

OECD QSAR Toolbox v.4.1

Example for predicting skin sensitisation potential of (2*E*,6*Z*)-2,6-nonadien-1-ol accounting for skin metabolism

Outlook

- **Background**
- Objectives
- The exercise
- Workflow
- Save the prediction

Background

- This is a step-by-step presentation designed to take the user through the Toolbox workflow for filling data gap for skin sensitization of trans-2,cis-6-nonadienol accounting for its skin metabolism

Outlook

- Background
- **Objectives**
- The exercise
- Workflow
- Save the prediction

Objectives

This presentation demonstrates a number of functionalities of the Toolbox:

- Simulating skin metabolism of target chemical
- Identify analogues for a selected active metabolite
- Filling data gaps for active metabolites by read across
- Assign prediction of metabolite to the parent chemical

Outlook

- Background
- Objectives
- **The exercise**
- Workflow
- Save the prediction

The Exercise

- In this exercise we will predict the skin sensitization potential for target chemical (**trans-2,cis-6-nonadienol**) [**CAS # 28069-72-9**].
- Profile the target chemical and identify no Protein binding alert for target chemical.
- Gather available experimental data for the target chemical and identify positive experimental data.
- Skin metabolism of target chemical will be accounted for.
- Read across prediction for active metabolite will be applied.
- The predicted result of metabolite will be assigned to the target chemical.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
- Save the prediction

Workflow

- **As you know the Toolbox has 6 modules which are typically used in sequence:**
 - Chemical Input
 - Profiling
 - Endpoint
 - Category Definition
 - Data Gap Filling
 - Report
- **In this example we will use the modules in a different order, tailored to the aims of the example.**

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 Chemicals

User Alternatives for input of Chemical:

A. Single target chemical

- Chemical Name
- Chemical Abstract Services (CAS) number (#)
- SMILES (simplified molecular information line entry system) notation/InChi
- Drawing chemical structure
- Select from User List/Inventory/Databases

B. Group of chemicals

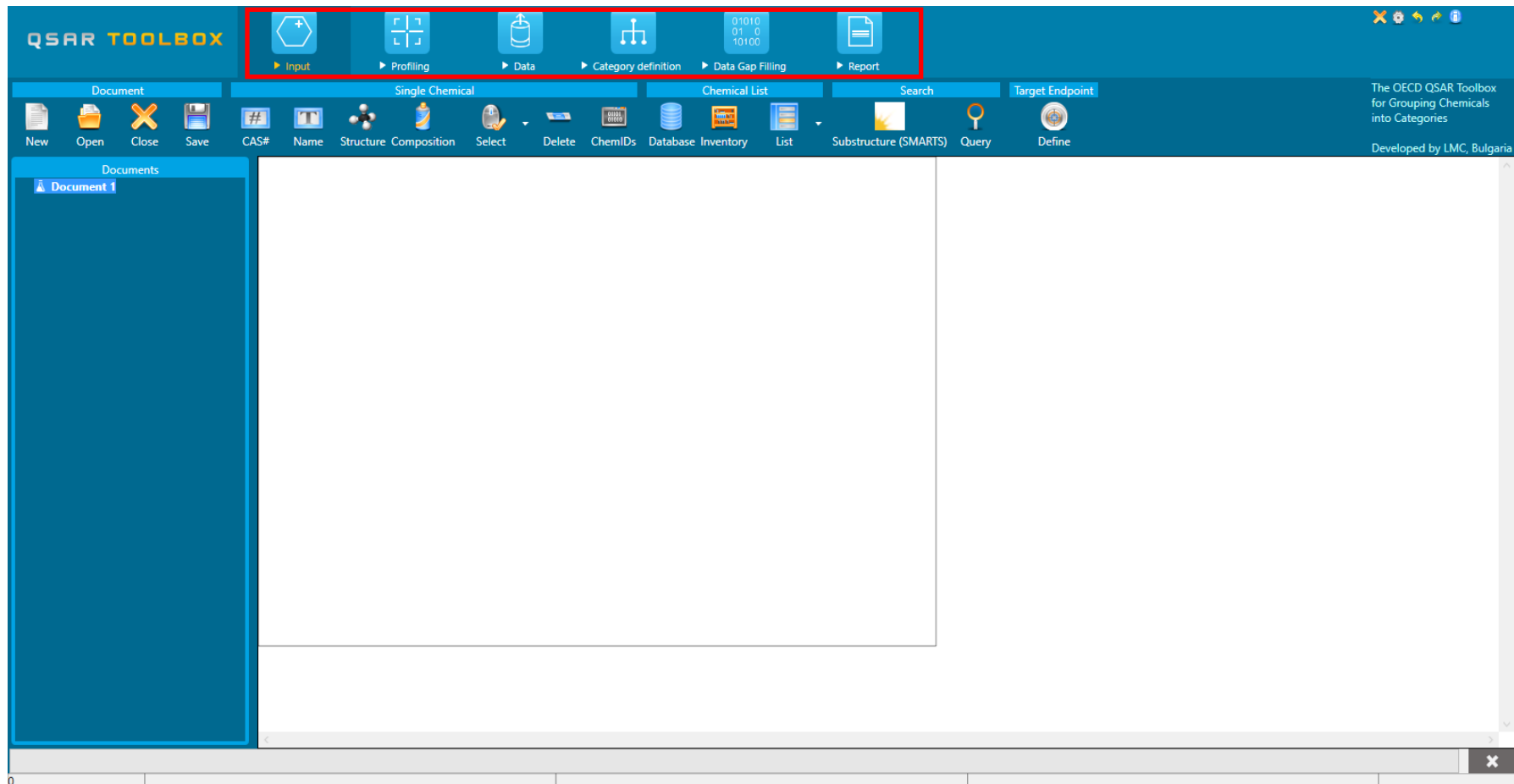
- User List/Inventory
- Specialized Databases

Chemical Input

Input Screen

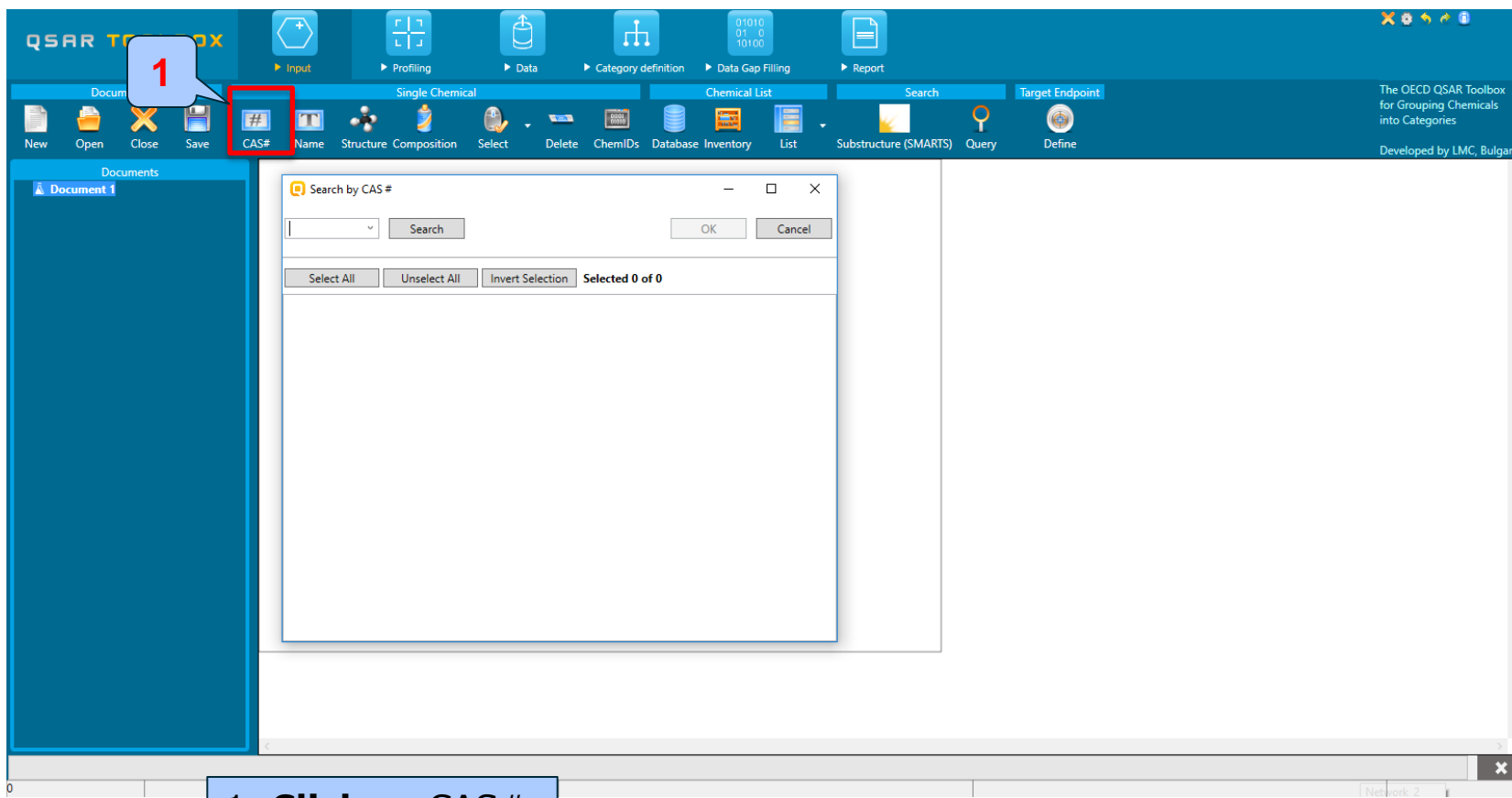
- Open the Toolbox.
- The six modules in the workflow are seen listed next to “QSAR TOOLBOX” title.
- Click on “Input” (see next screen shot)

Chemical Input Input Screen



Chemical Input

Input target chemical by CAS#



1. Click on CAS#

Chemical Input

Enter CAS# of trans-2,cis-6-nonadienol

Search by CAS #

28069729 Search OK Cancel

Select All Unselection Selected 1 of 1

1	CAS	28069-72-9
	SMILES	CCC=CCCC=CCO
	CS Relation	High
<input checked="" type="checkbox"/>	Substance	Mono constituent
	Composition	
	Name	(2E,6Z)-nona-2,6-dien-1-ol 2,6-Nonadien-1-ol, (2E,6Z)- 2,6-Nonadien-1-ol, (E,Z)-

H₃C-CH=CH-CH₂-CH=CH-CH₂-CH₂-OH

1. Enter the CAS# **28069-72-9** In the blank field;
2. Click Search button;
3. Press OK

Chemical Input

Target chemical identity

- Double click “CAS Smiles relation” displays the chemical identification information.
- This indicates the reliability of relation CAS-Name for the target chemical(see next screen shots).
- The workflow on the first module is now complete, and the user can proceed to the next module.

Chemical Input

Target chemical identity

The screenshot displays the QSAR Toolbox interface. On the left, there are panels for 'Documents', 'Databases', and 'Inventories'. The main area shows a 'Filter endpoint tree...' with a tree structure. A callout '1' points to the 'High' quality source for 'trans-2,cis-6-Nonadi'. A callout '2' points to a 'Relationships CAS-SMILES' dialog box showing a table of data sources for the chemical.

1. Double Click **2. Relationships CAS-SMILES**

Exist in data source	Data source type	Data source quality	Assigned SMILES in data source
Canada DSL	Inventory	Distribute to QA	no
DSSTOX	Inventory	High quality source	yes
ECHA PR	Inventory	Distribute to QA	yes
EINECS	Inventory	High quality source	yes
NICNAS	Inventory	Distribute to QA	yes
REACH ECB	Inventory	High quality source	yes
Skin Sensitization	Database	Distribute to QA	yes
TSCA	Inventory	Distribute to QA	yes

Chemical Input

Target chemical identity

The code indicates the reliability of the chemical identifier:

- **High:** This reliability corresponds to high reliability of CAS-SMILES relation. This label is assigned if the chemical belongs to at least one high quality data source (database or inventory)
- **Moderate:** This reliability corresponds to moderate reliability of CAS-SMILES relation. The moderate label is assigned if the chemical belongs to three "Distribute to QA" data sources.
- **Low:** This reliability corresponds to poor reliability of CAS-SMILES relation. This label is assigned if the chemical belongs to less than three, but at least one "Distribute to QA" data sources.

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 and clicking on “View” (see next screen shot)).

Profiling Side-Bar to Profiling

1 Highlight the profiler

2 Click View

3 Click Advance in order to see detailed description of highlighted category (in this case "Aldehydes")

List with categories

- Protein binding by OASIS
 - Acylation
 - Ionic interaction
 - Michael addition
 - Nucleophilic addition
 - Radical reactions
 - Schiiff base formation
 - Domain
 - Benzoyl Schiff base formation
 - Direct acting Schiff base formers
 - Schiiff base formation with carbonyl compounds
 - Mechanistic alert
 - Structural alert
 - Aldehydes
 - alpha-Ketoesters
 - Aromatic carbonyl compounds
 - Bis aldehydes
 - Schiiff base on pyrazolones and pyrazolidinones

Mechanistic Domain: Schiiff base formation

Mechanistic Alert: Schiiff base formation with carbonyl compounds

Structural Alert: Aldehydes

The chemical causes skin sensitization effect as a result of Schiiff base formation with aldehydes:

Structural alert

$$R-C(=O)H + Pr-NH_2 \rightarrow R-C(=N-Pr)H$$

R = H,alkyl

Aldehydes are highly reactive molecules, many of which are strong sensitizers and their direct conjugation to protein nucleophiles is thought to be responsible. Simple aldehydes react readily with the amino groups of lysine residues on proteins to form imines or Schiiff bases.

All aliphatic aldehydes can potentially undergo Schiiff base formation with a primary amine, which is a reversible reaction (optimal at pH 3-4) and proceeds in two stages via a tetrahedral intermediate.

1. **Highlight** the profiler

2. **Click** View

3. **Click** Advance in order to see detailed description of highlighted category (in this case "Aldehydes")

Profiling Side-Bar to Profiling

1. Highlight the profiler
2. Click View
3. Select "Aldehydes"

Profiling

Profiling the target chemical

- The outcome of the profiling determines the most appropriate way to search for analogues (detailed information about profilers could be found in “Manual for Getting started” (Chapter 4) published on the OECD website:

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

- Table 4 - 1 in chapter 4 (Manual for getting started) lists a selection of profilers and their relevance for different endpoints of regulatory relevance
- The following profiling schemes are relevant to the **Skin sensitization**:
 - Protein binding by OASIS – general mechanistic
 - Protein binding by OECD – general mechanistic
 - Protein Binding Potency – general mechanistic
 - Protein binding alerts for skin sensitization by OASIS - endpoint specific

Profiling

Profiling the target chemical

- This selects (a **green** check mark appears) or deselects (**green** check mark disappears) profilers.
- For this example, go through the general and endpoint specific profiling mechanisms and highlight those that apply to skin sensitization(see next screen shot).

Profiling

Profiling the target chemical

The screenshot shows the QSAR Toolbox interface with the following components:

- Top Menu:** Input, Profiling, Data, Category definition, Data Gap Filling, Report.
- Documents Panel:** Document 1 with chemical name: containing 'trans-2,cis-6-nonadienol'.
- Profilng methods Panel:**
 - Options: Select All, Unselect All, Invert.
 - Selected methods:
 - Protein binding by OASIS (checked)
 - Protein binding by OECD (checked)
 - Protein binding potency (checked)
 - Endpoint Specific:
 - Acute aquatic toxicity classification by Verhaar (unchecked)
 - Acute aquatic toxicity MOA by OASIS (unchecked)
 - Aquatic toxicity classification by ECOSAR (unchecked)
 - Bioaccumulation - metabolism alerts (unchecked)
- Filter endpoint tree... Panel:**
 - Structure: CCCCC=CCO
 - Structure info:
 - CAS Number: 28069-72-9
 - CAS Smiles relation: High
 - Chemical name(s): trans-2,cis-6-Nonadienol
 - Composition: C9H16O
 - Predefined substance type: Mono constituent
 - Structural Formula: CCC=CCCC=CCO
 - Parameters:
 - Physical Chemical Properties
 - Environmental Fate and Transport
 - Ecotoxicological Information
 - Human Health Hazards
- Bottom Panel:** Metabolism/Transformations.

1. Check protein binding profiles from **General Mechanistic** and **Endpoint specific** group: **Protein binding for skin sensitization by OASIS** profiler
2. Click Apply

Profiling

Profiling the target chemical

- The actual profiling will take up to several seconds depending on the number and type of profilers selected
- The results of profiling automatically appear as a dropdown box under the target chemical (see next screen shot)
- Please note the specific protein-binding profilers
- No protein binding alert has been found for the test compound (trans-2,cis-6-nonadienol).

Profiling

Profiling the target chemical

The target chemical has no protein binding alert. In this respect no skin sensitization effect is expected

1. Double click on "Profile" node to review the profiling results.

Outlook

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

Data Overview

- “Endpoint” refers to the electronic process of retrieving the environmental fate, ecotoxicity and toxicity data that are stored in the Toolbox.
- Data gathering can be executed in a global fashion (i.e., collecting all data for all endpoints) or on a more narrowly defined basis (e.g., collecting data for a single or limited number of endpoints).

Data Case study

- In this example, we limit our data gathering to a single toxicity endpoint (skin sensitization).
- In this example, we collect data from the databases containing experimental results for Skin sensitisation (Skin sensitisation and Skin sensitisation ECETOC).
- **Click** on “Endpoint” in the Toolbox workflow.
- **Expand the** “Human Health Hazards” section
- **Click** on the box to select the relevant databases.
- **Click** on “Gather data” (see next screen shot).

Data Gather data

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Data', 'Definition', 'Data Gap Filling', and 'Report'. The 'Data' menu is open, showing options like 'Gather', 'Import', 'IUCLID', and 'IUCLID6'. Callout 4 points to the 'Gather' button. Callout 1 points to the 'Data' menu. Callout 2 points to the 'Databases' section, where 'Skin Sensitization' and 'Skin sensitization ECETOC' are selected. The main window displays a chemical structure and its associated data fields, including CAS Number, CAS Smiles relation, Chemical name(s), Composition, Molecular Formula, and Predefined substance type. A table of data fields is shown on the right, with a 'No alert found' message.

1. Click Data

2. Select databases related to the target endpoint

3. Click Gather

Data

Gather data

- Toxicity information on the target chemical is electronically collected from the selected dataset(s)
- It should be kept in mind that the search for data and analogues is performed only among the chemicals which are listed in the selected databases, which in this example are **Skin sensitization** and **Skin sensitization ECETOC**
- In this example, there is Positive experimental data for the target chemical(see next screen shots)

Data

Gather data

Toxicity information on the target chemical is electronically collected from the selected datasets. A window with "Read data?" appears. Now the user could choose to collect "all" or "endpoint specific" data.

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes 'Data', 'Import', and 'Export'. The 'Data' menu is open, showing options like 'Gather', 'Import', 'IUCLID', and 'IUCLID6'. A 'Read data?' dialog box is overlaid on the main window, with a callout '1' pointing to the 'OK' button. The dialog box contains the following text:

Read data?
 All endpoints Choose... from Tautomers
 OK Cancel

The background window shows a 'Filter endpoint tree...' dialog with a search filter '1 [target]'. The main window displays a tree view of chemical properties, including 'Structure info', 'Parameters', 'Physical Chemical Properties', 'Environmental Fate and Transport', 'Ecotoxicological Information', 'Human Health Hazards', and 'Profile'. The 'Profile' section is expanded, showing 'General Mechanistic', 'Protein binding by OASIS', 'Protein binding by OECD', 'Protein binding potency', and 'Endpoint Specific'. The 'Endpoint Specific' section is further expanded, showing 'Protein binding alerts for skin sensitization by OASIS' with the result 'No alert found'.

1. Click OK to read all available data

Data Gather data

The screenshot displays the QSAR Toolbox software interface. The top navigation bar includes icons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below this, the 'Data' menu is open, showing options: Gather, Import, IUCLID, and IUCLID6. The main workspace is divided into several panels:

- Documents:** Shows a document titled 'Chemical names containing 'trans-2,cis-6-nonadienol''.
- Databases:** A list of databases with checkboxes. 'Skin Sensitization' and 'Skin sensitization ECETOC' are checked.
- Inventories:** A list of inventories with checkboxes.
- Filter endpoint tree...:** A hierarchical tree of toxicity endpoints. 'Sensitisation' is expanded, showing 'Skin' and 'in Vivo' sub-categories. 'GPMT' is selected under 'in Vivo'.
- 1 [target]:** A data matrix showing a single data point: 'M: Positive'. A callout box with the number '1' points to this entry.

At the bottom of the interface, a blue box contains the text: **1. Available experimental data appears on data matrix.**

Data Gather data

The screenshot shows the QSAR Toolbox interface with the 'Data' menu open. A 'Data points' window is displayed, showing a table of data points. A red callout '2' points to the 'X' button on the window's title bar. Another red callout '1' points to a cell in the table containing the text 'M: Positive'. The table has the following structure:

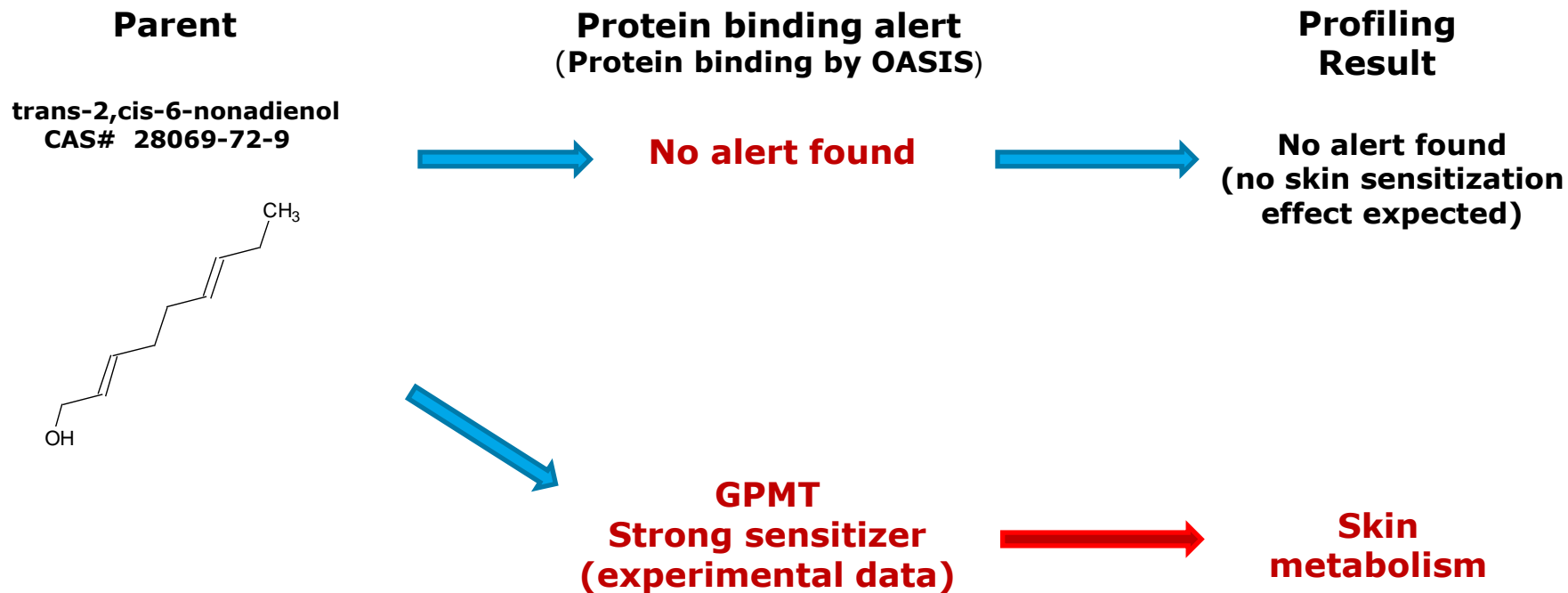
Datapoints	#	Value	Original value	Assay	Assigned SMILES	Author	Comments	Databases
Human Health Hazards;Sensitisation	1	M: Positive (Skin sensitisation II (ECETOC))	Strong sensitizer (Skin sensitisation IV (GPMT))	GPMT	False	Cronin M. T., Basketter D. A	DUV: vvsak sensitiser = Animals showing positive response > 1%	Skin Sensitization

1. **Double-click** on the cell displays metadata information for the observed data
2. **Click** on the X to close the window

Recap

- The first module, which introduces the target chemical, ensure correctness of the structure
- The second module shows that there is no protein binding alert for target chemical
- In the third module, you have found that the target chemical has positive skin sensitization data
- The positive experimental data could be due to skin metabolism
- The study continues with accounting for skin metabolism of target chemical(see next slides).

Recap



Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Data
 - **Handling of skin metabolism**

Handling of skin metabolism of target chemical

- Multiplying target chemical by skin metabolism simulator
- Multiplying target chemical by skin metabolism simulator is accomplished in section **Input**
- The generated metabolites appear in tree like form(see next screen shot)

Handling of skin metabolism of target chemical

Multiplication of target chemical

1. **Click** on the CAS of the target chemical and perform right click on it, then
2. **Select** Multiplication
3. **Select** Metabolism/Transformations
3. **Select** Skin metabolism simulator

Handling of skin metabolism of target chemical

Multiplication of target chemical

The screenshot shows the QSAR Toolbox interface with the 'Skin metabolism simulator' active. The main workspace displays a table with the following data:

Parent chemical [target]	metabolite #1	metabolite #2
28069-72-9	Invalid CAS number: 0-0	Invalid CAS number: 0-0
High	Not applicable	Not applicable
(2E,6Z)-nona-2,6-die		
C9H16O	C9H14O2	C9H14O
Mono constituent	Mono constituent	Mono constituent
CCC=CCCC=CCO	CCC=CCCC=CC(O)=O	CCC=CCCC=CC=O

Callout 1 points to the 'Structure' field in the left sidebar. Callout 2 points to the 'All generated metabolites' row, which is highlighted with a red bracket. A legend at the bottom identifies '1. Parent' and '2. Metabolites'.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - **Handling of skin metabolism of target chemical**
 - **Profiling of metabolites**

Handling of skin metabolism of target chemical

Profiling of metabolites

- This module identifies profilers of target chemical and its metabolites
- Protein binding profiles related to skin sensitization are applied on package of target and metabolites
- Profiling results of target and metabolites are illustrated
- Click on "**Profiling**" to go to the required module
- Apply **Protein binding profilers** (see next screen shot)

Handling of skin metabolism of target chemical

Protein binding result of parent and metabolites

The profiling results indicate no protein binding alerts of target chemical. However, one of the simulated metabolites exhibits interaction with proteins via two different protein binding mechanisms (Michael and Schiff base)

Parent chemical [target]	Metabolite #1	Metabolite #2
No alert found	No alert found	Michael addition
No alert found	No alert found	Michael addition
Not possible to classify	Not possible to classify	Highly reactive (GSH)
No alert found	No alert found	Michael Addition

1. **Go** to Profiling 2. **Check** the profilers related to the target endpoints 3. **Click** Apply

Handling of skin metabolism of target chemical

Recap

- The profiling results indicate no protein binding for target chemical
- One of the simulated skin metabolites has positive protein binding alerts
- This reactive metabolite is used for further read across analysis
- The next two parts of the exercise will focus on the reactive metabolite and find similar analogues of the reactive one (see next screenshot).

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - **Handling of skin metabolism of target chemical**
 - Profiling set of metabolites
 - **Focus of active metabolite**

Handling of skin metabolism of target chemical

Focus of active metabolite

This metabolite is selected for further read-across prediction

1

2

Filter endpoint use...	[target]	metabolite #1	metabolite #2
Structure	<chem>CCCCCCCCCO</chem>	<chem>CCCCCCCCCO</chem>	<chem>CCCCCCCCCO</chem>
Structure info			
Chemical Properties			
Fate and Transport			
Ecotoxicological Information			
Human Health Hazards			
Profile			
General Mechanistic			
Protein binding by OASIS	No alert found	No alert found	Michael addition
Protein binding by OECD	No alert found	No alert found	Michael addition
Protein binding potency	Not possible to classify	Not possible to classify	Highly reactive (GSH)
Endpoint Specific			
Protein binding alerts for skin sensitization ...	No alert found	No alert found	Michael Addition

Context menu for metabolite #2:

- Set as new target
- Edit and set as new target
- Chemical information
- Add in category
- Add target
- Delete
- Focus**
- Query tool matrix (Ctrl+F3)
- Set AOP target
- Copy

1. **Right click** over the active metabolite

2. **Select Focus** from the appeared menu

“Focus” functionality allows the selected metabolite to be used as post target representative of the target chemical

Handling of skin metabolism of target chemical

Focus of active metabolite

The screenshot displays the QSAR Toolbox interface. The top navigation bar includes 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The 'Profiling' section is active, showing 'Apply', 'View', 'New', and 'Delete' options. The 'Documents' panel on the left shows a tree structure for 'Document 1' with sub-items for 'Chemical name: 2,6-nonadien-1-ol', 'Skin metabolism simulator', 'metabolite #1', and 'metabolite #2'. The 'Filter endpoint tree...' panel shows a tree structure with 'General Mechanistic' expanded to 'Protein binding by OASIS'. The '1 [target]' panel shows a chemical structure and a list of endpoints. A red circle highlights the selected metabolite 'Michael Addition' in the '1 [target]' panel.

Filter endpoint tree...

- Structure
- Structure info
- Parameters
- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
- Human Health Hazards
- Profile
 - General Mechanistic
 - Protein binding by OASIS
 - Protein binding by OECD
 - Protein binding potency
 - Endpoint Specific
 - Protein binding alerts for skin sensitization by OASIS

1 [target]

Michael addition
 Michael addition >> Michael addition on alpha,beta-Unsaturated carbonyl compounds
 Michael addition >> Michael addition on alpha,beta-Unsaturated carbonyl compounds >> alpha,beta-Aldehydes
 Schiff base formation
 Schiff base formation >> Schiff base formation with carbonyl compounds
 Schiff base formation >> Schiff base formation with carbonyl compounds >> Aldehydes

Michael addition
 Michael addition >> Polarised Alkenes
 Michael addition >> Polarised Alkenes >> Polarised alkene - aldehydes
 Schiff Base Formers
 Schiff Base Formers >> Direct Acting Schiff Base Formers
 Schiff Base Formers >> Direct Acting Schiff Base Formers >> Mono-carbonyls

Highly reactive (GSH)
 Highly reactive (GSH) >> 2-Alken-1-als (MA)

Michael Addition
 Michael Addition >> Michael addition on alpha,beta-Unsaturated carbonyl compounds
 Michael Addition >> Michael addition on alpha,beta-Unsaturated carbonyl compounds >> alpha,beta-Aldehydes

The selected metabolite appears in a new data matrix.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Data
 - **Handling of skin metabolism of target chemical**
 - Profiling of metabolites
 - Focus of active metabolite
 - **Defining category for active metabolite**

Category Definition

Grouping methods

- The different grouping methods allow the user to group chemicals into chemical categories according to different measures of “similarity” so that within a category data gaps can be filled by read-across.
- 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**

Handling of skin metabolism of target chemical

Category definition for active metabolite

- In this exercise, the reactive metabolite is classified as: Aldehyde by Protein binding alerts for skin sensitization by OASIS category (phase I)
- The identified Protein binding profiler of the reactive metabolite is: Michael Addition >> Michael addition on alpha,beta-Unsaturated carbonyl compounds >> alpha,beta-Aldehydes (phase II)
- In order to expand the initial group of identified analogues the by Protein binding alerts for skin sensitization by OASIS category is used for categorization purposes (phase I).
- Searching for similar analogues is accomplished using the two skin sensitization databases: Skin sensitization and Skin sensitization ECETOC

Handling of skin metabolism of target chemical

Defining Protein binding alerts for skin sensitization by OASIS

- The category **Protein binding alerts for skin sensitization by OASIS** is used
- **Select Protein binding alerts for skin sensitization by OASIS**
- **Click** Define (see next screen shots)

Handling of skin metabolism of target chemical

Defining Protein binding alerts for skin sensitization by OASIS

The screenshot displays the QSAR Toolbox software interface. The top toolbar contains icons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report. The left sidebar shows a 'Documents' tree with 'metabolite #2' selected. The central pane shows a 'Filter endpoint tree...' with 'Structure' selected. The right pane shows a 'Target' dialog box with a list of profiles, including 'Protein binding alerts for skin sensitization by OASIS'. A red circle highlights this profile in the list (labeled '1'). A red circle highlights the 'Define' button in the top toolbar (labeled '2'). A red circle highlights the 'OK' button in the 'Target' dialog box (labeled '3').

1. **Highlight** "Protein binding alerts for skin sensitization by OASIS"
2. **Click** Define
3. **Click** OK.

Handling of skin metabolism of target chemical

Defining US-EPA category

The screenshot shows the QSAR Toolbox interface during the 'Category definition' step. The main window displays a 'Filter endpoint tree...' on the left and a grid of chemical structures on the right. A 'Grouping results' dialog box is open in the center, showing '16 chemicals found.' A red '1' in a blue callout bubble points to the 'OK' button in the dialog. A text box at the bottom of the screenshot reads '1. Click OK to confirm the name of the category'.

Documents

- Document 1
 - chemical name: 2,6-nonadien-1-ol
 - Skin metabolism simulator
 - metabolite #1
 - metabolite #2
 - Protein binding alerts for skin sensitization by OASIS

Filter endpoint tree...

Structure

- Structure info
- Parameters
- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
- Human Health Hazards
- Profile

Grouping results

16 chemicals found.

1

1. Click OK to confirm the name of the category

Handling of skin metabolism of target chemical

Category analogues

- The Toolbox now identifies all chemicals corresponding to *alpha, beta-Aldehydes* by Protein binding alerts for skin sensitization by OASIS listed in the skin sensitization databases.
- 16 analogues including the target chemical are identified; they form a mechanistic category named "*alpha, beta-Aldehydes*", which will be used for further data gap filling.
- The experimental data for analogues in the category appears on datamatrix

Handling of skin metabolism of target chemical

Summary information for Analogues

The experimental results for the analogues appeared on datamatrix

The screenshot displays the QSAR Toolbox interface. On the left, there are navigation tabs: 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. Below these are sub-tabs for 'Categorize' (Define, Define with metabolism, Subcategorize, Combine) and 'Documents' (Document 1: CAS: 28069729, Skin metabolism simulator, metabolite #1, metabolite #2, Protein binding alerts for skin sensitization by OASIS). A 'Filter endpoint tree...' is visible, with 'Sensitisation' selected. The main area is a datamatrix table with 8 columns representing different chemical analogues. The first column is labeled '1 [target]'. The table contains chemical structures in the top row and experimental data in the bottom row. A red box highlights the bottom row, and a blue callout bubble with the number '1' points to the first cell of this row. The highlighted cell contains the text '(2/2)', indicating 2 chemicals and 2 available experimental data points.

Structure	1 [target]	2	3	4	5	6	7	8
Structure info								
Parameters								
Physical Chemical Properties								
Environmental Fate and Transport								
Ecotoxicological Information								
Human Health Hazards								
Acute Toxicity								
Bioaccumulation								
Carcinogenicity								
Developmental Toxicity / Ter ...								
Genetic Toxicity								
Immunotoxicity								
Irritation / Corrosion								
Neurotoxicity								
Photoinduced toxicity								
Repeated Dose Toxicity								
Sensitisation								
in Vivo								
GPMT (2/2)						M: Positive		
HRIPT (6/8)						M: 1.2E+03 µg/cm2	M: Positive	M: Positi
LLNA (15/16)		M: Positive	M: Positive	M: Positive	M: Negative	M: Positive	M: Positive	M: Positi
Miscellaneous (3/3)						M: Positive		
Undefined Assay (2/2)						M: Positive		

Chemical statistic (1) presenting number of chemicals and the available experimental data

Recap

- In this case “Protein binding alerts for skin sensitization by OASIS-category *alpha, beta-Aldehydes* is used for categorization purposes.
- The defined category consists of 16 analogues along with the target chemical
- The available experimental data for these 16 analogues have been collected from two skin sensitization databases.
- But before the user can proceed with the “Filling Data Gap” module, he/she should navigate through the endpoint tree and find the specific gap that will be filled in (in this case Human Health Hazards#Sensitisation#Skin#In Vivo#LLNA#EC3
- In this case we mixed assays and endpoints (see slides #62-64).

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Data
 - **Handling of skin metabolism of target chemical**
 - Profiling of metabolites
 - Focus of active metabolite
 - Category definition for active metabolite
 - **Data gap filling**

Data gap filling

Apply Read across

The screenshot displays the QSAR Toolbox interface with the 'Data Gap Filling' workflow selected. The top toolbar includes buttons for 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The 'Data Gap Filling' button is circled in red and labeled with a '2'. Below the toolbar, the 'Filter endpoint tree...' panel is visible, showing a hierarchical list of endpoints. The 'Sensitisation/Skin/In vivo/LLNA/EC3' endpoint is circled in red and labeled with a '1'. The main table displays data for various target chemicals across different endpoints. The table has 8 columns representing target chemicals and rows representing different endpoints. The 'Sensitisation/Skin/In vivo/LLNA/EC3' row shows 'M: Positive' for targets 1, 3, 4, 5, 6, and 7, and 'M: Negative' for target 5.

Endpoint	1 [target]	2	3	4	5	6	7	8
Structure	<chem>CCCCCCCC</chem>	<chem>CCCCCCCC</chem>	<chem>CCCCCCCC</chem>	<chem>CCCCCCCC</chem>	<chem>CC(C)CCCC</chem>	<chem>CCCCCCCC</chem>	<chem>CCCCCCCC</chem>	<chem>CCCCCCCC</chem>
Ecotoxicological Information								
Human Health Hazards								
Acute Toxicity								
Bioaccumulation								
Carcinogenicity								
Developmental Toxicity / Teratogenicity								
Genetic Toxicity								
Immunotoxicity								
Irritation / Corrosion								
Neurotoxicity								
Photoinduced toxicity								
Repeated Dose Toxicity								
Sensitisation								
Skin								
in Vivo								
GPMT	(2/2)					M: Positive		
HRIPT	(6/8)					M: 1.2E+03 µg/cm2	M: Positive	M: P
LLNA								
EC3	(15/16)	M: Positive	M: Positive	M: Positive	M: Negative	M: Positive	M: Positive	M: P
Miscellaneous	(3/3)					M: Positive		
Undefined Assay	(2/2)					M: Positive		
ToxCast								
Toxicity to Reproduction								
Toxicokinetics, Metabolism and Distribution								

1. Click on the cell corresponding to "Sensitisation/Skin/In vivo/LLNA/EC3" for the target chemical(active metabolite) **2. Select** Read-across

Data gap filling for active metabolite

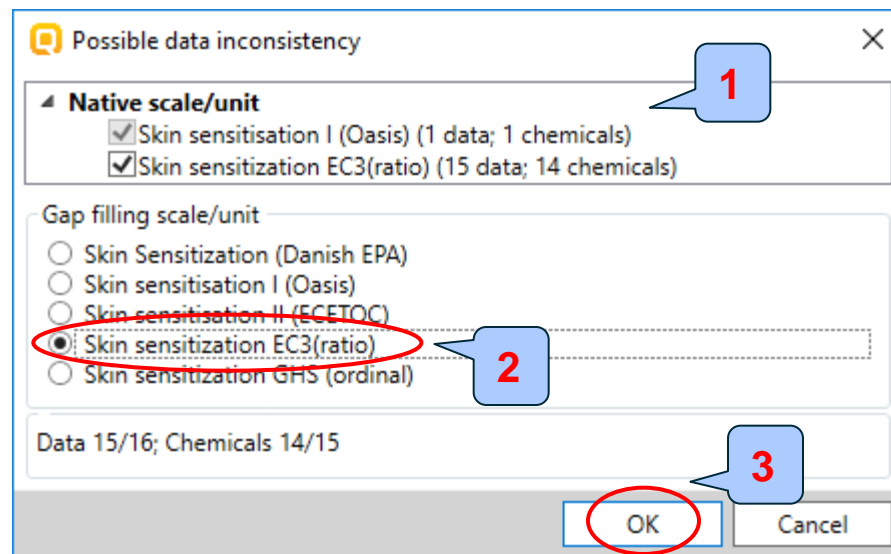
Scale definition

- Skin sensitisation is “qualitative” endpoint for which the results are presented with categorical data (for example: positive; negative; weak sensitizer; strong sensitizer ,etc).
- Skin sensitisation potential of the chemicals came from different authors coded with different names (for example: data from John Moores University of Liverpool are: *Strongly sensitizing, Moderately sensitizing etc.*; data from European centre for Ecotoxicology and Toxicology of chemicals are: *Positive, Negative, and Equivocal*).
- The main purpose of the scales is to unify all data available in the Toolbox databases for a certain endpoint.
- The default scale for Skin Sensitisation is “Skin Sensitisation ECETOC”. It converts all skin data into: Positive, Negative, and Equivocal.

Data gap filling for active metabolite

Scale definition

1. In this case all Scale/unit (1) related to skin sensitization are taken into account
2. Select Skin sensitization EC 3(ratio)
3. **Click** OK to enter data gap filling



Data gap filling for active metabolite

Read-across

- The resulting plot is experimental results of all analogues (Y axis) according to a descriptor (X axis) with the default descriptor of log Kow (see next screen shot).
- The **RED** dot represents predicted results for the target chemical .
- The **BROWN** dots represent the experimental results available for the analogues that are used for the read-across.
- The **BLUE** dot represents the experimental results available for the analogues but not used for read-across.

Data gap filling for active metabolite Read-across

QSAR TOOLBOX

Input | Profiling | Data | Category definition | **Data Gap Filling** | Report

Gap Filling | Workflow

Trend analysis Read across (QSAR) Standardized Automated

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

Documents

ame: 2,6-nonadien-1-ol
:abolism simulator
abolite #1
abolite #2
Protein binding alerts for skin sensitization by (...
Enter GF(RA) with 15 chemicals, 15 data
Protein binding alerts for skin sensitization by (...

Structure

EC3 (14/15)
Miscellaneous (3/3)
Undefined Assay (2/2)
ToxCast
Toxicity to Reproduction
Toxicokinetics, Metabolism and Distribution
Profile

1 [target]	2	3	4	6	7	8	9
	M: Positive	M: Positive	M: Positive	M: Positive	M: Positive	M: Positive	M: Positive
				M: Positive			
				M: Positive			

Descriptors

Prediction

Read-across prediction for EC3, based on 5 values
Predicted: 5.80 %

Active descriptor X log Kow

Select / filter data
Gap filling approach
Descriptors / data
Model/QSAR
Calculation options
Visual options
Information
Miscellaneous

Accept prediction

Data gap filling for active metabolite Subcategorization

- After the available data has been retrieved, the user can then further subcategorize the results according to the following endpoint-specific subcategorizations :
 - Protein binding by OASIS
 - Protein binding by OECD
 - Aquatic toxicity classification by ECOSAR
- These steps are summarized in the next screen shots.

Data gap filling for active metabolite

Subcategorization 1: Protein binding by OASIS

The screenshot displays the QSAR Toolbox software interface during a subcategorization task. The 'Filter endpoint tree...' window shows a list of endpoints, with 'Protein binding by OASIS' highlighted and circled in red, labeled with a blue callout '2'. The 'Subcategorization' window shows a list of categories, with 'Michael addition >> Michael addition on conjugated systems with electron withdrawing groups' circled in red, labeled with a blue callout '3'. The 'Select / filter data' window shows a 'Subcategorize' button circled in red, labeled with a blue callout '1'. The background shows a data table with columns for 'log Kow' and 'M: Positive'.

1. Select filter data/subcategorize
3. Eliminate dissimilar analogues.

2. Select Protein binding by OASIS.

Data gap filling for active metabolite

Subcategorization 2: Organic functional groups

The screenshot displays the QSAR Toolbox software interface during the 'Subcategorization' step. The main workspace shows a grid of chemical structures (11-15) with their corresponding predicted categories (e.g., 'M: Positive'). The 'Subcategorization' dialog box is open, showing a list of categories on the left and a list of selected categories on the right. A 'Select/filter data' panel is visible on the right side of the workspace. Three callout boxes with numbers 1, 2, and 3 highlight key steps: 1. The 'Subcategorize' button in the 'Select/filter data' panel. 2. The 'Organic functional groups' category in the 'Subcategorization' dialog. 3. The 'Remove selected' button in the 'Subcategorization' dialog.

1. **Open** Select/filter data/Subcategorize dissimilar chemicals
2. **Select** Organic functional groups
3. **Remove**

Data gap filling for active metabolite

Results after subcategorization

The screenshot displays the QSAR Toolbox interface during a data gap filling process. The top navigation bar shows the current step is 'Data Gap Filling'. The left sidebar contains a 'Documents' list and 'Data Gap Filling Settings' with options for endpoint and chemical relevance. The central workspace features a 'Filter endpoint tree' on the left, a table of results in the middle, and a scatter plot at the bottom. The table shows a 'Read-across prediction for EC3, based on 2 values' with a predicted value of 4.00%. The plot shows EC3 [%] on the y-axis (ranging from 3 to 5) and log Kow on the x-axis (ranging from 1.6 to 3.6). A callout box labeled '1' points to the 'R: Positive' cell in the table, and another callout box labeled '2' points to the 'Accept prediction' button.

Structure	1 [target]	4	7
<chem>CCCCCCCCCCCCCCCC</chem>	<chem>CCCCCCCCCCCCCCCC</chem>	<chem>CCCCCCCCCCCCCCCC</chem>	<chem>CCCCCCCCCCCCCCCC</chem>
	(1/1)		M: Positive
	(3/3)	R: Positive	M: Positive

EC3 [%]

log Kow

Read-across prediction for EC3, based on 2 values
Predicted: 4.00%

Accept prediction

1. The predicted EC3 is 4% 2. **Accept** prediction

Data gap filling for active metabolite

Read-across prediction

The screenshot displays the QSAR Toolbox interface during a data gap filling process. The top navigation bar includes 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. Below this, there are tabs for 'Gap Filling' and 'Workflow'. The main workspace is divided into several sections:

- Documents:** Lists various documents and data points.
- Filter endpoint tree...:** A hierarchical tree of endpoints. The 'Skin' endpoint is expanded, showing sub-endpoints like 'in Vivo', 'GPMT', 'HRIPT', 'LLNA', 'EC3', 'Miscellaneous', and 'Undefined Assay'. The 'Skin' endpoint is highlighted with a red circle and labeled 'I: Positive'.
- Data Gap Filling Settings:** Includes checkboxes for 'Only endpoint relevant' and 'Only chemical relevant', and a section 'At this position:' with a table of workflow counts.
- Data Matrix:** A table with 8 columns and multiple rows. The first column is labeled '1 [target]'. The matrix shows results for various endpoints across different chemical groups. A red circle highlights the 'I: Positive' result for the 'Skin' endpoint in column 1.

The data matrix is as follows:

Endpoint	1 [target]	2	3	4	5	6	7	8
GPMT (2/2)						M: Positive		
HRIPT (6/8)						M: 1.2E+03 µg/cm2	M: Positive	M: Positive
LLNA (16/17)	I: Positive	M: Positive	M: Positive	M: Positive	M: Negative	M: Positive	M: Positive	M: Positive
EC3 (3/3)						M: Positive		
Miscellaneous (2/2)						M: Positive		
Undefined Assay (2/2)						M: Positive		

The read-across prediction result for metabolite appears on data matrix

Data gap filling for active metabolite

Interpreting Read-across

- In this example, all analogues have same protein binding alerts
- All analogues exhibit positive skin sensitization
- The same positive sensitising potential is therefore predicted for the target (i.e. active metabolite).
- The prediction of metabolite is further transferred to the parent chemical using Independent MOA (see next screen shots)

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - **Handling of skin metabolism of target chemical**
 - Profiling set of metabolites
 - Focus of active metabolite
 - Category definition for active metabolite
 - Data gap filling
 - **Assigning prediction of active metabolite to parent**

Handling skin metabolism of target chemical

Assigning data to parent chemical

1. **Select** Skin metabolism simulator from document tree

2. **Select** The cell from data matrix corresponding to the parent chemical.

Handling skin metabolism of target chemical

Assigning data to parent chemical

The screenshot displays the QSAR Toolbox interface with the following components:

- Navigation Bar:** Includes buttons for Input, Profiling, Endpoint, Category Definition, Data Gap Filling (active), and Report.
- Left Panel:** Shows 'Data Gap Filling Method' with 'Independent MOA' selected (marked with a red '1'). Other options are 'Similar MOA' and 'Specific models'. Below is the 'Target Endpoint' section, currently set to 'Human Health Hazards Sensitisation Skin In Vivo'.
- Table:** A table with 6 columns representing different target chemicals. The first row shows chemical structures. The second row shows 'Ecotoxicological Information'. The third row shows 'Human Health Hazards' with a sub-tree expanded to 'Skin', which includes 'In Chemico', 'In Vitro', and 'In Vivo'. The 'In Chemico' cell in the third column is circled in red and contains the text '(3/3) M. Positive'. Other cells in the table contain 'M: Positive' or 'R: Positive'.
- Buttons:** An 'Apply' button (marked with a red '2') is located at the top left of the main workspace.
- Footer:** A blue banner at the bottom contains the instructions: '1. Check Independent MOA' and '2. Click Apply'.

Handling skin metabolism of target chemical

Assigning data to parent chemical

The screenshot displays the QSAR Toolbox software interface during a 'Data Gap Filling' session. The top navigation bar includes icons for 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The left sidebar shows a 'Documents' panel with a tree view of the 'Skin metabolism simulator' and a 'Data Gap Filling Settings' panel with checkboxes for 'Only endpoint relevant' and 'Only chemical relevant'. The main workspace is divided into a 'Filter endpoint tree...' on the left and a data table on the right. The table has columns for 'Parent Chemical', 'metabolite #1', and 'metabolite #2'. A 'Possible data inconsistency' dialog box is open, showing a list of 'Gap filling scale/unit' options, with 'Skin sensitization EC3(ratio)' selected and circled in red. Below the dialog, a table shows 'IMOA: Positive' and 'R: Positive' entries, both circled in red. A blue callout box at the bottom states: 'The prediction for the metabolite is assigned to the parent chemical'.

Handling skin metabolism of target chemical

Assigning data to parent chemical

- The following actions (steps) are used for assigning data to parent chemical:
 - Accept prediction
 - Return to matrix
- Final prediction for the parent compound labeled as CI (Component based Independent mode) is **positive** (see next screen shot)

Recap

- The target chemical **trans-2,cis-6-nonadienol** has been entered into the system.
- It has been profiled by Protein binding profilers; no protein binding has been found for target chemical.
- Positive experimental data has been retrieved for target chemical.
- Skin metabolism of target chemical is investigated. Two of simulated skin metabolites have positive protein binding alerts.
- These metabolites have similar protein binding alert: **α,β -unsaturated aldehydes**. One of the reactive metabolites is used for further read across analysis.
- No experimental data for the selected metabolite has been found, so the category of similar analogues has been defined.
- The initial group of analogues is defined by Protein binding alerts for skin sensitization by OASIS.
- 16 analogues including the target chemical are identified; they form a mechanistic category “ **α,β -unsaturated aldehydes**”, which will be used for gap filling.
- Read-across is used for data gap filling.
- Protein binding categories have been used for refining the initial category.
- Positive skin sensitization has been predicted for the active metabolite.
- Positive prediction for reactive metabolite has been transferred to the parent chemical using Independent MOA.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Handling of skin metabolism of target chemical
- **Report**

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

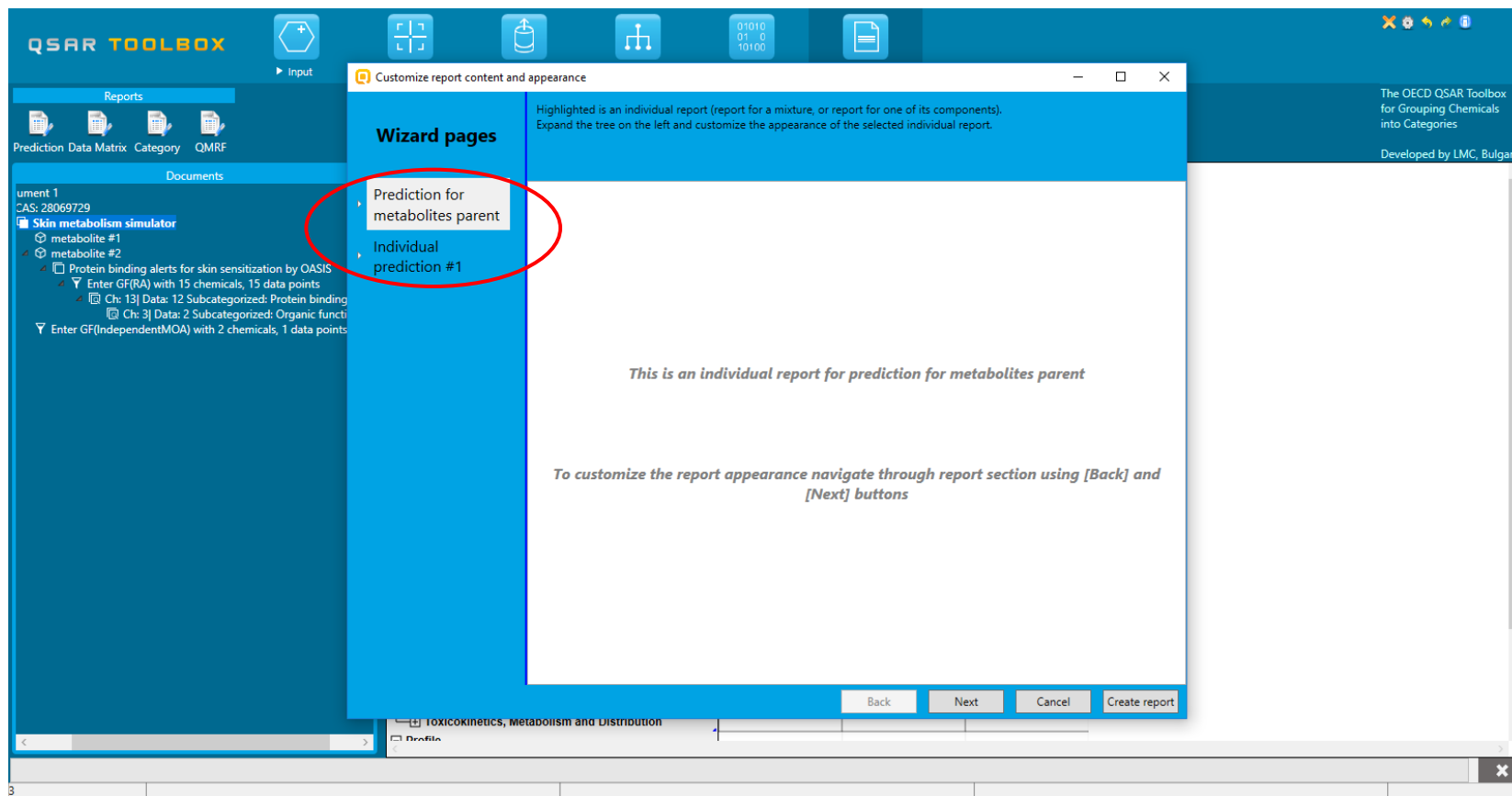
Report

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Report' (callout 2). The left sidebar has a 'Reports' section (callout 3). The main workspace displays a 'Filter endpoint tree...' on the left and a data table on the right. The table has columns for 'Parent chemical [target]', 'metabolite #1', and 'metabolite #2'. The 'IMOA: Positive' prediction is circled in red (callout 1).

Parent chemical [target]	metabolite #1	metabolite #2
<chem>CCCCCCCCCO</chem>	<chem>CCCCCCCCCO</chem>	<chem>CCCCCCCCCO</chem>
(1/1) IM: Positive		
(2/2) IMOA: Positive		R: Positive

1. **Select** prediction
2. Section Report
3. Select Prediction and create report

Report




The user could select the appropriate sections to create the report

Report

The screenshot displays the QSAR Toolbox software interface. On the left, a 'Documents' tree shows a project named 'Skin metabolism simulator' with sub-items for 'metabolite #1' and 'metabolite #2'. A 'Wizard pages' panel is open, with 'Prediction for metabolites parent' selected and circled in red. The main window shows a report titled 'QSAR Toolbox prediction for multicomponent substance' based on observed and predicted data for metabolites. The report includes a date of 11 Jul 2017 and a table of target information.

Target information

Structural information	Numerical identifiers	Chemical names
SMILES: <chem>CCC=CCCC=CCO</chem>	EC#: N/A CAS#: 28069-72-9 Other: N/A	(2E,6Z)-nona-2,6-die n-1-ol 2,6-Nonadien-1-ol, (2E,6Z)- 2,6-Nonadien-1-ol, (E,Z)-
Structure 		

Prediction for metabolites parent

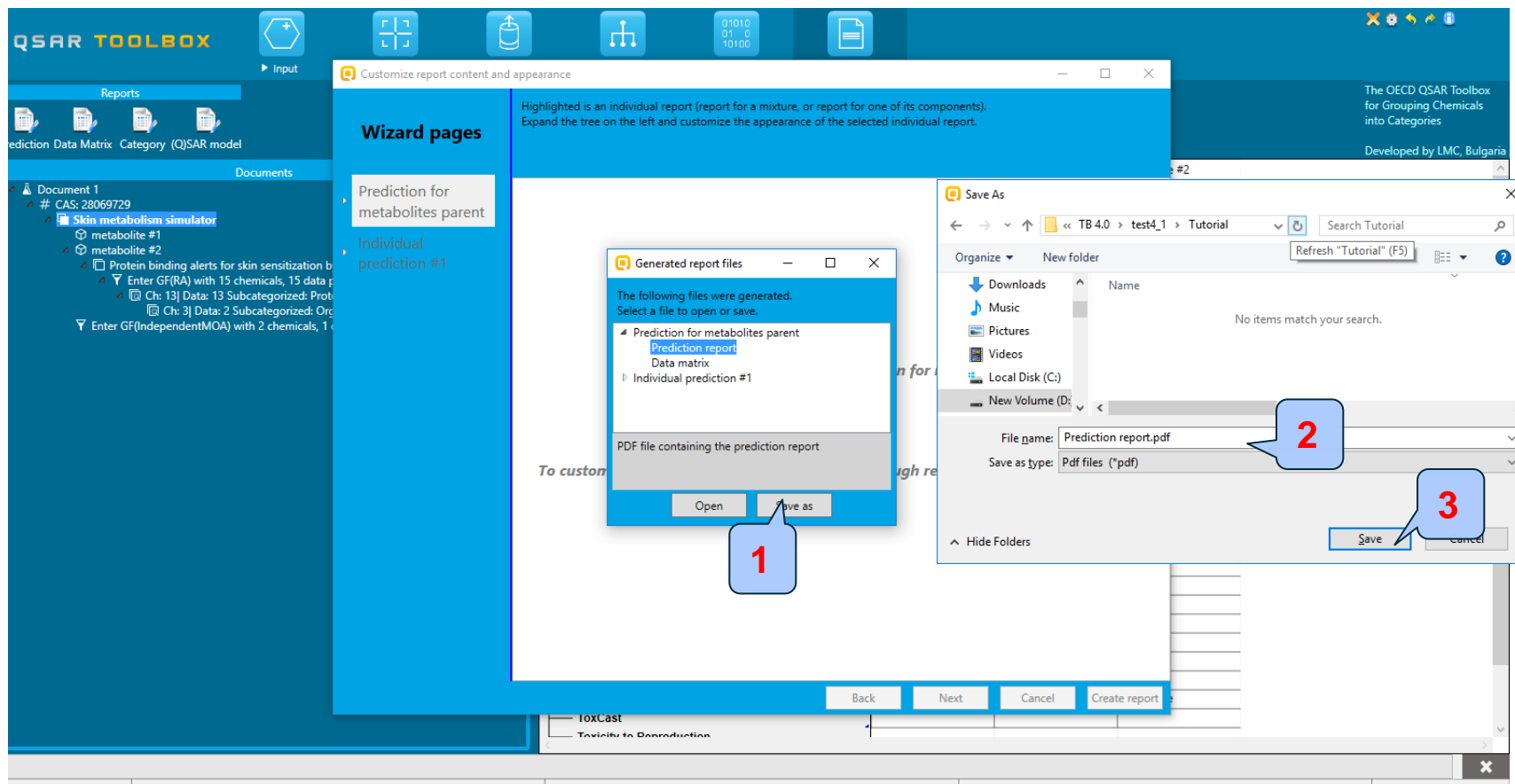
Outlook

- Background
- Objectives
- The exercise
- Workflow
 - Input
 - Profiling
 - Endpoint
 - Handling of skin metabolism of target chemical
 - Report
- **Save the prediction result**

Saving the prediction result

- This functionality allow storing/restoring the current state of Toolbox documents including loaded chemicals, experimental data, profiles, predictions etc, on the same computer. The functionality is implemented based on saving the sequence of actions that led to the current state of the Toolbox document and later executing these actions in the same sequence in order to get the same result(s).
- Saving/Loading the file with TB prediction is shown on next screenshots

Saving the prediction result



1. **Click** on Save as button; 2. **Define** name of the file; 3. **Click** Save button

Open saved file

Once the file has been saved **1. Go to Input**; **2. Click Open**; **3. Find and select file**; **4. Click Open**