No. 497

Toolbox supporting read-across

QSAR Toolbox can be used to support read-across:

- **Identification of source substances**
  Toolbox one of the best tools for identification of analogues based on structural and mechanistic similarity (including metabolism)
- **Identification of source studies**
  Finding source substances with available experimental data

Toolbox supporting read-across

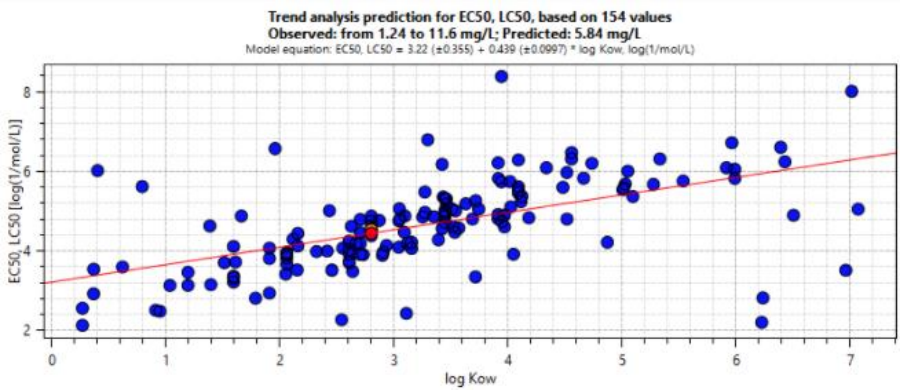
- **Support read-across hypothesis**
  Toolbox can be used to give supporting information for read-across hypothesis.
  For example, to prove similarity of sources and target
  - e.g. structure and mechanistic information

❌ Only basing similarity on use of profiler results not sufficient

Trend analysis and read-across

Data gap filling module predicts properties from analogues by:

- **Trend analysis**
  Endpoint correlates with a physicochemical property and follows a trend
- **Read-across**



Read-across predictions with the Toolbox

- ✗ Results from QSAR Toolbox often flagged as (Q)SARs in REACH registrations
  - If they rely on data from analogues: it is read-across
  - Then: results evaluated according to requirements for grouping and read-across adaptations
- ✗ Limitations in available information on experimental details and kinetic data (e.g. metabolic rates and quantities)
  - Toolbox read-across results not compliant without additional information

Read-across under REACH: requirements

Information requirements under REACH

- REACH standard information Annexes are tiered according to annual tonnage bands:1-10, 10-100, 100-1000 and >1000 tpa
  - **standard information requirements** refer to relevant test guidelines
- Standard information requirements can be **adapted**:
  - according to specific rules listed in Column 2 of respective information requirement
  - according to general rules listed in **Annex XI**

REACH Annex XI, Section 1.5
Grouping and read-across

Grouping and read-across: one option to adapt standard information requirements

- **Group or category:** substances whose physicochemical and (eco)toxicological properties likely to be similar or follow regular pattern as result of structural similarity
- **Read-across:** properties or effects of substances in the group may be predicted from data for reference substance(s) within the group
- **Similarities** may be based on:
  - Common functional group
  - Common precursors and/or breakdown products resulting in structurally similar chemicals
  - Constant pattern in changing of potency

REACH Annex XI, Section 1.5
Grouping and read-across

Recent update of REACH Annexes: clarification of information to be submitted

- Adequacy for the purpose of classification and labelling and/or risk assessment
- Adequate and reliable coverage of the key parameters of the
- Exposure duration comparable to/longer than

} corresponding study that is normally done for a particular information requirement

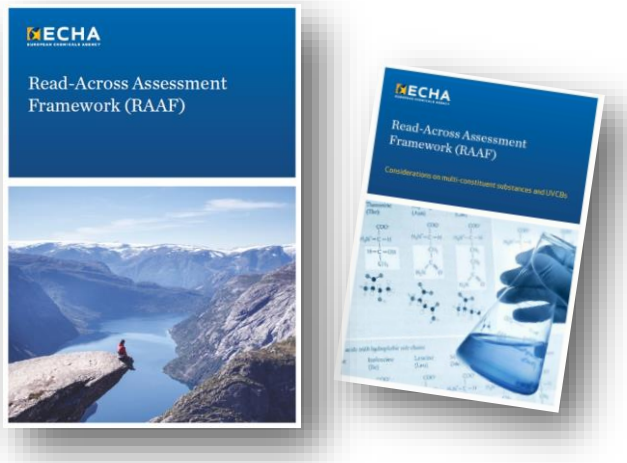
- Adequate and reliable documentation:
  - robust study summary for each source study
  - explanation why properties of registered substance may be predicted from other substances
  - supporting information to scientifically justify

Read-across under REACH

- Under REACH, a **substance** is defined by a set of chemical structures:
  - Constituents, impurities and/or additives
- **Structural similarity alone not sufficient** to justify prediction
  - Read-across hypothesis needed - establishes why a prediction is possible
- Due to different complexities (e.g. key parameters, biological targets) of each endpoint, read-across must be **specific to considered endpoint**
- Not the conclusion of a study (e.g. potential to be mutagenic), but the **results of a study** (e.g. results of *in vitro* mammalian gene mutation test) are **read across**

Read-across assessment framework (RAAF)

- Structures expert judgement of complex scientific questions on critical aspects of read-across
- Leading to improved, consistent and transparent assessment of read-across submissions
- Primarily designed for ECHA evaluators
- Registrants can use it to improve read-across adaptations



RAAF on ECHA website

RAAF PDF

RAAF assessment elements

- Assessment Elements (AE) describe scientific considerations deemed crucial to judge validity and reliability of read-across
  - Common AE's for analogue or category scenarios
  - Scenario-specific AE's
  - Both for human health and environmental/ fate effects

- Guidance for structured analysis of read-across submissions and justifications
  - Identifies strengths and weaknesses of read-across approach and justification
  - Characterises confidence of assessor in proposed read-across prediction

Toolbox to support read-across

Practical tips, limitations and best practice

Toolbox information supporting RAAF assessment elements

Common RAAF assessment element	Toolbox information
Characterisation of source and target	• CAS#, EC#, structure, name, composition
Category description and supporting information	• Category boundaries can be supported by structural profilers used for analogue selection • Category justification can be supported by explanation of profiling results
Link with structural similarity	• Structural profiler results
Impact of impurities	• Predicted mechanism, properties and experimental data for impurities
Consistency of properties in data matrix	• Experimental (or predicted) data among different properties relevant for prediction
Source data quality	• Metadata and original reference for experimental data
Bias	• Prediction report documenting all steps and manually removed data points

Toolbox information supporting RAAF assessment elements

Specific RAAF Assessment Element	Toolbox information
Formation of common and non-common compounds	• Observed and predicted transformation products (kinetics not included*)
Degradation	• Degradation related profilers, data and QSAR models
Bioaccumulation potential	• Bioaccumulation related profilers, data and QSAR models
Impact of common and non-common compounds	• Profilers, data and QSAR model results for common and non-common compounds
Common underlying mechanism - qualitative aspects	• Results from mechanistic and endpoint specific profilers related to the target endpoint
Common underlying mechanism - quantitative aspects	• Predicted and experimental data for the target endpoint

* experimental data must be provided (e.g. hydrolysis, toxicokinetics studies)

How to report read-across in IUCLID

- Follow instructions in Chapter 8.6.3 of our manual: *How to prepare registration and PPORD dossiers*
- **Source data:** one endpoint study record for each source substance
  - Complete as any other endpoint study record that documents an experimental study
- **Target data** endpoint study record
  - Endpoint: same as in source study/ies
  - Study type: read-across
  - Justification for type of information: read-across justification
  - Test material information: composition of read-across target material
  - Cross-reference: link to endpoint study record of source study/ies
  - Results and discussion: results predicted for read-across target

Video tutorials:

Analogue approach

Category approach

QSAR Toolbox
Limitations

- **Experimental data** does not include level of detail required for independent assessment of study reliability (i.e. robust study summaries)
- Toolbox can take metabolism into account but **metabolic and kinetic** information lacks details on metabolic rates and probability of reactions required for justification of hypothesis

Reasons for non-compliance

Non-compliance due to:

- **Limitations of the Toolbox** – can be avoided by retrieving additional (supporting) information from outside the QSAR Toolbox
- **Shortcomings in the use** – can be avoided with better use of QSAR Toolbox and better documentation

CAVEAT:

! Expert judgement required with critical assessment of identified analogues and relevance/reliability of the data

Common shortcomings

- Whole **composition** for both source and target substances to be addressed
- Selection of analogues not adequate
- No justification of impact of **structural differences** on prediction
- No **applicability domain** definition with relevant inclusion/exclusion criteria provided

Common shortcomings

- **No robust study summaries** for the source substances
- Under REACH whole study is read across, not only e.g. a NOAEL
  - Full study results provide key parameters, types of toxicity as basis for NOAELs

Note:

Profilers on their own, without e.g. bridging studies, do not prove toxicological similarity (endpoint-specific)

Must-have for read-across justification

- **Characterisation** of source and target substance composition
- **Hypothesis:** relationship between structural similarities and read-across prediction of a property, impact of dissimilarities on toxicokinetic/dynamic properties
- **Supporting information** to scientifically justify the hypothesis and give confidence in the prediction
- **Source study:** experimental data used as basis for the prediction; adequacy and reliability required as for standard information requirements

Conclusions

- Toolbox is valuable in supporting read-across
  - profiling of chemicals, retrieving experimental data and simulating metabolism
- For read-across under REACH, Toolbox can be used:
  - to identify analogues and source studies
  - to provide supporting information for read-across justification

  Note: not all necessary information can be covered with Toolbox alone
- Toolbox read-across is evaluated as any other read-across
  - adaptation of endpoint-specific information requirements according to REACH Annex XI
- Expert judgement and critical assessment required (analogues, data)
- Adequate justification, documentation, supporting information is key

Thank you!

andrea.richarz@echa.europa.eu

Final remarks

Webinar: New developments and practical regulatory applications of the QSAR Toolbox

28 April 2022

Tiago Pedrosa
Head of Unit - Computational Assessment
European Chemicals Agency

Looking forward

Next release of QSAR Toolbox (2023):

- Extension of IUCLID searches to endpoint study records
- Simplified selection and application of (Q)SAR models
- Updated reports

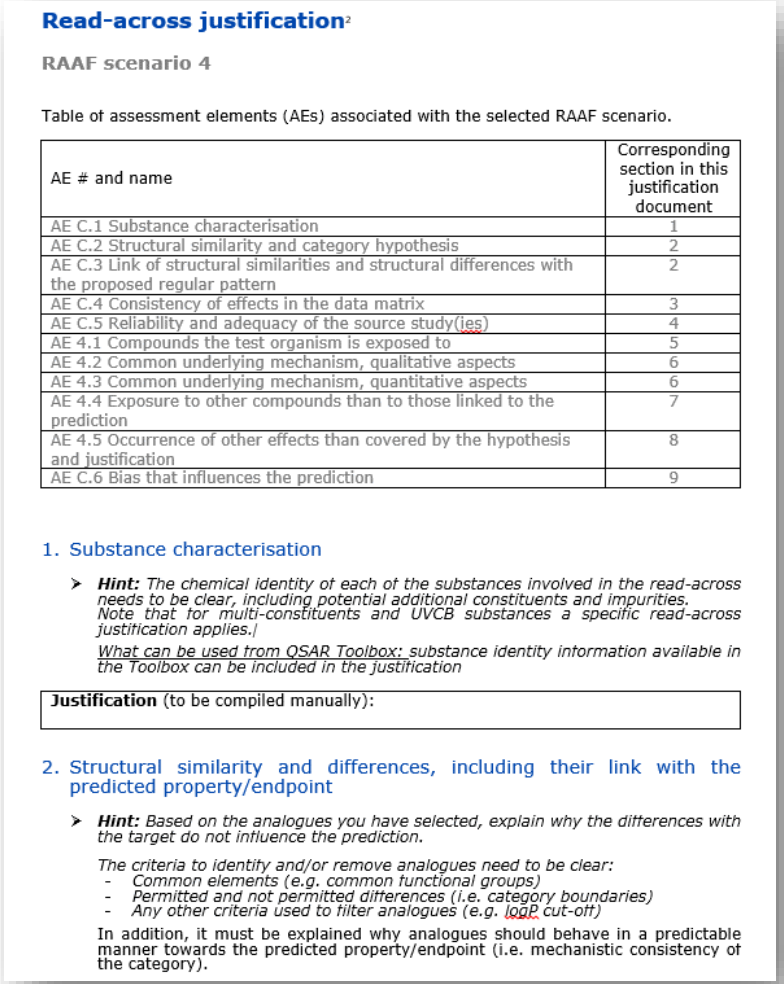
Updated reports

To increase regulatory acceptance:

- Link to **regulatory requirements**
- Improved **reproducibility**
- Increased **transparency**

Link to regulatory requirements

- New read-across and QSAR **justification documents**
- Completed manually, hints provided



Read-across justification[2]

RAAF scenario 4

Table of assessment elements (AEs) associated with the selected RAAF scenario.

AE # and name	Corresponding section in this justification document
AE C.1 Substance characterisation	1
AE C.2 Structural similarity and category hypothesis	2
AE C.3 Link of structural similarities and structural differences with the proposed regular pattern	2
AE C.4 Consistency of effects in the data matrix	3
AE C.5 Reliability and adequacy of the source study(ies)	4
AE 4.1 Compounds the test organism is exposed to	5
AE 4.2 Common underlying mechanism, qualitative aspects	6
AE 4.3 Common underlying mechanism, quantitative aspects	6
AE 4.4 Exposure to other compounds than to those linked to the prediction	7
AE 4.5 Occurrence of other effects than covered by the hypothesis and justification	8
AE C.6 Bias that influences the prediction	9

1. Substance characterisation

➢ ***Hint:*** *The chemical identity of each of the substances involved in the read-across needs to be clear, including potential additional constituents and impurities. Note that for multi-constituents and UVCB substances a specific read-across justification applies.*

What can be used from QSAR Toolbox: substance identity information available in the Toolbox can be included in the justification

Justification (to be compiled manually):

2. Structural similarity and differences, including their link with the predicted property/endpoint

➢ ***Hint:*** *Based on the analogues you have selected, explain why the differences with the target do not influence the prediction.*

The criteria to identity and/or remove analogues need to be clear:
- *Common elements (e.g. common functional groups)*
- *Permitted and not permitted differences (i.e. category boundaries)*
- *Any other criteria used to filter analogues (e.g. logP cut-off)*

In addition, it must be explained why analogues should behave in a predictable manner towards the predicted property/endpoint (i.e. mechanistic consistency of the category).

New report sections

Reproducibility

Prediction protocol

Input: CAS: 56-18-8

Database(s) used:
- Skin Sensitization

Selected endpoint: Human Health Hazards -> Sensitisation -> Skin -> in Vivo -> LLNA -> EC3

Categorisation:

Primary categorisation

Profiler: Protein binding alerts for skin sensitization by OASIS combined with Skin metabolism simulator (not strict)

Target: No alert found (parent) AND Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes (metabolites) AND Schiff base formation >> Schiff base formation with carbonyl compounds >> Bis aldehydes (metabolites)

Category: 10 analogues with 17 experimental data

Sub-categorization steps

- Step 1:

Profiler: Protein binding alerts for skin sensitization by OASIS combined with Autoxidation simulator

Target and metabolites: No alert found

Selection: Substances different from target are removed

Sub-category: 5 analogues with 13 experimental data

- Step 2:

Profiler: Protein binding alerts for skin sensitization by OASIS combined with Skin metabolism simulator

Target and metabolites: No alert found; Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes; Schiff base formation >> Schiff base formation with carbonyl compounds >> Bis aldehydes

Selection: Substances different from target are removed

Sub-category: 3 analogues with 11 experimental data

Data gap filling:

Calculation approach: takes the highest mode value from the 3 nearest neighbours, Active descriptor: log Kow (calculated), Data usage: All values

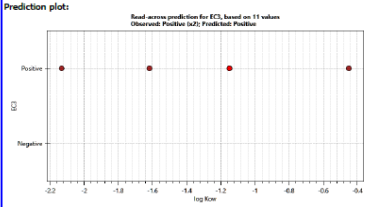
Transparency

Prediction details

Predicted value: Positive Skin sensitisation II (ECETOC)

Predicted endpoint: Human Health Hazards -> Sensitisation -> Skin -> in Vivo -> LLNA -> EC3

Prediction plot:



Values used for the prediction:

Structure	Experimental values used for the prediction (All)	log Kow
CAS: 111-40-0 SMILES: NCCNCCN Name: Diethylenetriamine	Positive [Skin sensitisation II (ECETOC)] (x4)	-2.13
CAS: 107-15-3 SMILES: NCCN Name: Ethylenediamine	Positive [Skin sensitisation II (ECETOC)] (x4)	-1.62
CAS: 109-55-7 SMILES: CN(C)CCCN Name: 3-aminopropyldimethylamine	Positive [Skin sensitisation II (ECETOC)] (x3)	-0.45

Calculation approach: takes the highest mode value from the 3 nearest neighbours

Active descriptor: log Kow (calculated)

Data usage: All values*

*When multiple values are available for the same chemical, all of them are taken individually in prediction calculations

References and explanations

Database information:
- Skin Sensitization

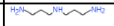
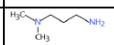
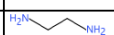
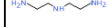
Profilers information:
- Protein binding alerts for skin sensitization by OASIS
- Protein binding alerts for skin sensitization according to GHS
- Protein binding by OASIS

Profilers result information:
- Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes (Protein binding alerts for skin sensitization by OASIS)
- Schiff base formation >> Schiff base formation with carbonyl compounds >> Bis aldehydes (Protein binding alerts for skin sensitization by OASIS)
- Skin sensitization Category 1A >> Bis Aldehydes (Protein binding alerts for skin sensitization according to GHS)

Appendix: Specific report explanations

Specific information regarding the prediction

Table with profiling results for "Organic functional groups"

CAS	Structure	Results
1 CAS# 56-18-8		Amine, primary Amine, secondary Aliphatic amine, primary Aliphatic amine, secondary
2 CAS# 109-55-7		Amine, primary Aliphatic amine, primary Amine, tertiary Aliphatic amine, tertiary
3 CAS# 107-15-3		Amine, primary Aliphatic amine, primary
4 CAS# 111-40-0		Amine, primary Amine, secondary Aliphatic amine, primary Aliphatic amine, secondary

Structural functionalities, different from the target are colored in red.

Take home messages

- We keep supporting OECD QSAR Toolbox to promote correct use of **non-animal methods** for screening and assessing chemicals
- We use our experience in assessing computational results for regulatory purposes to focus new developments on **improving regulatory acceptance**
- **Connection with IUCLID** enables the use of QSAR Toolbox functionalities for scientific operations on IUCLID structural and endpoint data
- In addition to OECD and ECHA, other **partners such as U.S. EPA** contribute to the success of the QSAR Toolbox by supporting its developments

Thank you!

tiago.pedrosa@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

Live Q&A panel

- Our panellists reply until **16:45 Helsinki time** (EEST, GMT+3)
- **All questions will be answered** in a Q&A document shortly after
- Send questions at slido.com, event code **qsartoolbox22**, or with QR code:
- Questions after the webinar? echa.europa.eu/contact

Material published

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