

OECD QSAR Toolbox v.3.4

Predicting skin sensitisation potential of eugenol (CAS 97-53-0) using a new categorization tool taking into account its abiotic activation

Outlook

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

Background

- This is a step-by-step presentation designed to take the user through the Toolbox workflow for predicting skin sensitization potential of eugenol using a newly implemented categorization tool taking into account its abiotic activation.

Outlook

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

Objectives

This presentation demonstrates a number of functionalities of the Toolbox:

- Profiling the target chemical.
- Identifying analogues of the target chemical.
- Filling data gaps for target chemical by read-across.
- Profiling target chemical taking into account its (a)biotic activation.
- Identifying analogues of the target using a new categorization functionality allowing (a)biotic activation to be taken into account.
- Filling data gaps by read across when (a)biotic activation is taken into account.

Outlook

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

The Exercise

- In this exercise we will predict the skin sensitization potential of target chemical **Eugenol** [**CAS# 97-53-0**].
- Profile and gather data for the target chemical.
- Two types categorizations are applied:
 - Identifying analogues using well-known categorization group
 - Identifying analogues based on autoxidation activation of the target illustrating new categorization functionality
- Filling data gap by read-across.

Outlook

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

Workflow

- **As you know the Toolbox has 6 modules which are typically used in sequence:**
 - Chemical Input
 - Profiling
 - Endpoint
 - Categorization
 - 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
- Chemical IDs such as EC number, ENECS number

B. Group of chemicals

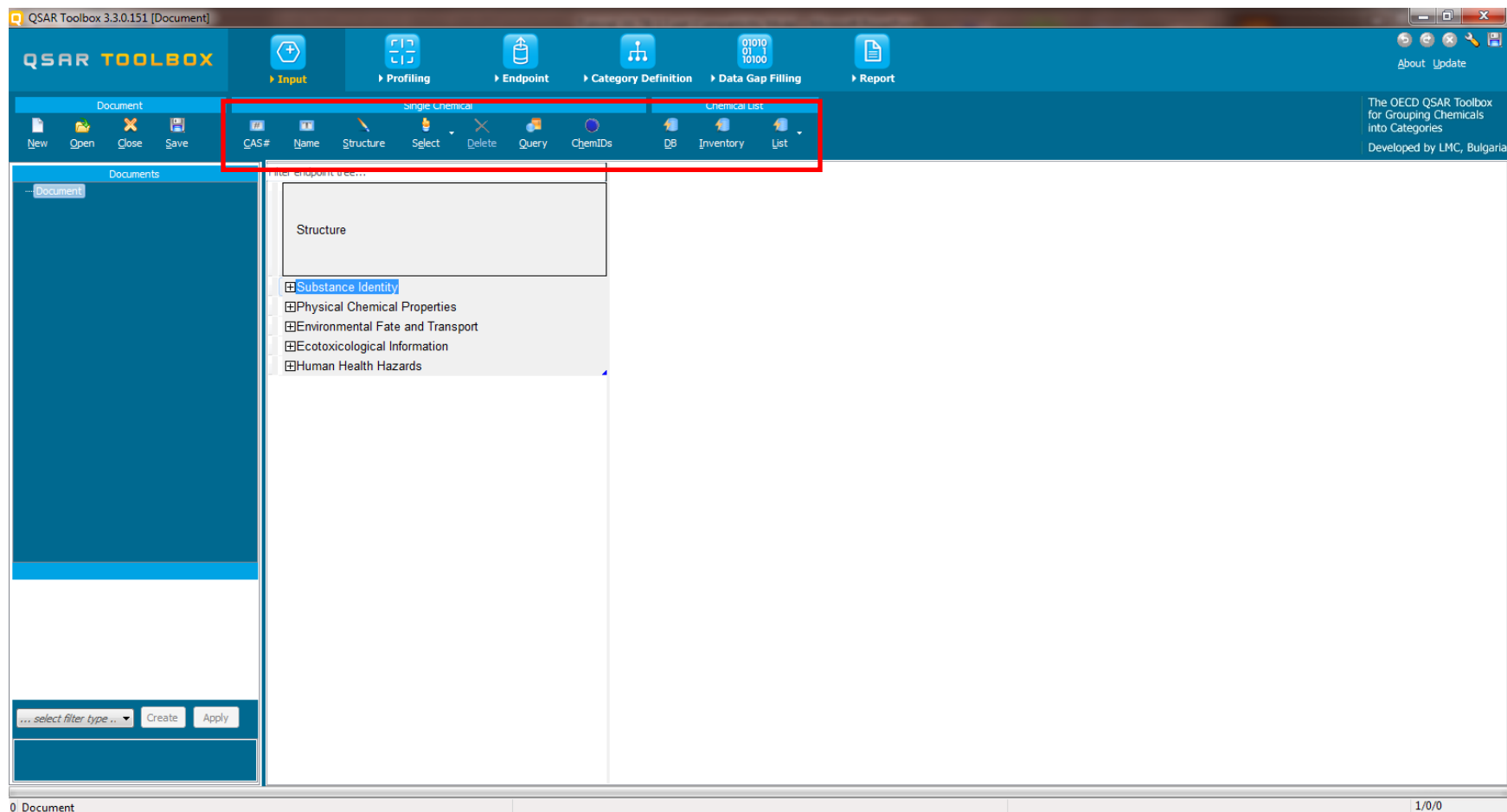
- User List/Inventory
- Specialized Databases

Chemical Input

Input Screen

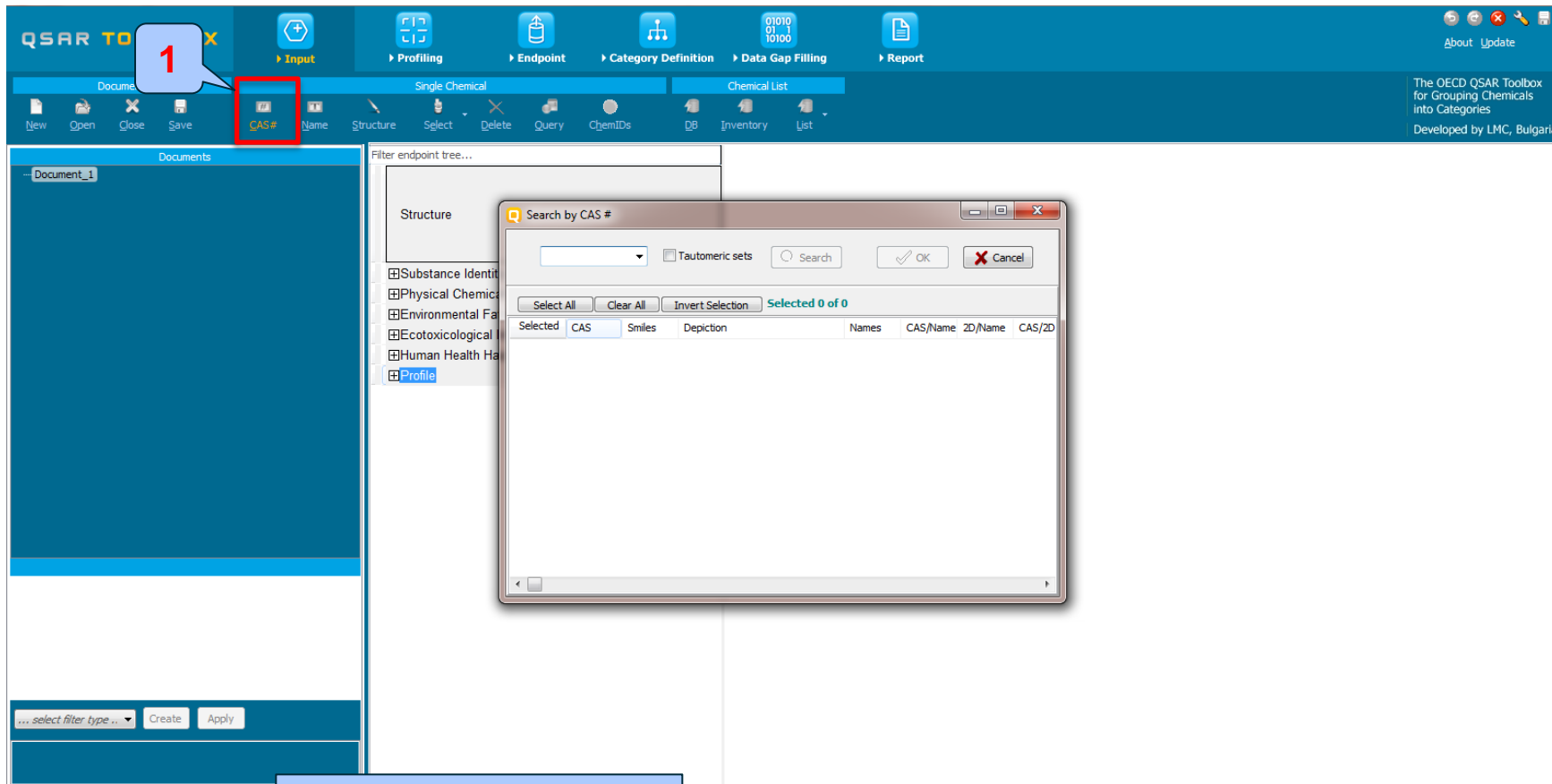
- Open the Toolbox.
- The six modules in the workflow are seen listed next to “QSAR TOOLBOX” title.
- **Click** on “Input” (see next screen shot)

Chemical Input Input Screen



Chemical Input

Input target chemical by CAS#



1. Click on "CAS#"

Chemical Input

Enter CAS# of 2-methoxy-4-(2-propenyl)phenol (Eugenol)

Search by CAS #

97-53-0 ☐ Tautomeric sets

Select All Clear **1** Invert Selection Selected 1 of 1 **2**

Selected	CAS	Str	Depiction	N	CAS/Name	2D/Name	CAS/2D
1. Yes	97-53-0	COc1cc(C=C)ccc1O			1:: Low (1:: Low (1:: High		

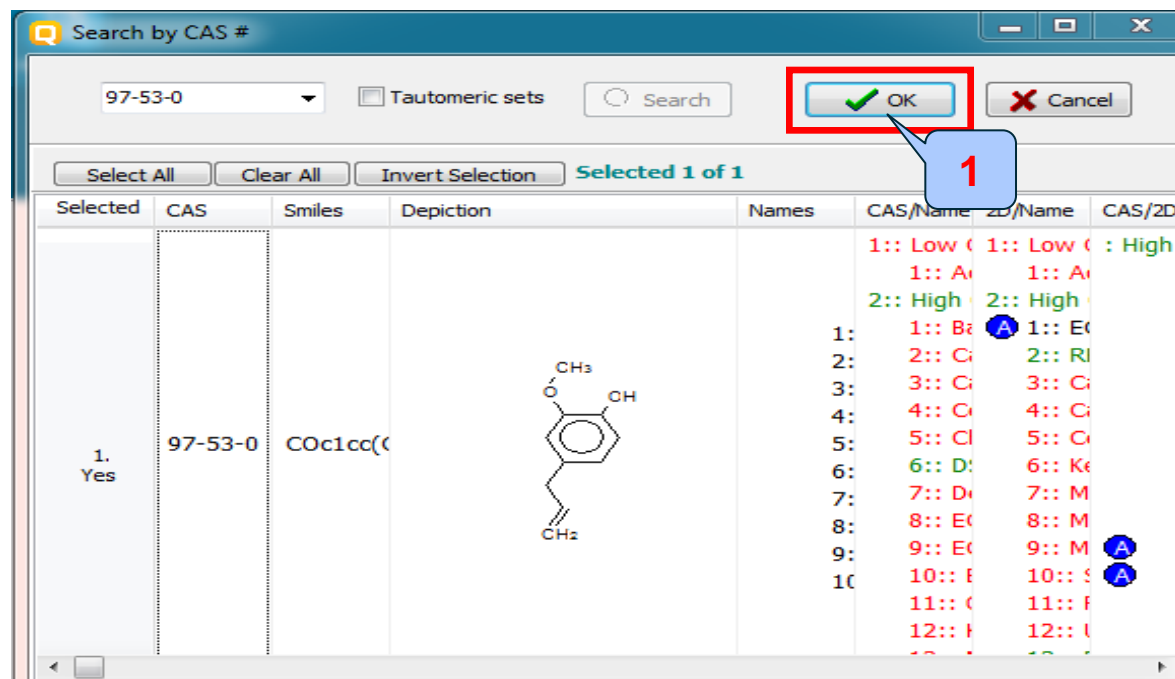
3

1. **Enter** the CAS# in the blank field; 2. **Click** "Search" button; 3. **Press** "OK"

Chemical Input

Target chemical identity

The Toolbox now searches the databases to find out if the CAS# you entered is linked to a molecular structure stored in the Toolbox. It is displayed as a 2-dimensional depiction.



1. Click "OK" to enter the target structure into data matrix

Chemical Input

Target chemical identity

- **Double click** "Substance Identity" displays the chemical identification information.
- The user should note that existing names of the target chemical are presented in different colours. This indicates the reliability of relation CAS-Name-SMILES 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 with the 'Input' tab selected. The 'Chemical List' section is open, showing a list of chemicals. A red circle highlights the 'Substance Identity' section, which contains the following information:

- CAS Number: 97-53-0
- EINECS: 2025891
- Chemical Name: eugenol (4-allyl-2-methoxyphenol)
- Molecular Formula: C10H12O2
- Structural Formula: COc1cc(CC=C)ccc1O

The chemical structure of eugenol is shown above the list. The 'Filter endpoint tree...' section on the left shows the 'Structure' endpoint selected. The 'Documents' section on the left shows a document named 'Document_1' with the CAS number 97-53-0.

Chemical Input

Target chemical identity

The colour code indicates the reliability of the chemical identifier:

- **Green:** There is a high reliability between the identifier and the structure. This colour is applied if the identifier is the same in several quality assured databases.
- **Yellow:** There is only a moderate reliability between the identifier and the structure. The colour is applied if the identifier is the same in several databases for which the quality assurance could not be established.
- **Red:** There is a poor reliability between the identifier and the structure. The colour is applied if the identifier is allocated to different structures in different databases.

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 alerts for skin sensitization by OASIS v1.4 and clicking on “View” (see next screen shot)).

Profiling Side-Bar to Profiling

The screenshot shows the QSAR Toolbox Profiling Scheme Browser. The interface is divided into several sections:

- Left Panel (Profiling methods):** Contains a list of various profiling methods. A red circle labeled '1' highlights the 'Protein binding alerts for skin sensitization' method. A red circle labeled '2' highlights the 'View' button.
- Top Panel (List with categories):** Displays a hierarchical list of categories. A red circle labeled '3' highlights the 'Aldehydes' category under 'Schiff base formation'.
- Right Panel (Textual description):** Provides a detailed description for the selected category. It includes:
 - Mechanistic Domain:** Schiff base formation
 - Mechanistic Alert:** Schiff base formation with carbonyl compounds
 - Structural Alert:** Aldehydes
 - Chemical reaction:** A chemical reaction showing the formation of a Schiff base from an aldehyde and a primary amine: $R-C(=O)H + Pr-NH_2 \rightarrow R-C(=N-Pr)H$.
 - Textual description:** A paragraph explaining that aldehydes are highly reactive molecules that can form Schiff bases with amino groups on proteins, leading to skin sensitization.

Red boxes and arrows highlight specific elements: 'Domain' points to 'Schiff base formation', 'Mechanistic alert' points to 'Schiff base formation with carbonyl compounds', and 'Structural alert' points to 'Aldehydes'.

1. **Highlight** the profiler
2. **Click** "View"
3. **Click** over "Aldehydes" to see textual description associated with the category. In order to see more details about structural boundaries coding the rule you should click "Advanced" button (4) (see next slide)

Profiling Side-Bar to Profiling

The screenshot displays the QSAR Toolbox Profiling Scheme Browser. The left sidebar contains the 'Profiling' tab with a tree of methods. The main window shows a selected rule for 'Michael addition on phosphoranylidene compounds'. A red circle labeled '1' highlights the 'Structural boundaries' section, and another red circle labeled '2' highlights the 'Structural fragment' section, which shows a chemical structure of a carbonyl group (R₂C=O).

1. Illustrates structural boundary coding the rule
2. Illustrates structural fragment used for defining the rule

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 OECD – general mechanistic
 - Protein Binding Potency – general mechanistic
 - Protein binding alerts for skin sensitization by OASIS v1.4 – endpoint specific

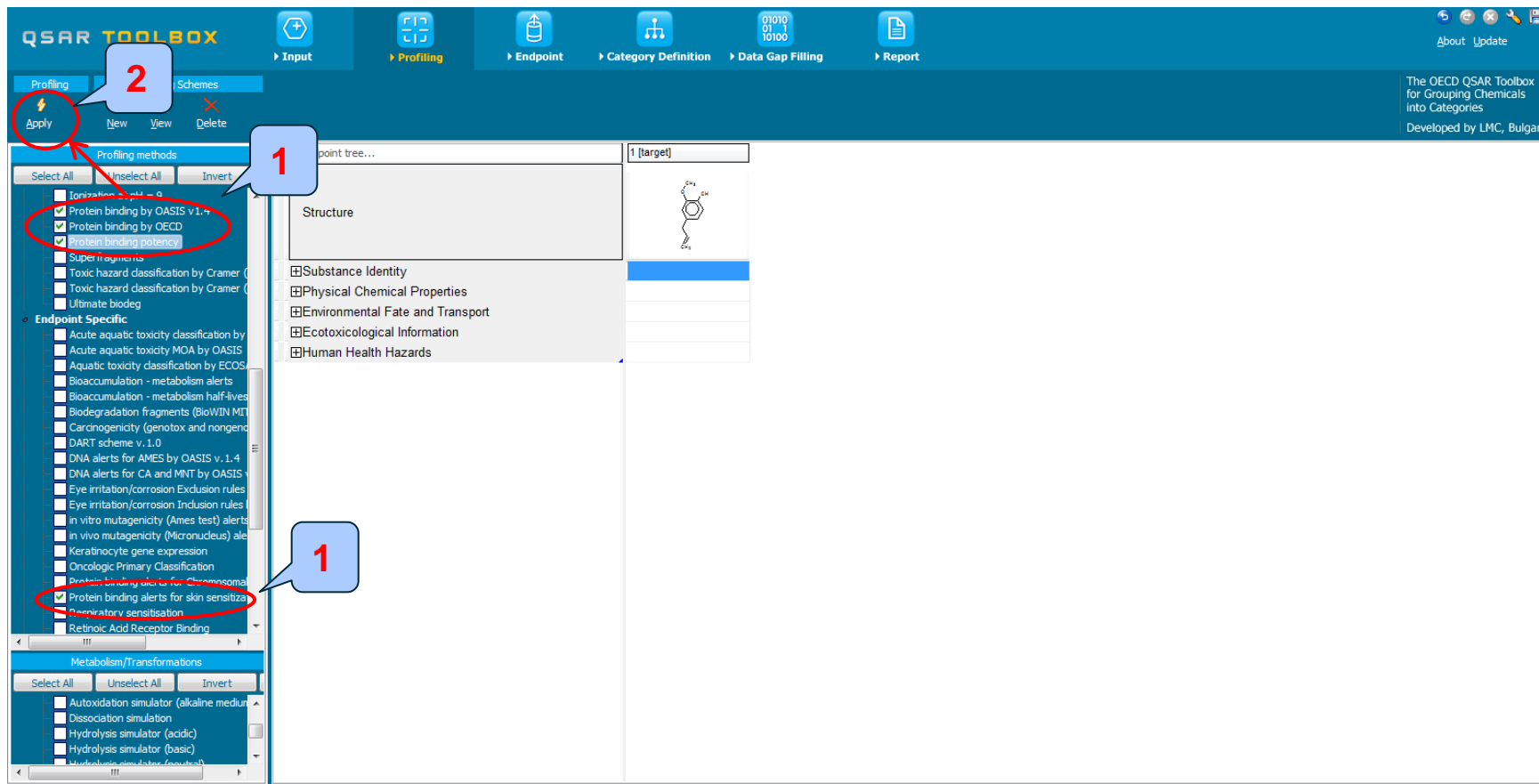
Profiling

Profiling the target chemical

- **Click** in the box next to the name of the profiling methods related to the target endpoint.
- 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 are relevant to skin sensitization effect (see next screen shot).

Profiling

Profiling the target chemical



1. **Select** protein binding profiles from “General Mechanistic” and “Endpoint specific” group mentioned on slide 26
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 result obtained by the specific protein-binding profilers.
- No protein binding alert has been found for the target compound (eugenol) based on three protein binding profilers. Therefore no skin sensitization effect is expected.

Profiling

Profiling the target chemical

QSAR Toolbox 3.3.0.151 [Document_1]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Profiling Profiling Schemes

Apply New View Delete

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

Profilig methods

Select All Unselect All Invert About

- ☐ Ionization at pH = 4
- ☐ Ionization at pH = 7.4
- ☐ Ionization at pH = 9
- ☒ Protein binding by OASIS v1.3
- ☒ Protein binding by OECD
- ☒ Protein binding potency
- ☐ Superfragments
- ☐ Toxic hazard classification by Cramer (extension)
- ☐ Toxic hazard classification by Cramer (original)
- ☐ Ultimate biodeg

Endpoint Specific

- ☐ Acute aquatic toxicity classification by Verhaar (Modif)
- ☐ Acute aquatic toxicity MOA by OASIS
- ☐ Aquatic toxicity classification by ECOSAR
- ☐ Bioaccumulation - metabolism alerts
- ☐ Bioaccumulation - metabolism half-lives
- ☐ Biodegradation fragments (BioWIN MITI)
- ☐ Carcinogenicity (genotox and nongenotox) alerts by ISS
- ☐ DART scheme v.1.0
- ☐ DNA alerts for Ames, MN and CA by OASIS v.1.3
- ☐ Eye irritation/corrosion Exclusion rules by BIR
- ☐ Eye irritation/corrosion Inclusion rules by BIR
- ☐ In vitro mutagenicity (Ames test) alerts by ISS
- ☐ In vivo mutagenicity (Micronucleus) alerts by ISS
- ☐ Keratinocyte gene expression
- ☐ Oncologic Primary Classification
- ☐ Protein binding alerts for Chromosomal aberration by ISS
- ☒ Protein binding alerts for skin sensitization by OASIS v1.3

Metabolism/Transformations

Select All Unselect All Invert About

Documented

- ☐ Observed Mammalian metabolism
- ☐ Observed Microbial metabolism
- ☐ Observed Rat In vivo metabolism
- ☐ Observed Rat Liver S9 metabolism

Simulated

Autocatalytic simulation

Filter endpoint tree...

Structure

1 [target]

Substance Identity

Physical Chemical Properties

Environmental Fate and Transport

Ecotoxicological Information

Human Health Hazards

Profile

- General Mechanistic
 - Protein binding by OASIS v1.3
 - Protein binding by OECD
 - Protein binding potency
- Endpoint Specific
 - Protein binding alerts for skin sensitization by OASIS v1.3

No alert found

No alert found

Not possible to cla...

No alert found

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

Outlook

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

Endpoint 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).

Endpoint

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).

Endpoint Gather data

1

4

2

3

1. **Click** on "Endpoint"

2. **Expand** the "Human Health Hazards" section

3. **Select** databases related to the target endpoint

4. **Click** "Gather"

Endpoint

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).

Endpoint 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', 'Export', 'Delete', and 'Tautomerize'. The 'Endpoint' menu is open, showing a list of databases and inventories. A 'Read data?' dialog box is overlaid on the interface, with a red '1' in a blue circle pointing to the 'OK' button. The dialog box contains the text 'Read data?' and three radio buttons: 'All endpoints' (selected), 'Choose...', and 'from Tautomers'. The 'OK' button is highlighted with a green checkmark, and the 'Cancel' button is highlighted with a red X. Below the dialog box, a blue box contains the text '1. Click "OK" to read all available data'.

Endpoint Gather data

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Data', 'Import', 'Export', 'Delete', and 'Tautomerize'. The 'Endpoint' tab is selected. The 'Filter endpoint tree...' panel on the left shows a tree structure of endpoints. The 'Structure' panel on the right shows the chemical structure of the target chemical. The '1 [target]' panel on the right shows the results of the search, with a callout box highlighting the 'Skin' endpoint.

Filter endpoint tree...

- Environmental Fate and Transport
- Ecotoxicological Information
- Human Health Hazards
 - Acute Toxicity
 - Bioaccumulation
 - Carcinogenicity
 - Developmental Toxicity / Teratogenicity
 - Genetic Toxicity
 - Immunotoxicity
 - Irritation / Corrosion
 - Neurotoxicity
 - Photoinduced Toxicity
 - Repeated Dose Toxicity
 - Sensitisation
 - Respiratory Tract (1/1) M: Negative
 - Skin
 - In Chemico (1/1) M: Positive
 - In Vitro (1/3) M: Positive, Positive
 - In Vivo (1/1) M: Positive
 - GPMT (1/1) M: Positive
 - HRIPPT (1/1) M: Positive
 - LLNA (1/1) M: Positive
 - Undefined Assay (1/1) M: Positive
 - Toxicity to Reproduction
 - Toxicokinetics, Metabolism and Distribution

1. Positive experimental data for skin sensitization is found for the target chemical.

Endpoint Gather data

The screenshot shows the QSAR Toolbox software interface. The 'Endpoint' tab is selected in the top menu. The left sidebar shows a tree view of databases, with 'Human Health Hazards' expanded. The main window displays a 'Filter endpoint tree...' on the left and a 'Structure' view on the right. A 'Data points' dialog box is open, showing a table of data points. A red circle with the number '1' points to the 'EC3' cell in the 'Endpoint' column of the first row.

#	Endpoint	Value	Original value	Organ	Reference source	Phylum (common name)	Phylum	Test method / Data source	Type of method	Year	Test organisms
1	EC3	Positive (Skin sensitisation II (ECETOC))	13 % (Skin sensitization EC3 (ratio))	Skin	Dermatitis, 16 (4): 1-46	Vertebrates	Chordata	LLNA	in Vivo	2005	mouse

Below the table, there are two rows of data points:

- (1/1) M: Positive
- (1/1) M: Positive

1. **Double-click** on the cell displays metadata information for the observed data

Recap

- The first module, introduces the target chemical, ensure for correctness of the structure.
- The second module shows that there is no protein binding alert for the target chemical.
- In the third module, you have found that the target chemical has positive skin sensitization data.
- In the further read-across analysis we will try to reproduce positive skin sensitization data.
- The study continues with identifying analogues and applying read-across.

Outlook

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

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>

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.

Basic guidance for category formation and assessment

Suitable categorization phases:

1. Structure-related profilers (for primary categorization).
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**

Category Definition

Grouping methods – phase I

Suitable Categorization/Assessment Phases

Phase I. Structure based

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

Broad grouping Endpoint Non-specific

Each of the above grouping method is applied to the target chemical and number of the identified analogue is provided below. In order to preserve the basic functional groups available within the molecule: Allyl, Ether and Phenol, OFG is used for categorization purposes. US-EPA and ECOSAR are not used because they omit the other two important functionalities: Allyl and Ether. Str. similarity identifies small set of analogues and apparently could not be used for categorization.

Phase I categorization in Toolbox

Filter endpoint tree...

Structure

1 [target]

Substance Identity

Physical Chemical Properties

Environmental Fate and Transport

Ecotoxicological Information

Human Health Hazards

Profile

Predefined

OECD HPV Chemical Categories

US-EPA New Chemical Categories

Endpoint Specific

Aquatic toxicity classification by ECOSAR

Empiric

Organic Functional groups

Organic Functional groups (nested)

Not categorized

Phenols (Acute toxicity)

Phenols

Alkene

Allyl

Aryl

Ether

Phenol

Precursors quinoid compounds

Allyl

Ether

Overlapping groups

Precursors quinoid compounds

128 analogues are identified.

102 analogues are identified

7 analogues are identified (in case all categories are preserved)
11 analogues are identified (in case Allyl, Ether and Phenol are preserved)

21 analogues are identified

Structural similarity, Dice ACF, 50%

Category Definition

Define category by OFG

1. Select "OFG"

2. Click "Define"

Combination of all seven organic functional groups identified seven analogues only (3). In order to expand the initial group the categories "Allyl", "Ether" and "Phenol" are used only.

4. Click "Cancel". See next slide

Category Definition

Define category by OFG

2

Define

Define with metabolism

Subcategorize

Combine

Clustering

Delete

Delete All

1

Organic Functional groups

Target(s) profiles

Allyl

Ether

Phenol


3

4

5

6

All groups except "Allyl", "Ether" and "Phenol" are removed from the selection

1. **Select** "OFG";
2. **Click** "Define" button;
3. **Select** "Alkenes", "Alkoxy" "Aryl" and "Precursors quinoid compounds" (highlighted in blue) and **click** arrow down  to remove them. They are moved in the panel down called "All profiles";
4. Arrow down
5. "Allyl", "Ether" and "Phenol" should remain in the upper panel only
6. **Click** "OK" button

Category Definition

Define category by OFG

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Category Definition' menu is open, showing options like 'Define', 'Define with metabolism', 'Subcategorize', 'Combine', 'Clustering', 'Delete', and 'Delete All'. The 'Filter endpoint tree...' dialog is open, showing a chemical structure and a list of categories. A 'Warning' dialog (1) asks if the user wants to continue with different categories. A 'Define category name' dialog (2) shows the category name being defined as '>Ether<AND>Phenol (Organic Functional groups)'. A 'Read data?' dialog (3) shows the 'All endpoints' option selected and the 'OK' button highlighted.

1. A message informs for different categories from the target have been selected. **Click "Yes"**
2. **Click "OK"** to confirm the name of the category
3. **Click "OK"** to read all available data

Category Definition

Gather data for analogues chemicals

The screenshot shows the QSAR Toolbox interface with the 'Category Definition' tab selected. The left sidebar lists various grouping methods and defined categories. The main workspace displays a list of chemical structures (analogues) and a data matrix. A red oval highlights the data matrix row for 'LLNA' (10/10) M: Positive, indicating experimental data for analogues.

Structure	1 [target]	2	3	4	5	6	7	8
Structure								
Substance Identity								
Human Health Hazards								
Sensitisation								
Skin								
In Vivo								
LLNA	(10/10)	M: Positive	M: Positive	M: Positive	M: Positive	M: Positive	M: Positive	M: Positive
EC3								

The analogues along with their experimental data appears on data matrix

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Categorization
 - **Data gap filling without taken into account metabolism**

Data Gap Filling Overview

- “Data Gap Filling” module gives access to three different data gap filling tools:
 - Read-across
 - Trend analysis
 - Q)SAR models
- Depending on the situation, the most relevant data gap mechanism should be chosen, taking into account the following considerations:
 - Read-across is the appropriate data-gap filling method for “qualitative” endpoints like skin sensitisation or mutagenicity for which a limited number of results are possible (e.g. positive, negative, equivocal). Furthermore read-across is recommended for “quantitative endpoints” (e.g., 96h-LC50 for fish) if only a low number of analogues with experimental results are identified.
 - Trend analysis is the appropriate data-gap filling method for “quantitative endpoints” (e.g., 96h-LC50 for fish) if a high number of analogues with experimental results are identified.
 - “(Q)SAR models” can be used to fill a data gap if no adequate analogues are found for a target chemical.
- In this example, we use read-across.

Data gap filling

Apply Read-across

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling

File Apply

Data Gap Filling Methods

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

Target Endpoint

Human Health Hazards Sensitisation Skin In Vivo LLNA EC3

Substance Identity

Human Health Hazards

Sensitisation

Skin

In Vivo

LLNA

EC3

(10/10) M: Positive M: Positive

Possible data inconsistency

Scale/Unit

Gap filling scale/unit

- ☐ Skin sensitization EC3(ratio)
- ☐ Skin sensitization I (Oasis)
- ☒ Skin sensitization II (ECETOC)
- ☐ Skin Sensitization (Danish EPA)

converted data

9 from scale Skin sensitization EC3(ratio)

1 from scale Skin sensitization I (Oasis)

Selected [10/10] points

OK Cancel

1. **Click** on the cell corresponding to "Sensitization>>Skin>>In Vivo>>LLNA>>EC3"
2. **Select** "Read-across"
3. **Click** "Apply"

Additional window informing for more than one data/scale has been used appears.
More details about scale definitions is provided on next slide.

Data gap filling

Scale definition

- Skin sensitisation is a “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

Scale definition

Back to our example

Possible data inconsistency

Scale/Unit

- ☒ Skin sensitization EC3(ratio) (9 points)
- ☒ Skin sensitisation I (Oasis) (1 points)

Gap filling scale/unit

- ☐ Skin sensitization EC3(ratio)
- ☐ Skin sensitisation I (Oasis)
- ☒ Skin sensitisation II (ECETOC)
- ☐ Skin Sensitization (Danish EPA)

converted data

9 from scale Skin sensitization EC3(ratio)
1 from scale Skin sensitisation I (Oasis)

Selected [10/10]

OK Cancel

Verify that the default scale "Skin sensitisation II (ECETOC)" is selected
1. **Click "OK"**

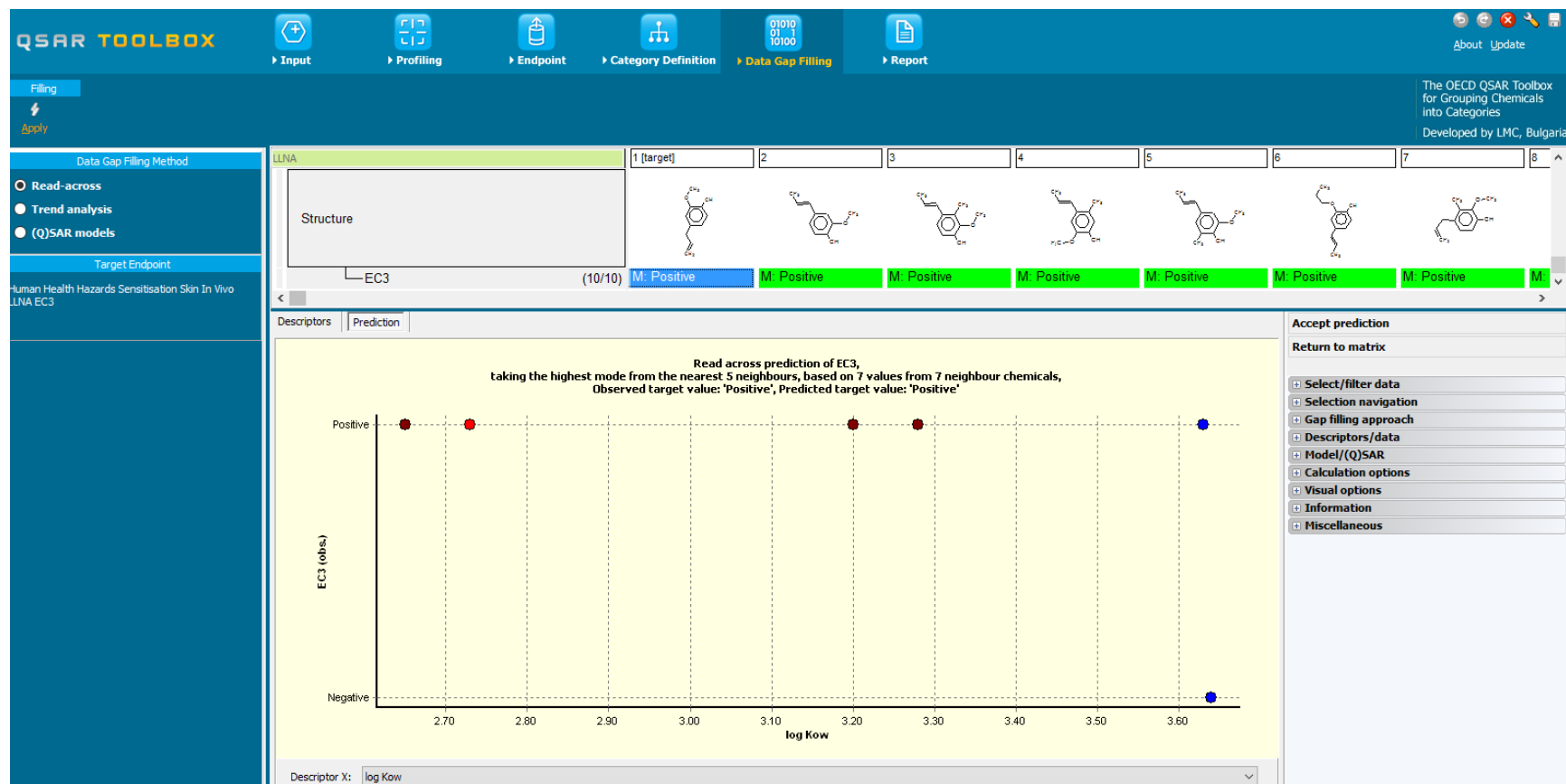
Data gap filling

Read-across

- The resulting plot place the experimental results of all analogues (Y axis) according to a descriptor (X axis) which by default is 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 represent the experimental results available for the analogues but not used for read-across.

Data gap filling

Read-across



Initial graph without any subcategorizations

Data Gap Filling Subcategorizations

In this example, the following subcategorizations are applied in order to eliminate dissimilar analogues (see slide 44):

- Organic functional group (US-EPA) – phase I is repeated in order to eliminate multifunctional analogues (subcategorization 1)
- Protein binding alerts for skin sensitization by OASIS v1.4 (subcategorization 2)

See next screen shots.

Data gap filling

Subcategorization 1: Organic functional groups (US EPA)

The screenshot shows the 'Subcategorization' window of the OECD QSAR Toolbox. The 'Grouping methods' pane on the left lists various methods, with 'Organic functional groups (US EPA)' highlighted under the 'Empiric' section. The 'Adjust options' pane shows the 'Target' as 'Alcohol, olefinic attach [-OH]' and 'Differ from target by' set to 'At least one category'. The 'Analogues' list shows several chemical structures and their predicted values. The main pane displays a scatter plot of log Kow vs. log Kow, with a red dot representing the target chemical and several blue dots representing the identified analogues. The right pane shows the 'Accept prediction' and 'Return to matrix' options, with 'Select/filter data' and 'Subcategorize' highlighted.

Read across prediction of EC3, taking the highest mode from the nearest 5 neighbours, based on 7 values from 7 neighbour chemicals, Observed target value: 'Positive', Predicted target value: 'Positive'

The identified analogues are similar to target chemical with respect to Organic functional groups (US EPA)

1. Open "Select/filter data/Subcategorize" 2. Select "Organic functional groups (US EPA)"

Data gap filling

Subcategorization 2: Protein binding alerts for skin sensitization by OASIS v.1.4

Subcategorization

Grouping methods

- Acute aquatic toxicity classification
- Acute aquatic toxicity MOA by
- Aquatic toxicity classification by
- Bioaccumulation - metabolism
- Bioaccumulation - metabolism
- Biodegradation fragments (Bio)
- Carcinogenicity (genotox and n
- DART scheme v.1.0
- DNA alerts for AMES
- Eye irritation/corrosion
- Eye irritation/corrosion
- in vitro mutagenicity
- in vivo mutagenicity
- Keratinocyte gene expression
- Oncologic Primary Classification
- Protein binding alerts for skin
- Respiratory sensitization
- Retinoic Acid Receptor Binding
- rER Expert System ver.1 - US
- Skin irritation/corrosion Exclusion
- Skin irritation/corrosion Inclusion

Empiric

- Chemical elements
- Groups of elements
- Lipinski Rule Oasis
- Organic Functional groups
- Organic Functional groups (nes
- Organic functional groups (US
- Organic functional groups, North

Metabolism/Transformations

Do not account metabolism

Documented

- Observed Mammalian metabolism
- Observed Microbial metabolism
- Observed Rat In vivo metabolism
- Observed Rat Liver S9 metabolism

Simulated

- Autoxidation simulator
- Autoxidation simulator (alkaline
- Disassociation simulation
- Hydrolysis simulator (acidic)
- Hydrolysis simulator (basic)
- Hydrolysis simulator (neutral)
- Microbial metabolism simulator

Adjust options

Target

No alert found

Differ from target

- At least one chemical
- All categories

Correlation

Analogues

(9) No alert found

Profiling **Endpoint** **Category Definition** **Data Gap Filling** **Report**

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

1 [target]	2	3	4	5	6	7	8
<chem>O=C1C=CC(=O)C=C1</chem>	<chem>O=C1C=CC(=O)C=C1</chem>	<chem>O=C1C=CC(=O)C=C1</chem>	<chem>O=C1C=CC(=O)C=C1</chem>	<chem>O=C1C=CC(=O)C=C1</chem>	<chem>O=C1C=CC(=O)C=C1</chem>	<chem>O=C1C=CC(=O)C=C1</chem>	<chem>O=C1C=CC(=O)C=C1</chem>
(10/10)	M. Positive	M. Positive	M. Positive	M. Positive	M. Positive	M. Positive	M. Positive

Read across prediction of EC3, taking the highest mode from the nearest 5 neighbours, based on 7 values from 7 neighbour chemicals, Observed target value: 'Positive', Predicted target value: 'Positive'

No protein binding alerts are identified for target and analogues, which can not be explained by positive experimental data found. In this respect metabolism should be taken into account (see next slide)

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

10 Allyl<AND> Ether<AND> Phenol (Organic Functional groups)

Create prediction by gap filling

1/1/0

1. Select "Protein binding alerts for skin sensitization by OASIS v1.4"

Data gap filling

Subcategorization when metabolism is taken into account

1

2

Read across prediction of EC3,
taking the highest mode from the nearest 5 neighbours, based on 7 values from 7 neighbour chemicals,
Observed target value: 'Positive', Predicted target value: 'Positive'

The metabolites of target chemical and its analogues possess same distribution of protein binding alerts. This could explain positive experimental data and respectively positive read-across prediction

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

Now subcategorization will be applied accounting for autoxidation simulation in combination with "Protein binding alerts" on the target and its analogues. Follow the steps:

1. The "Protein binding alerts for skin sensitization by OASIS v1.4" has been already selected;
2. **Click over** "Autoxidation simulator"

Data gap filling

Interpreting Read-across

- In this example the target and all analogues have no protein binding alerts.
- All analogues along with the target possess same distribution of positive protein binding alerts when autooxidation is taken into account.
- The latter could explain the positive experimental data of the target compound.
- Once ready go back to data matrix, when click on "Return to matrix" button (see next slide).

Data gap filling

Return to data matrix

The screenshot shows the QSAR Toolbox software interface. The 'Data Gap Filling' tab is active. The left sidebar contains a list of grouping methods and empiric methods. The main area displays a table of chemical structures and their predicted values. A callout box labeled '1' points to the 'Return to matrix' button in the 'Accept prediction' section. Below the table, a scatter plot shows 'log Kow' on the x-axis and 'log Kow' on the y-axis. The status bar at the bottom indicates '10 Allyl<AND>Ether<AND>Phenol (Organic Functional groups)' and 'Create prediction by gap filling'.

1. Click "Return to matrix"

10 Allyl<AND>Ether<AND>Phenol (Organic Functional groups)

Create prediction by gap filling

1/1/0

Data gap filling

Next actions

- The study continues with second data gap filling where a category of analogues is defined by using new categorization functionality allowing to define category accounting for (a)biotic activation of the target.
- Before proceeding with Data gap filling the following two items will be illustrated intended to explain and support the analysis. Following the steps is not necessary.
 - Multiplication of the target chemical
 - Profiling the parent and metabolites based on (a)biotic activation

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Categorization
 - Data gap filling without taken into account metabolism
 - **Multiplication of the target chemical**

Multiplication of the target chemical

- Multiplication of the target chemical could be accomplished by two ways:
 - In the **Input** section outside data gap filling module (scenario 1)
 - slide 66
 - In the **Profiling** section (scenario 2) – slide 69
- Both scenarios will be demonstrated on next few slides

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Categorization
 - Data gap filling without taken into account metabolism
 - **Multiplication of the target chemical**
 - **In the Input section (scenario 1)**

Multiplication of target chemical in the Input section (scenario 1)

The screenshot shows the QSAR Toolbox software interface. The 'Input' menu is highlighted with a red circle and the number 1. The 'Document' menu is open, and the 'Multiplication' option is highlighted with a red circle and the number 2. The 'Metabolism/Transformations' submenu is open, and the 'Autoxidation simulator' option is highlighted with a red circle and the number 4. The 'Document_2' list is open, and the 'Autoxidation simulator' option is highlighted with a red circle and the number 5. The 'Autoxidation simulator' option is also highlighted with a red circle and the number 3. The 'Autoxidation simulator' option is also highlighted with a red circle and the number 4. The 'Autoxidation simulator' option is also highlighted with a red circle and the number 5.

1. Go to "Input"
2. Click over the SMILES of the target chemical and perform **right click** on it, then
3. Select "Multiplication-Metabolism/Transformations"
4. Select "Autoxidation simulator"
5. Generated metabolites appeared in a tree-like form. They could be visualized in two modes. See next slide

Multiplication of target chemical in the Input section (scenario 1)

Visualization the set of parent and metabolites

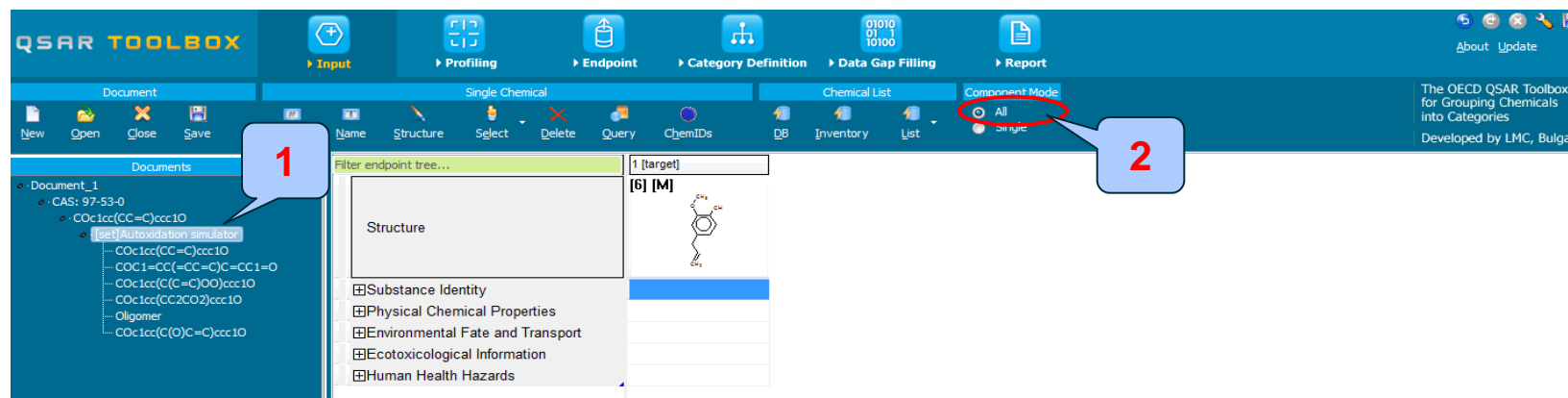
- Two component modes are implemented:
 - **Set Mode** - all metabolites are analysed as a package
 - **Individual Component Mode** - each metabolite is analysed individually

(graphical illustration of both modes is provided next screenshot)

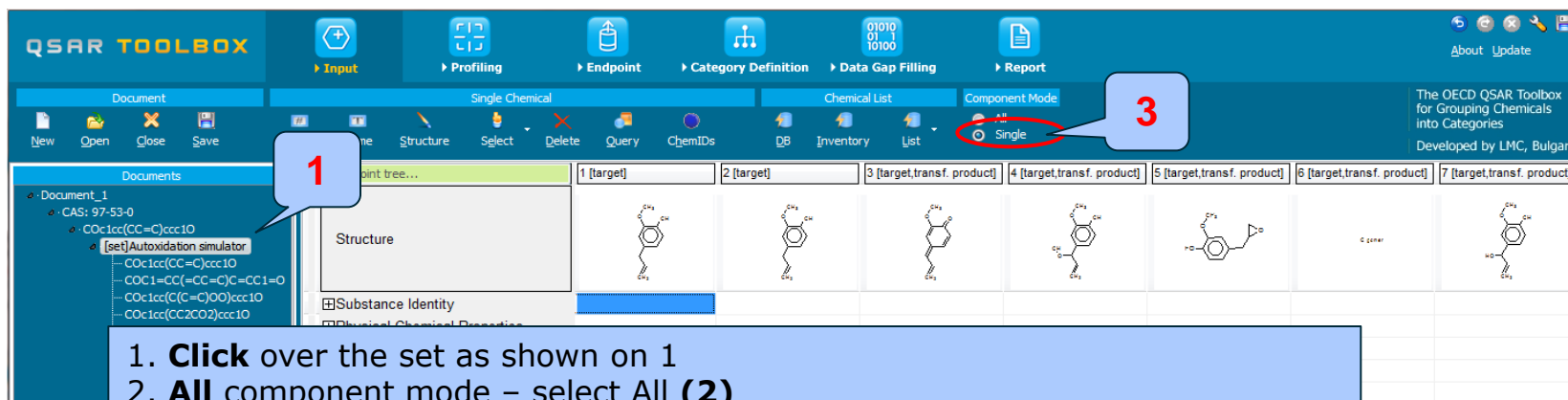
Multiplication of target chemical in the Input section (scenario 1)

Visualization the set of parent and metabolites

- All Component Mode – all metabolites are analyzed as a package



- Single Mode – each metabolite is analyzed individually



1. Click over the set as shown on 1
2. **All** component mode – select All (2)
3. **Single** component mode – select Single (3)

Protein binding result for parent and metabolites multiplied in the Input section

Autoxidation simulator

The profiling result indicates no protein binding alerts for target chemical. However, three of simulated AO metabolites exhibit interaction with proteins via three different protein binding mechanisms (Michael Addition, Radical reactions, and SN2).

The screenshot shows the QSAR Toolbox software interface. The 'Profiling' tab is selected. In the left sidebar, under 'Endpoints', 'Protein binding alerts for skin sensitization by OASIS v1.4' is checked. In the main table, the results for the seventh target (labeled '7 [target,transf. product]') are circled in red. The results show alerts for Michael Addition, Radical reactions, and SN2.

1 [target]	2 [target]	3 [target,transf. product]	4 [target,transf. product]	5 [target,transf. product]	6 [target,transf. product]	7 [target,transf. product]
<chem>CC1=CC=C(C=C1)C=C</chem>	<chem>CC1=CC=C(C=C1)C=C</chem>	<chem>CC1=CC=C(C=C1)C=C</chem>	<chem>CC1=CC=C(C=C1)C=C</chem>	<chem>CC1=CC=C(C=C1)C=C</chem>	<chem>CC1=CC=C(C=C1)C=C</chem>	<chem>CC1=CC=C(C=C1)C=C</chem>
No alert found	No alert found	Michael Addition Michael Addition >... Michael Addition >...	Radical reactions Radical reactions ... Radical reactions ...	SN2 SN2 >> Ring openi... SN2 >> Ring openi...	(N/A)	No alert found

Once the chemical is multiplied in the Input section and metabolites are visualized (distributed on data matrix) via "Single mode" 1. **Go** to "Profiling"; 2. **Check** "Protein binding alerts for skin sensitization by OASIS v1.4"; 3. **Click** "Apply"

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Categorization
 - Data gap filling without taken into account metabolism
 - **Multiplication of the target chemical**
 - In the Input section (scenario 1)
 - **In the Profiling section (scenario 2)**

Multiplication of target chemical in the Profiling section (scenario 2)

1. Once the chemical is entered into the system into a new document, **go** to "Profiling"

2. **Select** "Protein binding for skin sensitization by OASIS v1.4" profiler from **Endpoint specific** group

3. **Select** "Autoxidation simulator" from **Metabolism/Transformations** menu

4. **Click** "Apply"

5. **Double click** over the cell corresponding to **5 metabolites** to see the generated metabolites

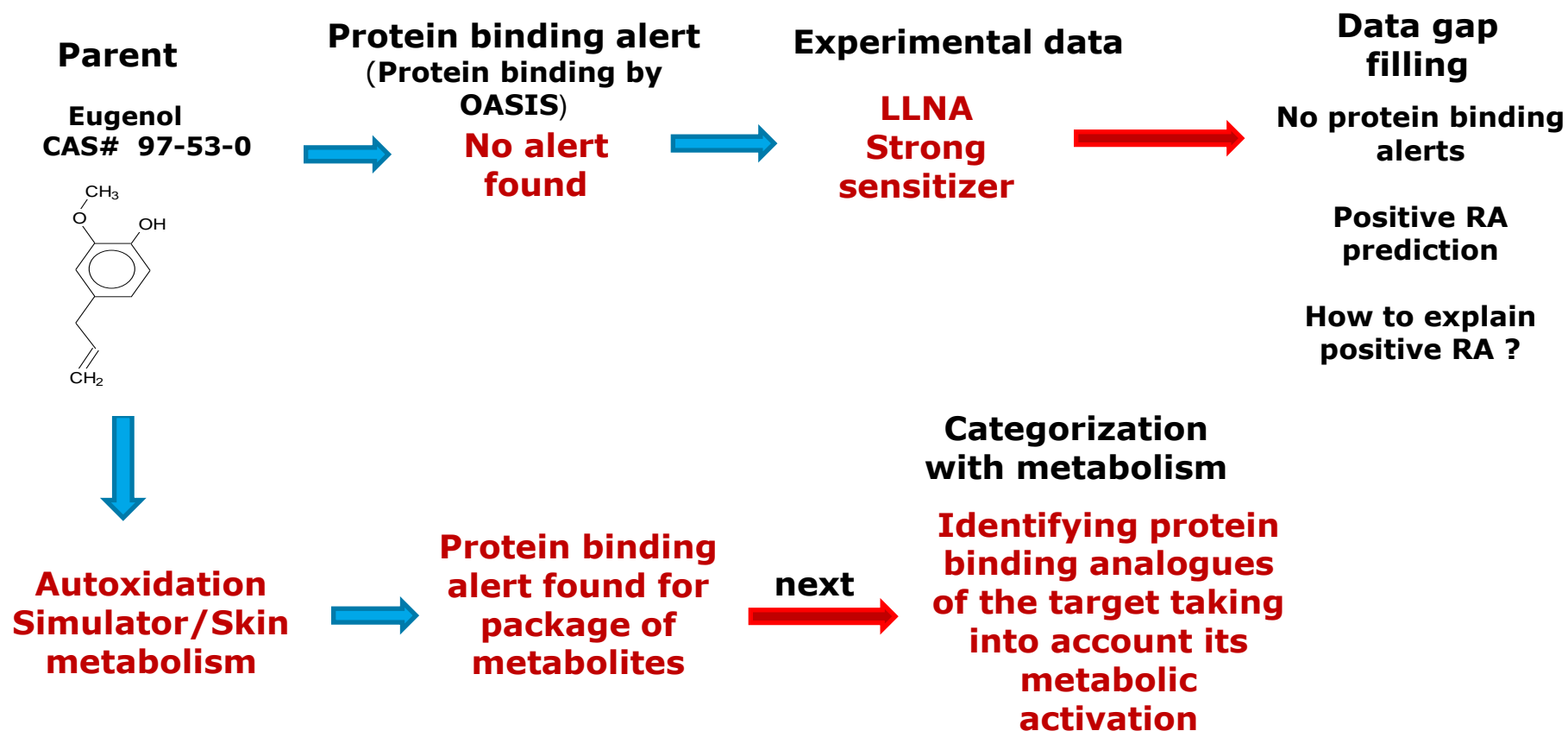
Protein binding result for parent and metabolites multiplied in the Profiling section

The screenshot shows the QSAR Toolbox interface in the Profiling section. The 'Filter endpoint tree...' panel on the left shows a tree structure where the 'Autoxidation simulator' node is selected. A red circle and a blue callout with the number '1' highlight this node. The main panel on the right displays the results for the selected node, showing a chemical structure of the parent compound and a list of 5 metabolites. A blue callout with the number '2' points to the list of metabolites and their associated reactions.

1. Open node "Autoxidation simulator"

2. Double click over the cell to investigate the profiling results obtained for the metabolites

Recap



Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Categorization
 - Data gap filling without taken into account metabolism
 - Multiplication of the target chemical
 - **Categorization applying metabolism**

Categorization applying metabolism

The advantages of the new functionality are:

- Application of metabolism for analogues identification during process of categorization. Metabolism could be used for primary categorization.
- Possibility to expand the chemical domain of the category and to identify analogues based on metabolism approach.
- Before proceeding with categorization accounting for (a)biotic activation of the target input the target in a new document (see next slide).

Categorization applying metabolism

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The left sidebar shows 'Document' and 'Chemical List' tabs. The central panel displays 'Structure' and 'Substance' information. A 'Search by CAS #' dialog box is open, showing the search results for CAS 97-53-0. The dialog box includes a table with columns: Selected, CAS, Smiles, Depiction, Names, CAS/Name, 2D/Name, and CAS/2D. The search results show a list of 10 items, with the first item (CAS 97-53-0) selected. The 'OK' button is highlighted in green.

1. Go to "Input" section; number of the target
2. Click on "New" button;
3. Click on "CAS#" button;
4. Enter the CAS
5. Click "OK"

Categorization applying metabolism

1. Click on "Category Definition" section;

2. Click on "Protein binding alerts for skin sensitization by OASIS v1.4";

3. Click on "Define with metabolism" button;

4. Select "Autoxidation simulator";

5. Click "OK" (additional window appears, see next slide)

Note: In some cases this process may take longer time, due to not indexed results for the rest of the simulators.

Categorization applying metabolism

The screenshot shows the QSAR Toolbox software interface. The 'Input' tab is active, displaying a chemical structure of a target molecule. A red box highlights the 'Package of Protein profiling result for parent and its autooxidation products' in the 'Filter endpoint tree...'. A dialog box titled 'Autoxidation simulator<WITH>Protein binding alerts for skin sensitization by OASIS v1.3' is open, showing a list of profiles. A red box highlights the 'OK' button in the dialog box.

Grouping methods

- Bioaccumulation - metabolism half-lives
- Biodegradation fragments (BioWIN MITI)
- Carcinogenicity (genotox and nongenotox)
- DART scheme v.1.0
- DNA alerts for AMES, MN and CA by OASIS
- Eye irritation/corrosion Exclusion rules by
- Eye irritation/corrosion Inclusion rules by
- in vitro mutagenicity (Ames test) alerts by
- in vivo mutagenicity (Micronucleus) alerts
- Keratinocyte gene expression
- Oncologic Primary Classification
- Protein binding alerts for Chromosomal ab
- Protein binding alerts for skin sens
- Respiratory sensitisation
- Retinoic Acid Receptor Binding
- rtER Expert System ver. 1 - USEPA
- Skin irritation/corrosion Exclusion ru
- Skin irritation/corrosion Inclusion ru

Filter endpoint tree...

Structure

Substance Identity

Physical Chemical Properties

Environmental Fate and Transport

Ecotoxicological Information

Target(s) profiles

(N/A)

Michael Addition

Michael Addition >> Quinoid type compounds

Michael Addition >> Quinoid type compounds >> Quinone methide(s)/mines; Quinoid oxime structure; Nitroquinones, Na

No alert found

Radical reactions

Radical reactions >> Free radical formation

Radical reactions >> Free radical formation >> Hydroperoxides

SN2

SN2 >> Ring opening SN2 reaction

SN2 >> Ring opening SN2 reaction >> Epoxides, Aziridines and Sulfuranes

All profiles

Acylation

Acylation >> Acyl transfer via nucleophilic addition reaction

Acylation >> Acyl transfer via nucleophilic addition reaction >> Carbodimides

Acylation >> Acyl transfer via nucleophilic addition reaction >> Isocyanates, Isothiocyanates

Acylation >> Direct acylation involving a leaving group

Acylation >> Direct acylation involving a leaving group >> (Thio)Acetates

Acylation >> Direct acylation involving a leaving group >> (Thio)Acyl and (thio)carbamoyl halides and cyanides

Acylation >> Direct acylation involving a leaving group >> Anhydrides (sulphur analogues of anhydrides)

Acylation >> Direct acylation involving a leaving group >> Azlactones and unsaturated lactone derivatives

Acylation >> Direct acylation involving a leaving group >> Carbamates

Acylation >> Direct acylation involving a leaving group >> Diacyl peroxides

Acylation >> Direct acylation involving a leaving group >> N-Acylated heterocycles

Acylation >> Direct acylation involving a leaving group >> N-Acylloxysuccinimides

Acylation >> Direct acylation involving a leaving group >> N-Acylsulfonyl

Combine profiles logically

☒ AND ☐ OR

☐ Invert result

☐ Strict

OK

Cancel

1. Select "OK"

Categorization applying metabolism

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Categorize Delete

Define Define with metabolism Subcategorize Combine Clustering Delete Delete All

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

Grouping methods

- Acute aquatic toxicity MOA by OASIS
- Aquatic toxicity classification by ECOSAR
- Bioaccumulation - metabolism alerts
- Bioaccumulation - metabolism half-lives
- Biodegradation fragments (BioWIN MITI)
- Carcinogenicity (genotox and nongenotox)
- DART scheme v.1.0
- DNA alerts for Ames by OASIS v.1.4
- DNA alerts for CA and MNT by OASIS v.1.4
- Eye irritation/corrosion Exclusion rules by
- Eye irritation/corrosion Inclusion rules by
- in vitro mutagenicity (Ames test) alerts by
- in vivo mutagenicity (Micronucleus) alerts
- Keratinocyte gene expression
- Oncologic Primary Classification
- Protein binding alerts for Chromosomal ab
- Protein binding alerts for skin sensitization
- Respiratory sensitization
- Retinoic Acid Receptor Binding
- rTER Expert System ver. 1 - USEPA
- Skin irritation/corrosion Exclusion rules by
- Skin irritation/corrosion Inclusion rules by

Empiric

- Chemical elements
- Groups of elements
- Lipinski Rule Oasis
- Organic Functional groups
- Organic Functional groups (nested)
- Organic functional groups (US EPA)
- Organic functional groups, Norbert Haider
- Structural similarity

Filter endpoint tree...

Structure

1 [target] 2 3 4 5

Ecotoxicological Information

Human Health Hazards

Acute Toxicity

Bioaccumulation

Carcinogenicity

Developmental Toxicity / Teratogenicity

Genetic Toxicity

Immunotoxicity

Irritation / Corrosion

Neurotoxicity

Photoinduced Toxicity

Repeated Dose Toxicity

Sensitisation

Respiratory Tract

Skin

In Chemico

In Vitro

In Vivo

GPMT

HRIPT

LLNA

EC3

Undefined Assay

Toxicity to Reproduction

Toxicokinetics, Metabolism and Distribution

Category of 3 analogues has been defined with EC3 data

(1/1) M: Negative

(1/1) M: Positive

(2/4) M: Positive, Positiv...

M: 8E3 µg/cm2

(4/4) M: Positive

M: Positive

M: Positive

M: Positive

(1/1) M: Positive

The forthcoming two slides illustrates how consistent is the identified category with respect to protein binding alerts when metabolism is taken into account

Categorization applying metabolism

Profiling results for parent and metabolites

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Profiling Profiling Schemes

Apply New View Delete

Filter endpoint tree...

Structure

1 [target] 2 3 4 5

1 (1/1) M: Positive
2 (2/4) M: Positive, Positiv...
4 (4/4) M: Positive
1 (1/1) M: Positive

M: 8E3 µg/cm2

The profiling results indicates no protein binding alerts for target chemical. There are protein binding alerts identified in the autoxidation products.

Metabolism/Transformations

Select All Unselect All Invert

Documented

Observed Mammalian metabolism
Observed Microbial metabolism
Observed Rat In vivo metabolism
Observed Rat Liver S9 metabolism

Simulated

Autoxidation simulator
Autoxidation simulator (alkaline medium)
Dissoication simulator
Hydrolysis simulator (acidic)

Protein binding alerts for skin sensitization by OA...

Protein binding alerts for skin sensitization by...

5 metabolites 5 metabolites 5 metabolites 5 metabolites 5 metabolites

1 x (N/A)
1 x Michael Addition
1 x Michael Addition
1 x Michael Addition
1 x Michael Addition
1 x No alert found
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x SN2
1 x SN2 >> Ring o...
1 x SN2 >> Ring o...

1 x (N/A)
1 x Michael Addition
1 x Michael Addition
1 x Michael Addition
1 x Michael Addition
1 x No alert found
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x SN2
1 x SN2 >> Ring o...
1 x SN2 >> Ring o...

1 x (N/A)
1 x Michael Addition
1 x Michael Addition
1 x Michael Addition
1 x Michael Addition
1 x No alert found
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x SN2
1 x SN2 >> Ring o...
1 x SN2 >> Ring o...

1 x (N/A)
1 x Michael Addition
1 x Michael Addition
1 x Michael Addition
1 x Michael Addition
1 x No alert found
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x SN2
1 x SN2 >> Ring o...
1 x SN2 >> Ring o...

1 x (N/A)
1 x Michael Addition
1 x Michael Addition
1 x Michael Addition
1 x Michael Addition
1 x No alert found
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x Radical reactions
1 x SN2
1 x SN2 >> Ring o...
1 x SN2 >> Ring o...

1. Go to "Profiling";
2. Check "Protein binding alerts for skin sensitization by OASIS v1.4"
3. Check "Autoxidation simulator";
4. Click "Apply"

Categorization applying metabolism

Profile statistic

The screenshot displays the QSAR Toolbox interface. The 'Profile statistic' window is open, showing a table of categories and their counts. The category '(N/A)+Michael Addition+Michael Addition >>' has a count of 5. Below this, a bar chart shows the distribution of protein binding alerts for skin sensitization by OASIS. A red circle highlights the 'Protein binding alerts for skin sensitization by OASIS' category, which shows a count of 5 for all metabolites. A red arrow points from this category to the 'All chemicals' list, where five chemical structures are shown, each with a blue box indicating they all belong to the same category.

Profile statistic

#	Category	Count
1	(N/A)+Michael Addition+Michael Addition >>	5

Protein binding alerts for skin sensitization by OASIS

Count	Category
5	(N/A)+Michael Addition+Michael Addition >>

All chemicals

CAS#	Chemical Structure
1 CAS# 97-53-0	<chem>CC1=CC=C(C=C1)C=C</chem>
2 CAS# 186743-26-0	<chem>CC1=CC=C(C=C1)C=C</chem>
3 CAS# 186743-25-9	<chem>CC1=CC=C(C=C1)C=C</chem>
4 CAS# 186743-24-8	<chem>CC1=CC=C(C=C1)C=C</chem>
5 CAS# 97-53-1	<chem>CC1=CC=C(C=C1)C=C</chem>

Protein binding alerts for skin sensitization by OASIS

Category	Count
1 x (N/A)	1
1 x Michael Addition	1
1 x Michael Addition	1
1 x Michael Addition	1
1 x No alert found	1
1 x Radical reactions	1
1 x Radical reactions	1
1 x Radical reactions	1
1 x SN2	1
1 x SN2 >> Ring o...	1
1 x SN2 >> Ring o...	1

All metabolites have same distribution of the protein binding alerts after metabolic transformation

Categorization applying metabolism

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Categorize Delete

Define Define with metabolism Subcategorize Combine Clustering Delete Delete All

Grouping methods

- Bioaccumulation - metabolism half-lives
- Biodegradation fragments (BioWIN MITI)
- Carcinogenicity (genotox and nongenotox)
- DART scheme v. 1.0
- DNA alerts for Ames by OASIS v. 1.4
- DNA alerts for CA and MNT by OASIS v. 1.4
- Eye irritation/corrosion Exclusion rules by OECD
- Eye irritation/corrosion Induction rules by OECD
- in vitro mutagenicity (Ames test) alerts by OECD
- in vivo mutagenicity (Micronucleus) alerts by OECD
- Keratinocyte gene expression
- Oncologic Primary Classification
- Protein binding alerts for Chromosomal aberrations
- Protein binding alerts for skin sensitization
- Respiratory sensitisation
- Retinoic Acid Receptor Binding
- rTER Expert System ver. 1 - USEPA
- Skin irritation/corrosion Exclusion rules by OECD
- Skin irritation/corrosion Induction rules by OECD

Empiric

- Chemical elements
- Groups of elements
- Lipinski Rule Oasis
- Organic Functional groups
- Organic Functional groups (nested)
- Organic functional groups (US EPA)
- Organic functional groups, Norbert Halder
- Structural similarity
- Tautomers unstable

Toxicological

- Repeated dose (HES)

Defined Categories

- Document
- [5] (N/A) <AND> Michael Addition <AND> Michael Addition

Filter endpoint tree...

Structure

- Bioaccumulation
- Carcinogenicity
- Developmental Toxicity / Teratogenicity
- Genetic Toxicity
- Immunotoxicity
- Irritation / Corrosion
- Neurotoxicity
- Photoinduced Toxicity
- Repeated Dose Toxicity
- Sensitisation
 - Respiratory Tract
 - Skin
 - In Chemico
 - In Vitro
 - In Vivo
 - GPMPT
 - HRIPT
 - LLNA
 - EC3
 - Undefined Assay
- Toxicity to Reproduction
- Toxicokinetics, Metabolism and Distribution
 - Basic Toxicokinetics
 - Dermal Absorption
- Profile

1 [target] 2 3 4 5

Chemical structures: 1. c1ccc(cc1)/C=C/C 2. c1ccc(cc1)/C=C/C 3. c1ccc(cc1)/C=C/C 4. c1ccc(cc1)/C=C/C 5. c1ccc(cc1)/C=C/C

Next action: Apply read-across for EC3 LLNA data for 3 analogue chemicals

(1/1) M: Positive				
(2/4) M: Positive, Positiv...				M: 8E3 µg/cm2
(4/4) M: Positive	M: Positive	M: Positive	M: Positive	
(1/1) M: Positive				

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Categorization
 - Data gap filling without taken into account metabolism
 - Categorization applying metabolism
 - **Data gap filling handling metabolism of the target chemical**

Data gap filling

Apply Read across

The screenshot shows the QSAR Toolbox interface with the 'Data Gap Filling' tab selected. The left sidebar contains a 'Filter endpoint tree...' section with a tree view of endpoints. The main area displays a table of results for five target chemicals.

Endpoint Tree (Left):

- Bioaccumulation
- Carcinogenicity
- Developmental Toxicity / Teratogenicity
- Genetic Toxicity
- Immunotoxicity
- Irritation / Corrosion
- Neurotoxicity
- Photoinduced Toxicity
- Repeated Dose Toxicity
- Sensitisation
 - Respiratory Tract
 - Skin
 - In Chemico
 - In Vitro
 - In Vivo
 - GPMT
 - HRIPT
 - LLNA
 - EC3
 - Undefined Assay
- Toxicity to Reproduction
- Toxicokinetics, Metabolism and Distribution
 - Basic Toxicokinetics
 - Dermal Absorption
- Profile

Target Chemicals (Top):

- 1 [target]
- 2
- 3
- 4
- 5

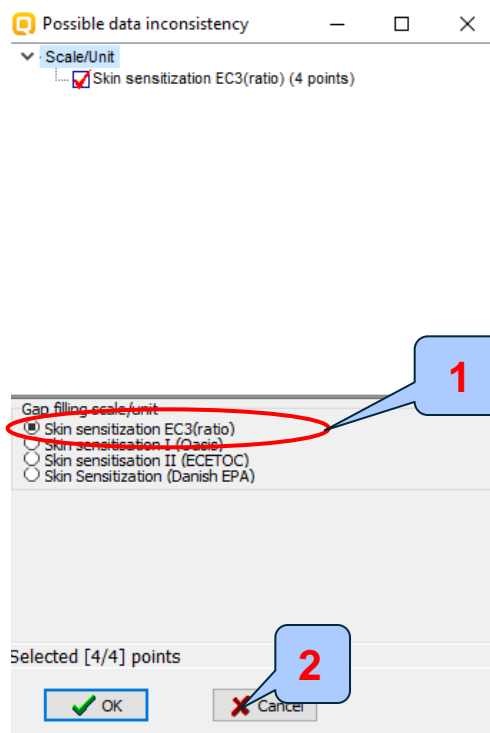
Results Table (Main):

Endpoint	1 [target]	2	3	4	5
Sensitisation	(1/1) M: Negative				
Skin					
In Vivo	(1/1) M: Positive				
LLNA	(2/4) M: Positive, Positiv...				M: 8E3 µg/cm2
EC3	(4/4) M: Positive	M: Positive	M: Positive		
Undefined Assay	(1/1) M: Positive				

1. **Click** on the cell corresponding to "Sensitization>>Skin>>In Vivo>>LLNA>>EC3" for the target chemical
2. **Select** "Read-across"
3. **Click** "Apply"

Data gap filling

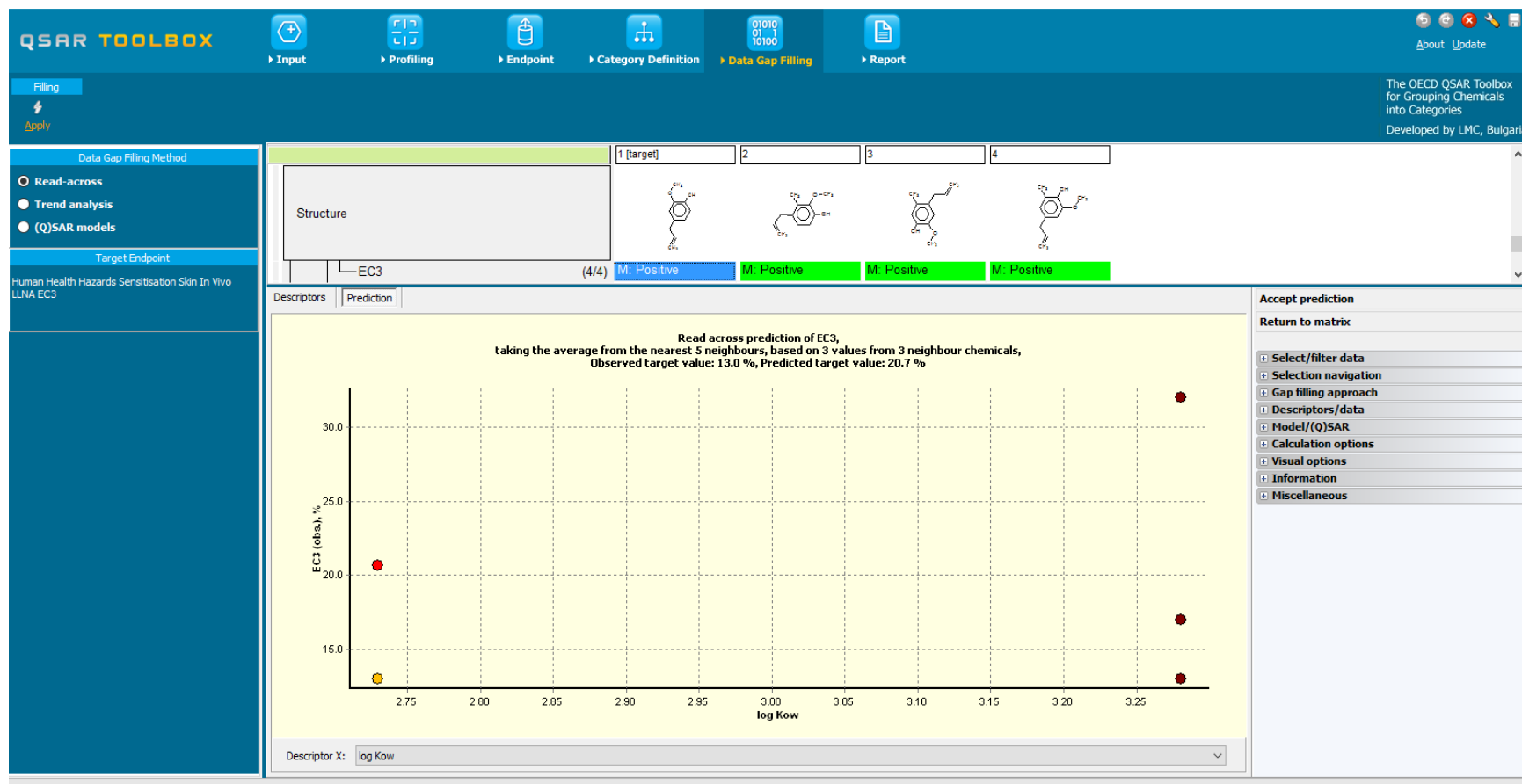
Scale definition



1. **Select** scale "Skin sensitisation EC3 (ratio)"
2. **Click** "OK"

Data gap filling

Read-across



Initial graph without any subcategorizations

Data Gap Filling Subcategorizations

In this second data gap filling, the following subcategorizations are applied (see slide 44):

- Protein binding alerts for skin sensitization by OASIS v1.4 (subcategorization 1)
- Protein binding alerts skin sensitization by OASIS v1.4 taking into account autoxidation metabolism (subcategorization 2)

See next screen shots.

Data gap filling

Subcategorization 1: Protein binding alerts for skin sensitization by OASIS v1.4

The screenshot displays the OECD QSAR Toolbox v1.4 interface during a data gap filling process. The workflow is as follows:

- Endpoint:** Protein binding alerts for skin sensitization
- Category Definition:** At least one category
- Data Gap Filling:** No alert found (indicated by a red circle and a blue callout '2')
- Report:** Read across prediction of EC3, taking the average from the nearest 5 neighbours, based on 3 values from 3 neighbour chemicals. Observed target value: 13.0 %, Predicted target value: 20.7 %.

The 'Select/filter data' panel on the right shows the 'Subcategorize' option selected (indicated by a red circle and a blue callout '1').

There are no protein binding alerts found for target chemical and its analogues

1. **Open** "Select filter data/subcategorize";
2. **Select** "Protein binding alerts for skin sensitization by OASIS v1.4"

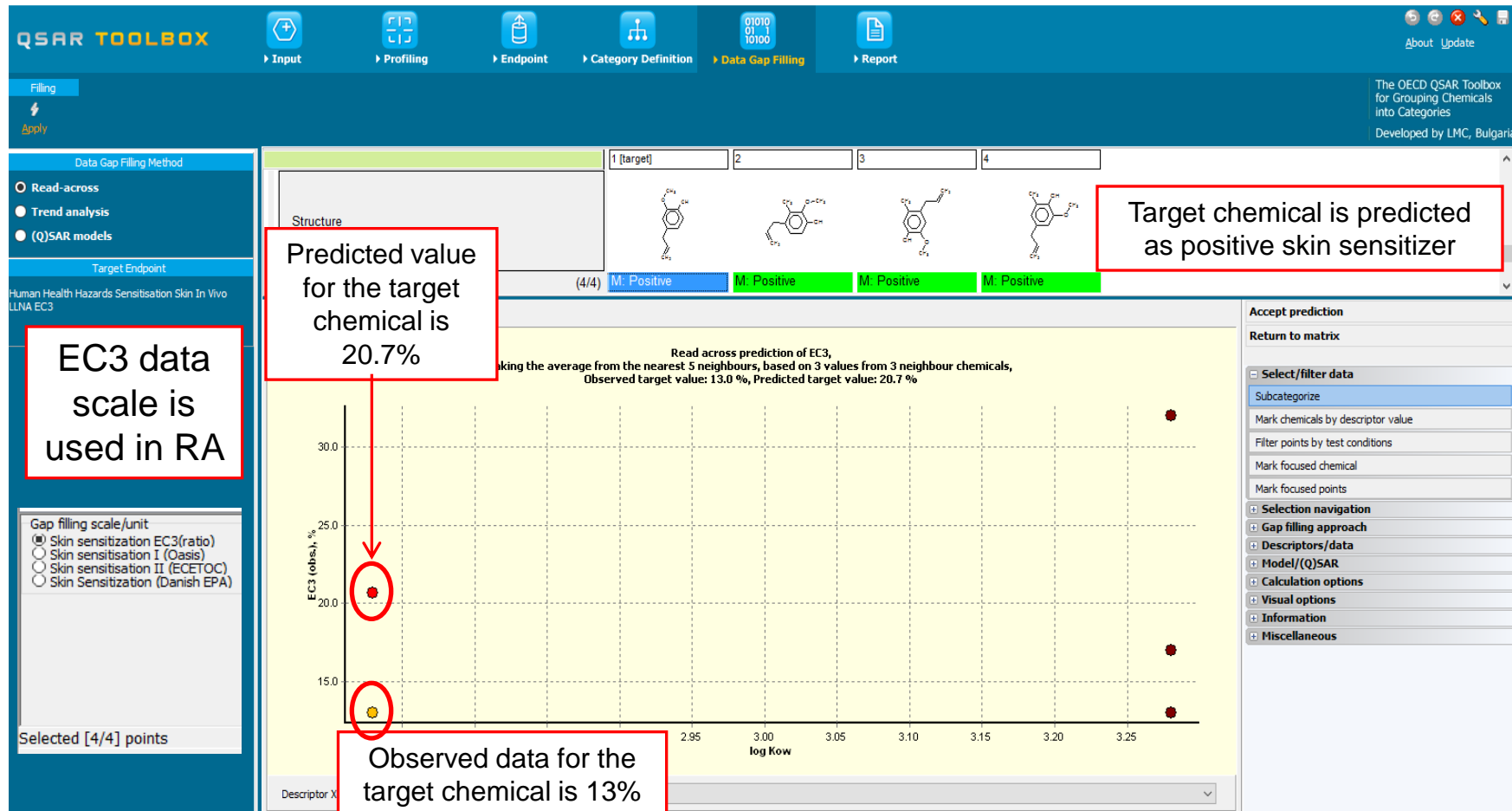
Data gap filling

Subcategorization 2: Protein binding alerts for SS when AO is taken into account

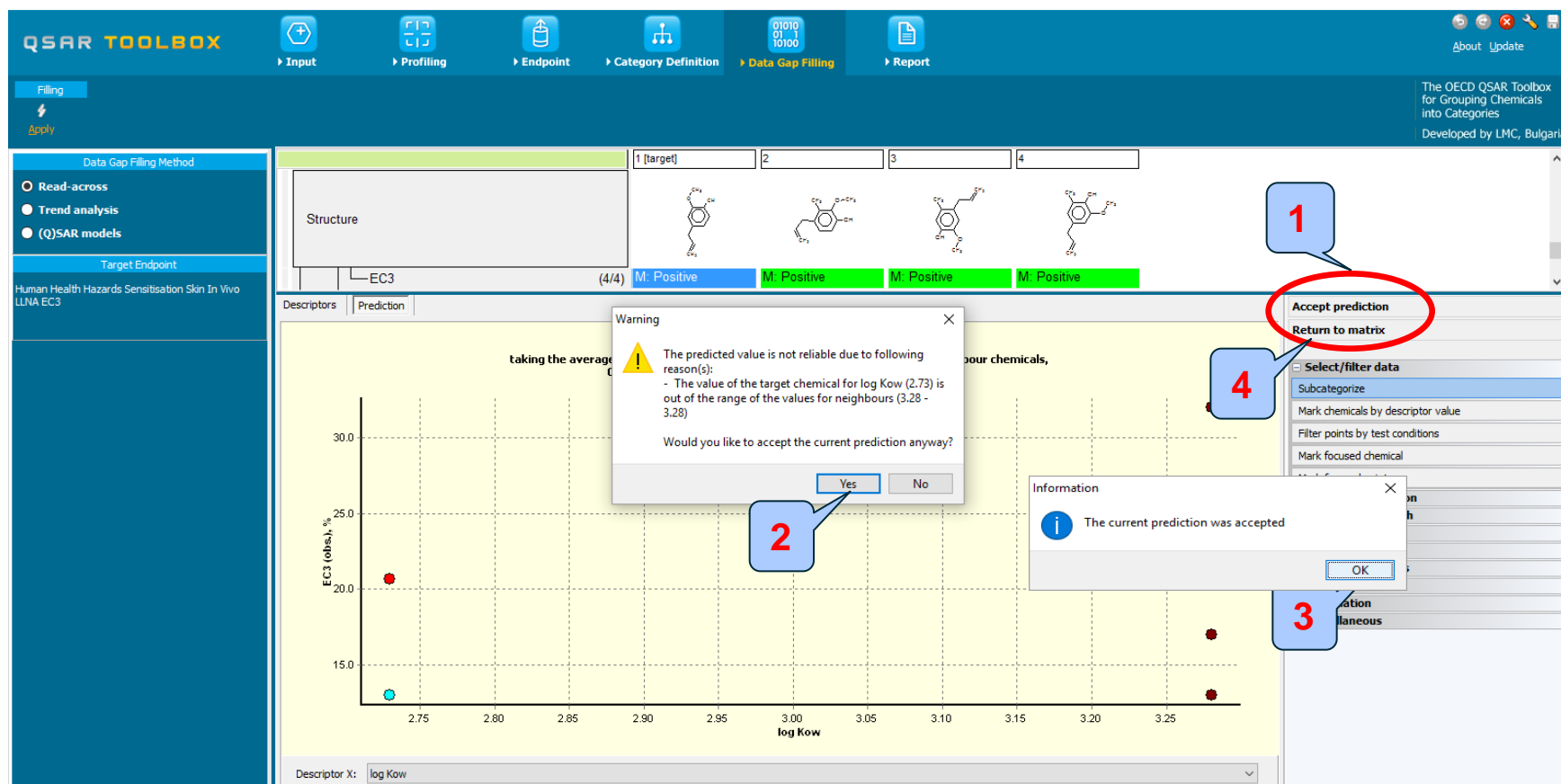
The screenshot displays the OECD QSAR Toolbox interface during a data gap filling process. The left sidebar lists various grouping methods and endpoint-specific alerts. The main panel shows the 'Data Gap Filling' workflow, including a 'Target' section with chemical structures and a 'Read across prediction of EC3' graph. A red text overlay states: "The same autoxidation products of target chemical and its analogues explain the positive experimental data". Three numbered callouts highlight key steps: 1. "Select/filter data" and "Subcategorize" in the right sidebar; 2. "Protein binding alerts for skin sensitization by OASIS v1.4" in the left sidebar; 3. "Autoxidation simulator" in the bottom left sidebar.

1. **Open** "Select filter data", **click** "Subcategorize";
2. **Select** "Protein binding alerts for skin sensitization by OASIS v1.4";
3. **Select** "Autoxidation simulator"

Data gap filling results

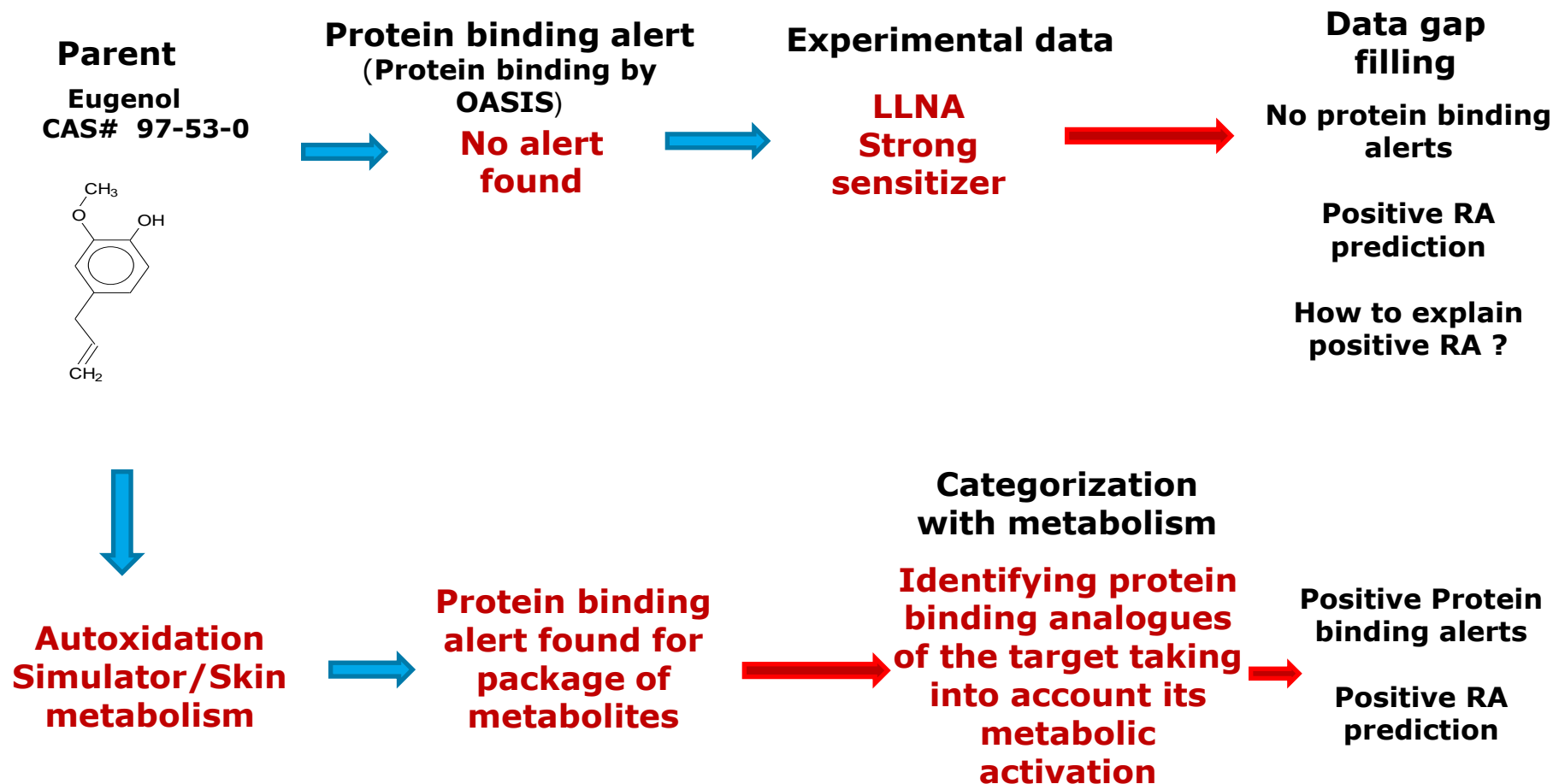


Data gap filling results



1. **Click** "Accept prediction"; appears. **Click** "Yes";
2. A message informing the user that the target is out of parametric domain
3. **Click** "OK" on the appeared message
4. **Click** "Return to matrix"

Recap



Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Categorization
 - Data gap filling without taken into account metabolism
 - Categorization applying metabolism
 - Data gap filling handling metabolism of the target chemical
- **Report**

Report

- The report module allows you to generate a report on the predictions obtained 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 bar contains icons for Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. The 'Report' tab is active. On the left, the 'Data Gap Filling Method' section shows 'Read-across' selected. Below it, the 'Target Endpoint' is 'Human Health Hazards Sensitisation Skin In Vivo LLNA EC3'. The main area displays a 'Filter endpoint tree...' on the left and a table of predictions on the right. A context menu is open over the table, with 'Report' highlighted. Two callouts are present: '1' points to a prediction entry in the table, and '2' points to the 'Report' option in the context menu.

1. **Select** prediction
2. **Right click** and **Select "Report"**

Report

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes options like Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. The left sidebar shows a tree view with 'Available data to report' (Predictions, QSARs, Categories) and 'Available report templates' (Standard predefined, Custom user defined). The main window displays a prediction report for 'Prediction [1]'. The report title is 'QSAR Toolbox prediction based on read-across'. The subtitle is 'Prediction of EC3 for eugenol (4-allyl-2-methoxyphenol)'. A callout box with the number '1' points to the 'Summary' section of the report.

Summary

Toxicity of the target chemical (20.7 %) is predicted from category members using read-across based on 3 values within the range 13.0 - 32.0 % from 3 nearest neighbours compared by prediction descriptors. Category members are single chemicals or mixtures and are selected based on the profile of the target chemical. Only chemicals having experimental data are listed in the category.

The target chemical DOES NOT FALL within applicability domain of the prediction (see Section 4.3 for details).

The data used for calculating the current prediction is taken from 3 experimental values selected from the following database(s):

1. Skin sensitization

Below is a summary table for endpoint & descriptor values for the target chemical and the category members.

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)	Descriptor(s)
	Human Health Hazards; Sensitisation	log Kow
	%	-
Target chemical	13.0	2.73
Cat. member No. 1	32.0	3.28
Cat. member No. 2	13.0	3.28
Cat. member No. 3	17.0	3.28

1. Summary information for prediction

Report

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes options like Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. The left sidebar shows 'Available data to report' (Predictions, QSARs, Categories) and 'Available report templates' (Standard and Custom). The main window shows a prediction for 'eugenol (4-allyl-2-methoxyphenol)' with a diagram of its chemical structure and associated boundaries. A red circle highlights the '1M) Metabolism' section, which states: 'In boundary [1] profiling scheme "Protein binding alerts for skin sensitization by OASIS v1.4" was combined with "Autoxidation simulator"'. Other sections include '1) Referential boundary' and '2) Parametric boundary'.

Prediction [1] Prediction of EC3 for eugenol (4-allyl-2-methoxyphenol) 6 / 18

1) Referential boundary:
The target chemical should be classified as (N/A) AND Michael Addition AND Michael Addition >> Michael addition on qinoide type compounds AND Michael Addition >> Michael addition on qinoide type compounds >> Quinone methide(s)/imines; Quinoide oxime structure; Nitroquinones, Naphthoquinone(s)/imines AND No alert found AND Radical reactions AND Radical reactions >> Free radical formation AND Radical reactions >> Free radical formation >> Hydroperoxides AND SN2 AND SN2 >> Ring opening SN2 reaction AND SN2 >> Ring opening SN2 reaction >> Epoxides, Aziridines and Sulfuranes by Protein binding alerts for skin sensitization by OASIS v1.4

1M) Metabolism:
In boundary [1] profiling scheme "Protein binding alerts for skin sensitization by OASIS v1.4" was combined with "Autoxidation simulator"

Additional info:
Process :All, Match : Accumulative. ~~Separate query criteria may be met by different metabolites (or the parent)~~

2) Parametric boundary:
The target chemical should have a value of log Kow which is ≥ 3.28

3) Parametric boundary:

1. Information that metabolism was taken into account when predicting skin sensitization is available

Report

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Reports: Create Print Close Save as

Repository: Register Unregister Update Clone Design

Available data to report: Predictions (Q)SARs Categories

Available report templates: Standard (predefined) Custom (user defined)

Prediction [1]

f. Input for prediction (target chemical):
SMILES

g. Descriptor and endpoint values for target chemical (if applicable):

Descriptor(s)	
log Kow	2.73
Endpoint (dep. variable)	
Human Health Hazards; Sensitisation	13.0 %

h. Additional chemicals eliminations (not determined by domain):
Not available

i. Additional data eliminations (not determined by domain):
Not available

j. Predicted value (model result):
20.7 %

k. Predicted value (comments):
Not provided by the user *manually editable field*

4.3. Applicability domain (OECD Principle 3):
The target chemical DOES NOT FALL within applicability domain (see Section 3.1.b for detailed description of the domain)

1

1. Predicted value

Report

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes options like Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. The left sidebar shows a tree view with Predictions, (Q)SARs, and Categories. The main window displays a prediction report for Sensitisation. The report includes sections for additional chemical and data eliminations, predicted values, and applicability domain. A red circle highlights the applicability domain section, which states the target chemical does not fall within the domain. A blue callout box with the number 1 points to this section.

h. Additional chemicals eliminations (not determined by domain):
Not available

i. Additional data eliminations (not determined by domain):
Not available

j. Predicted value (model result):
20.7 %

k. Predicted value (comments):
Not provided by the user *manually editable field*

4.3. Applicability domain (OECD Principle 3):
The target chemical DOES NOT FALL within applicability domain (see Section 3.1.b for detailed description of the domain)

4.4. Uncertainty of the prediction (OECD Principle 4): *manually editable field*

QSAR Toolbox 3.4.0.17
Database version: 3.8.8/2.1.2

TPRF v.3.4.1.34101

1. Applicability domain

The target chemical is "Out of domain", because does not fall within parametric range of Log Kow[3.28-4.75]

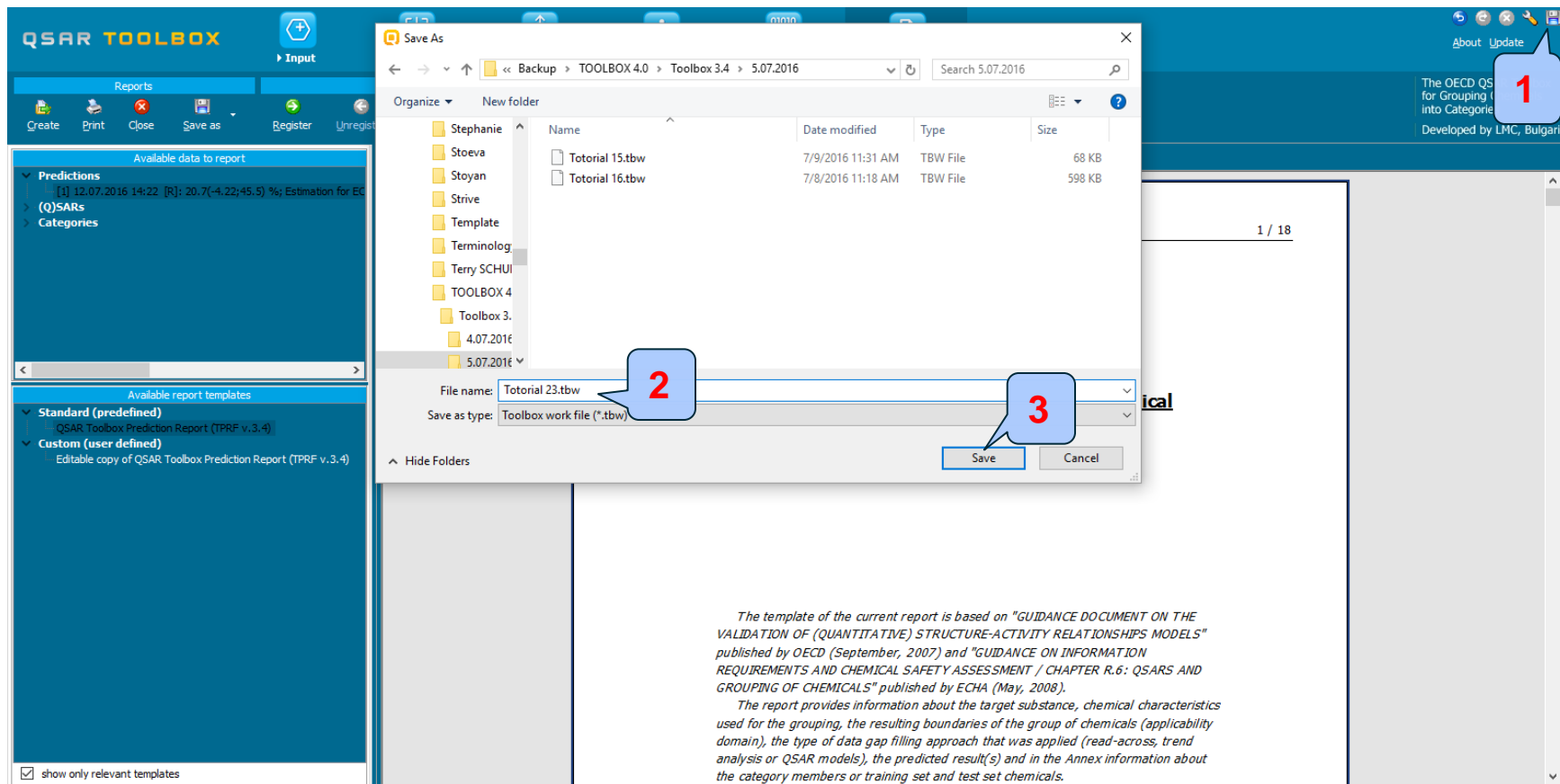
Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Categorization
 - Data gap filling without taken into account metabolism
 - Categorization applying metabolism
 - Data gap filling handling metabolism of the target chemical
 - Report
 - **Save the prediction**

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" button; 2. **Browse** and **put** name of the file; 3. **Click** Save button

Open saved file

The screenshot illustrates the steps to open a saved file in the QSAR Toolbox software. The interface is divided into several panels:

- Top Menu Bar:** Contains icons for 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'.
- Document Panel (Left):** Shows a list of documents. The 'Open' button is highlighted with a red circle labeled '3'.
- Select File Dialog (Center):** A file explorer window showing the 'Backup' folder. The file 'Tutorial 23.tbw' is selected, highlighted with a red circle labeled '4'.
- Main Workspace (Right):** Displays a chemical structure and a table of results. The 'Open' button in the dialog is highlighted with a red circle labeled '5'.

The table of results shows the following data:

Assay	Result	Endpoint	Category	Value
In Vivo				
GPMT	(1/1) M: Positive			
HRIPT	(2/4) M: Positive, Positiv...			
LLNA	(4/5) M: Positive	M: Positive	M: Positive	M: Positive
EC3	(1/1) M: Positive			
Undefined Assay				
Toxicity to Reproduction				
Toxicokinetics, Metabolism and Distribution				
Basic Toxicokinetics				
Dermal Absorption				
Profile				

Once the file has been saved 1. **Go** to "Input"; 2. **Create** new document 3. **Click** "Open"; 4. **Browse** and **select the file**; 5. **Click** "Open" button