

## OECD QSAR Toolbox v.3.4

Example for predicting Skin Sensitization of  
mixture with known components

# Outlook

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

# Background

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

# Outlook

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

# Objectives

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

# Outlook

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

## Exercise

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

# Outlook

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



# Workflow

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

# Outlook

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

# Chemical Input Overview

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

# Chemical Input

## Ways of Entering a mixture

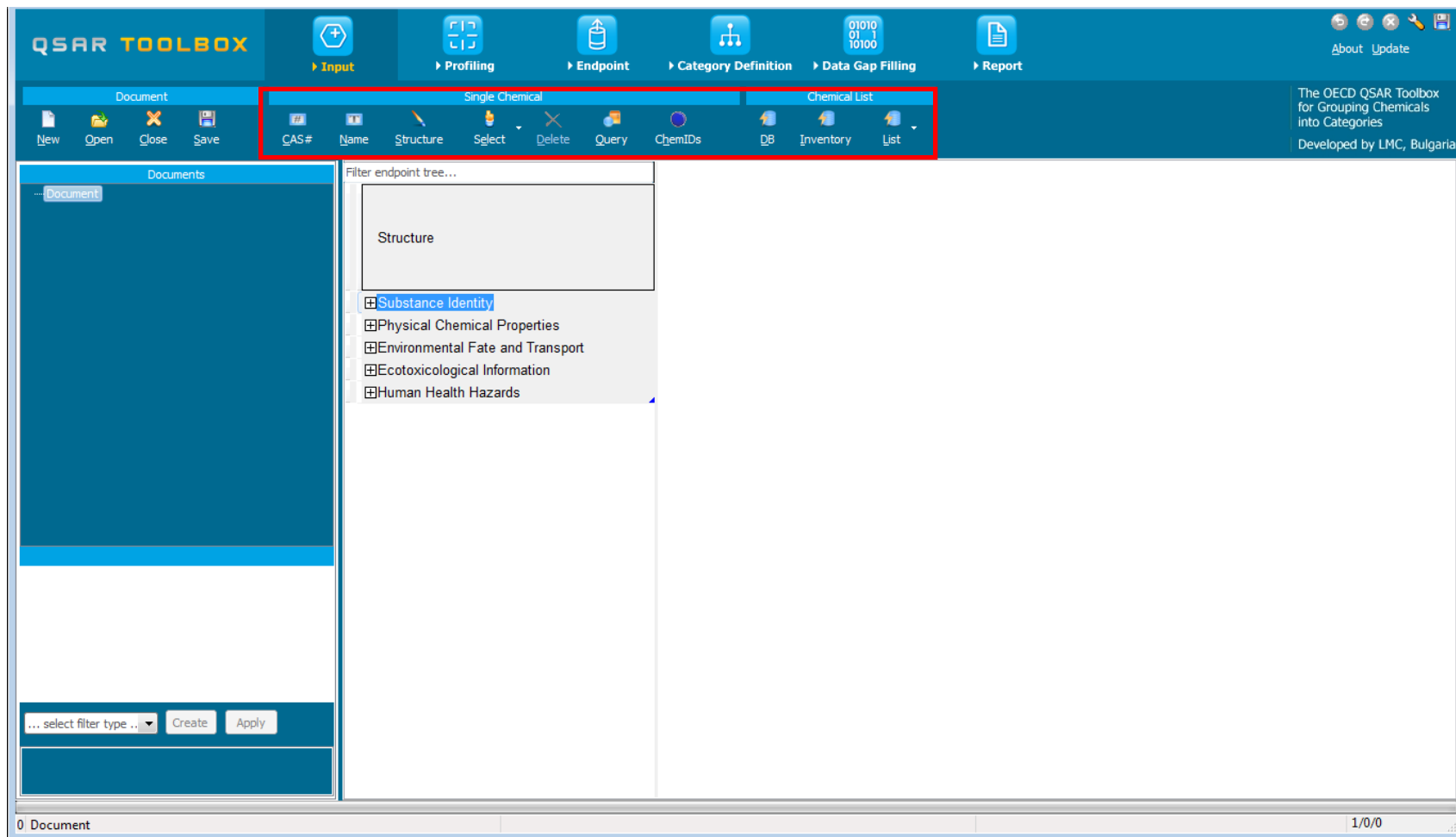
### **User alternatives for defining mixtures with known compositions:**

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

# Getting Started

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

# Chemical Input Input Screen



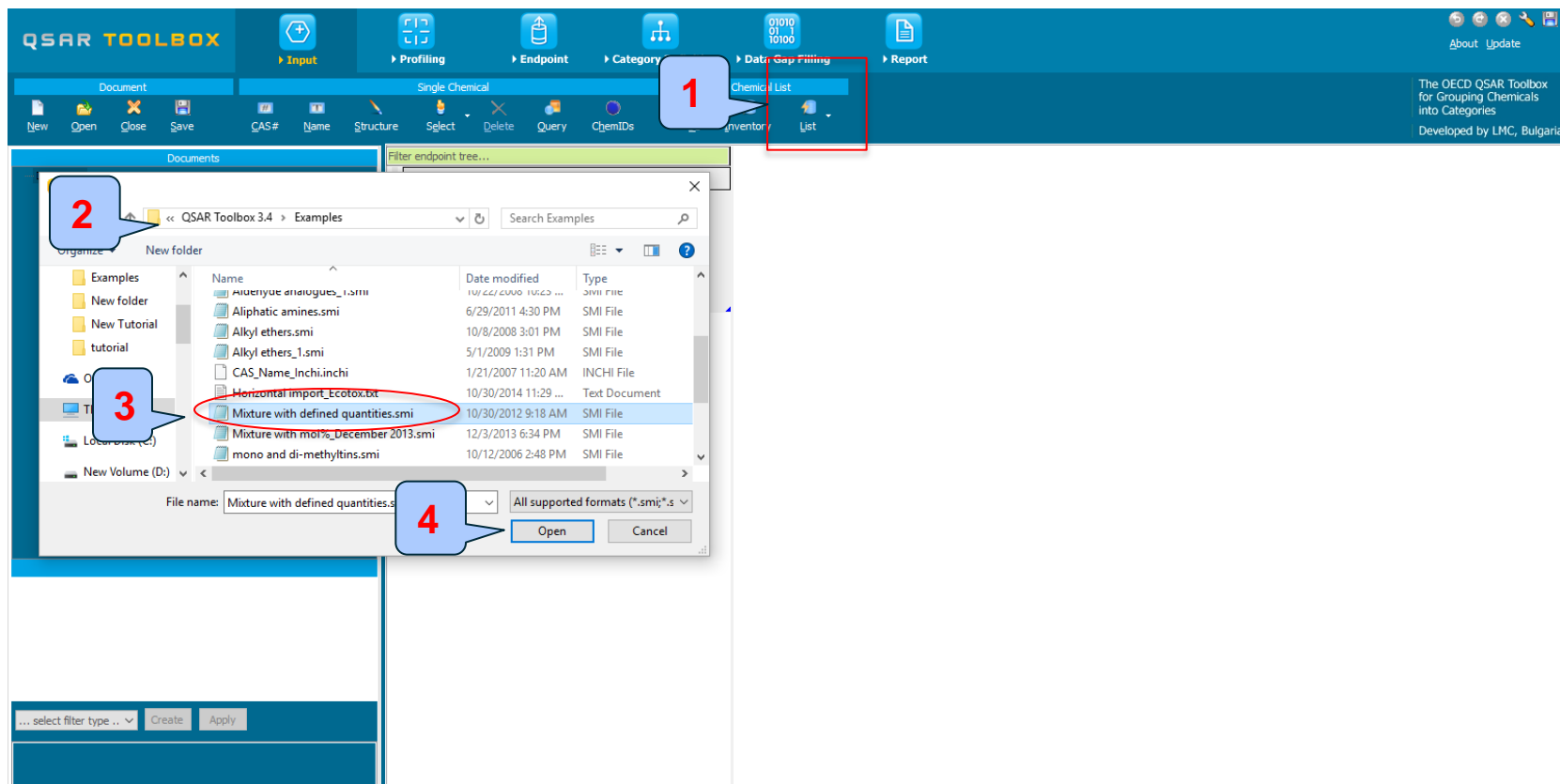
# Chemical input

## Load list with chemical mixture

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

# Chemical input

## Load list with chemical mixture

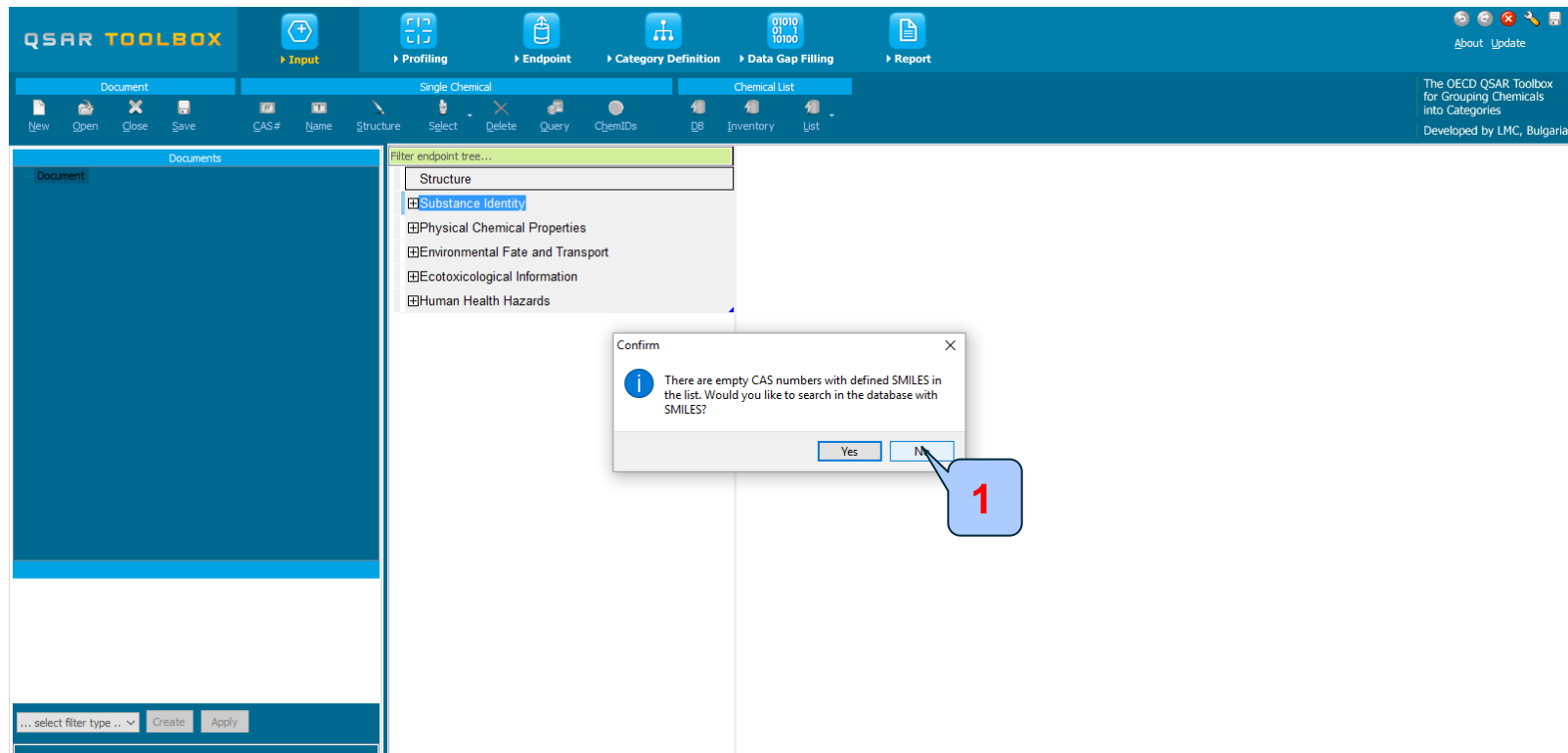


1. **Click** on Chemical list
2. **Browse** and find the file with mixture located at Examples directory
3. **Select** the file
4. **Open** the file "Mixture with defined quantities.smi"



# Chemical input

## Load list with chemical mixture



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

1. **Select No**

# Chemical Input

## Target chemical identity

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

# Chemical Input

## Target chemical identity

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

- Callout 1:** Points to the entry `[set][Mix]{X=1/Miligrams...}` in the Documents list.
- Callout 2:** Points to the 'Component Mode' dropdown menu, where 'Single' is selected.
- Callout 3:** Points to the entry '1 [target]' in the Filter endpoint tree.

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

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

# Chemical Input

## Target chemical identity

The screenshot displays the QSAR Toolbox interface in 'Component Mode'. The menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Component Mode' option is highlighted in the 'Report' menu, and the 'Single' radio button is selected in the toolbar. The main workspace shows a table of chemical components with their structures and quantities. A callout bubble with the number '1' points to the 'Component Mode' button. A text box at the bottom explains that 'Single' component mode visualizes all components of the mixture and allows working with each as an individual substance.

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

# Outlook

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

# Profiling Overview

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

# Profiling

## Side-Bar to Profiling

- For most of the profilers, background information can be retrieved by highlighting one of the profilers (for example, Protein binding by OASIS v1.4 and clicking on “View” (see next screen shot)).





# Profiling

## Profiling the target chemical

- Select the “Profiling methods” related to the target endpoint by clicking on the box next to the profilers name.
- This selects (a green check mark appears) or deselects (green check disappears) profilers.
- For this example, select the following profilers relevant to the Skin sensitization (see next screenshot):
  - Protein binding by OASIS v1.4 – general mechanistic
  - Protein binding by OECD – general mechanistic
  - Protein binding potency – general mechanistic
  - Protein binding alerts for skin sensitization by OASIS v1.4 – endpoint specific

# Profiling Side-Bar to Profiling

The screenshot shows the QSAR Toolbox Profiling interface. The 'Profiling methods' sidebar on the left contains a list of methods, with 'Protein binding alerts for skin sensitization by OASIS v1.4' checked. The 'Profiling Schemes' toolbar at the top left has an 'Apply' button circled in red. The main window displays a 'Filter endpoint tree...' on the left and a table of results on the right. The table has four columns representing different target mix components. The first column shows 'No alert found'. The second column shows 'SNAr >> Nucleoph...'. The third and fourth columns show 'No alert found'. A red box highlights the 'SNAr >> Nucleoph...' alert with the text 'This component has positive protein binding alert'.

1. Check the profilers related to the target endpoint;

2. Click Apply.

# Outlook

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

## Endpoint

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

# Endpoint

The screenshot shows the QSAR Toolbox interface. In the 'Databases' list on the left, 'Skin sensitization ECETOC' is selected with a green checkmark, indicated by callout 1. The 'Gather' button in the top toolbar is circled in red, indicated by callout 2. The main window displays a filter tree on the left and a table of results for four target endpoints on the right.

Filter endpoint tree...	1 [target]	2 [target,mix.component]	3 [target,mix.component]	4 [target,mix.component]
Structure				
	Qty: 1 Milligrams	Qty: 10 Milligrams	Qty: 100 Milligrams	
Substance Identity				
Physical Chemical Properties				
Environmental Fate and Transport				
Ecotoxicological Information				
Human Health Hazards				
Profile				
General Mechanistic				
Protein binding by OASIS v1.4	No alert found SNAr SNAr >> Nucleoph...	SNAr SNAr >> Nucleoph...	No alert found	No alert found
Protein binding by OECD	No alert found	No alert found	No alert found	No alert found
Protein binding potency	Not possible to cla...	Not possible to cla...	Not possible to cla...	Not possible to cla...
Endpoint Specific				
Protein binding alerts for skin sensitization by O...	No alert found SNAr SNAr >> Nucleoph...	SNAr SNAr >> Nucleoph...	No alert found	No alert found

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

# Endpoint

## Process of collecting data

The screenshot displays the QSAR Toolbox software interface during the 'Endpoint' process. The main workspace is divided into several panels:

- Left Panel (Databases):** A list of databases with checkboxes, including 'Skin sensitization' and 'Skin sensitization ECETOC' which are checked.
- Middle-Left Panel (Filter endpoint tree...):** A tree view of endpoints, with 'Skin' expanded to show 'In Chemico', 'In Vitro', and 'In Vivo' sub-categories. 'In Vivo' is further expanded to show 'GPMT', 'LLNA', 'Miscellaneous', and 'Undefined Assay'. 'LLNA' and 'Miscellaneous' are both marked with '(1/1)'.
- Middle-Right Panel (Structure):** A grid showing chemical structures for four components. Component 1 is a mixture of three structures. Component 2 is a benzene ring with a chlorine atom and a methyl group. Component 3 is a benzene ring with a chlorine atom and a methyl group. Component 4 is a benzene ring with a chlorine atom and a methyl group. The quantities are: Qty: 1 Milligrams, Qty: 10 Milligrams, and Qty: 100 Milligrams.
- Right Panel (Data Table):** A table showing experimental data for two components. The table has columns for 'M: Negative' and 'M: Negative'. The data is circled in red.

In this example, negative skin sensitization experimental data have been found for the two components

## Recap

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

# Outlook

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



# Read across prediction of constituent without data

## Focus constituent

1

2

This constituent is selected for further read-across prediction

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

# Read across prediction of constituent without data

## Focus constituent

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'. Below this, there are sub-menus for 'Data', 'Import', 'Export', 'Delete', and 'Tautomerize'. The main workspace is divided into several panels:

- Databases:** A tree view on the left showing various databases under 'Human Health Hazards', such as 'Acute Oral Toxicity database (ChemIDplus)', 'Bacterial mutagenicity ISSSTY', etc.
- Filter endpoint tree...:** A tree view in the center showing a hierarchical structure of endpoints. The 'Structure' panel is selected, showing a chemical structure of a chlorinated compound.
- Structure:** A panel on the right showing the chemical structure of the focused constituent, labeled '1 [target,mix.component]'. Below the structure, it indicates 'Qty: 1 Milligrams'.
- Data Matrix:** A table on the right showing the data for the focused constituent. A red box highlights the entry '1 [target,mix.component]' in the 'Structure' column, with a callout bubble containing the number '1'. A text box below the table states: 'This focused component appeared in separate data matrix'.

# Outlook

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

# Category Definition Overview

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

# Category Definition

## Grouping methods

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

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

## Basic guidance for category formation and assessment

### Suitable categorization phases:

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

### Performing categorization:

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

**Graphical illustration of suitable categorization phases is shown on next slide**

## Suitable Categorization/Assessment Phases

### Phase I. Structure based

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

**Broad grouping  
Endpoint Non-specific**

**Repeating Phase I due to Multifunctionality of chemicals**

### Phase II. Mechanism based

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

**Subcategorization  
Endpoint Specific**

**Metabolism accounted for**

### Phase III. Eliminating dissimilar chemicals

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

**Subcategorization  
Endpoint Specific**

# Read across prediction of constituent without data

## Forming category for studied endpoint

### Suitable Categorization/Assessment Phases

#### Phase I. Structure based

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

Repeating Phase I due to Multifunctionality of chemicals

**Broad grouping**  
**Endpoint Non-specific**

### Phase I categorization in Toolbox

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

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

46 analogues are identified as Aryl halides by OFG

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



## Read across prediction of constituent without data

### Forming category for studied endpoint

- Based on the above recommendations and classifications from structurally similar profilers the OFG is used as initial categorization group
- Refinement of the initial group is based on endpoint-specific protein binding profiler:
  - Protein binding alerts for skin sensitization by OASIS v1.4.

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

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

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

See next slides

# Read across prediction of constituent without data

## Define category by OFG

**2**

Define Define with metabolism Subcategorize Combine Clustering Delete Delete All

Grouping methods

- Bioaccumulation - metabolism half-lives
- Biodegradation fragments (BioWIN MITI)
- Carcinogenicity (genotox and nongenotox) alerts by ISS
- DART scheme v.1.0
- DNA alerts for Ames by OASIS v.1.4
- DNA alerts for CA and MNT by OASIS v.1.1
- Eye irritation/corrosion Exclusion rules by BfR
- Eye irritation/corrosion Inclusion rules by BfR
- 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 OASIS v.1.2
- Protein binding alerts for skin sensitization by OASIS v1.4
- Respiratory sensitisation
- Retinoic Acid Receptor Binding
- rTER Expert System ver. 1 - USEPA
- Skin irritation/corrosion Exclusion rules
- Skin irritation/corrosion Inclusion rules

**1**

Empiric

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

Toxicological

- Repeated dose (HESS)

Defined Categories

Document

Structure

1 [[target.mix.component]]

Qty: 1 Milligrams

Ecotoxicological Information

- Human Health Hazards
  - Acute Toxicity
    - Bioaccumulation
    - Organic Functional groups
  - Carcinogenicity
  - Developmental Toxicity
  - Genetic Toxicity
  - Immunotoxicity
  - Irritation / Corrosion
  - Neurotoxicity
  - Photoirritation
  - Repeated Dose
  - Sensitisation
  - Skin
  - In Chem
  - In Vitro
  - In Vivo
    - GPM
    - LLNA
    - Miscellaneous
    - Undefined Assay
  - ToxCast
  - Toxicity to Reproduction
  - Toxicokinetics, Metabolism and Distribution

Target(s) profiles

- Aryl halide
- Ketone

All profiles

- Acetal
- Acetoxyl
- Acid anhydride
- Acid anhydride, mixed phosphonic
- Acridine

Combine profiles logically  Invert result  OK

**3**

Define category name

Category name (1 chemicals) halide<AND>Ketone (Organic Functional groups)

OK Cancel

Based on above recommendations the OFG is used as an initial group (phase I)

**1. Select OFG**

**2. Click Define**

**3. Combination of three organic functional groups identified a single analogue (3). In order to expand the category Aryl halide is used only. See next slide**

# Read across prediction of constituent without data

## Define category by OFG

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

**1. Select** Aryl and Ketone (hold Ctrl button) and remove them by the arrow down **3. Aryl halide** should be remained only **4. Click** OK

# Outlook

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

# Read across prediction of constituent without data

## Gather data for analogues chemicals

The screenshot displays the QSAR Toolbox interface. On the left, a list of grouping methods is visible, including 'Organic Functional groups (nested)'. The main window shows a tree view of endpoints under 'Human Health Hazards'. Two dialog boxes are overlaid: 'Define category name' (labeled with a red '1') and 'Read data?' (labeled with a red '2'). The 'Define category name' dialog has 'Aryl halide (Organic Functional groups)' entered in the 'Category name' field. The 'Read data?' dialog has 'All endpoints' selected and 'from Tautomers' checked. A chemical structure of 2,4-dichloroacetic acid is shown in the background.

1. Click OK

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

# Read across prediction of constituent without data

## Gather data for analogues chemicals

The screenshot displays the QSAR Toolbox interface. At the top, there are navigation tabs: Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. Below these are sub-tabs for Categorize and Delete. The main workspace shows a chemical structure of a polychlorinated biphenyl (PCB) with the label 'Qty: 1 Milligrams'. To the left is a 'Filter endpoint tree...' panel with a tree view of various endpoints. The 'In Vivo' endpoint is expanded, showing a sub-entry '(70/93)'. Below the tree is a data matrix table with 9 columns. The first column is highlighted with a red oval, and its value 'M: Positive' is also highlighted with a red box. A larger red box encompasses the text 'The experimental data for the identified analogues appears on data matrix'.

Endpoint	1 [target,mix.component]	2	3	4	5	6	7	8	9
In Vivo (70/93)	M: Positive	M: Positive, Positive	M: Positive	M: Positive, Positive	M: Positive	M: Negative	M: Positive, Positiv...	M: Ne	

# Outlook

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

# Read across prediction of constituent without data

## Apply read across

The screenshot shows the QSAR Toolbox interface with the 'Data Gap Filling' tab active. A 'Possible data inconsistency' dialog box is open, listing various assays and endpoints. Red callouts 1-4 highlight the steps: 1. Clicking the 'Skin sensitization in vivo' cell in the table; 2. Selecting 'Read-across' in the 'Data Gap Filling Method' dropdown; 3. Clicking the 'Apply' button; 4. Clicking 'OK' in the dialog box. A text box on the right states 'Assays and Endpoints are mixed'.

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



# Read across prediction of constituent without data

## Apply read across

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

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

Filling Apply

Data Gap Filling Method

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

Target Endpoint

Human Health Hazards Sensitisation Skin In Vivo

Structure

Qty: 1 Milligrams

In Vivo (70/93)

M. Positive M. Positive, Positive M. Positive M. Positive, Positive M. Positive M. Negative M. Positive, Positiv M. Ne

Descriptors Prediction

Read across prediction of A B C, EC3, NOEL, S M W N, S W A N, Skin sensitisation, taking the highest mode from the nearest 5 neighbours, based on 7 values from 6 neighbour chemicals, Observed target value: N/A, Predicted target value: 'Positive'

Descriptor X: log Kow

Accept prediction

Return to matrix

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

71 Aryl halide (Organic Functional groups) Create prediction by gap filling 0/1 1/10

# Read across prediction of constituent without data

## Subcategorization

- The initial category could be refined by subcategorizing the analogues according to the following endpoint specific profiler (phase II, slide #37):
  - Protein binding alerts for skin sensitization by OASIS v1.4.
- These steps are summarized in the next screen shots.

# Read across prediction of constituent without data Subcategorization by Protein binding alert for SS

The screenshot displays the QSAR Toolbox software interface. On the left, the 'Grouping methods' sidebar is visible, with 'Protein binding alerts for skin sensitization by OASIS v1.4' selected and circled in red, with a callout '2'. The main window shows a 'Target' selection menu with 'SNAR >> Nucleophilic aromatic substitution' selected. Below this, the 'Differ from target by' section has 'All categories' selected. The 'Correlation' section shows a list of analogues. The 'Read across prediction' plot shows a scatter plot of log Kow values for various chemicals, with a callout '3' pointing to the 'Remove' button. The 'Accept prediction' panel on the right has the 'Subcategorize' button highlighted in red, with a callout '1'.

1. Select filter data/subcategorize
2. Select Protein binding alerts for SS by OASIS v1.4.
3. Click Remove to eliminate dissimilar chemicals.

# Read across prediction of constituent without data

## Apply read across

The screenshot displays the QSAR Toolbox interface with the 'Data Gap Filling' method selected. The 'Read-across' option is chosen under the 'Data Gap Filling Method' section. The target endpoint is 'Human Health Hazards Sensitisation Skin In Vivo'. The main workspace shows a grid of chemical structures and their predicted outcomes for 'Skin sensitisation'. A red oval highlights a specific row of data points in a scatter plot, with a text overlay explaining the read-across prediction logic.

**Read across prediction of A B C, EC3, NOEL, S M W N, S W A N, Skin sensitisation, taking the highest mode from the nearest 5 neighbors, based on 6 values from 6 weight matrix chemicals. Observed target value: N/A, Predicted target value: 'Positive'**

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

Descriptor X: log Kow

# Read across prediction of constituent without data

## Apply read across

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

# Outlook

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

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

The screenshot shows the QSAR Toolbox interface with the 'Data Gap Filling' tab selected. The main window displays a table with the following columns: '1 [target]', '2 [target,mix.component]', '3 [target,mix.component]', and '4 [target,mix.component]'. The table contains three rows of chemical structures and their corresponding quantities: 10 Milligrams, 10 Milligrams, and 100 Milligrams. A filter endpoint tree on the left shows 'Skin' selected under 'Sensitisation'. A table at the bottom shows the results for skin sensitization: (3/3) R: Positive, M: Negative, and M: Negative. Red boxes and arrows highlight the 'R: Positive' prediction and the 'M: Negative' experimental data.

1 [target]	2 [target,mix.component]	3 [target,mix.component]	4 [target,mix.component]
[3] [Mix]	<chem>CC(=O)c1ccc(Cl)c(Cl)c1</chem>	<chem>O=C(c1ccccc1)c1ccccc1</chem>	<chem>CCCCO</chem>
	Qty: 1 Milligrams	Qty: 10 Milligrams	Qty: 100 Milligrams
(3/3)	R: Positive	M: Negative	M: Negative

**Here is the Read across prediction for Skin sensitization of the constituent without data**

**Here are the experimental data for Skin sensitization**

**Next step: Applying Read across for the mixture**

# Filling data gap for skin sensitization of mixture

## Applying Independent Mode of Action

The screenshot shows the QSAR Toolbox software interface during the 'Data Gap Filling' process. The top navigation bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The left sidebar shows 'Data Gap Filling Method' with 'Independent MOA' selected. The central panel displays a 'Filter endpoint tree...' with a table of endpoints. A blue callout '1' points to the 'Skin' cell in the table. A blue callout '2' points to 'Independent MOA' in the sidebar. A blue callout '3' points to the 'Apply' button. A blue callout '4' points to the 'OK' button in the 'Possible data inconsistency' dialog box. The dialog box shows a list of assays with 'Skin sensitisation II (ECETOC)' selected. A 'Starting gap filling...' message is visible in the center of the interface.

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

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



# Filling data gap for skin sensitization of mixture

## Applying Independent Mode of Action

The screenshot displays the QSAR Toolbox interface for filling data gaps. The top navigation bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling' (highlighted), and 'Report'. The left sidebar shows 'Data Gap Filling Method' with 'Independent MOA' selected. The central workspace shows four target components: 1 [target], 2 [target,mix.component], 3 [target,mix.component], and 4 [target,mix.component]. Below these are chemical structures and their predicted outcomes: 'R: Positive' for component 2, and 'M: Negative' for components 3 and 4. The right sidebar contains a 'Prediction' section with a plot titled 'Empiric calculation of A B C, EC3, taking the maximal from the component values, based on 3 values from 3 target components, Observed target value: N/A, Predicted target value: 'Positive''. The plot shows 'log Kow' on the x-axis and 'A B C, EC3 (obs.)' on the y-axis, with points for 'Positive' and 'Negative' outcomes. The right sidebar also has a 'Set empiric calculations options:' section with 'Approximation type: Maximal' selected. Red boxes and callouts labeled '1' highlight the 'Mixture models' and 'Set empiric calculations options:' sections.

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

# Filling data gap for skin sensitization of mixture

## Applying Independent Mode of Action

**1. Accept prediction**

**2. Return to matrix**

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

# Filling data gap for skin sensitization of mixture

## Applying Independent Mode of Action

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

- Navigation Bar:** Input, Profiling, Endpoint, Category Definition, **Data Gap Filling**, Report.
- Left Panel:**
  - Data Gap Filling Method:
    - Independent MOA (Selected)
    - Similar MOA
    - Specific models
  - Target Endpoint: Human Health Hazards Sensitisation Skin In Vivo
- Filter endpoint tree:**
  - Ecotoxicological Information
  - Human Health Hazards
    - Acute Toxicity
    - Bioaccumulation
    - Carcinogenicity
    - Developmental Toxicity
    - Genetic Toxicity
    - Immunotoxicity
    - Irritation / Corrosion
    - Neurotoxicity
    - Photoinduced Toxicity
    - Repeated Dose Toxicity
    - Sensitisation
      - Respiratory Tract
      - Skin
        - In Chemico
        - In Vitro
        - In Vivo
    - Toxicity to Reproduction
    - Toxicokinetics, Metabolism and ...
    - Profile

- Main Data Matrix:**

Structure	1 [target]	2 [target,mix.component]	3 [target,mix.component]	4 [target,mix.component]
[3] [Mix]	<chem>ClC1=CC=C(C=C1)C2=CC=CC=C2</chem>	<chem>ClC1=CC=C(C=C1)C2=CC=CC=C2</chem>	<chem>ClC1=CC=C(C=C1)C2=CC=CC=C2</chem>	<chem>ClC1=CC=C(C=C1)C2=CC=CC=C2</chem>
	Qty: 1 Milligrams	Qty: 10 Milligrams	Qty: 100 Milligrams	
	Cl: Positive	R: Positive	M: Negative	M: Negative

# Outlook

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

## Report

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

# Report

The screenshot shows the QSAR Toolbox software interface during the 'Report' step. The top navigation bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The left sidebar shows 'Data Gap Filling Method' (Independent MOA, Similar MOA, Specific models) and 'Target Endpoint' (Human Health Hazards Sensitisation Skin In Vivo). The main area displays a 'Filter endpoint tree...' on the left and a table of predictions on the right. The table has columns for '1 [target]', '2 [target,mix.component]', '3 [target,mix.component]', and '4 [target,mix.component]'. Each column shows a chemical structure and a quantity: 'Qty: 1 Milligrams', 'Qty: 10 Milligrams', and 'Qty: 100 Milligrams'. A context menu is open over the first prediction, with 'Report' highlighted. A red circle highlights the prediction 'Cl: Positive' in the table, and a blue callout '1' points to it. Another blue callout '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 navigation bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Report' window is open, showing a prediction for a multicomponent substance. The prediction text is: 'Prediction of EC3; Skin sensitisation; S M W N; A B C; NOEL; S W A N for {X=1/Milligrams}CC(=O)c1ccc(Cl)c(Cl)c1Cl\_{X=10/Milligrams}O=C(c1cccc1)c1cccc1\_{X=100/Milligrams}CCCCO'. The prediction is highlighted with a red oval. Below the prediction, it states: 'QSAR Toolbox prediction for multicomponent substance (uses single component mode for handling of target mixture and its components)'. At the bottom of the report, it mentions: 'The template of the current report is based on "GUIDANCE DOCUMENT ON THE VALIDATION OF (QUANTITATIVE) STRUCTURE-ACTIVITY RELATIONSHIPS MODELS" published by OECD (September, 2007) and "GUIDANCE ON INFORMATION REQUIREMENTS AND CHEMICAL SAFETY ASSESSMENT / CHAPTER R.6: QSARS AND GROUPING OF CHEMICALS" published by ECHA (May, 2008)'. The sidebar on the left shows 'Available data to report' and 'Available report templates'.

Toolbox report for mixture

# Report

The screenshot shows the QSAR Toolbox software interface. The main window displays a report titled "QSAR Toolbox prediction for mixture based on independent mode of action for mixture components". The report content includes:

- Prediction of EC3; Skin sensitisation; S M W N; A B C; NOEL; S W A N** for {X=1/Milligrams}ccc(Cl)c(Cl)c1Cl\_{X=10/Milligrams}O=C(c1cccc1\_{X=100/Milligrams})CCCCO
- Summary** (circled in red):
 

Toxicity of the target mixture (Positive) is predicted from its components using estimation based on 3 values (Positive x1, Negative x2) from 3 components having independent mode of action. Both experimental and predicted values for mixture components are used in predicting the target toxicity.

The components of a mixture are handled with the functionality for category. The same approach can be applied for mixtures, but also for mono-constituent substances with impurities, multi-constituent substances and UVCBs with identified constituents.

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

The data used for calculating the current prediction is taken from 1 Toolbox prediction and 2 experimental values selected from the following database(s):

  - Skin sensitization

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

At the bottom of the report, a table is partially visible with the following structure:

	Qty, Milligrams	Endpoint(s)
		Human Health Hazards; Sensitisation

1. Summary information for mixture prediction



# Outlook

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

The screenshot shows the QSAR Toolbox software interface. The main window displays a chemical structure and a table of results. A 'Save As' dialog box is open, showing the file name 'Tutorial 13.tbw' and the save type 'Toolbox work file (\*.tbw)'. Red callout boxes with numbers 1 through 4 highlight the 'Save' button in the toolbar, the 'Save' button in the dialog, the file name field, and the 'Save' button in the dialog respectively.

**1. Go to Input section** **2. Click on Save button** **3. Define name of the file;** **4. Click Save button**

# Open saved file

The screenshot shows the QSAR Toolbox software interface. The 'Document' menu is open, and the 'Open' option is highlighted. A 'Select file' dialog box is open, showing a list of files. The file 'Tutorial 13.tbw' is selected. The 'Open' button in the dialog is highlighted.

1. **Create** new document
2. **Click** Open;
3. **Find** and **select** file;
4. **Click** Open

# Open saved file

The screenshot displays the QSAR Toolbox interface. At the top, there is a navigation bar with icons for Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. Below this is a toolbar with various file and analysis actions. The main workspace is divided into several panes:

- Documents:** A list of open files with their respective SMILES strings and file names (e.g., 13.tbw, int\_1).
- Filter endpoint tree...:** A tree view showing various endpoints such as Substance Identity, Physical Chemical Properties, Environmental Fate and Tran..., Ecotoxicological Information, Human Health Hazards (4/4), Profile, General Mechanistic, Protein binding by OASIS..., Protein binding by OECD, Protein binding potency, and Endpoint Specific.
- Table:** A table with columns for different target mixtures (1 [target], 2 [target,mix.component], 3 [target,mix.component], 4 [target,mix.component]) and rows for different endpoints. The table contains chemical structures and quantities (e.g., Qty: 1 Milligrams, Qty: 10 Milligrams, Qty: 100 Milligrams).
- Information Dialog:** A small window in the center with an information icon and the text "The file was executed successfully" and an "OK" button.

1. The file is opened successfully      1. **Click OK**