

OECD QSAR Toolbox v.4.1

Tutorial on how to predict skin sensitisation potential by automated workflow

Outlook

- **Background**
- Objectives
- Specific Aims
- Automated workflow
- The exercise

Background

- This is a step-by-step presentation designed to take the Toolbox user through the filling of skin sensitization data gaps using an automated workflow.

Outlook

- Background
- **Objectives**
- Specific Aims
- Automated workflow
- The exercise

Objectives

This presentation demonstrates how to :

- Fill data gaps for a single chemical or batch of chemicals by automated workflow for skin sensitization (SS)

Outlook

- Background
- Objectives
- **Specific Aims**
- Automated workflow
- The exercise

Specific Aims

- To introduce to the user the automated workflow for predicting of skin sensitization potential of chemicals (single chemical or batch of chemicals)
- To familiarize the user with the new Toolbox interface and new notification messages;
- To explain to the user the rationale behind each step of the exercise.

Outlook

- Background
- Objectives
- Specific Aims
- **Automated workflow**
- The exercise

Automated workflow for skin sensitization

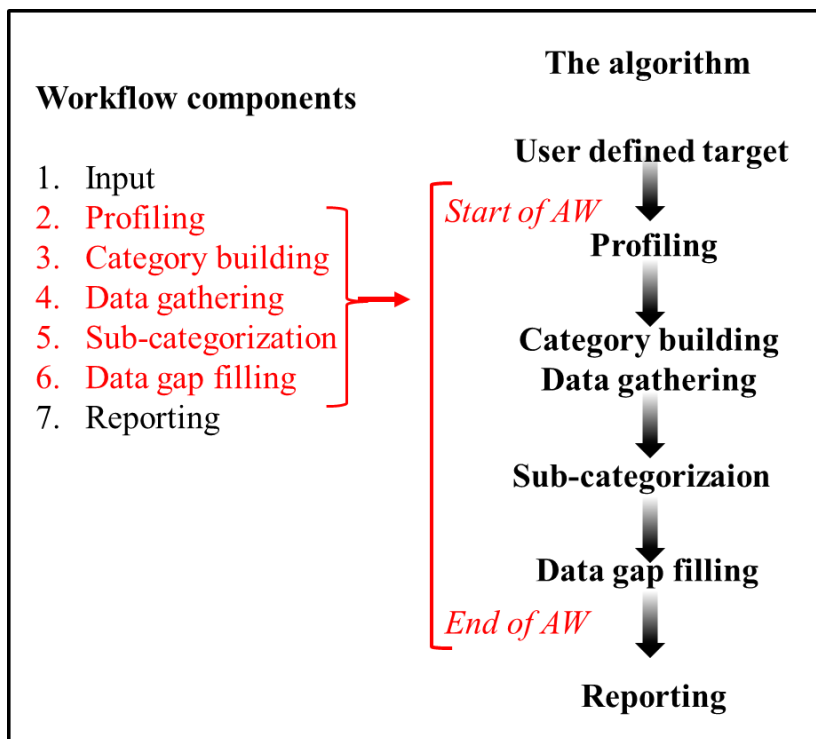
Mechanistic understanding

- Skin sensitization effect is a complex endpoint resulting in a multifactorial sequence of events.
- The initiating event of sensitization process is the covalent binding of a target chemical to the skin proteins.
- The target chemical may have inherent reactivity towards skin proteins or can be activated following metabolic or abiotic transformations.

Automated workflow for skin sensitization

Workflow components

- Except *Input* and *Reporting*, the rest of the Toolbox modules are part of the automated workflow (AW).

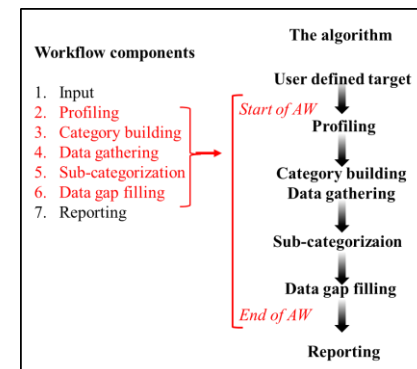


Automated workflow for skin sensitization

Workflow components

• Profiling

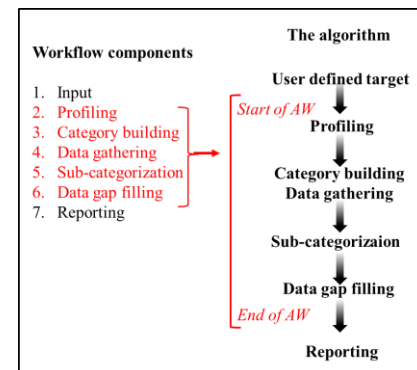
- ✓ Profilers for primary grouping
 - US EPA New Chemical Categories
 - Aquatic toxicity classification by ECOSAR
 - Protein binding alerts for Skin sensitization effect
 - Organic Functional Groups
 - Organic Functional Groups by US EPA
 - Organic Functional Groups by N. Haider
- ✓ Supporting profilers for further improvement of the category
 - Substance type
 - Protein binding potency
 - Keratinocyte gene expression
 - Structure Similarity
- ✓ Abiotic and biotic activation of chemical is accounted by application of respective Autoxidation (AU) and Skin metabolism (SM) simulators



Automated workflow for skin sensitization

Workflow components

- **Data** – SS predictions are based on LLNA and GPMT exp. data.
 - ✓ LLNA potency categories are Extreme, Strong, Moderate, Weak and Non sensitizer
 - ✓ GPMT potency categories are Strong, Moderate, Weak and Non sensitizer
 - ✓ The dichotomous scale converting LLNA and GPMT potency categories into **Positive** and **Negative** SS is preferred and used for prediction purposes



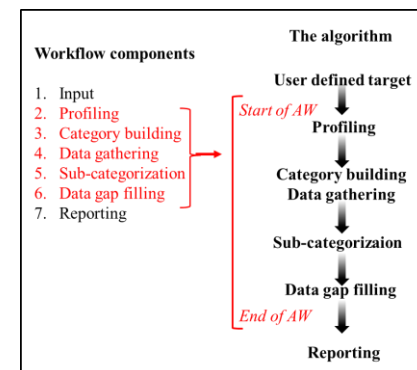
Unified skin sensitization scale in Toolbox	LLNA	GPMT
Positive	Extreme, Strong, Moderate, Weak	Strong & Moderate, Weak
Negative	Non	Non

Automated workflow for skin sensitization

Workflow components

- **Category definition**

- ✓ If protein binding alert (**PBA**) is identified in the target structure then the primary category is based on this alert
- ✓ If **PBA** is identified after AU or SM simulation then the primary category is defined accounting the metabolic simulation.



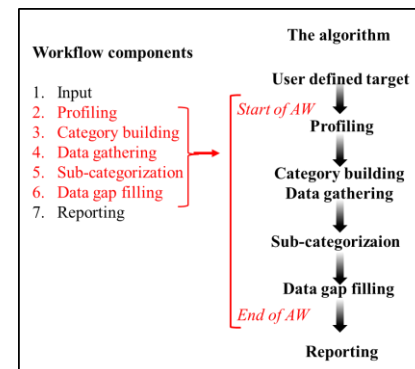
- ✓ If **more than one PBA** are identified in the parent structure or in the generated metabolites, then:
 - the category is defined based on all available PBA as presented in the target structure
 - In case no analogues found, the selection of alert is based on the criteria for **reliability** of alerts, i.e. most reliable alert is selected (see next slide)
- ✓ If **No PBA** is identified in the parent structure and in the generated metabolites, then the primary category is defined on global molecular features by using:
 - OFGs
 - *Acute aquatic classification by ECOSAR*
 - *US-EPA New Chemical categories*

In this case, the most appropriate category is the collection of a **broader group** of analogues

Automated workflow for skin sensitization

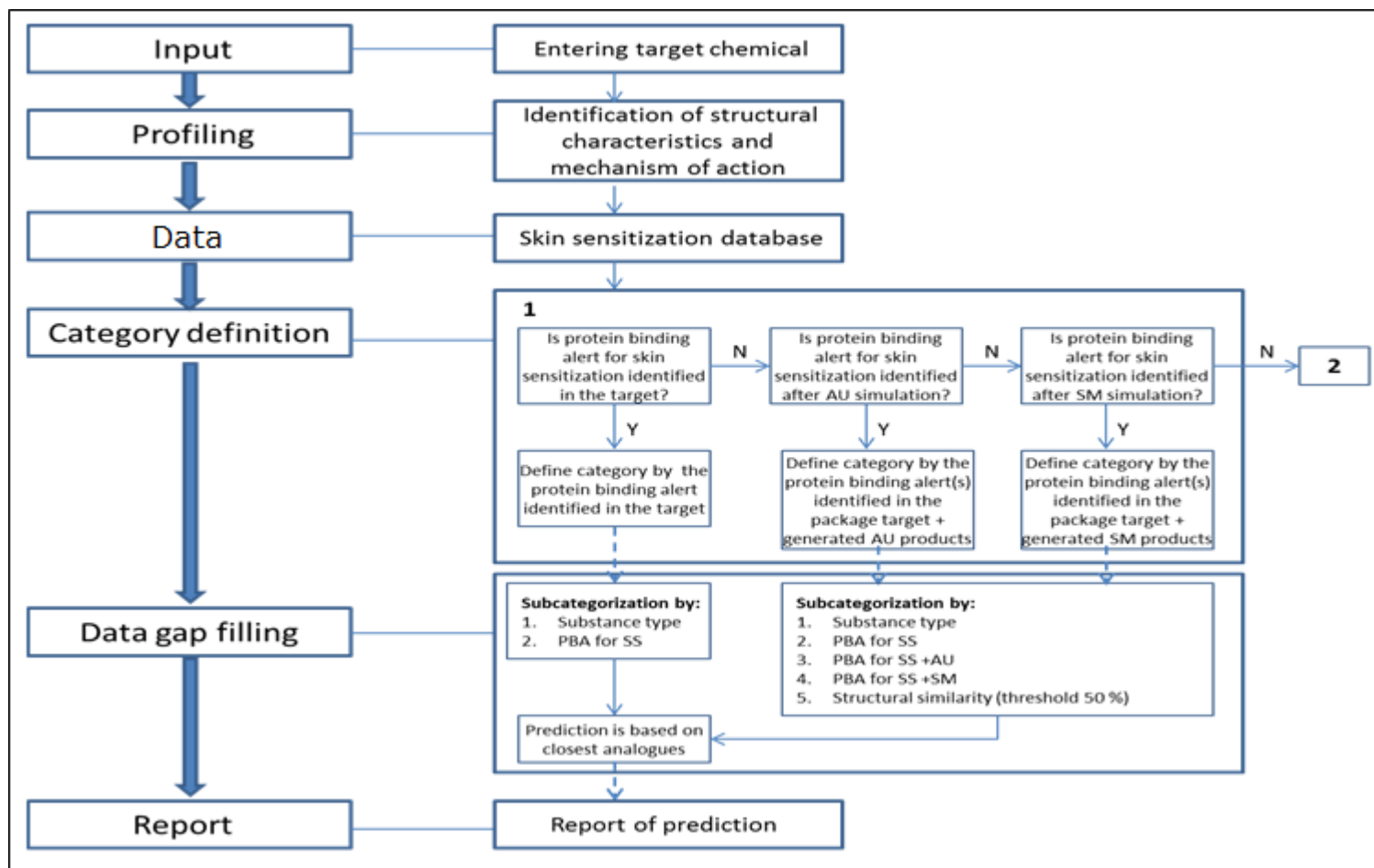
Workflow components

- **Data gap filling – prediction is based on up to five closest analogues with respect to logKow.**
 - ✓ Read across is applied as default gap filling approach
 - ✓ Specific subcategorizations are applied depending on the profiling result and subsequent primary group formation (see next slide for more information)



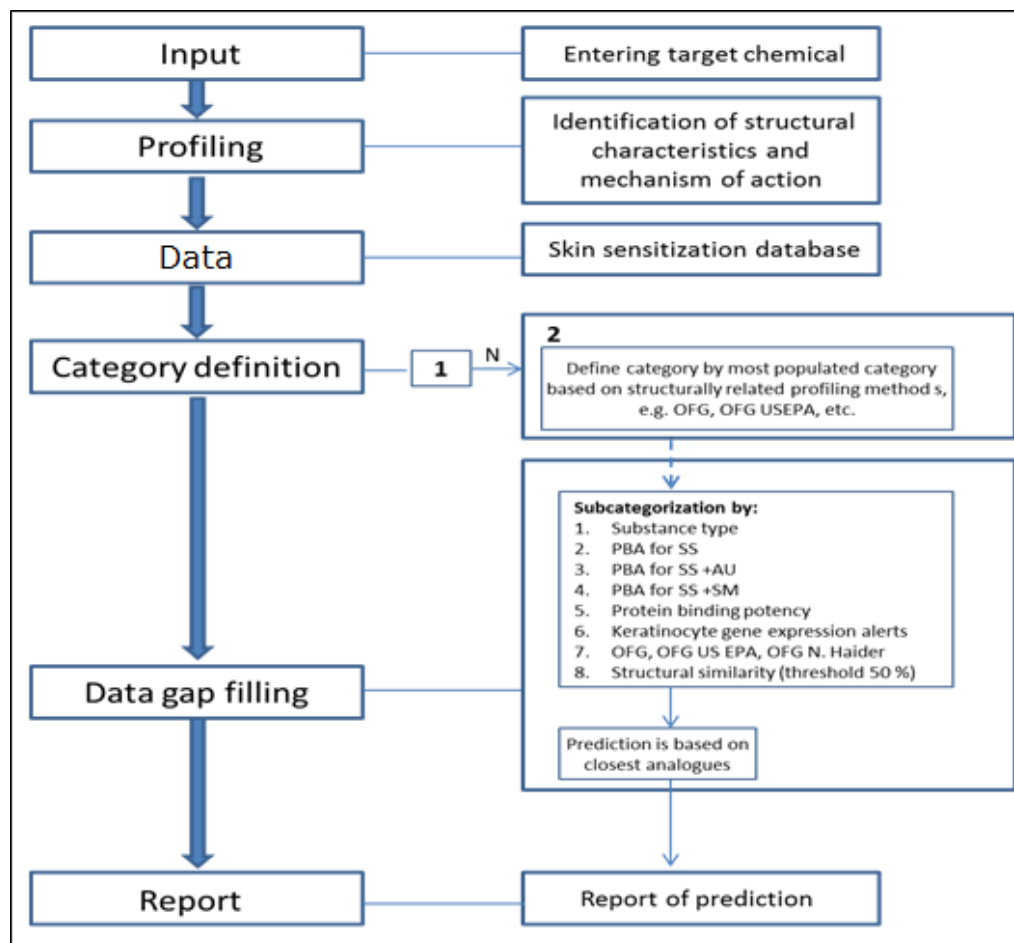
Automated workflow for skin sensitization

Algorithm of the workflow – Part 1



Automated workflow for skin sensitization

Algorithm of the workflow – Part 2



Automated workflow for skin sensitization

Overview

- The automated workflow for skin sensitization is based on LLNA and GPMT experimental data, only
- The AW follows the implemented logic and finished with prediction;
- The automated workflow (AW) is designed to apply data gap filling for discrete chemicals, only;
- The AW does not allow interactions during the workflow process;
- Analogous sets of chemicals are selected based on the hypothesis that the toxicological effects of each member of the category will show a common behaviour;
- AW can be executed for one chemical as well as for a batch of chemicals.

Outlook

- Background
- Objectives
- Specific Aims
- Automated workflow
- **The exercise**

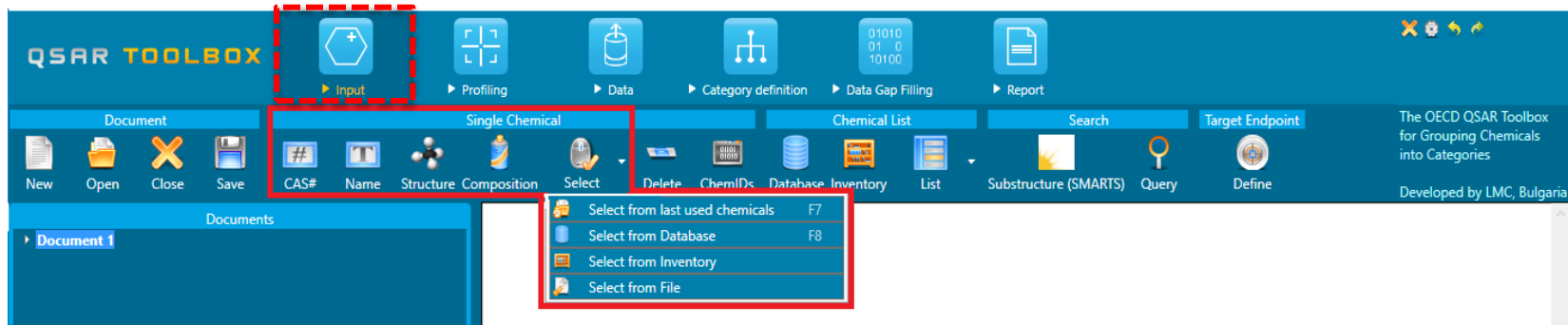
The Exercise

- In this exercise we will predict the skin sensitization effect for:
 - **Single chemical – Chlorpyrifos [CAS# 2921-88-2];**
 - Batch of chemicals - the first five substances of the Skin sensitization ECETOC database.
- This prediction will be accomplished by using of the automated workflow for skin sensitization.

Execution of AW for Skin sensitization

Single chemical, CAS # 2921-88-2

Input: Ways of Entering a Single Chemical



- Chemical Name
- Chemical Abstract Services (CAS) number (#)
- Drawing chemical structure with or without additional defining of the composition
- Select from User List/Inventory/Databases

Execution of AW for Skin sensitization

Single chemical, CAS # 2921-88-2

Input by CAS number

The screenshot shows the QSAR Toolbox software interface. The 'Input' menu is highlighted with a red dashed box. The 'CAS#' option in the 'Input' menu is highlighted with a red box and labeled with a callout box '1'. The 'Search by CAS #' dialog box is open, showing the CAS number '2921882' entered in the input field and the 'Search' button highlighted with a callout box '3'. The dialog box also shows 'Selected 0 of 0' and buttons for 'Select All', 'Unselect All', and 'Invert Selection'. A callout box '2' points to the input field in the dialog box.

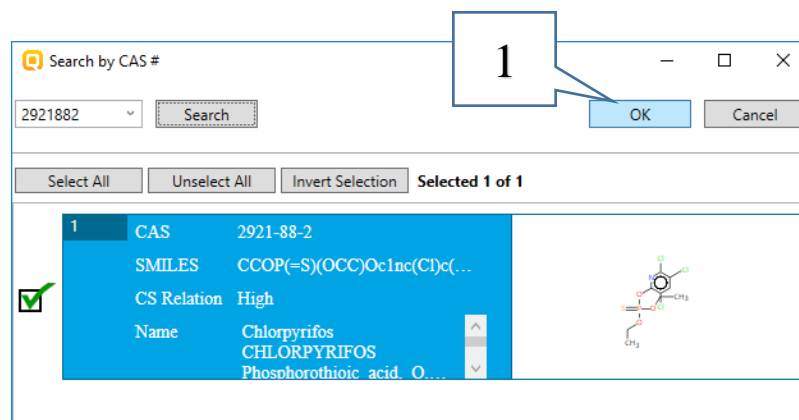
1. Click on **CAS#**;
2. Enter the CAS# of Chlorpyrifos (CAS **2921-88-2**);
3. Click on **Search**.

Execution of AW for Skin sensitization

Single chemical, CAS # 2921-88-2

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. Click on **OK** (1).



In case the entered CAS# corresponds to more than one structure or to one structure but with different predefined substance type, more than one chemical identity could be retrieved. In this case the user can decide which substance to be retained for the subsequent workflow.

Execution of AW for Skin sensitization

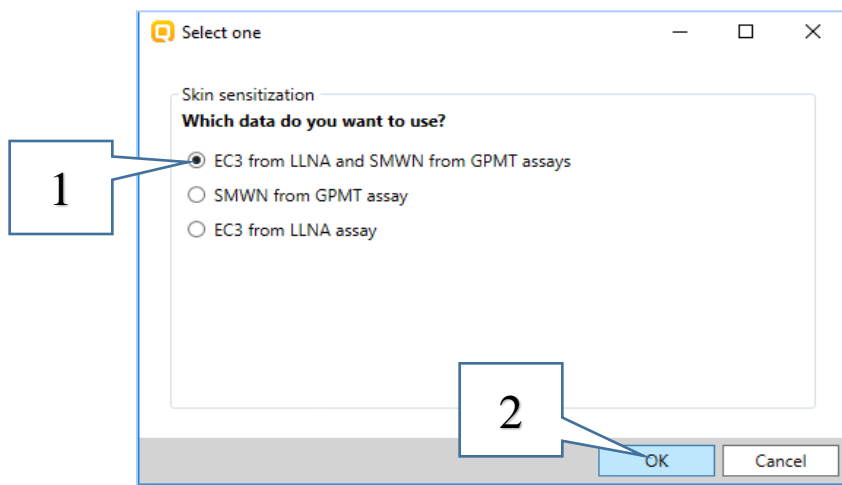
Location of the Automated workflow for skin sensitization

The screenshot displays the QSAR Toolbox software interface. The top menu bar contains several options: 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling' (highlighted with a red dashed box and labeled '1'), and 'Report'. Below the menu bar, the 'Workflows' section is active, with the 'Automated' button highlighted (labeled '2'). On the left, the 'Filter endpoint' list includes 'Structure', 'Structure info', 'Parameters', 'Physical Chemical Properties', 'Environmental Fate and Transport', 'Ecotoxicological Information', and 'Human health hazards'. The 'Skin sensitization' option is selected (labeled '3'). A 'Select workflow' dialog box is open, showing 'Ecotoxicological Endpoint' and 'Skin sensitization' as choices, with 'Skin sensitization' selected. The 'OK' button is highlighted (labeled '4').

1. Go to **Data Gap Filling**; 2. Click on **Automated** button; 3. Select **Skin sensitization**; 4. Confirm with **OK**

Execution of AW for Skin sensitization

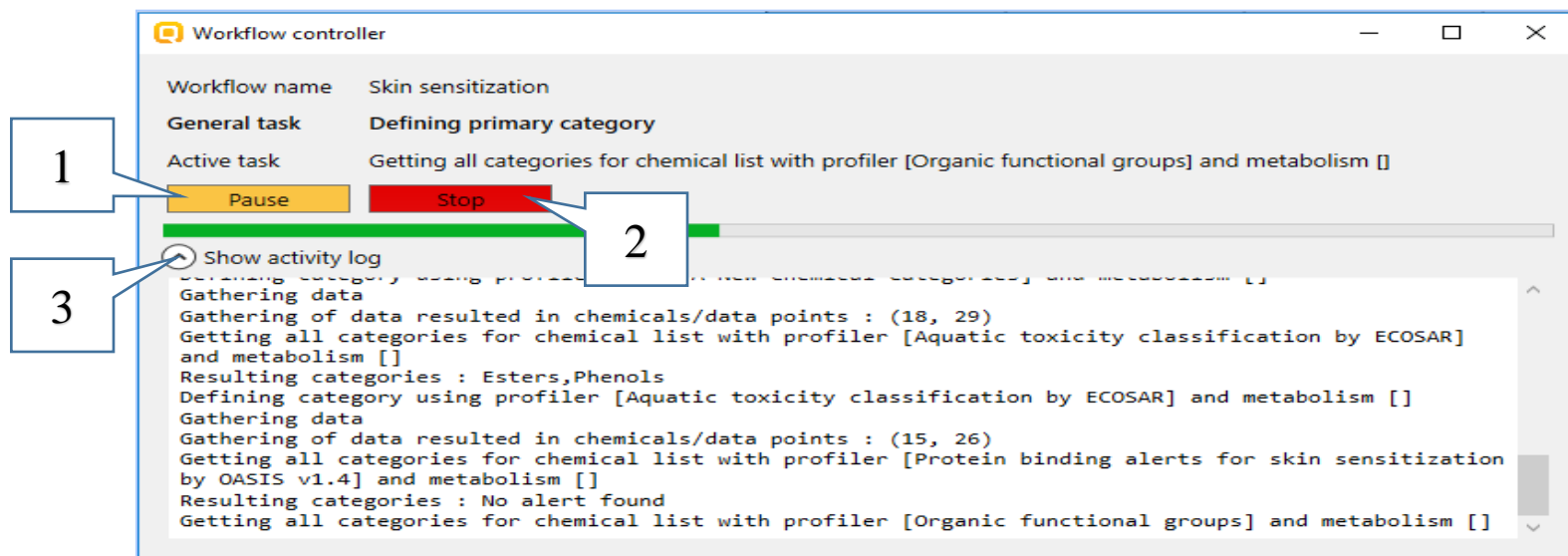
There are three options for endpoint selection and the user should select one of them.



1. Select the first combined endpoint – ***EC3 from LLNA and SMWN from GPMT assays***; 2. Confirm with ***OK***

Execution of AW for Skin sensitization

Once the workflow is started, the **Workflow controller** appears. It has two main buttons **Continue/Pause** (1) allowing to continue or pause and **Stop** (2) – which stops the workflow. Furthermore, all actions that have been done during the execution of the workflow are tracked down and could be seen from the **Show activity log** (3) part of the Workflow controller.



Execution of AW for Skin sensitization

During the process, the user can see so called “Helpers” (1). This is a new functionality intended to help the user with specific information associated with analogues used in gap filling approach.

Execution of AW for Skin sensitization

Information messages and functionalities

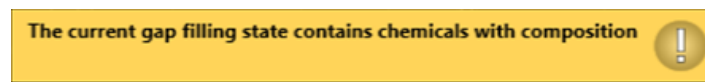
The helpers provide different type of information, e.g.:

- warning messages, alerting the user:

- for qualifiers availability



- for analogue(s) with composition



- notification messages:

- if the read-across prediction is obtained by more analogues than the default 5. This could be due to chemicals with equal descriptor values (logKow values)

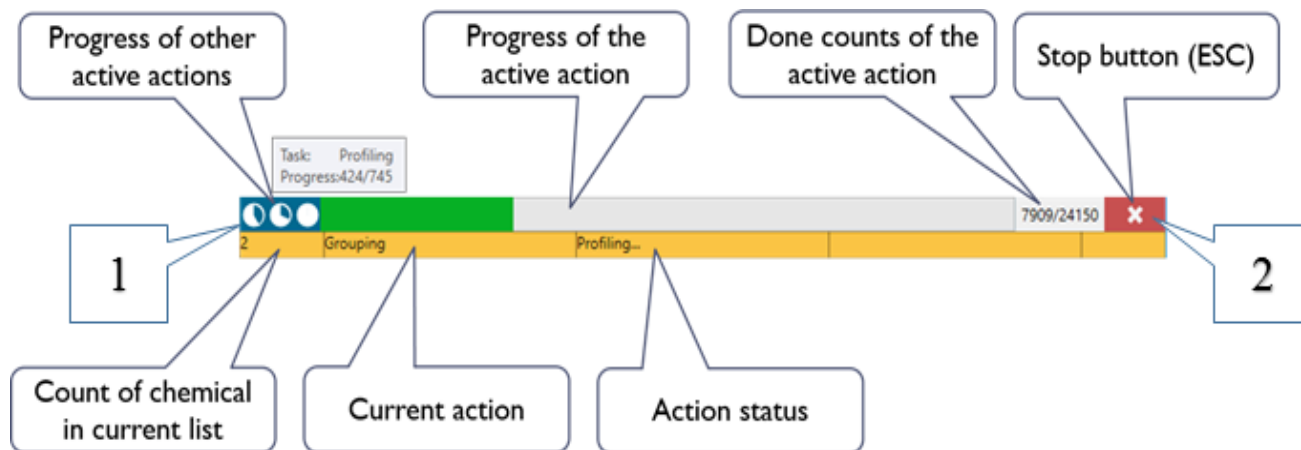


Execution of AW for Skin sensitization

Information messages and functionalities

Functionality used to follow the progress of the active action or to cancel an active action when an Automated workflow is run.

- To see the progress of the active action click on the clock (1)
- To cancel an active action click on Stop button (2)



Execution of AW for Skin sensitization

The automated workflows follow the implemented logic (see slides 10-16) and finish with accepted prediction.

The screenshot shows the QSAR Toolbox interface. On the left, a 'Filter endpoint tree...' sidebar lists various endpoints, with 'Sensitisation' expanded to show 'EC3 <OR> S M W N'. The main area displays a data matrix with columns numbered 1 to 6. The first column is labeled '1 [target]'. A 'Success' dialog box is open in the center, displaying 'Prediction accepted successfully' and an 'OK' button. A red box highlights the 'OK' button with a callout '2'. Another red box highlights a cell in the matrix (row 'EC3 <OR> S M W N', column 1) with a callout '1'. The cell contains '(18/33) M: Positive R: Positive'. The matrix has a yellow background for the first few columns.

When the prediction is accepted successfully, it appears on the matrix (1). In this case the predicted result (R:) corresponds to the experimental data (M:). Click on “OK” (2)

Execution of AW for Skin sensitization

The screenshot displays the QSAR Toolbox interface. At the top, there are tabs for 'Gap Filling' and 'Workflow'. Below this is a 'Filter endpoint tree...' section showing a list of endpoints with corresponding chemical structures. The main area is a 'Data points' table with the following columns: Datapoints, #, Value, Original value, Assay, Assigned SMILES, Author, Comments, Database, Endpoint, Institution and country, Organ, and Prediction approach.

Datapoints	#	Value	Original value	Assay	Assigned SMILES	Author	Comments	Database	Endpoint	Institution and country	Organ	Prediction approach
Human health hazards;Sensitisation	1	M: Positive (Skin sensitisation II (ECETOC))	6.92 % (Skin sensitization EC3 (ratio))	LLNA	False	Submitted by Dow AgroSciences	Strong sensitizer = EC3 < 10%; Weak sensitizer = EC3 >=10% < 50%; Non sensitizer = EC3 >=50%	Skin Sensitization	EC3	LMC,BUL	Skin	
Human health hazards;Sensitisation	2	R: Positive (Skin sensitisation II (ECETOC))	Positive (Skin sensitisation II (ECETOC))						EC3		Skin	Read-across, executed by AW 'Skin sensitization'

Below the table is a 'Data Gap Filling Settings' panel with options for 'Only endpoint relevant' and 'Only chemical relevant'. At the bottom, a tree view shows the workflow structure, with a callout '1' pointing to a cell containing 'M: Positive' and 'R: Positive'. Callout '2' points to the 'Data points' window, and callout '3' points to the 'Read-across' prediction approach in the table.

Double click on the cell with the prediction (1) opens the *Data points* window. Here the user can see more information for their prediction. It is seen that only EC3 data is used for the prediction (2) and the prediction approach is read across executed by AW “Skin sensitization”.

Execution of AW for Skin sensitization

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes options like Input, Profiling, Data, Category definition, Data Gap Filling, and Report. The main window displays the 'Workflow Controller [Finished workflow]' dialog box. The dialog box contains the following information:

- Workflow name: Skin sensitization
- General task: Acceptance resolution
- Active task: Show activity log

The 'Show activity log' button is highlighted with a blue box and the number '1'. The log content is as follows:

```

Initial data collection
Select databases
Setting selected databases to [Skin Sensitization, Skin sensitization ECETOC] and inventories to []

Alert detection and primary category definition
Retrieving alerts
Retrieving alerts with profiler : [Protein binding alerts for skin sensitization by OASIS v1.4] and metabolism []
Getting all categories for chemical list with profiler [Protein binding alerts for skin sensitization by OASIS v1.4] and metabolism []
Resulting categories : SNAr, SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds, SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds >> Activated aryl and heteroaryl compounds, SN2, SN2 >> Nucleophilic substitution at sp3 carbon atom, SN2 >> Nucleophilic substitution at sp3 carbon atom >> (Thio)Phosphates
Active alerts found : [SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds >> Activated aryl and heteroaryl compounds, SN2 >> Nucleophilic substitution at sp3 carbon atom >> (Thio)Phosphates]
Defining category using profiler [Protein binding alerts for skin sensitization by OASIS v1.4] and metabolism []
Computing alert performance for profiler : [Protein binding alerts for skin sensitization by OASIS v1.4] and metabolism [] : scale [Skin sensitisation II (ECETOC)] and category [Positive]
Defining category using profiler [Protein binding alerts for skin sensitization by OASIS v1.4] and metabolism []
Using category : SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds >> Activated aryl
    
```

The close button (X) in the top right corner of the dialog box is highlighted with a red box and the number '2'. The background shows a table of results with columns for 'R: Positive', 'M: Positive', and 'M: Positive'. The first row of the table is highlighted in yellow.

The *Workflow Controller* **does not** close itself automatically. The user can expand the *activity log* (1) and to examine all performed steps during the Automated workflow execution. After that the controller have to be closed by click on the close button (2).

Execution of AW for Skin sensitization

The screenshot displays the QSAR Toolbox interface. At the top, there is a toolbar with icons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below this, there are tabs for Gap Filling and Workflow. The 'Documents' panel on the left is highlighted with a red box and labeled with a '1' in a white speech bubble. It shows a tree view of documents, with the following structure:

- Document 3
 - # CAS: 2921882
 - Protein binding alerts for skin sensitization by OASIS
 - Enter GF (AW by read-across) with 21 Chemicals with 37 data
 - Data usage options are changed to: Maximal
 - Ch: 20| Data: 34 Subcategorized: Substance type
 - Ch: 15| Data: 28 Subcategorized: Protein binding alerts for skin

Red arrows point to the 'Ch: 15| Data: 28 Subcategorized: Protein binding alerts for skin' entry. To the right of the 'Documents' panel is a 'Filter endpoint tree...' section with a list of categories: Structure, Structure info, Parameters, Physical Chemical Properties, Environmental Fate and Transport, Ecotoxicological Information, and Human Health Hazards. Further right is a table with 4 columns and 1 row, containing chemical structures and data.

1 [target]	2	3	4

The new organization of the Toolbox allows the *Documents* panel (1) to be consistent in each of the toolbar sections. This gives the possibility to the user to go back/forward through the subcategorization steps by moving up/down to the document tree. In this way the user can follow each step of the executed automated workflow.

Execution of AW for Skin sensitization

Report: Generation of report

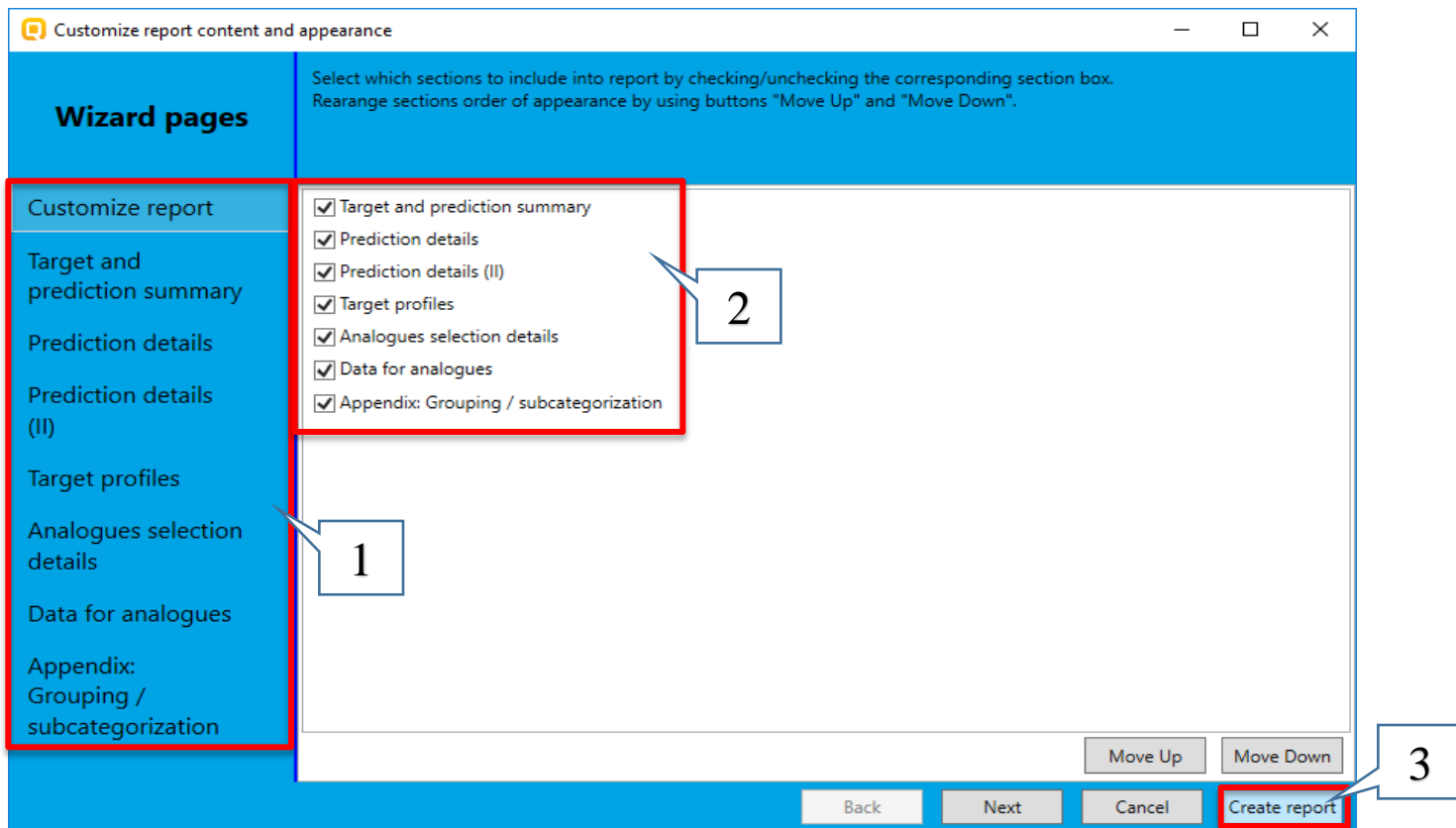
The screenshot displays the QSAR Toolbox interface. The top navigation bar includes buttons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report. The Report button is highlighted with a red dashed box and labeled '1'. Below the navigation bar, the Reports menu is open, showing options for Prediction, Data Matrix, Category, and QMRF. The Prediction option is highlighted with a red dashed box and labeled '3'. The main workspace shows a 'Filter endpoint tree...' on the left, with 'Sensitisation' expanded to 'Skin' and 'in Vivo' selected. The 'AW SW AOP' is set to '(15/29)'. The main table displays chemical structures and their predictions for skin sensitization across columns 1 through 11. The prediction for column 1 is 'M: Positive' and 'R: Positive', which is highlighted with a blue background and labeled '2'. Other columns show 'M: Positive' predictions with yellow backgrounds.

1 [target]	7	8	9	10	11
M: Positive R: Positive	M: Positive M: Positive	M: Positive M: Positive	M: Positive	M: Positive	M: Positive

1. Go to the **Report** module;
2. Click on the cell with the prediction;
3. Click on **Prediction**

Execution of AW for Skin sensitization

Report: Generation of report



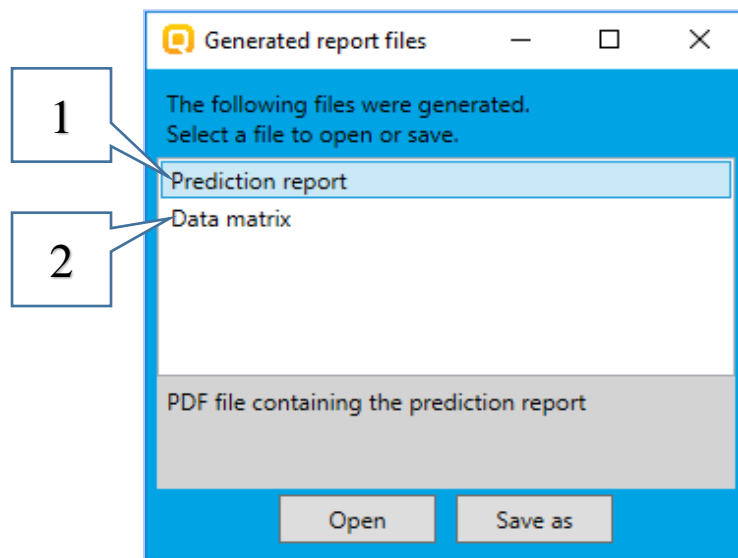
The user can customize the report content (1) and appearance (2). Generation of the reports happens by click on the **Create report** button (3).

Execution of AW for Skin sensitization

Report: Generation of report

After the click on the Create report button, *Generated report files* window appears. It contains two type of files:

- 1) **Prediction report** - a PDF file containing the prediction information related to the target.
- 2) **Data matrix** - a MS Excel file containing chemicals used for prediction along with their data for selected parameters, profiles and endpoint tree positions.



Execution of AW for Skin sensitization

Report: Generation of report

Prediction of EC3 for LORSBAN

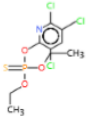
1 / 8

QSAR Toolbox prediction for single chemical

Date: 1 abr 2017

Author(s):

Contact details:

Target information		
Structural information SMILES: <chem>CCOP(=S)(OCC)Oc1nc(Cl)c(Cl)cc1Cl</chem> Structure 	Numerical identifiers EC#: N/A CAS#: 2921-88-2 Other: N/A	Chemical names Chlorpyrifos CHLORPYRIFOS (DURSBA N) chlorpyrifos (dursba n) (cpf) (o,o-diethy l-o-(3,5,6-trichloro -2-pyridyl)phosphoro thiolate) (chlorpyri phos)

Usage of the automated workflow for predicting of skin sensitization potential is noted in the *Prediction report*.

Prediction summary
Predicted endpoint: EC3; No effect specified; No species specified; No duration specified; No guideline specified Predicted value: Positive Unit/scale: Skin sensitisation II (ECETOC) Data gap filling method: Read-across analysis, executed via AW "Skin sensitization" Summary: manually editable field Not provided by the user

Execution of AW for Skin sensitization

Report: Generation of report

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1			Target chemical			Neighbour #1			Neighbour #2			Neighbour #3			Neighbour #4			Neighbour #5			
2	Substance identity																				
3	Structure																				
4	CAS number																				
5	Chemical name																				
6	Other identifier																				
7	SMILES																				
8	Parameters																				
9	unit																				
10	Profilers																				
11	Profiles used for grouping/subcategorization																				
12	SN2; SN2 >> Nucleophilic substitution at sp3 carbon atom; SN2 >> Nucleophilic substitution at sp3 carbon atom >> (Thio)Phosphates; SNAr; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds >> Activated aryl and heteroaryl compounds																				
13	SNAr; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds >> Activated aryl and heteroaryl compounds																				
14	SNAr; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds >> Activated aryl and heteroaryl compounds																				
15	SNAr; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds >> Activated aryl and heteroaryl compounds																				
16	SNAr; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds; SNAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds >> Activated aryl and heteroaryl compounds																				
17	Measured and predicted data																				
18	Data used for prediction																				
19	environment	endpoint	value	unit	reference, species, duration, test type, type of method	value	unit	reference, species, duration, test type, type of method	value	unit	reference, species, duration, test type, type of method	value	unit	reference, species, duration, test type, type of method	value	unit	reference, species, duration, test type, type of method	value	unit	reference, species, duration, test type, type of method	
20	Sensitisation	EC3			Submitted by Dow AgroSciences	6.15	%	Submitted by Dow AgroSciences	6.14	%	Submitted by Dow AgroSciences	6.59	%	Submitted by Dow AgroSciences	0.004	%	Submitted by Dow AgroSciences	0.004	%	Submitted by Dow AgroSciences	0.004
21	Sensitisation	EC3			Submitted by Mouse	7.31	%	Submitted by Mouse	7.31	%	Submitted by Mouse	7.31	%	Submitted by Mouse	7.31	%	Submitted by Mouse	7.31	%	Submitted by Mouse	7.31

Analogues used for the target prediction can be seen in the **Data matrix** report. Their selected profiling results, experimental data and/or parameters are also shown.

Execution of AW for Skin sensitization

Recap

The application of the automated workflow aims to facilitate the Toolbox users` work limiting their activities. The AW for SS requires only user activity such as:

- defining the target molecule by entering it in Toolbox by using either CAS#, name or smiles
- activate the automated workflow
- specification of the endpoint

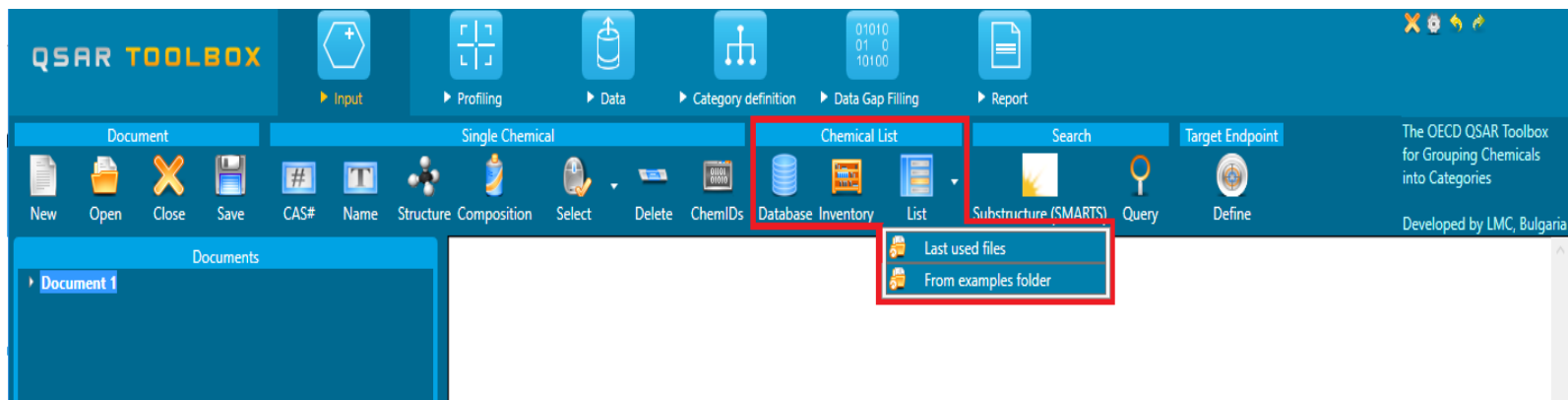
The Exercise

- In this exercise we will predict the skin sensitization effect for:
 - Single chemical – Chlorpyrifos [CAS# 2921-88-2];
 - **Batch of chemicals - the first five substances of the Skin sensitization ECETOC database.**
- This prediction will be accomplished by using of the automated workflow for skin sensitization.

Execution of AW for Skin sensitization

List with chemicals (batch work)

Input: Ways of Entering a Chemical List



- Database
- Inventory
- List: Last used files/ From examples folder

Execution of AW for Skin sensitization

List with chemicals (batch work)

Input: Load a database

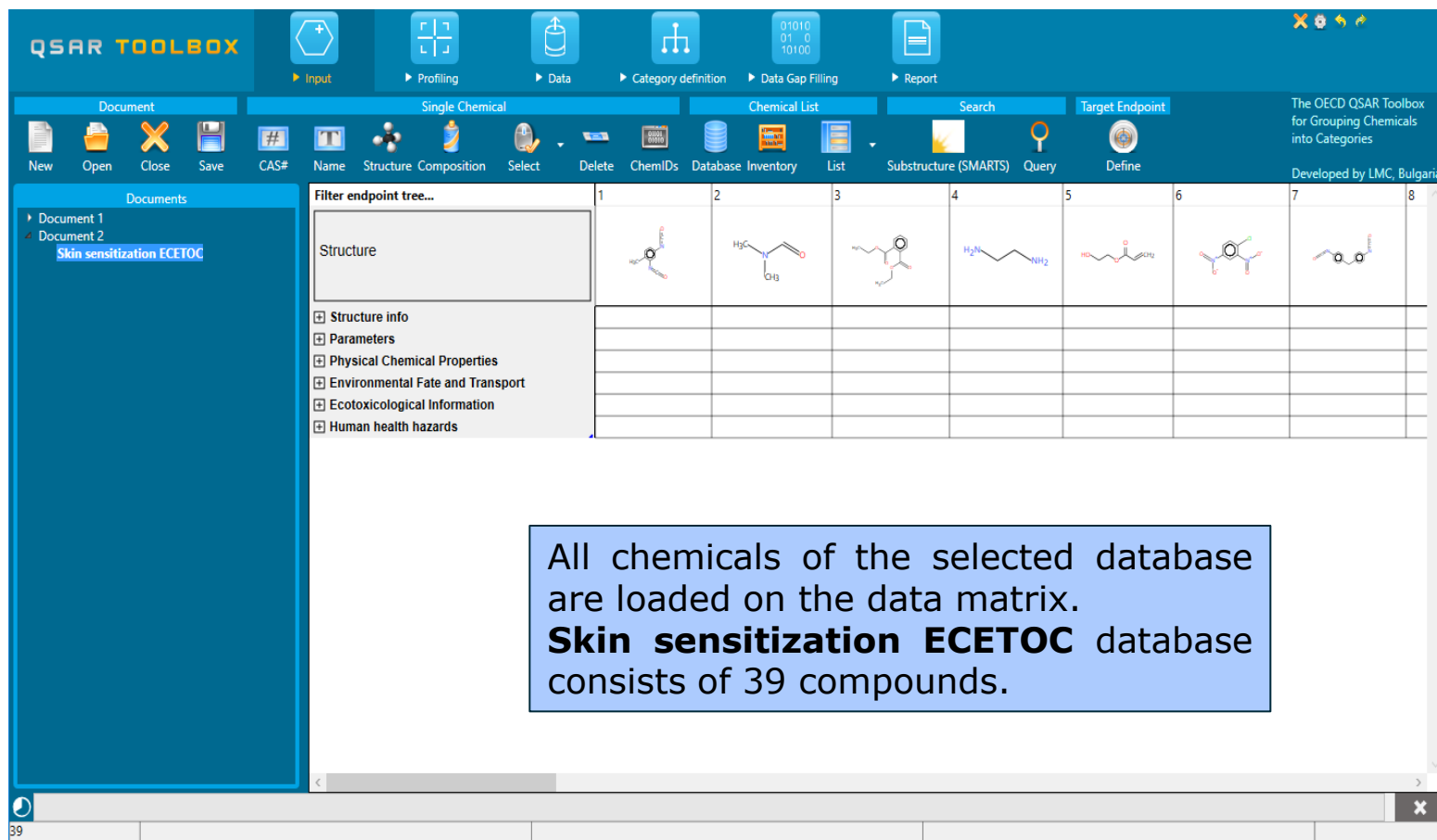
The screenshot shows the QSAR Toolbox software interface. The top toolbar contains several icons for different functions: Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below the toolbar is a ribbon with tabs for Document, Single Chemical, Chemical List, Search, and Target Endpoint. The 'Database' button in the Chemical List tab is highlighted with a red box and a callout box labeled '1'. A 'Select database' dialog box is open, showing a list of databases. The 'Skin sensitization ECETOC' database is selected, highlighted with a blue box and a callout box labeled '2'. The 'OK' button at the bottom right of the dialog box is highlighted with a red box and a callout box labeled '3'. A blue box on the left contains the following instructions:

1. Click on **Database**;
2. Select **Skin sensitization ECETOC**;
3. Click on **OK**.

Execution of AW for Skin sensitization

List with chemicals (batch work)

Input: Load a database



The screenshot displays the QSAR Toolbox interface. The top menu bar includes options like Document, Single Chemical, Chemical List, Search, and Target Endpoint. Below the menu is a toolbar with icons for various functions. The main workspace is divided into a left sidebar and a central table. The sidebar shows a document tree with 'Skin sensitization ECETOC' selected. The central table has 8 columns and 10 rows. The first row contains chemical structures for columns 1 through 8. The remaining rows are empty. A text box at the bottom right of the screenshot states: "All chemicals of the selected database are loaded on the data matrix. **Skin sensitization ECETOC** database consists of 39 compounds."

Execution of AW for Skin sensitization

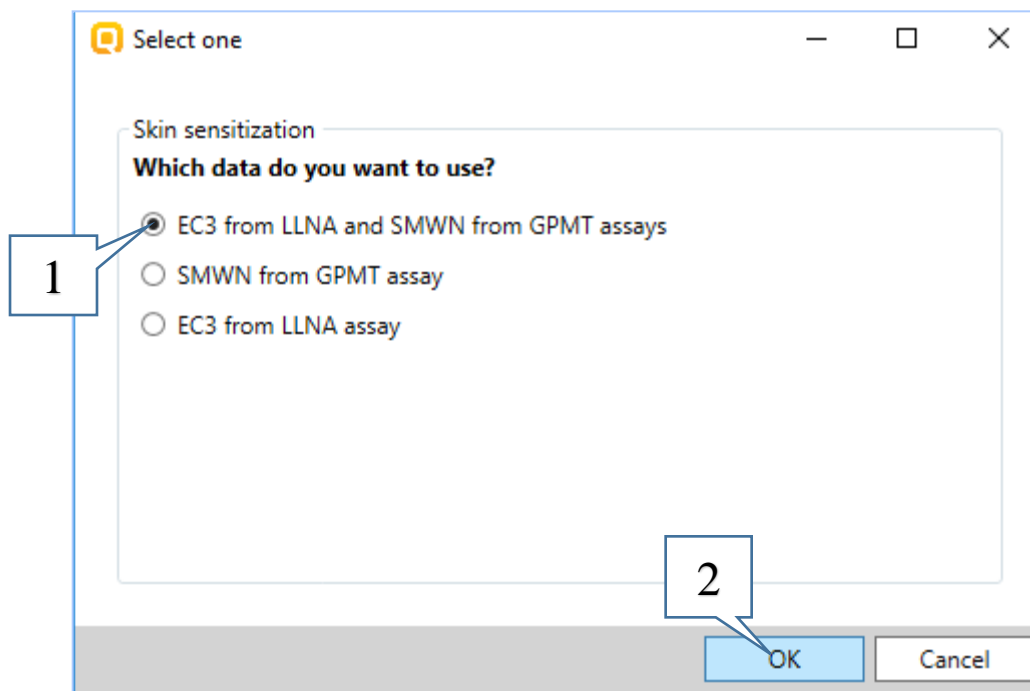
List with chemicals (batch work)

The screenshot displays the QSAR Toolbox software interface. The top toolbar contains icons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report. The 'Data Gap Filling' icon is highlighted with a red dashed box and labeled '1'. Below the toolbar, the 'Automated' button is highlighted with a red box and labeled '2'. A 'Data Gap Filling Settings' panel on the left shows checkboxes for 'Only endpoint relevant' and 'Only chemical relevant', both checked. Below this, the 'At this position:' section shows 'Automated workflows' and 'Standardized workflows' both set to '0'. A 'Select workflow' dialog box is open in the center, with 'Skin sensitization' selected under the 'Choices' section, highlighted with a red box and labeled '3'. The 'OK' button in the dialog is highlighted with a red box and labeled '4'. The background shows a table with chemical structures in columns 1 through 7.

1. Go to **Data Gap Filling**; 2. Click on **Automated** button; 3. Select **Skin sensitization**; 4. Confirm with **OK**

Execution of AW for Skin sensitization

List with chemicals (batch work)



1. Select the first combined endpoint – EC3 from LLNA and SMWN from GPMT assays; 2. Confirm with OK

Execution of AW for Skin sensitization

List with chemicals (batch work)

When the targeted endpoint is selected, a new dialog related to the application range of the workflow appears.

The user can define the range(s) of chemicals, on which the automated workflow will be executed. The numbering of chemicals corresponds to their number on the data matrix.

The whole range is set by default.

Execution of AW for Skin sensitization

List with chemicals (batch work)

The screenshot displays the QSAR Toolbox software interface. At the top, there are tabs for 'Gap Filling' and 'Workflow'. Below these, a 'Filter endpoint tree...' window shows a list of endpoints: Developmental toxicity, Genetic Toxicity, Immunotoxicity, Irritation / Corrosion, Neurotoxicity, Photoinduced toxicity, Repeated Dose Toxicity, Sensitisation, Skin, In Vivo, EC3 <OR> S M W N, and ToxCast. The 'Sensitisation' endpoint is selected, and a sub-tree shows 'Skin' and 'In Vivo' options. A 'Workflow controller' window is open, showing the workflow name 'Skin sensitization' and buttons for 'Pause' and 'Stop'. A 'Select range' dialog box is prominently displayed in the center, with a red border. It contains the text 'Add range' and two input fields: '>= 1' and '<= 39'. An 'Ok' button is located at the bottom of the dialog. The background shows a grid of chemical structures in columns 1 through 7, with a blue bar under column 1 and yellow bars under columns 2 through 7. The bottom status bar shows the number '39'.

Execution of AW for Skin sensitization

List with chemicals (batch work)

The screenshot displays the QSAR Toolbox software interface. The top bar shows 'Gap Filling' and 'Workflow' tabs. The left sidebar contains 'Documents' (with 'Skin sensitization ECETOC' selected) and 'Data Gap Filling Settings'. The main area shows a 'Filter endpoint tree...' with a list of endpoints including 'Developmental toxicity / teratogenicity', 'Genetic Toxicity', 'Immunotoxicity', 'Irritation / Corrosion', 'Neurotoxicity', 'Photoinduced toxicity', 'Repeated Dose Toxicity', 'Sensitisation', 'Skin', 'In Vivo', 'EC3 <OR> S M W N', and 'ToxCast'. A 'Workflow controller' dialog is open, showing 'Skin sensitization' as the workflow name and 'General task' as the active task. A 'Select range' dialog is also open, showing two ranges: '>= 1 <= 2' and '>= 4 <= 6'. A red dashed box highlights the first two and fourth to sixth chemicals in the list. A blue box at the bottom contains the following instructions: '1. Select to execute the automated workflow for the first two chemicals and for chemicals #4 to #6; 2. Confirm with **OK**'.

Execution of AW for Skin sensitization

List with chemicals (batch work)

The screenshot displays the QSAR Toolbox interface during the execution of an Automated Workflow (AW) for skin sensitization. The 'Documents' panel on the left shows a tree view with 'Skin sensitization ECETOC' and '4/5 NCCN' selected. A 'Filter endpoint tree...' window is open, showing a list of endpoints with 'Sensitisation' selected. A 'Workflow controller' dialog box is displayed in the center, showing the workflow name 'Skin sensitization', the general task 'Performing data gap filling', and the active task 'Performing gap filling, approach: [ReadAcross]'. The dialog has 'Pause' and 'Stop' buttons and a progress bar. The main workspace shows a table with 6 columns and 1 row of chemical structures. A status bar at the bottom indicates 'Batch execution Automated Skin sensitization'.

1 [target]	2	3	4	5	6
<chem>NCCN</chem>	<chem>C1CCNCC1</chem>	<chem>C1CCNCC1</chem>	<chem>C1CCNCC1</chem>	<chem>C1CCNCC1</chem>	<chem>C1CCNCC1</chem>

At this position:
 Select a cell with a rigid (bold) path
 Automated workflows 0
 Standardized workflows 0

All chemicals in the selected ranges are predicted automatically in a row. The user can see the number of the running chemical out of all chemicals for prediction (1).

Execution of AW for Skin sensitization

List with chemicals (batch work)

- The workflow in a batch mode finishes when automated workflow is executed for last chemical in the specified range(s).
- No messages for the individual predictions appear during the workflow.
- In some cases following the implemented logic in the workflow, no analogues are found or they are removed during the subcategorization process. Then no prediction result appears in the data matrix cell (see on the next slide).

Execution of AW for Skin sensitization

List with chemicals (batch work)

The screenshot shows the QSAR Toolbox interface with a workflow controller dialog box. The dialog box displays the following information:

- Workflow name: Skin sensitization
- General task: Acceptance resolution
- Active task: (empty)
- Buttons: Pause, Stop
- Show activity log: (checked)

The main interface shows a table with 5 columns representing different chemicals. The 'Filter endpoint tree...' section on the left lists various endpoints, with 'Sensitisation' expanded to show 'Skin' and 'In Vivo'. The 'In Vivo' section is further expanded to show 'EC3 <OR> S M W N'. The results for this endpoint are shown in a table below:

Chemical	1	2	3	4	5
EC3 <OR> S M W N (4/4)	R: Positive	R: Negative	R: Positive	R: Negative	

A callout box points to the 'R: Positive' result for the first chemical, stating: "There is no prediction".

Execution of AW for Skin sensitization

List with chemicals (batch work)

If automated workflow in single mode is executed for such kind of chemicals, message advising to try the standardized workflow appears.

The screenshot displays the QSAR Toolbox software interface. On the left, a 'Filter endpoint tree...' shows a selected endpoint: 'EC3 <OR> S M W N'. The 'Structure' panel shows a chemical structure of 2-methylacrolein. The 'Workflow Controller [Automated mode]' window is open, showing the workflow name 'Skin sensitization', general task 'Acceptance resolution', and active task 'Accepting prediction'. A red box highlights a 'Message' dialog box that appears, containing the text: 'Couldn't make a prediction using automated workflow, please try using standardized workflow instead'. Below the workflow controller, a 'Read-across prediction for EC3 <OR> S M W N' plot is visible, showing a single data point at approximately log Kow = -1.1. The plot title indicates 'Observed: not converted (x2); Predicted: N/A'. The bottom right corner of the interface shows a 'Accent predicti...' button.

Execution of AW for Skin sensitization

List with chemicals (batch work)

Report: Generation of report

On this stage, it is not possible to generate report for chemicals predicted in batch mode.

The system kindly informs you that *Predictions from batch mode cannot be reported.*

Congratulation

- You have completed the tutorial on the automated workflow for skin sensitization data gap filling.
- You have been introduced to the automated workflow in a single and in a batch mode.
- You have now been introduced to the consecutive steps of the AW for SS and the rationale behind each step.
- Note, proficiency comes with practice!