

OECD QSAR Toolbox v.3.1

Example for predicting Skin Sensitization of
mixture with known components

Outlook

- **Background**
- Objectives
- The exercise
- Workflow

Background

- This is a step-by-step presentation designed to take the user of the Toolbox through the workflow for prediction skin sensitization of mixture with known components

Outlook

- Background
- **Objectives**
- The exercise
- Workflow

Objectives

- **This presentation reviews a number of functionalities of the Toolbox:**
 - 2D editor for defining Mixture components
 - Filling data gaps by Independent mode approach

Outlook

- Background
- Objectives
- **The exercise**
- Workflow

Exercise

- In this exercise we will predict the skin sensitization of mixture, which is the “target” chemical.
- Investigate the mode of action for each components of the mixture
- Gather available experimental data for target chemical
- Investigate skin sensitization of non-tested component
- Applying read across for non-tested component
- Predict skin sensitization potential of mixture based on experimental data of tested compounds and predicted data of non-tested one.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**

Workflow

- **The Toolbox has six modules which are used in a sequential workflow:**
 - Chemical Input
 - Profiling
 - Endpoints
 - Category Definition
 - Filling Data Gaps
 - Report

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - **Input**

Chemical Input Overview

- This module provides the user with several means of entering the chemical of interest or the target chemical.
- Since all subsequent functions are based on chemical structure, the goal here is to make sure the molecular structure assigned to the target chemical is the correct one.

Chemical Input

Ways of Entering a mixture

User alternatives for defining mixtures with known compositions:

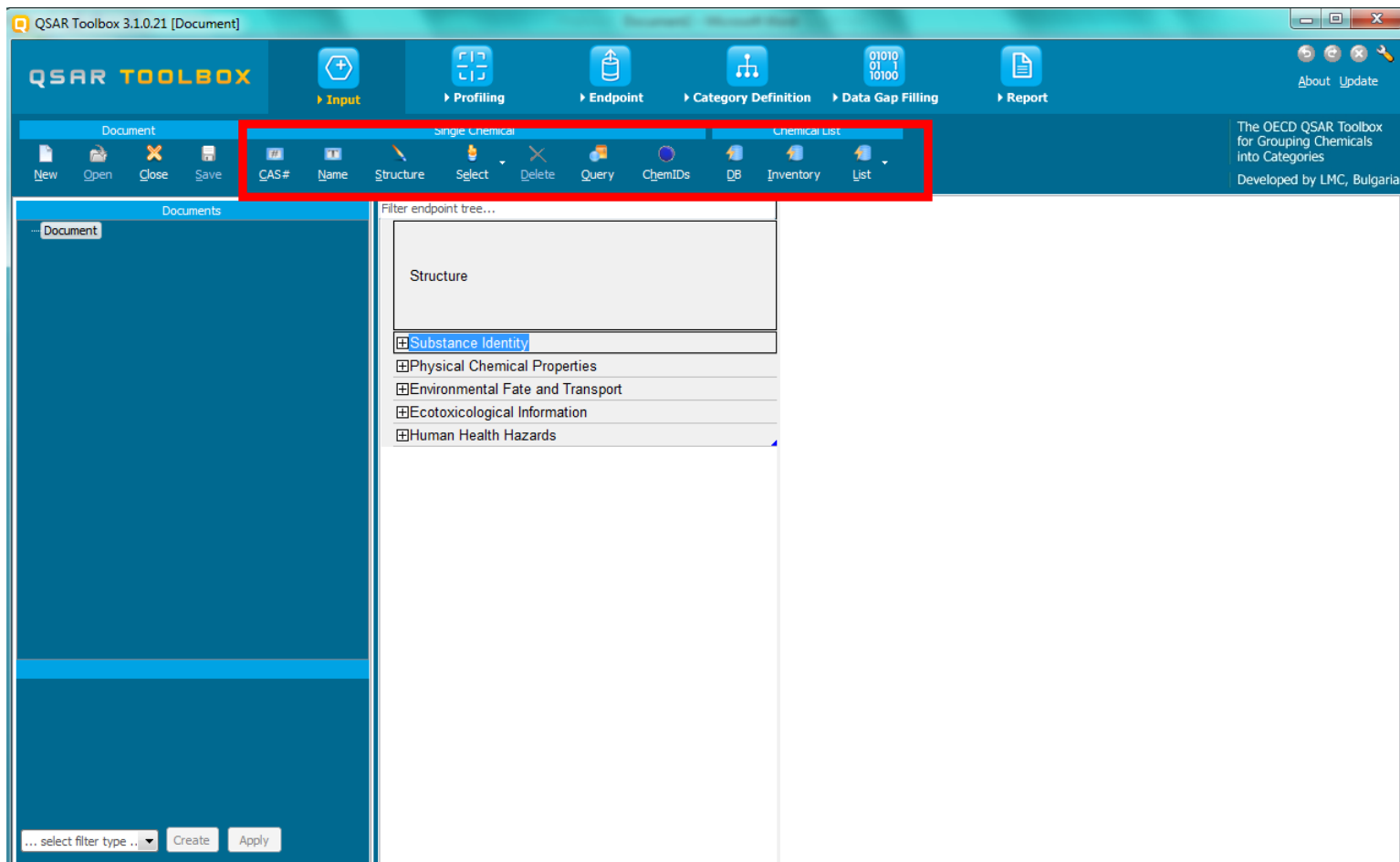
- Chemical Name
- Chemical Abstract Services (CAS) number (#)
- SMILES (simplified molecular information line entry system) notation/InChi
- Drawing mixture constituents and defining their quantities
- Select from User List/Inventory/Databases
- Chemical IDs such as EC number, EINECS number
- Load file with mixture

Getting Started

- Open the Toolbox.
- The six modules in the workflow are seen listed next to "QSAR TOOLBOX".
- **Click** on "Input" (see next screen shot)

Chemical Input Screen

Input screen



Chemical input

Load list with chemical mixture

- Toolbox allows to enter target chemicals through tab delimited file
- This requires mixture with defined components to be previously defined in a tab delimited file
- The subsequent series of screen shots will take you through the process of entering the target chemical via tab delimited file
- In this particular case, the example file with mixture is available in the Example directory of Toolbox installation (C:\Program Files (x86)\QSAR Toolbox\QSAR Toolbox 3\Examples)

Chemical input

Load list with chemical mixture

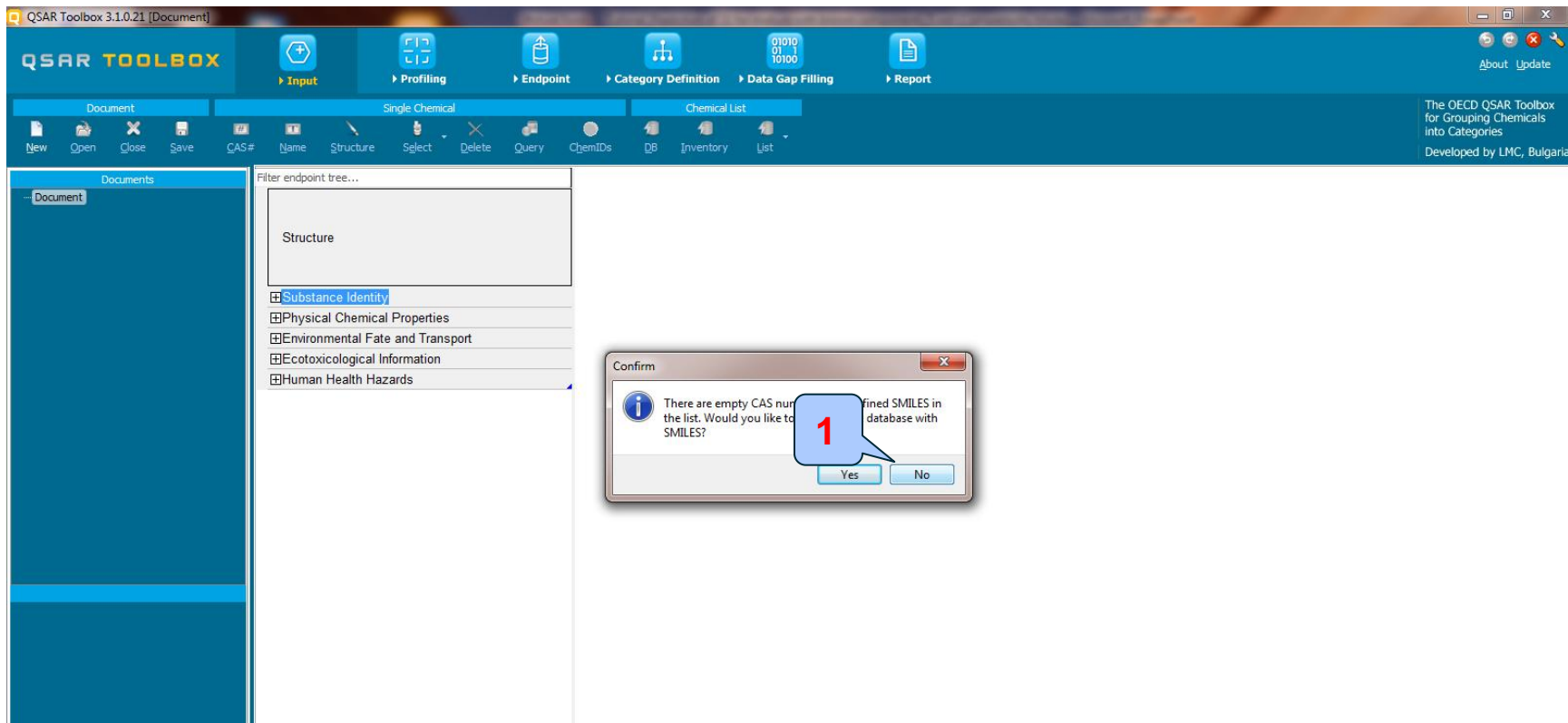
The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Input' menu is open, and the 'Chemical list' option is highlighted with a red box and a callout '1'. A file explorer window is open, showing the 'Examples' directory. The file 'Mixture with defined quantities.smi' is selected, highlighted with a red oval and a callout '3'. The 'Open' button is highlighted with a callout '4'. The file name 'Mixture with defined quantities.smi' is entered in the 'File name' field. The file explorer also shows a callout '2' pointing to the 'Examples' directory path.

Name	Date modified	Type	Size
Alkyl ethers_1.smi	5/1/2009 1:31 PM	SMI File	1 KB
CAS_Name_Inchi.inchi	1/21/2007 11:20 AM	INCHI File	7 KB
Horizontal import_Ecotox.txt	10/22/2012 3:32 PM	TXT File	62 KB
Mixture with defined quantities.smi	10/30/2012 9:18 AM	SMI File	1 KB
Mixture with mol %s.smi	12/4/2012 4:13 PM	SMI File	1 KB
mono and di-methyltins.smi	10/12/2006 2:48 PM	SMI File	1 KB
Multifunctional acrilates and methacrylat...	1/3/2008 6:15 PM	SMI File	1 KB
OECD Mock Inventory.smi	7/14/2011 1:57 PM	SMI File	3 KB
Phenols_EPA.smi	8/29/2007 7:27 AM	SMI File	14 KB
Phenols_EPA_short.smi	8/29/2007 7:27 AM	SMI File	3 KB
Primary amines.smi	1/3/2008 5:03 PM	SMI File	1 KB

1. **Click** on Chemical list
2. **Browse** and find the file with mixture located at Examples directory
3. **Select** the file
4. **Open** the file "Mixture with defined quantities.smi"

Chemical input

Load list with chemical mixture



The notification message appears, informing the user that there are structures without CAS numbers. If you want the software to search databases for their CAS numbers, click Yes, otherwise click No.

1. **Select No**

Chemical Input

Target chemical identity

- The already drawn target structure automatically appears on the data matrix
- Note that no CAS number or name is displayed for this chemical. This means the target chemical is not listed in the chemical inventories/databases available in Toolbox(see next slide).
- Visualization of components of the mixture is possible when user selects Single Component Mode

Chemical Input

Target chemical identity

1. Select "[set][Mix]{X=1/Miligrams.....}" of mixture

2. Component Mode functionality appears. All components mode is selected by default (3)

1. **Select** "[set][Mix]{X=1/Miligrams.....}" of mixture
2. Component Mode functionality appears. All components mode is selected by default (3)

Chemical Input

Target chemical identity

QSAR Toolbox 3.1.0.21 [Mixture with defined quantities.smi]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Document Single Chemical Chemical List **Component Mode**

New Open Close Save CAS# Name Structure Select Delete Query ChemIDs DB Inventory List

Documents

- Document
 - [set][Cust]Mixture with defined quantity
 - [set][Mix] (X=1 Milligrams) C(C)(=O)c1c(C)c(C)c(C)cc1
 - C(C)(=O)c1c(C)c(C)c(C)cc1
 - c1(C(=O)c2ccccc2)ccccc1
 - C(O)CCC

Filter endpoint tree...

1 [target]	2 [target, mix, component]	3 [target, mix, component]	4 [target, mix, component]
[3] [Mix]			
	Qty: 1 Milligrams	Qty: 10 Milligrams	Qty: 100 Milligrams

Substance Identity
Physical Chemical Properties
Environmental Fate and Transport
Ecotoxicological Information
Human Health Hazards

C(C)(C)cc1_(X=10/Milligrams)c1(C(=O)c2ccccc2)c

“Single” component mode visualize all components of the mixture and allows the user to work with each of the components as individual substance (1)

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - **Profiling**

Profiling Overview

- “Profiling” refers to the electronic process of retrieving relevant information on the target compound, other than environmental fate, ecotoxicity and toxicity data, which are stored in the Toolbox database.
- Available information includes likely mechanism(s) of action, as well as observed or simulated metabolites.

Profiling

Side-Bar to Profiling

- For most of the profilers, background information can be retrieved by highlighting one of the profilers (for example, Protein binding by OASIS v1.1 and clicking on “View” (see next screen shot)).

Profiling Side-Bar to Profiling

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. In the 'Profiling methods' sidebar, the 'View' button is highlighted with a red circle and a callout '2'. The 'Protein binding by OASIS v1.1' method is selected in the list, highlighted with a red circle and a callout '1'. The 'Profiling Scheme Browser' window is open, displaying a tree view of categories on the left and a detailed textual description on the right. Red boxes and arrows highlight specific items in the tree: 'Domain' (under Acylation and Schiff base formation), 'Mechanistic alert' (under Schiff base formation), and 'Structural alert' (under Aldehydes). The textual description includes the title 'Mechanistic Domain: Schiff base formation', a 'Mechanistic Alert' description, a 'Structural Alert' description, a chemical reaction diagram showing the reaction of an aldehyde with a primary amine to form an imine, and a reference to Camilla K. Smith et al. (2001).

1. **Highlight** the profiler
2. **Click** View

Profiling Side-Bar to Profiling

The screenshot displays the QSAR Toolbox 3.1.0.21 interface. The top toolbar includes buttons for Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. The 'Profiling' button is circled in red with a '2' next to it. Below the toolbar, the 'Profiling methods' list on the left has 'Aldehydes' circled in red with a '1' next to it. The main window shows the 'Protein binding by OASIS v1.1' Profiling Scheme Browser. The 'Structural boundaries' panel on the right shows a logical expression: $\text{AND}(\text{OR}(\text{Aldehydes}, \dots))$. The 'Aldehydes' term is circled in red, and an arrow points from this circle to a chemical structure of an aldehyde (C=O) labeled 'Structural fragment'. The 'Boundary Options' panel shows the fragment CC(=O)=O.

1. **Highlight** the profiler
2. **Click** View

Profiling

Profiling the target chemical

- Select the “Profiling methods” related to the target endpoint by clicking on the box next to the profilers name.
- This selects (a green check mark appears) or deselects (green check disappears) profilers.
- For this example, select the following profilers relevant to the Skin sensitization (see next screenshot):
 - Protein binding by OASIS v1.1 – general mechanistic
 - Protein binding by OECD – general mechanistic
 - Protein binding potency – general mechanistic
 - Protein binding alerts for skin sensitization by OASIS v1.1 – endpoint specific

Profiling Side-Bar to Profiling

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The top navigation bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Profiling' side-bar is active, showing a list of profiling methods. Two callouts are present: '1' points to the 'Protein binding alerts for skin sensitization by OAS' checkbox, and '2' points to the 'Apply' button. The main window displays a table of results for four different components. A red box highlights the second component, which has a positive protein binding alert.

1 [target]	2 [target,mix.component]	3 [target,mix.component]	4 [target,mix.component]
[3] [Mix] <chem>ClC1=CC=C(C=C1)C(=O)O</chem>	<chem>ClC1=CC=C(C=C1)C(=O)O</chem> Qty: 1 Milligrams	<chem>O=C1C=CC(=C1)C(=O)O</chem> Qty: 10 Milligrams	<chem>CCO</chem> Qty: 100 Milligrams
No alert found SNAr SNAr >> Nucleophi... SNAr >> Nucleophi...	SNAr SNAr >> Nucleophi... SNAr >> Nucleophi...	No alert found	No alert found
No alert found	No alert found	No alert found	No alert found
Not possible to cla...	Not possible to cla...	Not possible to cla...	Not possible to cla...
No alert found SNAr SNAr >> Nucleophi... SNAr >> Nucleophi...	SNAr SNAr >> Nucleophi... SNAr >> Nucleophi...	No alert found	No alert found

1. **Check** the profilers related to the target endpoint;
2. **Click** Apply.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - **Endpoint**

Endpoint

- “Endpoint” refer to the electronic process of retrieving the environmental fate, ecotoxicity and toxicity data that are stored in the Toolbox database.
- Data gathering can be executed in a global fashion (i.e., collecting all data of all endpoints) or on a more narrowly defined basis (e.g., collecting data for a single or limited number of endpoints).
- In this example, we limit our data gathering to the common Skin endpoints from databases containing Skin Sensitization data

Endpoint

QSAR Toolbox 3.1.0.21 [Mixture with defined quantities.smi]

Menu: Data (2) | Import | Export | Delete | Tautomerize

Sub-menu: Gather | Import | IUCLID5 | Export | IUCLID5 | Database | Inventory | Database

Databases List:

- Carcinogenic Potency Database (CPDB)
- Carcinogenicity&mutagenicity ISSCAN
- Cell Transformation Assay ISSCTA
- Dendritic cells COLIPA
- Developmental toxicity ILSI
- ECHA CHEM
- Estrogen Receptor Binding Affinity OASIS
- Eye Irritation ECETOC
- Genotoxicity OASIS
- Keratinocyte gene expression Givaudan
- Micronucleus ISSMIC
- Micronucleus Oasis
- MUNRO non-cancer EFSA
- Rep Dose Tox Fraunhofer I
- Repeated Dose Toxicity HE
- Rodent Inhalation Toxicity
- Skin irritation
- Skin sensitization
- Skin sensitization ECETOC (1)
- Terrestrial US EPA ECOTOX
- Toxicity Japan MHLW
- ToxRefDB US-EPA
- Verical import_ BOD and Ames
- Yeast estrogen assay database University of Ten

Filter endpoint tree...

- Structure
- Substance Identity
- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
- Human Health Hazards
- Profile
 - General Mechanistic
 - Protein binding by OASIS v1.1
 - Protein binding by OECD
 - Protein binding potency
 - Endpoint Specific
 - Protein binding alerts for skin sensitization b...

	1 [target]	2 [target,mix.component]	3 [target,mix.component]	4 [target,mix.component]
Structure	[3] [Mix] 	 Qty: 1 Milligrams	 Qty: 10 Milligrams	 Qty: 100 Milligrams
No alert found	No alert found	SNAr SNAr >> Nucleophi...	No alert found	No alert found
Protein binding by OASIS v1.1	SNAr SNAr >> Nucleophi...	SNAr >> Nucleophi...	SNAr >> Nucleophi...	SNAr >> Nucleophi...
Protein binding by OECD	No alert found	No alert found	No alert found	No alert found
Protein binding potency	Not possible to cla...	Not possible to cla...	Not possible to cla...	Not possible to cla...
Protein binding alerts for skin sensitization b...	No alert found	SNAr SNAr >> Nucleophi... SNAr >> Nucleophi...	No alert found	No alert found

- Select** databases related to the target endpoint by adding a green check in the box before the database name.
- Click** Gather

Endpoint Process of collecting data

QSAR Toolbox 3.1.0.21 [Mixture with defined quantities.smi]

QSAR TOOLBOX
Input
Profiling
Endpoint
Category Definition
Data Gap Filling
Report

Data
Import
Export
Delete
Tautomerize

Gather
Import
IUCALID5
Export
IUCALID5
Database
Inventory
Database

Databases

Select All Unselect All Invert About

- Carcinogenic Potency Database (CPDB)
- Carcinogenicity&mutagenicity ISSCAN
- Cell Transformation Assay ISSCTA
- Dendritic cells COLIPA
- Developmental toxicity ILSI
- ECHA CHEM
- Estrogen Receptor Binding Affinity OASIS
- Eye Irritation ECETOC
- Genotoxicity OASIS
- Keratinocyte gene expression Givaudan
- Micronucleus ISSMIC
- Micronucleus Oasis
- MUNRO non-cancer EFSA
- Rep Dose Tox Fraunhofer ITEM
- Repeated Dose Toxicity HESS
- Rodent Inhalation Toxicity Database
- Skin irritation
- Skin sensitization
- Skin sensitization ECETOC
- Terrestrial US-EPA ECOTOX
- Toxicity Japan MHLW
- ToxRefDB US-EPA
- Verical import_ BOD and Ames
- Yeast estrogen assay database University of Ten

Inventories

Select All Unselect All Invert About

- AICS
- Canada DSL
- COSING
- DSSTOX
- ECHA PR

Filter endpoint tree...

	1 [target]	2 [target,mx.component]	3 [target,mx.component]	4 [target,mx.component]
Structure				
		Qty: 1 Milligrams	Qty: 10 Milligrams	Qty: 100 Milligrams
<input type="checkbox"/> Substance Identity				
<input type="checkbox"/> Physical Chemical Properties				
<input type="checkbox"/> Environmental Fate and Transport				
<input type="checkbox"/> Ecotoxicological Information				
<input type="checkbox"/> Human Health Hazards				
<input type="checkbox"/> Acute Toxicity				
<input type="checkbox"/> Carcinogenicity				
<input type="checkbox"/> Developmental Toxicity / Teratogenicity				
<input type="checkbox"/> Genetic Toxicity				
<input type="checkbox"/> Immunotoxicity				
<input type="checkbox"/> Irritation / Corrosion				
<input type="checkbox"/> Neurotoxicity				
<input type="checkbox"/> Repeated Dose Toxicity				
<input type="checkbox"/> Sensitisation				
<input type="checkbox"/> Skin				
<input type="checkbox"/> In Chemico				
<input type="checkbox"/> In Vitro				
<input type="checkbox"/> In Vivo				
<input type="checkbox"/> Toxicity to Reproduction				
<input type="checkbox"/> Toxicokinetics, Metabolism and Distribution				
<input type="checkbox"/> Profile				

The third component without experimental data will be used for further read-across analysis

(2/2)

			M: Negative	M: Negative
--	--	--	-------------	-------------

Recap

- You have entered the mixture with defined components
- You have profiled the target chemical mixture and found no protein binding alerts for two of the mixture constituents. The third constituent has positive protein binding alerts and could elicit skin sensitization effect
- Negative experimental data has been found for two of mixture components. No experimental data has been found for the third constituent
- The constituent without experimental data and positive protein binding alert has been used for further read across analysis. Then, all of the available data – experimental and predicted will be used for SS prediction of the mixture.
- Now you are ready to continue with “Read across prediction of constituent without data”.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - **Read across prediction of constituent without data**
 - **Focus constituent without experimental data**

Read across prediction of constituent without data

Focus constituent

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The main window displays a table with columns for chemical structures and their weights. A context menu is open over a chemical structure, with the 'Focus' option highlighted. A red box surrounds the 'Focus' button, and callout boxes '1' and '2' indicate the steps: right-clicking and selecting 'Focus'. A text box at the bottom states: 'This constituent is selected for further read-across prediction'.

1. **Right click** over the chemical without experimental data
2. **Select Focus**

Read across prediction of constituent without data

Focus constituent

1

Structure

Qty: 1 weight %

Substance Identity

- CAS Number: N/A
- Chemical IDs: NA
- Chemical Name
- Structural Formula: C(C)(=O)c1c(Cl)c(Cl)c(Cl)c1...

Physical Chemical Properties

Environmental Fate and Transport

Ecotoxicological Information

Human Health Hazards

Profile

Defined Categories

Document_1[C(C)(=O)c1c(Cl)c(Cl)c(Cl)c1]

This focused component appeared in separate data matrix

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - **Read across prediction of constituent without data**
 - Focus constituent without experimental data
 - **Define category**

Category Definition Overview

- This module provides the user with several means of grouping chemicals into a toxicologically meaningful category that includes the target molecule.
- This is the critical step in the workflow.
- Several options are available in the Toolbox to assist the user in refining the category definition.

Category Definition

Grouping methods

- The different grouping methods allow the user to group chemicals into chemical categories according to different measures of “similarity”.
- Detailed information about grouping chemical (Chapter 4) could be found in document “Manual for Getting started” published on OECD website:

<http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm>

Basic guidance for category formation and assessment

Suitable categorization phases:

1. Structure-related profilers
2. Endpoint specific profilers (for sub-cat)
3. Additional structure-related profilers, if needed to eliminate dissimilar chemicals (to increase the consistency of category) (e.g. chemical elements)

Performing categorization:

1. The categorization phases should be applied successively
2. The application order of the phases depend on the specificity of the data gap filling
3. More categories of same Phase could be used in forming categories
4. Some of the phases could be skipped if consistency of category members is reached

Graphical illustration of suitable categorization phases is shown on next slide

Suitable Categorization/Assessment Phases

Phase I. Structure based

- US EPA Categorization
- OECD Categorization
- Organic functional group
- Structural similarity
- ECOSAR

**Broad grouping
Endpoint Non-specific**

Repeating Phase I due to Multifunctionality of chemicals

Phase II. Mechanism based

- DNA binding mechanism
- Protein binding mechanism
- Genotoxicity/carcinogenicity
- Cramer rules
- Verhaar rule
- Skin/eye irritation corrosion rules

**Subcategorization
Endpoint Specific**

Metabolism accounted for

Phase III. Eliminating dissimilar chemicals

**Apply Phase I – for structural dissimilarity
Filter by test conditions – for Biological dissimilarity**

**Subcategorization
Endpoint Specific**

Read across prediction of constituent without data

Forming category for studied endpoint

Suitable Categorization/Assessment Phases

Phase I. Structure based

- US EPA Categorization
- OECD Categorization
- Organic functional group
- Structural similarity
- ECOSAR

Repeating Phase I due to Multifunctionality of chemicals

Broad grouping
Endpoint Non-specific

Phase I categorization in Toolbox

The screenshot shows the 'Filter endpoint tree...' on the left and a list of categories on the right. The chemical structure of 2,4-dichlorobenzaldehyde is shown at the top right. The 'Organic functional groups' category is expanded, showing 'Aryl halide' selected. Other categories include 'Substance Identity', 'Physical Chemical Properties', 'Environmental Fate and Transport', 'Ecotoxicological Information', 'Human Health Hazards', 'Profile', 'Predefined', 'US-EPA New Chemical Categories', 'Endpoint Specific', 'Aquatic toxicity classification by EC...', 'Empiric', and 'Organic functional groups'.

It is not recommended to use "Neutral organic" * as phase I

46 analogues are identified as Aryl halides by OFG

*Neutral organic category include chemicals having different functionalities as alcohols, ketones, ethers etc. In this respect the basic principle that structurally similar chemicals may elicit similar effects would not be preserved, because Neutral organic mixed many different functionalities

Read across prediction of constituent without data

Forming category for studied endpoint

- Based on the above recommendations and classifications from structurally similar profilers the OFG is used as initial categorization group
- Refinement of the initial group is based on endpoint-specific protein binding profiler:
 - Protein binding alerts for skin sensitization by Oasis v1.1.

Category definition is a tool for grouping chemicals, which allows to group chemicals based on different measures of "similarity". For more details see tutorials posted on LMC and OECD website:

<http://www.oecd.org/env/ehs/risk-assessment/theoecdqsartoolbox.htm>

<http://superhosting.oasis-lmc.org/products/software/toolbox/toolbox-support.aspx>

See next slides

Read across prediction of constituent without data

Define category by OFG

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The 'Define' button is circled in red and labeled with a blue callout '2'. In the 'Grouping methods' list on the left, 'Organic functional groups' is selected and labeled with a blue callout '1'. A text box on the right states: 'Based on above recommendations the OFG is used as an initial group (phase I)'. A dialog box titled 'Organic functional groups' is open, showing a list of profiles. A smaller dialog box in the foreground displays the error message: 'Grouping by Organic functional groups: no chemicals found!' with an 'OK' button, which is labeled with a blue callout '3'. The main interface also shows a 'Structure' window with a chemical structure and a 'Filter endpoint tree...' window.

1. Select Protein binding by OASIS v1.1
 Combination of four organic functional group do not identify similar analogues (3).
 In this respect Aryl halide is used only. See next slide

2. Click Define

Read across prediction of constituent without data

Define category by OFG

1. Select OFG

2. Click Define

3. Select Ketone group and remove it when

4. Click arrow down

5. Aryl halide should be remained only

6. Click OK

Except Aryl halide all other groups are removed by selecting the categories and clicking on arrow down

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - **Read across prediction of constituent without data**
 - **Focus constituent without experimental data**
 - **Define category**
 - **Gather data for analogues**

Read across prediction of constituent without data

Gather data for analogues chemicals

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The main window displays a filter endpoint tree on the left and a chemical structure in the center. Two dialog boxes are overlaid on the interface:

- Define category name:** A dialog box with the title "Define category name" and a close button. It contains the text "Category name (46 chemicals)" and "Aryl halide (Organic functional groups)". There are "OK" and "Cancel" buttons. A red callout box with the number "1" points to the "OK" button.
- Read data?:** A dialog box with the title "Read data?". It has three radio buttons: "All endpoints" (selected), "Choose...", and "from Tautomers" (checked). There are "OK" and "Cancel" buttons. A red callout box with the number "2" points to the "OK" button.

At the bottom of the screenshot, there is a blue banner with the following text:

1. Click OK **2. Click OK in order to read data for all endpoints**

Read across prediction of constituent without data

Gather data for analogues chemicals

The screenshot displays the QSAR Toolbox 3.1.0.21 interface. The main window shows a data matrix with columns representing different target molecules (labeled 1 through 8) and rows representing various chemical endpoints. A red oval highlights the 'Repeated Dose Toxicity' row, which contains data for targets 2 through 8. A red box at the bottom of the screenshot contains the text: 'The experimental data for the identified analogues appear on data matrix'.

Filter endpoint tree...	1 [target]	2	3	4	5	6	7	8
Structure								
Substance Identity								
Physical Chemical Properties								
Environmental Fate and Transport								
Ecotoxicological Information								
Human Health Hazards								
Acute Toxicity								
Carcinogenicity								
Developmental Toxicity / Teratogenicity								
Genetic Toxicity								
Immunotoxicity								
Irritation / Corrosion								
Neurotoxicity								
Repeated Dose Toxicity								
Sensitisation		M: Positive	M: Positive	M: Negative	M: Negative, Positiv...	M: Negative, Positive	M: Negative, Negative	M: Positive
Toxicity to Reproduction								
Toxicokinetics, Metabolism and Dietri								
Profile								

The experimental data for the identified analogues appear on data matrix

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - **Read across prediction of constituent without data**
 - **Focus constituent without experimental data**
 - **Define category**
 - **Gather data for analogues**
 - **Apply read across**

Read across prediction of constituent without data

Apply read across

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The main window displays a table with columns for target, structure, and various endpoints. A callout '1' points to the 'Skin Sensitization in Vivo' cell, which contains '(45/56) M Positive'. A callout '2' points to the 'Read-across' option in the left sidebar. A callout '3' points to the 'Apply' button. A callout '4' points to the 'OK' button in the 'Possible data inconsistency' dialog box. The dialog box lists the following endpoints and assays:

- A B C (12 points)
- EC3 (21 points)
- S M A N (13 points)
- S M W N (9 points)
- Skin Sensitisation (1 points)
- GPMT (9 points)
- LLNA (21 points)
- Miscellaneous (25 points)
- Undefined Assay (1 points)
- Skin sensitisation V (BFR) (12 points)
- Skin sensitisation I (Oasis) (8 points)
- Skin sensitisation II (ECETOC) (1 points)
- Skin Sensitization (Danish EPA)
- Skin sensitisation IV (GPMT) (9 points)
- Skin sensitization EC3(ratio) (12 points)
- Skin sensitisation III (LJMU) (13 points)

A red box highlights the dialog box content with the text "Endpoints and Assays are mixed".

1. **Click** on the cell corresponding to Skin Sensitization in Vivo
2. **Select** Read-across
3. **Click** Apply
4. **Click** OK (in this case we mixed all endpoints and assays)

Read across prediction of constituent without data

Apply read across

QSAR Toolbox 3.1.0.21 [Mixture with mol %sm]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

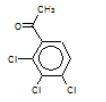
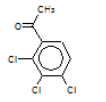

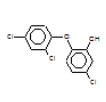
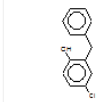
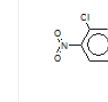
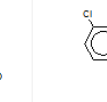
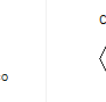
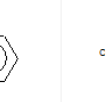
The OECD QSAR Toolbox for Grouping Chemicals into Categories
Developed by LMC, Bulgaria

Data Gap Filling Method

- Read-across
- Trend analysis
- (Q)SAR models

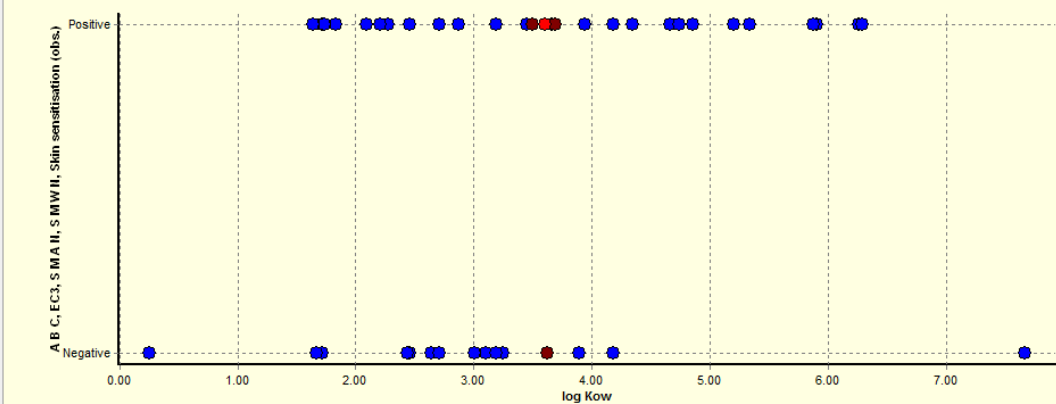
Target Endpoint

Human Health Hazards Sensitisation Skin In Vivo

Structure	1 [target]	2	3	4	5	6	7	8
								
In Vivo (45/56)	M. Positive	M. Positive	M. Negative	M. Positive, Positiv	M. Negative, Positive	M. Negative, Negative	M. Posi	

Descriptors Prediction

Read across prediction of A B C, EC3, S M A N, S M W H, Skin sensitisation, taking the highest mode from the nearest 5 neighbours, based on 5 values from 5 neighbour chemicals, Observed target value: N/A, Predicted target value: 'Positive'



Descriptor X: log Kow

Accept prediction
Return to matrix

- Select/filter data
 - Subcategorize
 - Mark chemicals by descriptor value
 - Filter points by test conditions
 - Mark focused chemical
 - Mark focused points
 - Remove marked chemicals/points
 - Clear existing marks
- Selection navigation
- Gap filling approach
- Descriptors/data
- Model/(Q)SAR
- Calculation options
- Visual options
- Information
- Miscellaneous

46_Aryl halide (Organic functional groups) Data gap filling 3/1/0

Read across prediction of constituent without data Subcategorization

- The initial category could be refined by subcategorizing the analogues according to the following endpoint specific profiler (phase II, slide #37):
 - Protein binding alerts for skin sensitization by Oasis v1.1.
- These steps are summarized in the next screen shots.

Read across prediction of constituent without data

Subcategorization by Protein binding alert for SS

The screenshot displays the QSAR Toolbox interface during a subcategorization task. The left sidebar shows the 'Grouping methods' list, with 'Protein binding alerts for SS' selected (callout 2). The top navigation bar includes 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The central area shows a table of chemical structures with their predicted categories: (45/56), M. Positive, M. Positive, M. Negative, M. Positive, Positive, M. Negative, Positive, M. Negative, Negative, M. Positive. The bottom plot shows 'Skin sensitisation (obs.)' vs 'log Kow', with a title: 'Read across prediction of A B C, EC3, S M A N, S M W N, Skin sensitisation, taking the highest mode from the nearest 5 neighbours, based on 5 values from 5 neighbour chemicals, Observed target value: N/A, Predicted target value: Positive'. The 'Accept prediction' panel on the right has 'Subcategorize' selected (callout 1). The 'Remove' button at the bottom is highlighted (callout 3).

1. Select filter data/subcategorize
2. Select Protein binding alerts for SS by OASIS v1.1.
3. Click Remove to eliminate dissimilar chemicals.

Read across prediction of constituent without data

Apply read across

QSAR Toolbox 3.1.0.21 [Mixture with mol %smi]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Filling Apply

The OECD QSAR Toolbox for Grouping Chemicals into Categories
Developed by LMC, Bulgaria

Data Gap Filling Method

- Read-across
- Trend analysis
- (Q)SAR models

Target Endpoint

Human Health Hazards Sensitisation Skin In Vivo

Structure	1 [target]	5	8	11	14	21	22	27
In Vivo (45/56)	M: Positive, Positiv...	M: Positive	M: Positive	M: Positive	M: Positive	M: Negative	M: Negative	M: Posi

Descriptors Prediction

Read across prediction of A B C, EC3, S M A N, S M W N, Skin sensitisation, taking the highest mode from the nearest 5 neighbours, based on 5 values from 5 neighbour chemicals, Observed target value: NA, Predicted target value: Positive

Almost all analogs have been found to be positive.
Predicted SS effect of the target is positive

Descriptor X: log Kow

Accept prediction
Return to matrix

- Select/filter data
 - Subcategorize
 - Mark chemicals by descriptor value
 - Filter points by test conditions
 - Mark focused chemical
 - Mark focused points
 - Remove marked chemicals/points
 - Clear existing marks
- Selection navigation
 - Gap filling approach
 - Descriptors/data
 - Model/(Q)SAR
 - Calculation options
 - Visual options
 - Information
 - Miscellaneous

46 Ayl halide (Organic functional groups) Data gap filling 8/1/0

Read across prediction of constituent without data

Apply read across

QSAR Toolbox 3.1.0.21 [Mixture with mol %.smi]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Filling Apply

The OECD QSAR Toolbox for Grouping Chemicals into Categories
Developed by LMC, Bulgaria

Structure	1 [target]	5	8	11	14	21	22	27
In Vivo (45/56)	M: Positive, Positiv...	M: Positive	M: Positive	M: Positive	M: Positive	M: Negative	M: N	M: Posi

Descriptors Prediction

Read across prediction of A B C, EC3, S M A N, S M W N, Skin sensitisation, taking the highest mode from the nearest 5 neighbours, based on 5 values from 5 neighbour chemicals, Observed target value: N/A, Predicted target value: 'Positive'

Information: The current prediction was accepted

log Kow

3.00 4.00

Positive

C, EC3, S M A N, S M W N, Skin sensitisation (obs.)

Accept prediction
Return to matrix

Select/filter data
Selection navigation
Gap filling approach
Descriptors/data
Model/(Q)SAR
Calculation options
Visual options
Information
Miscellaneous

1. Click Accept prediction
2. Click OK
3. Return to matrix

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Read across prediction of constituent without data
 - **Filling data gap for skin sensitization of mixture**

Filling data gap for skin sensitization of mixture Applying Independent Mode of Action

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The 'Data Gap Filling Method' is set to 'Independent MOA'. The 'Target Endpoint' is 'Human Health Hazards Toxicity to Reproduction'. The 'Filter endpoint tree...' panel shows a tree structure with 'Skin Sensitization' expanded. The data table shows a mixture with a positive result for skin sensitization. A dialog box titled 'Possible data inconsistency' is open, showing 'Selected [3/3] points' and 'OK' and 'Cancel' buttons. An 'Apply' button is highlighted in the top left corner.

1. **Click** on the cell corresponding to Skin Sensitization for mixture
2. **Select** Independent MOA

3. **Click** Apply
4. **Click** OK

Filling data gap for skin sensitization of mixture

Applying Independent Mode of Action

QSAR Toolbox 3.1.0.21 [Mixture with defined quantities.sm]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Filling Apply

Data Gap Filling Method

- Independent MOA
- Similar MOA
- Specific models

Target Endpoint

Human Health Hazards Sensitisation Skin In Vivo

	1 [target]	2 [target,mix. component]	3 [target,mix. component]	4 [target,mix. component]
Structure	[3] [Mix]	<chem>ClC1=CC=C(C=C1)C(=O)C</chem>	A <chem>O=C1C=CC=CC=C1</chem>	A <chem>CCO</chem>
In Vivo	(3/3)	Qty: 1 Milligrams R: Positive	Qty: 10 Milligrams M: Negative	Qty: 100 Milligrams M: Negative

Prediction

Empiric calculation of A B C, EC3, S M A II, taking the maximal from the component values, based on 3 values from 3 target components, Observed target value: N/A, Predicted target value: 'Positive'

Descriptor X: log Kow

Accept prediction

Return to matrix

- Select/filter data
- Selection navigation
- Descriptors/data
- Calculation options

Data usage

- Mixture models
- Visual options
- Information
- Miscellaneous

Set empiric calculations option

Approximation type: Maximal

OK Cancel

Read across is applied for the mixture (assuming Independent Mode of Action)
 "Maximal" approximation type is set by default for categorical endpoints (worst case scenario)(see 1)

Filling data gap for skin sensitization of mixture

Applying Independent Mode of Action

QSAR Toolbox 3.1.0.21 [Mixture with defined quantities.sm]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Filling

Apply

Data Gap Filling Method

- Independent MOA
- Similar MOA
- Specific models

Target Endpoint

Human Health Hazards Sensitisation Skin In Vivo

1 [target]	2 [target,mix.component]	3 [target,mix.component]	4 [target,mix.component]
[3] [Mix]		A	A
	Qty: 1 Milligrams	Qty: 10 Milligrams	Qty: 100 Milligrams
	R: Positive	M: Negative	M: Negative

1

2

Prediction

Empiric calculation of A B C, EC3, S M A II, taking the maximal from the component values, based on 3 values from 3 target components, Observed target value: N/A, Predicted target value: 'Positive'

Positive

Negative

A B C, EC3, S M A II (obs.)

log Kow

1.00 2.00 3.00

Based on the positive skin sensitization value for one of the mixture components the prediction for the mixture is positive

Accept prediction

Return to matrix

- Select/filter data
- Selection navigation
- Descriptors/data
- Calculation options
- Data usage
- Mixture models
- Visual options
- Information
- Miscellaneous

Set empiric calculations options:

Approximation type: Maximal

OK Cancel

1. Accept prediction
2. Return to matrix

Filling data gap for skin sensitization of mixture

Applying Independent Mode of Action

QSAR Toolbox 3.1.0.21 [Mixture with defined quantities.smi]

Input
Profiling
Endpoint
Category Definition
Data Gap Filling
Report

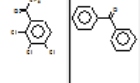
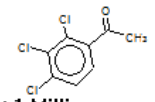
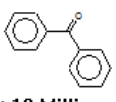
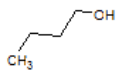
Filling
 ⚡
 Apply

Data Gap Filling Method

- Independent MOA
- Similar MOA
- Specific models

Target Endpoint

Human Health Hazards Sensitisation Skin In Vivo

Filter endpoint tree...	1 [target]	2 [target,mix.component]	3 [target,mix.component]	4 [target,mix.component]
Structure	[3] [Mix] 	 Qty: 1 Milligrams	A  Qty: 10 Milligrams	A  Qty: 100 Milligrams
<input type="checkbox"/> Substance Identity				
<input type="checkbox"/> Physical Chemical Properties				
<input type="checkbox"/> Environmental Fate and Transport				
<input type="checkbox"/> Ecotoxicological Information				
<input type="checkbox"/> Human Health Hazards				
<input type="checkbox"/> Acute Toxicity				
<input type="checkbox"/> Carcinogenicity				
<input type="checkbox"/> Developmental Toxicity / Teratogenicity				
<input type="checkbox"/> Genetic Toxicity				
<input type="checkbox"/> Immunotoxicity				
<input type="checkbox"/> Irritation / Corrosion				
<input type="checkbox"/> Neurotoxicity				
<input type="checkbox"/> Repeated Dose Toxicity				
<input type="checkbox"/> Sensitisation				
<input type="checkbox"/> Respiratory Tract				
<input type="checkbox"/> Skin				
<input type="checkbox"/> In Chemo				
<input type="checkbox"/> In Vitro				
<input type="checkbox"/> In Vivo				
<input type="checkbox"/> Toxicity to Reproduction				
<input type="checkbox"/> Toxicokinetics, Metabolism and Distribution				
	(4) Cl: Positive	R: Positive	M: Negative	M: Negative

Read across prediction for the mixture based on predicted and experimental data of mixture constituents appears on datamatrix

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Read across prediction of constituent without data
 - Filling data gap for skin sensitization of mixture
 - **Generating report for mixture**

Report

- Remember the report module allows you to generate a report on the predictions performed with the Toolbox. This module contains predefined report templates as well as a template editor with which users can define their own user defined templates. The report can then be printed or saved in different formats.
- Generating the report is shown on next screenshots

Report

QSAR Toolbox 3.1.0.21 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Reports Repository

Create Print Close Save as Register Unregister Update Clone Design

Available data to report

- Predictions
- (Q)SARs
- Categories

Prediction [2]

Prediction of EC3, S M A N for {X=1/weight %}C(C)(=O)c1c(Cl)c(Cl)c(Cl)cc1_{X=9/weight %}c1(C(=O)c2cccc2)cccc1_{X=90/weight %}C(O)CCC 1 / 34

QSAR Toolbox prediction for multicomponent substance
 (uses single component mode for handling of target mixture and its components)

Available report templates

- Standard (predefined)
 - QSAR Model Reporting Format (
 - QSAR Toolbox Category Report
 - QSAR Toolbox Prediction Report
- Custom (user defined)
 - Editable copy of QSAR Model Re
 - Editable copy of QSAR Toolbox (
 - Editable copy of QSAR Toolbox f

Toolbox report for mixture

The template of the current report is based on "GUIDANCE DOCUMENT ON THE VALIDATION OF (QUANTITATIVE) STRUCTURE-ACTIVITY RELATIONSHIPS MODELS" by OECD (September, 2007) and "GUIDANCE ON INFORMATIONMENTS AND CHEMICAL SAFETY ASSESSMENT / CHAPTER R.6: QSARS AND ICG OF CHEMICALS" published by ECHA (May, 2008).

Report

Available data to report

- Predictions
- (Q)SARs
- Categories

Available report templates

- Standard (predefined)
 - QSAR Model Reporting Format (
 - QSAR Toolbox Category Report
 - QSAR Toolbox Prediction Report
- Custom (user defined)
 - Editable copy of QSAR Model Re
 - Editable copy of QSAR Toolbox (
 - Editable copy of QSAR Toolbox f

Summary

Toxicity of the target mixture (Positive) is predicted from its components using estimation based on 1 values (Positive x1) from 1 components having independent mode of action. Both experimental and predicted values for mixture components are used in predicting the target toxicity. The components of a mixture are handled with the functionality for category. The same approach can be applied for mixtures, but also for mono-constituent substances with impurities, multi-constituent substances and UVCBs with identified constituents.

The target mixture FALLS within applicability domain of the prediction (see Section 4.3 for details).

The endpoint data is selected from the following database(s):

- Skin sensitization
- Skin sensitization ECETOC

Below is a summary table for endpoint & descriptor values for the target mixture and the mixture components. Experimental values from data matrix are presented in bold font. Recalculated endpoint values (if required by selected data usage option in Gap Filling) are presented in italic font. Recalculated endpoint values based on experimental data only are presented in bold and italic font.

	Endpoint(s)
Qty, weight	Sensitisation

1. Summary information for mixture prediction