QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

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

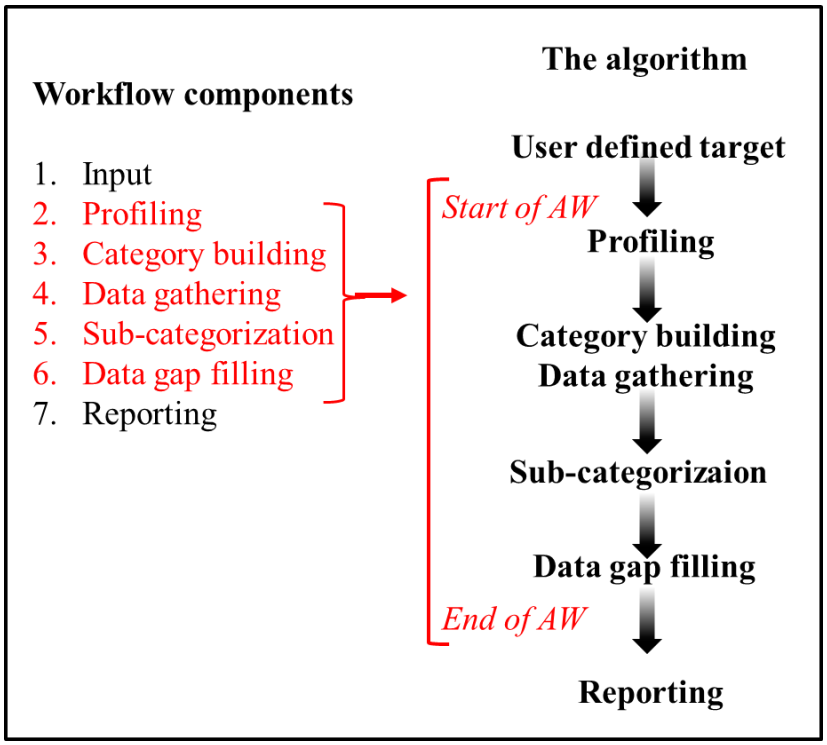
Mechanistic understanding

- 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

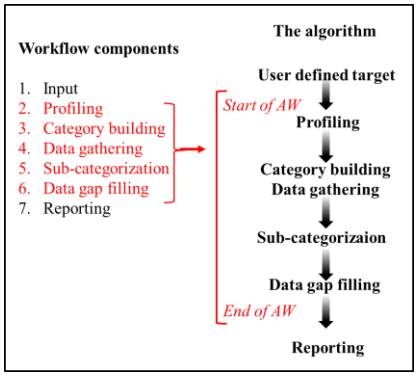
Workflow components

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



Automated workflow for skin sensitization

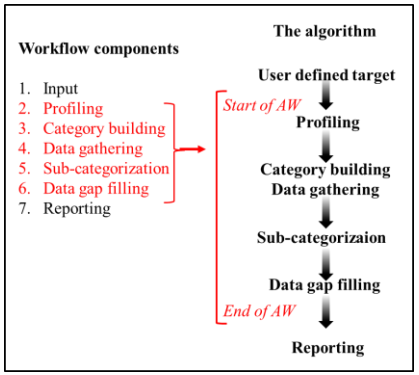
Workflow components



- **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

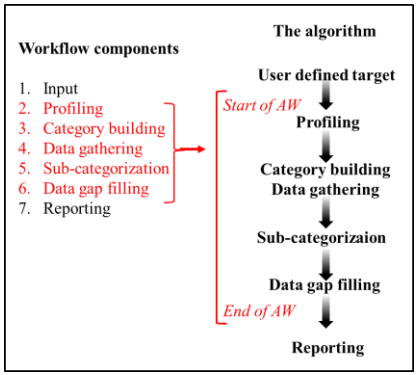


- **Data** – SS predictions are based on LLNA and GPMT exp. data.
  - ✓ LLNA potency categories are Extreme, Strong, Moderate, Weak and Non sensitizer
  - ✓ GPMT potency categories are Strong, Moderate, Weak and Non sensitizer
  - ✓ The dichotomous scale converting LLNA and GPMT potency categories into **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:

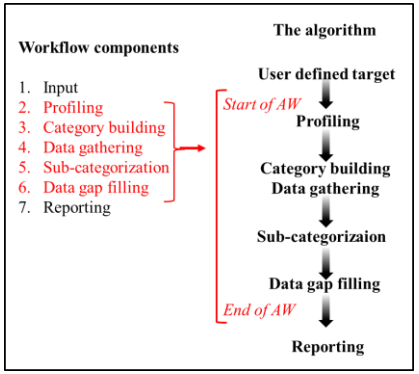
- OFGs
- *Acute aquatic classification by ECOSAR*
- *US-EPA New Chemical categories*

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

Automated workflow for skin sensitization

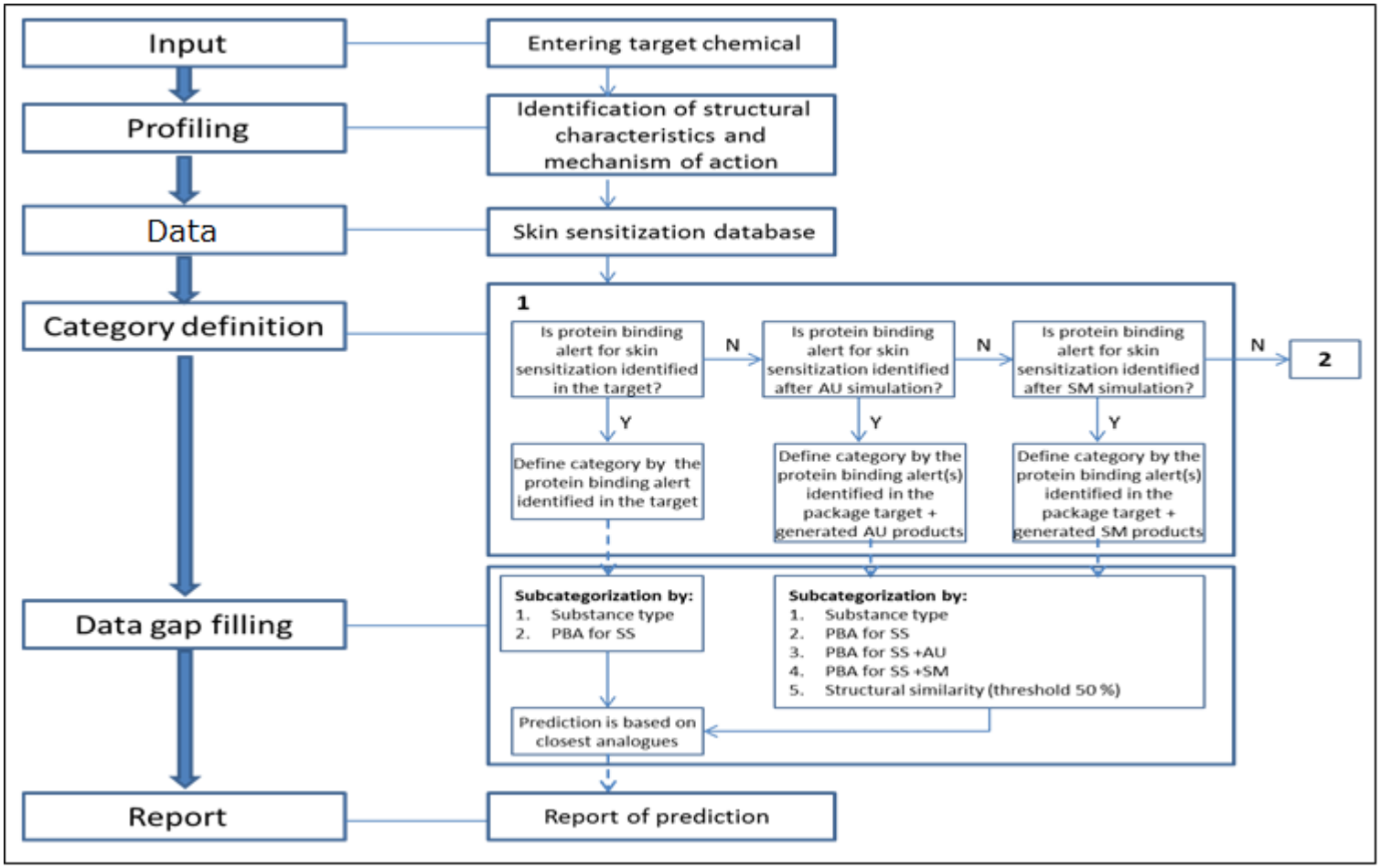
Workflow components

- **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)



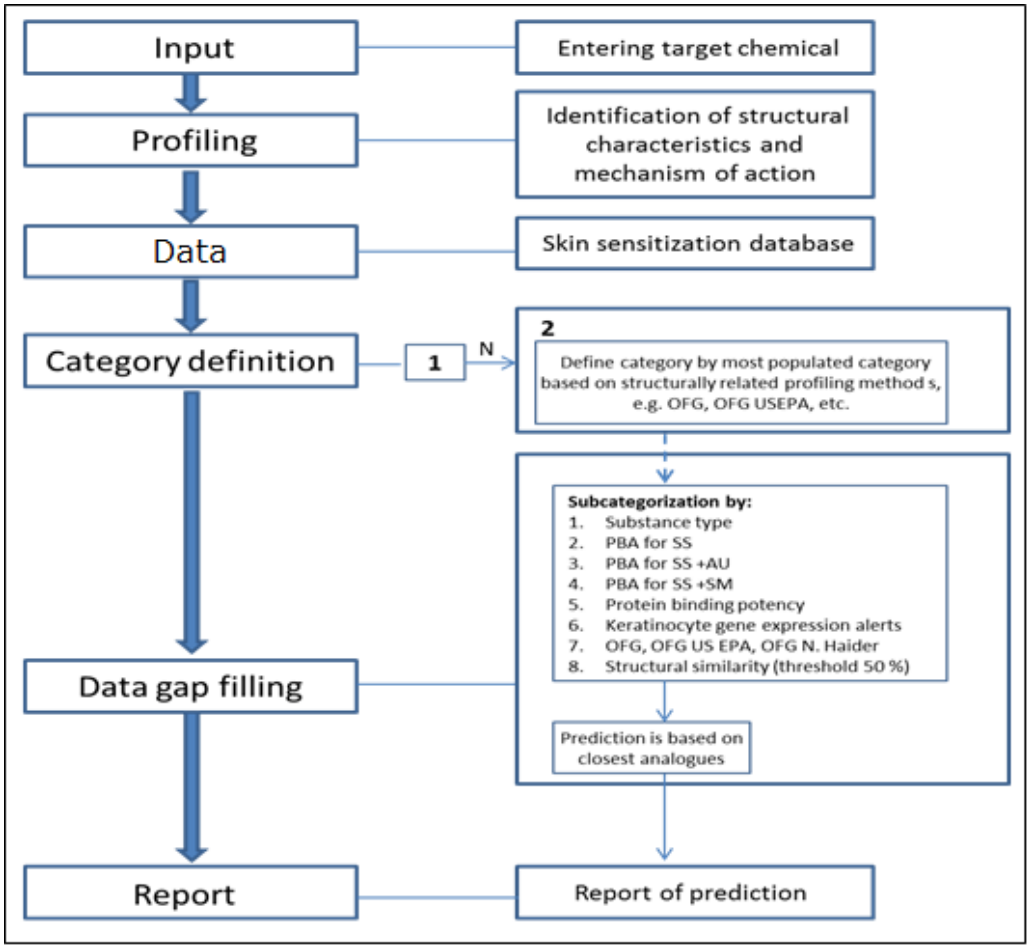
Automated workflow for skin sensitization

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

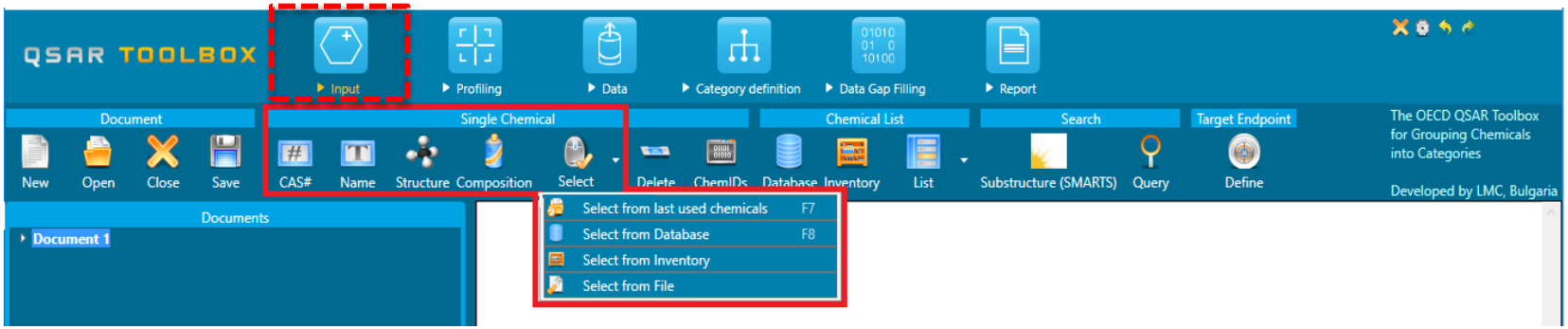
- In this exercise we will predict the skin sensitization effect for:
  - **Single chemical – Chlorpyrifos [CAS# 2921-88-2];**
  - Batch of chemicals - the first five substances of the Skin sensitization ECETOC database.

- 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

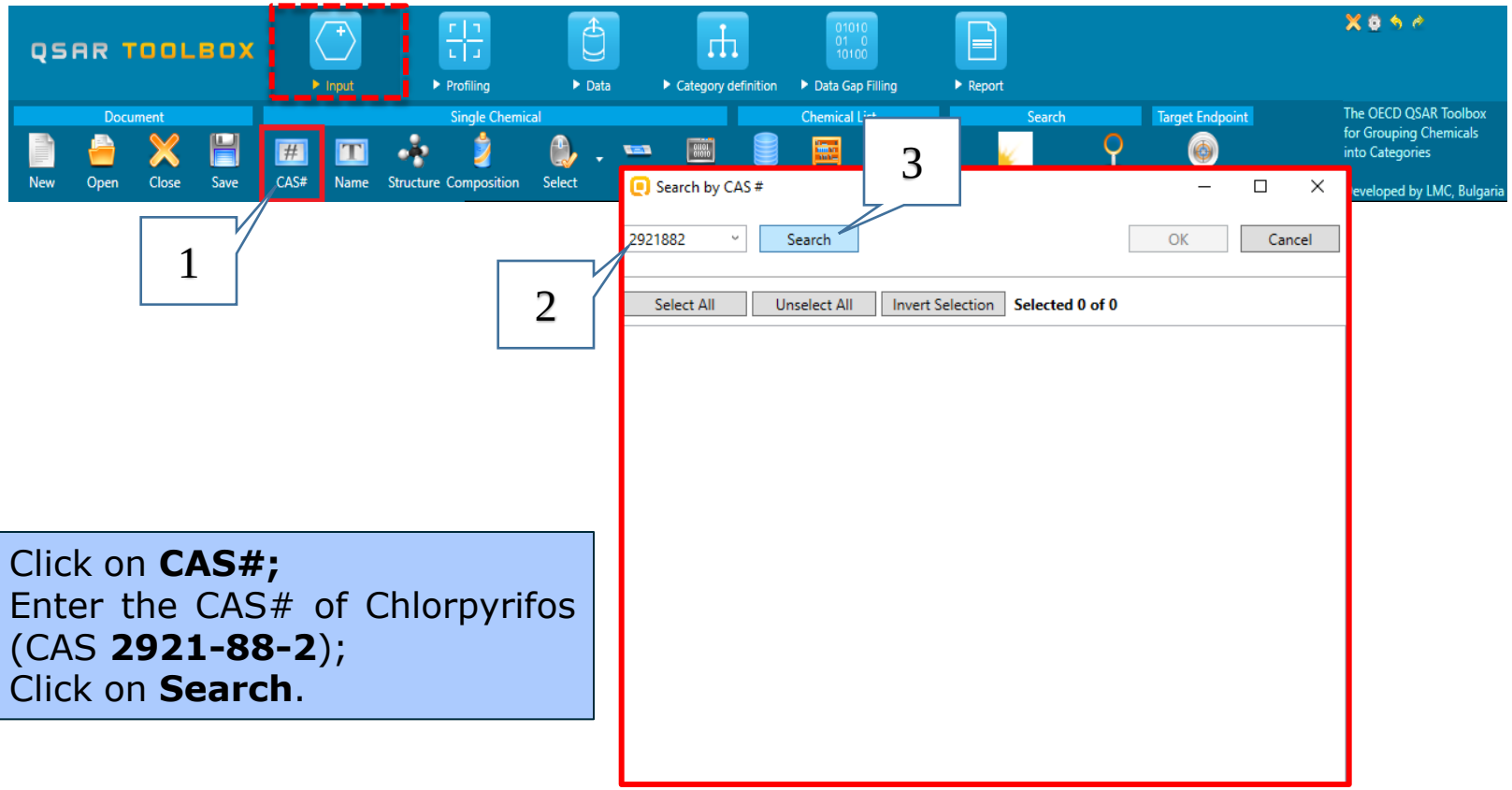


- Chemical Name
- Chemical Abstract Services (CAS) number (#)
- Drawing chemical structure with or without additional defining of the composition
- Select from User List/Inventory/Databases

Execution of AW for Skin sensitization

Single chemical, CAS # 2921-88-2

Input by CAS number



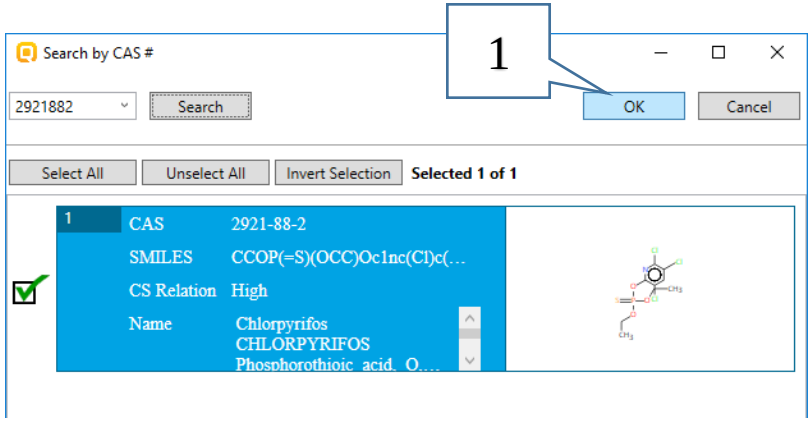
1. Click on **CAS#;**
2. Enter the CAS# of Chlorpyrifos (CAS **2921-88-2**);
3. Click on **Search**.

Execution of AW for Skin sensitization

Single chemical, CAS # 2921-88-2

Input - Target chemical identity

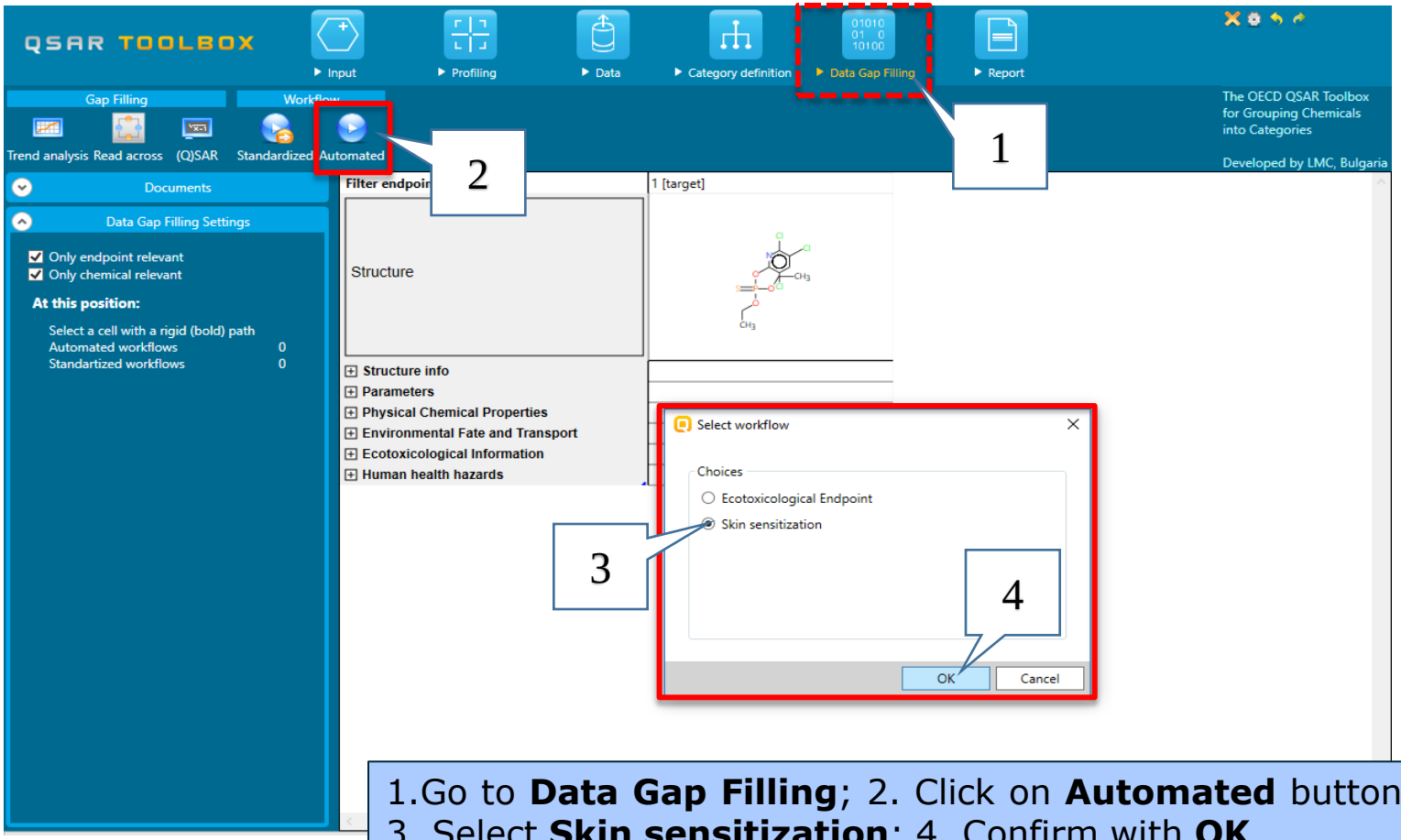
The Toolbox now searches the databases to find out if the CAS# you entered is linked to a molecular structure stored in the Toolbox. It is displayed as a 2-demensional depiction. Click on **OK** (1).



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

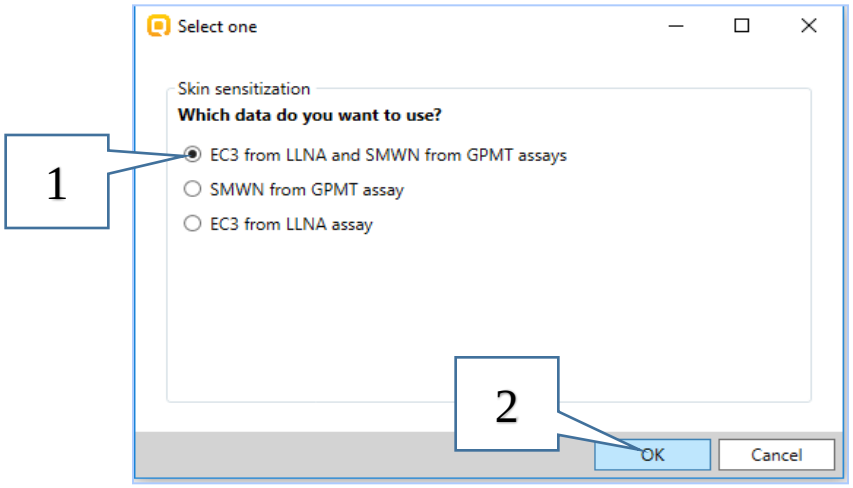
Location of the Automated workflow for skin sensitization



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

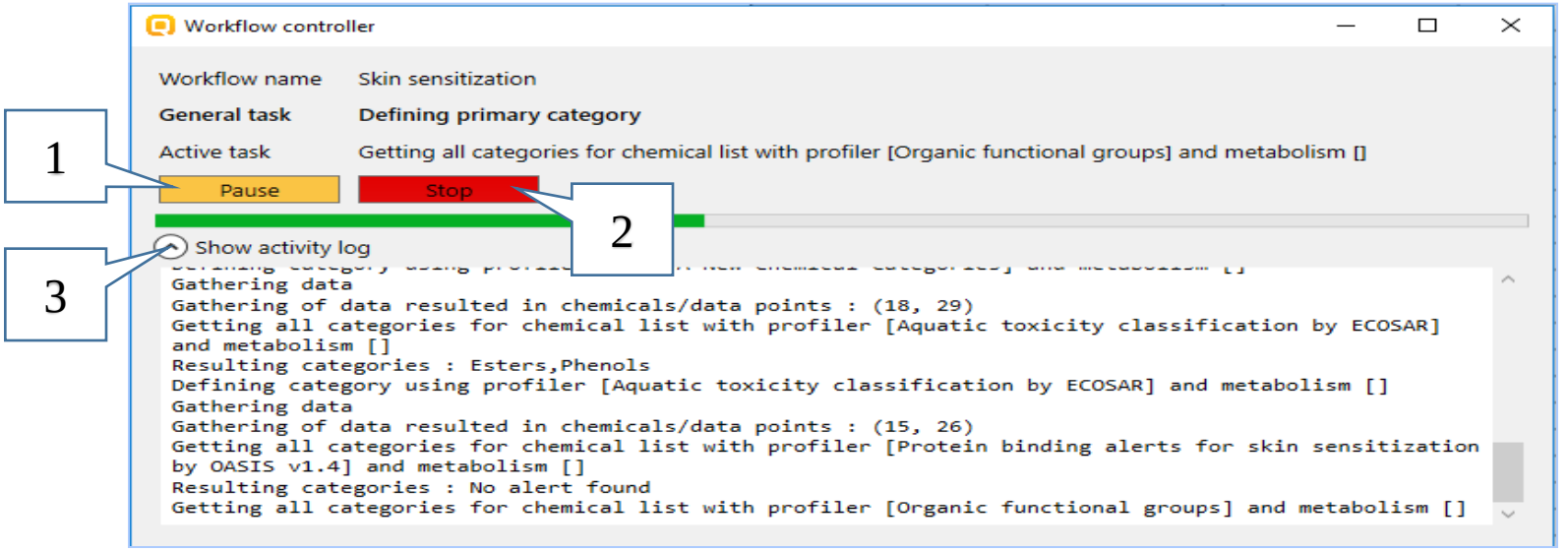
There are three options for endpoint selection and the user should select one of them.



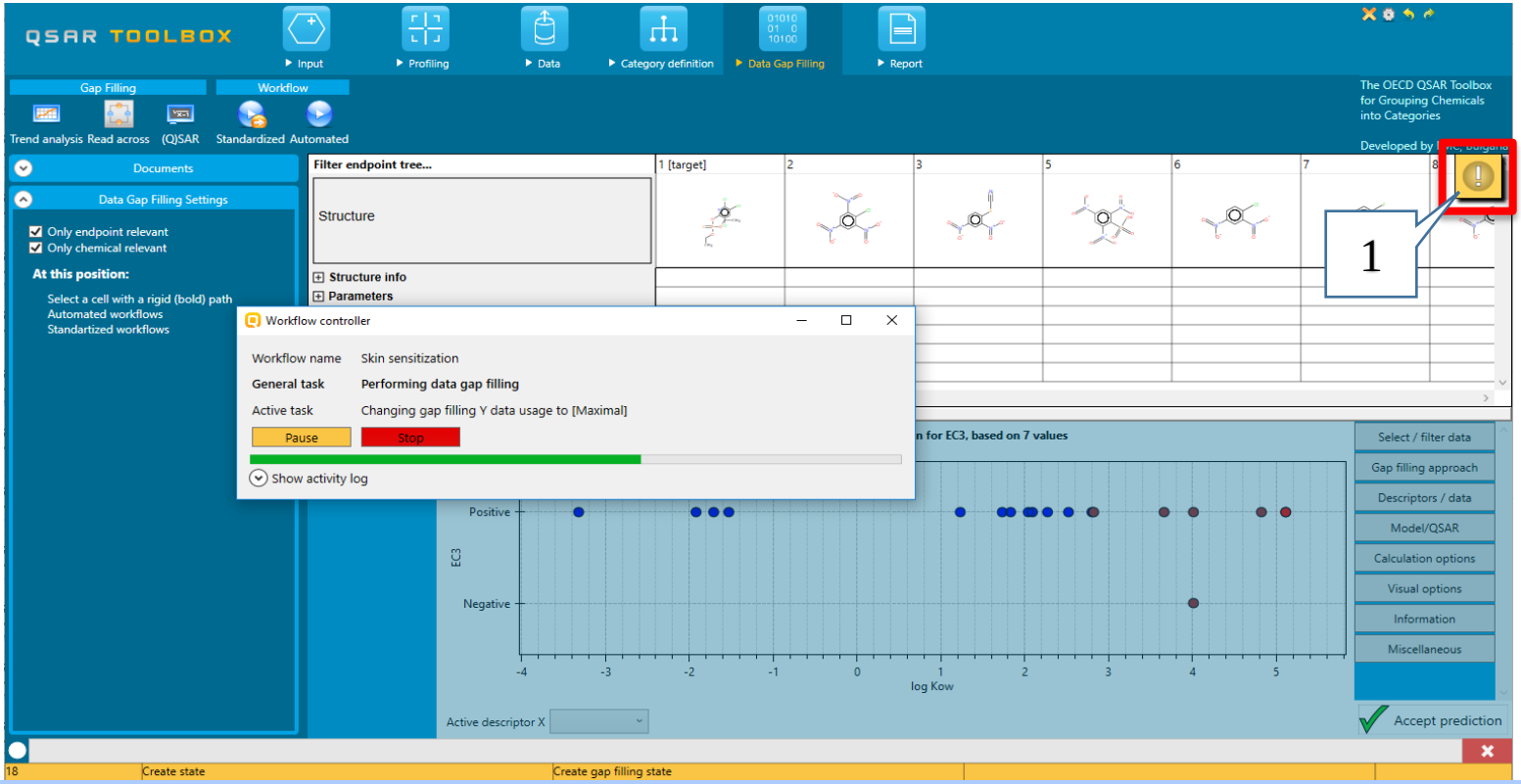
1. Select the first combined endpoint – ***EC3 from LLNA and SMWN from GPMT assays***; 2. Confirm with ***OK***

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

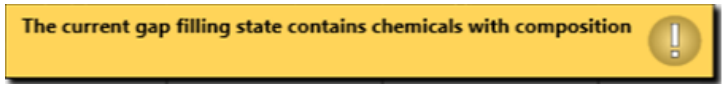
Information messages and functionalities

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

- warning messages, alerting the user:

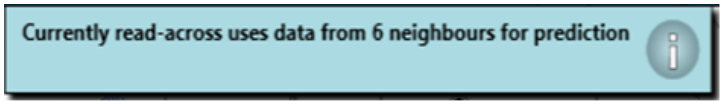
  - for qualifiers availability

  - 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)

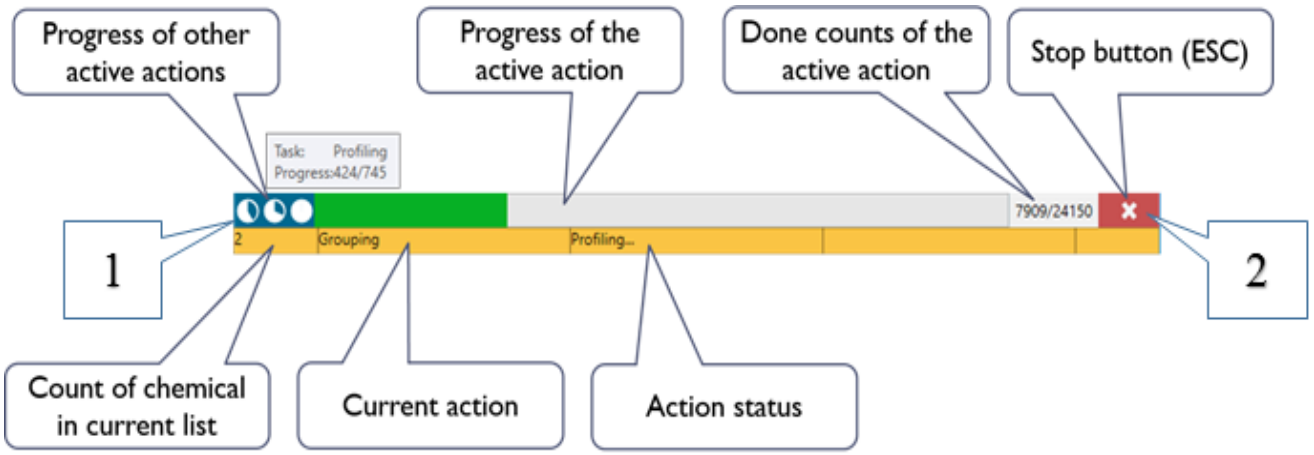


Execution of AW for Skin sensitization

Information messages and functionalities

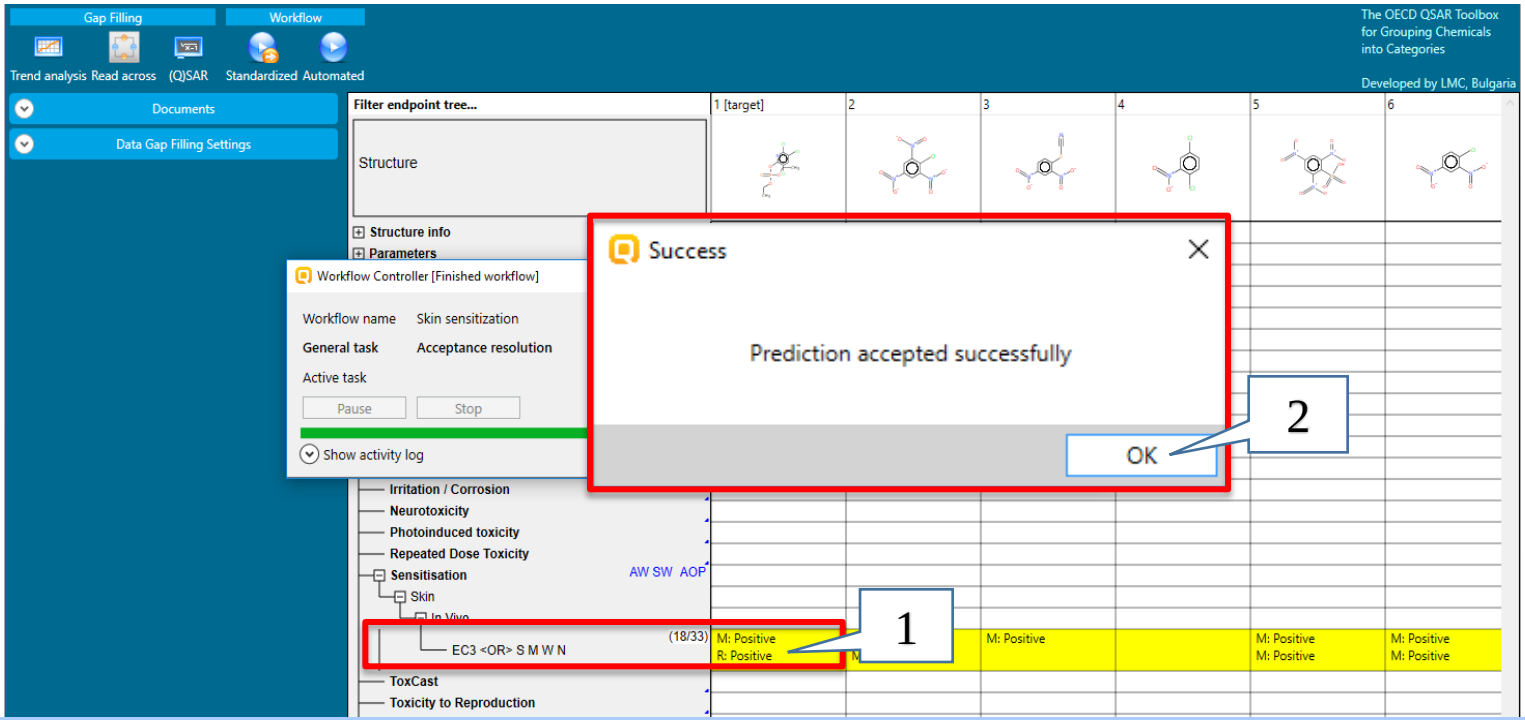
Functionality used to follow the progress of the active action or to cancel an active action when an Automated workflow is run.

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



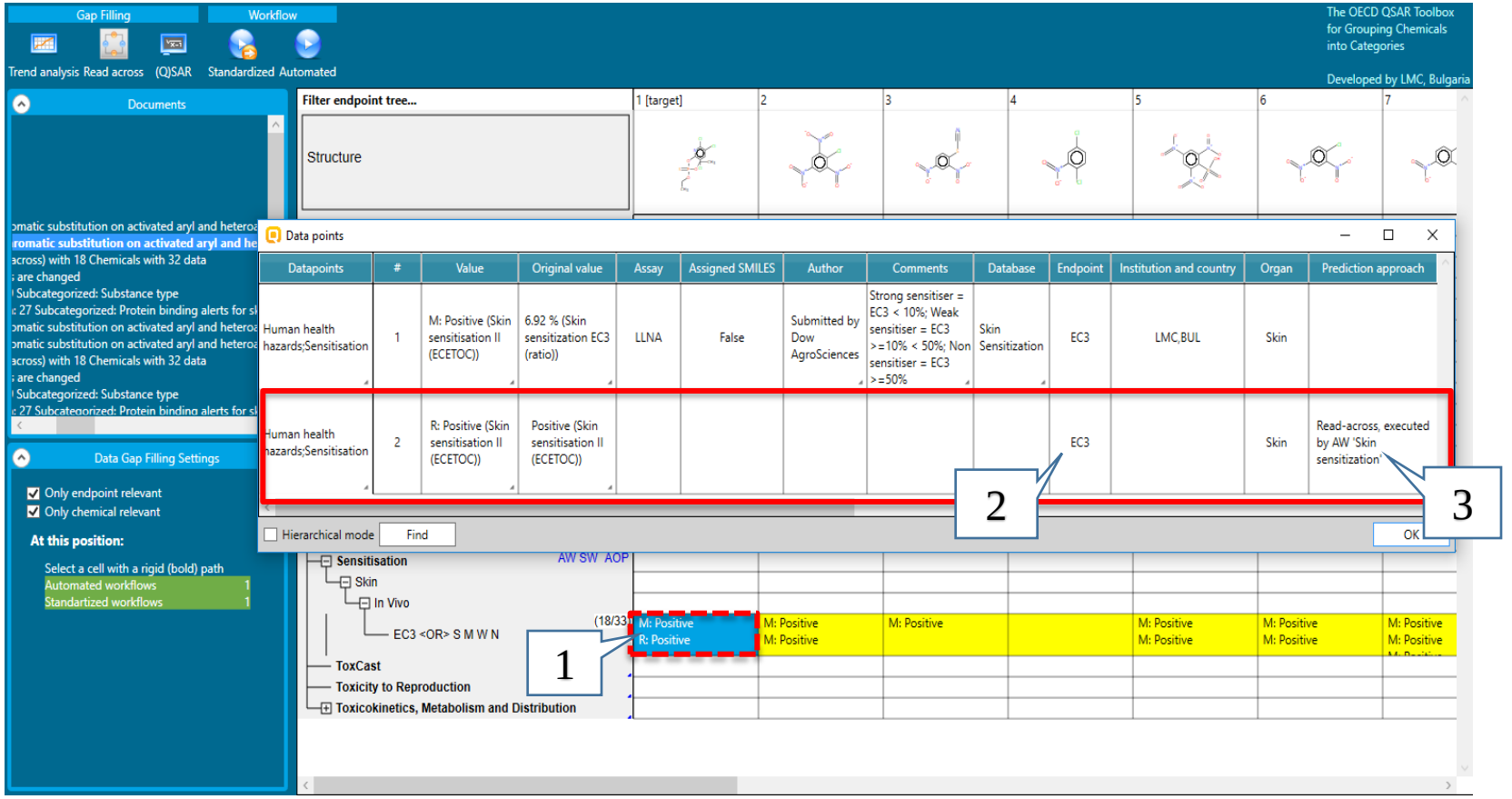
Execution of AW for Skin sensitization

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



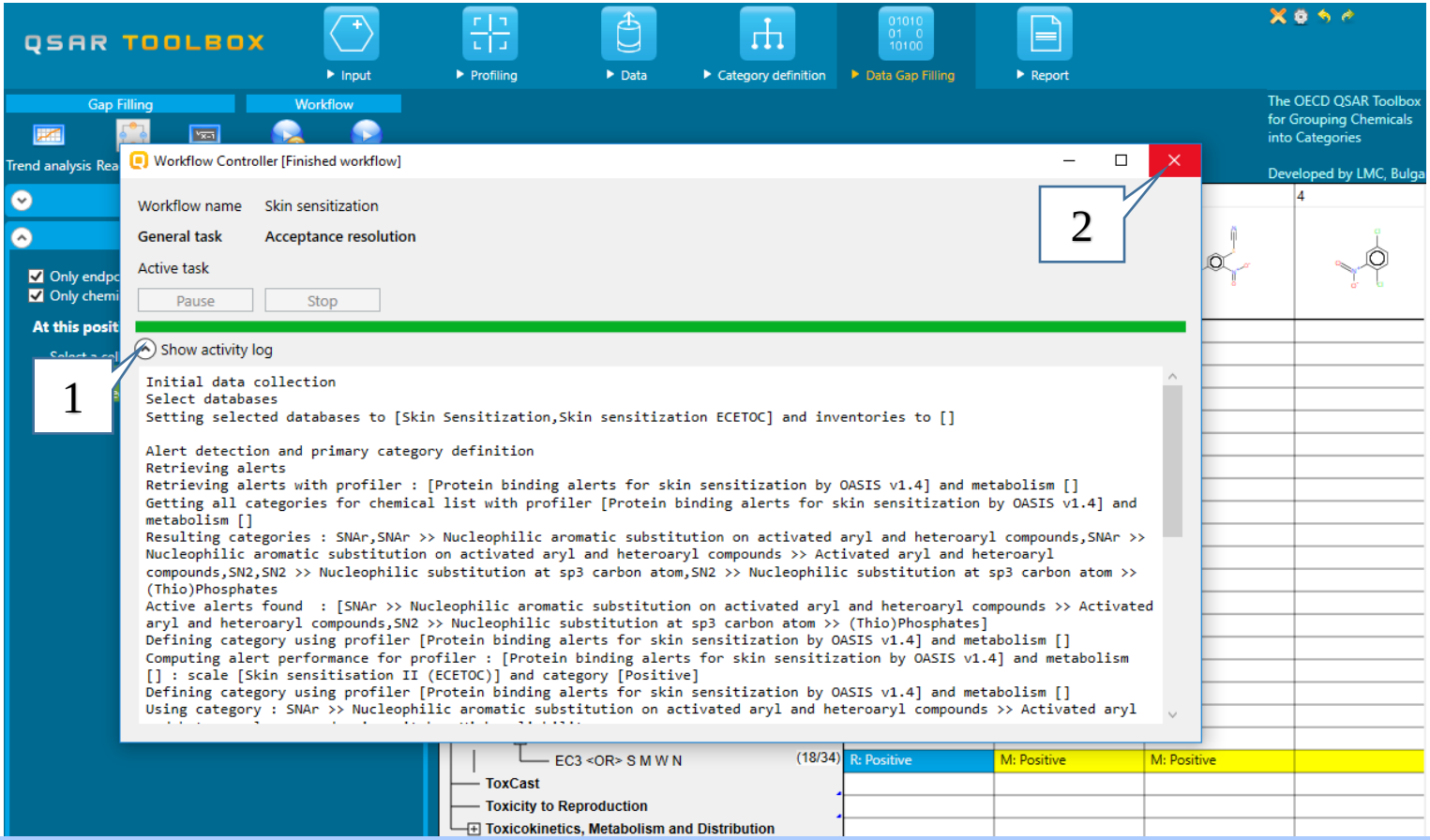
When the prediction is accepted successfully, it appears on the matrix (1). In this case the predicted result (R:) corresponds to the experimental data (M:). Click on "**OK**" (2)

Execution of AW for Skin sensitization



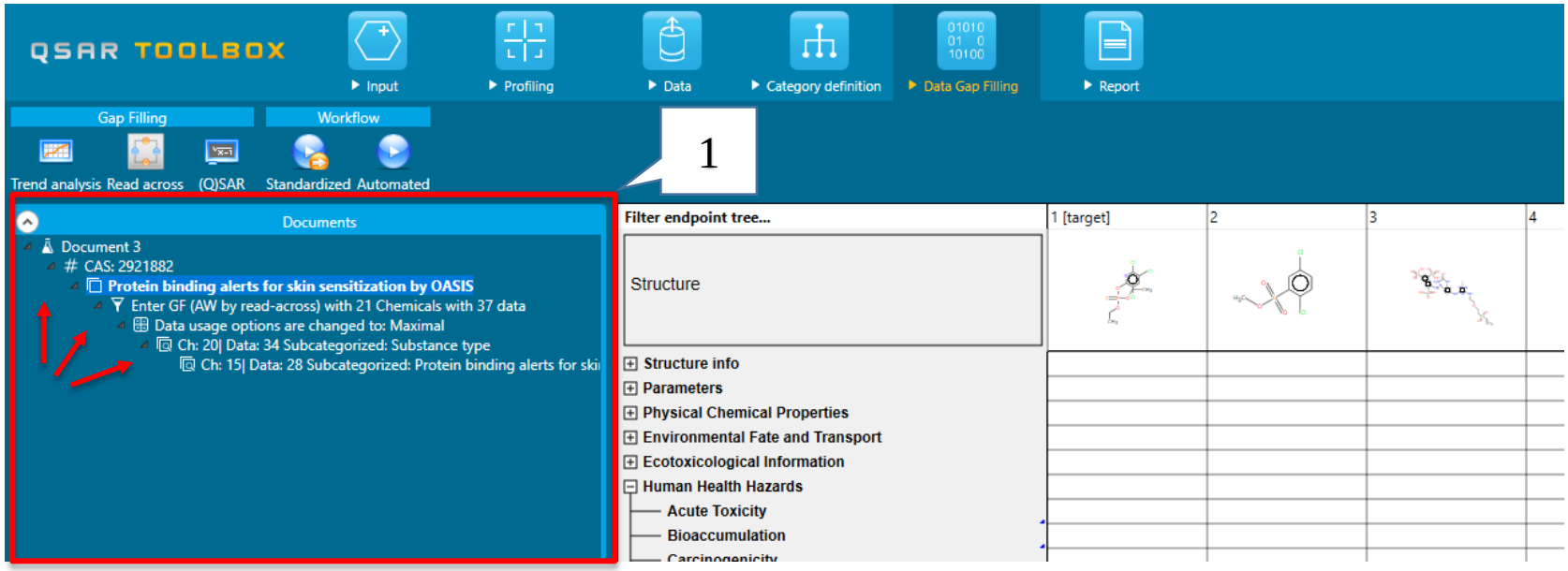
Double click on the cell with the prediction (1) opens the *Data points* window. Here the user can see more information for their prediction. It is seen that only EC3 data is used for the prediction (2) and the prediction approach is read across executed by AW “Skin sensitization”.

Execution of AW for Skin sensitization



The *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).

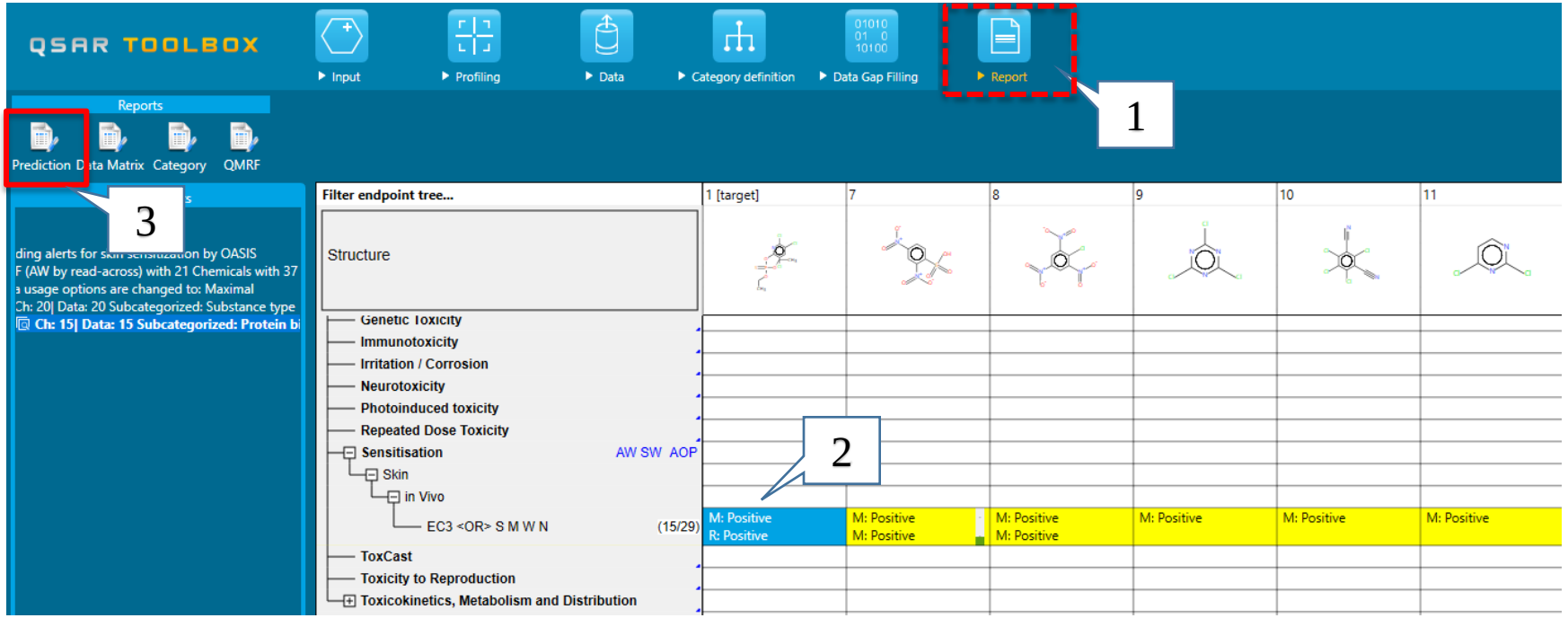
Execution of AW for Skin sensitization



The 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

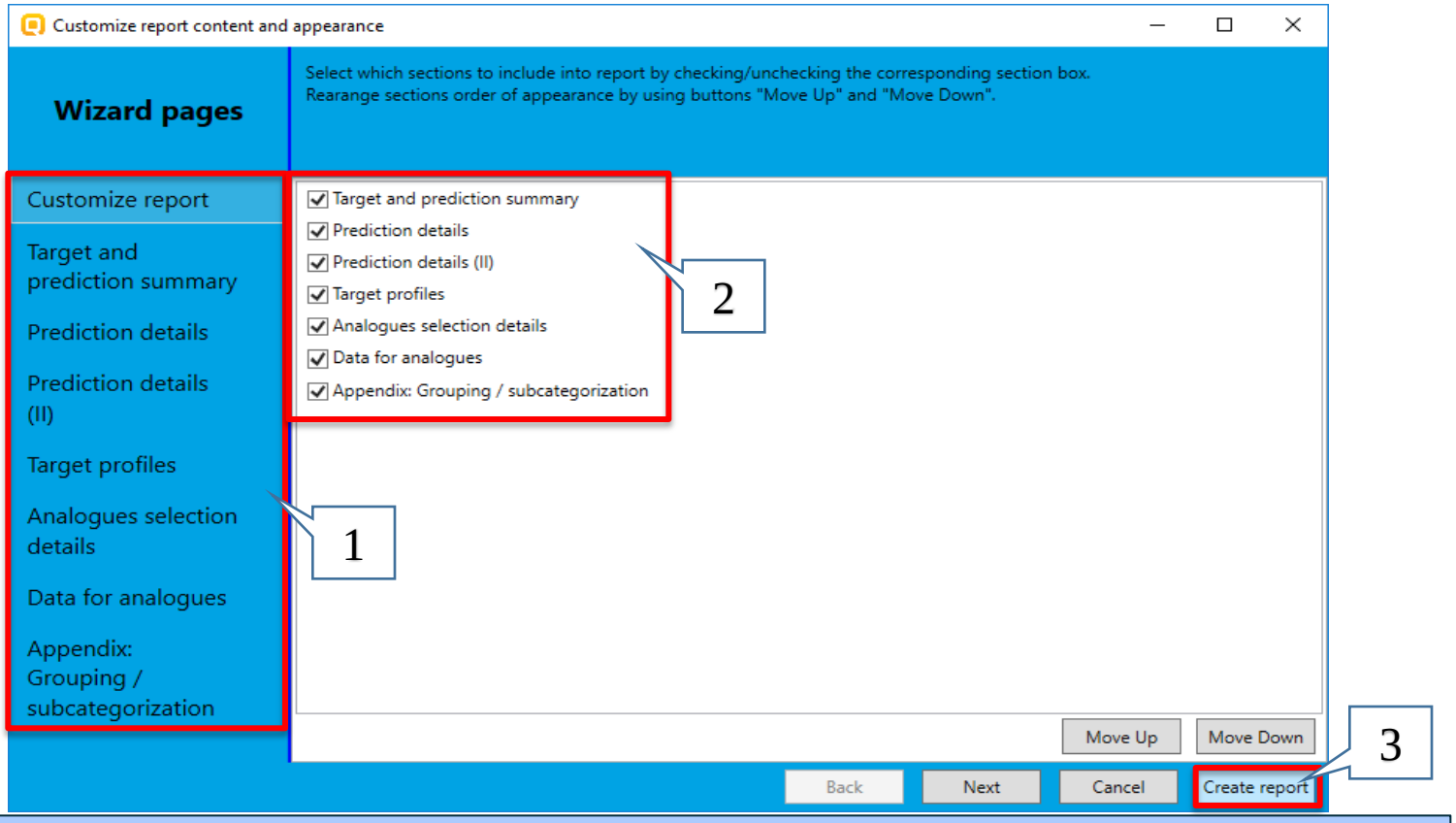
Report: Generation of report



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



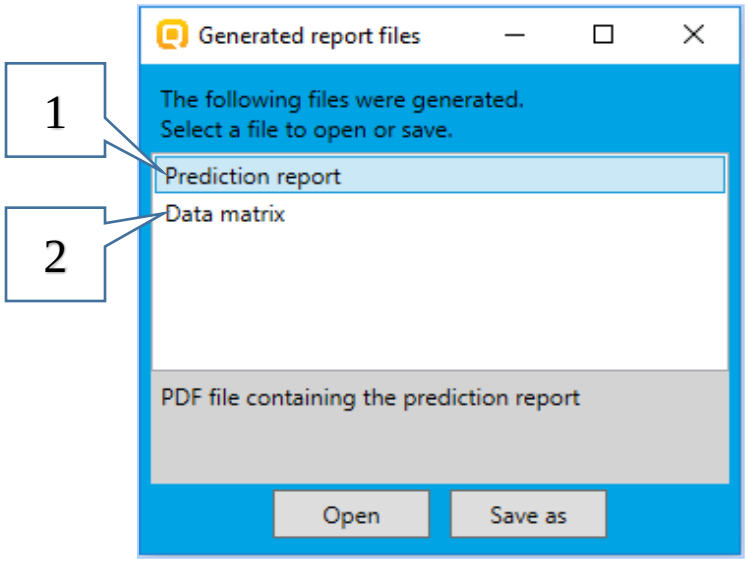
The user can customize the report content (1) and appearance (2). Generation of the reports happens by click on the **Create report** button (3).

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.

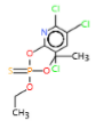


Execution of AW for Skin sensitization

Report: Generation of report

QSAR Toolbox prediction for single chemical

Date: 1 авг 2017
Author(s):
Contact details:

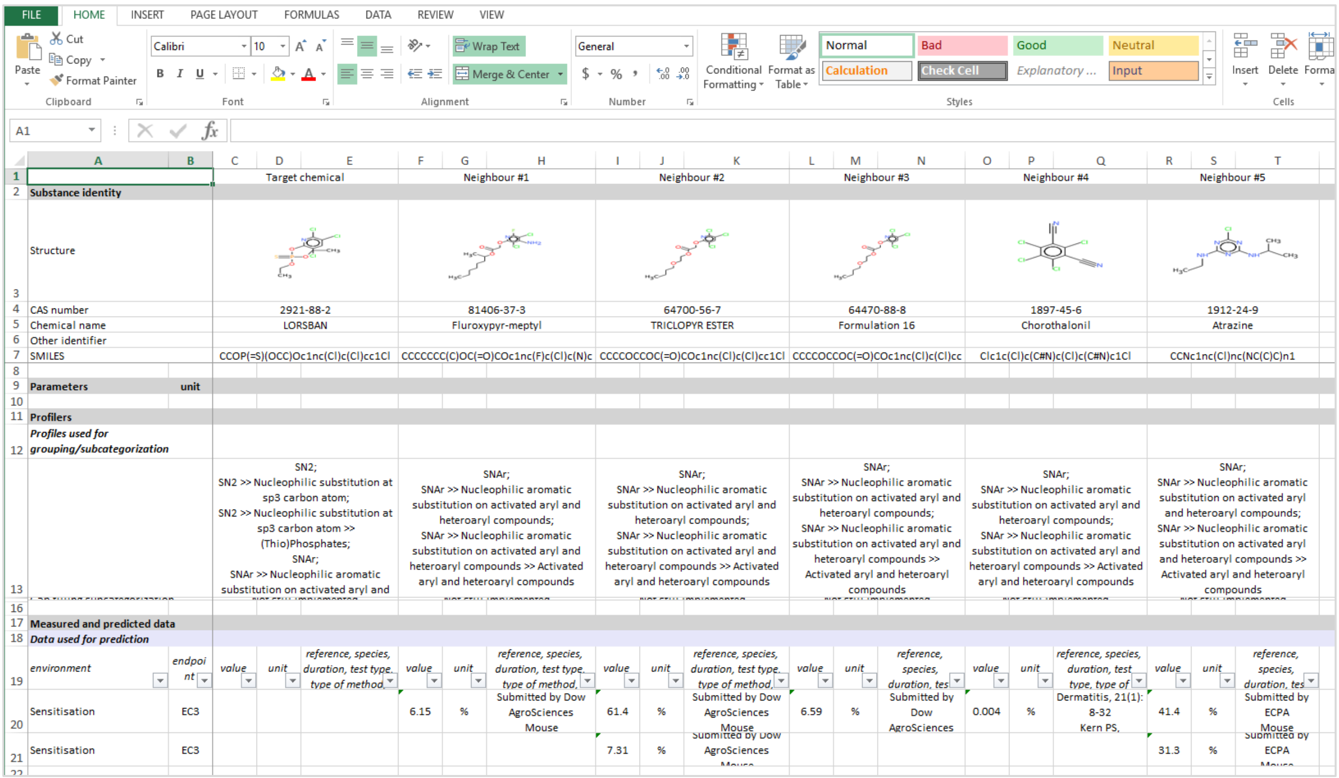
Target information		
Structural information	**Numerical identifiers**	**Chemical names**
SMILES: CCOP(=S)(OCC)Oc1nc(Cl)c(Cl)cc1Cl Structure 	EC#: N/A CAS#: 2921-88-2 Other: N/A	Chlorpyrifos CHLORPYRIFOS (DURSBAN) chlorpyrifos (dursban) (cpf) (o,o-diethyl-o-(3,5,6-trichloro-2-pyridyl)phosphorothiolate) (chlorpyriphos)

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

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

Execution of AW for Skin sensitization

Report: Generation of report



Row	A	B	Target chemical	Neighbour #1	Neighbour #2	Neighbour #3	Neighbour #4	Neighbour #5
2	**Substance identity**							
3	Structure							
4	CAS number		2921-88-2	81406-37-3	64700-56-7	64470-88-8	1897-45-6	1912-24-9
5	Chemical name		LORSBAN	Fluroxypyr-meptyl	TRICLOPYR ESTER	Formulation 16	Chorothalonil	Atrazine
6	Other identifier							
7	SMILES		CCOP(=S)(OCC)Oc1nc(Cl)c(Cl)cc1Cl	CCCCCCC(C)OC(=O)COc1nc(F)c(Cl)c(N)c	CCCCOCCOC(=O)COc1nc(Cl)c(Cl)cc1Cl	CCCCOCCOC(=O)COc1nc(Cl)c(Cl)cc	Clc1c(Cl)c(C#N)c(Cl)c(C#N)c1Cl	CCNc1nc(Cl)nc(NC(C)C)n1
8								
9	**Parameters**	**unit**						
10								
11	**Profilers**							
12	***Profiles used for grouping/subcategorization***							
13			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	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	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	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	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	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								
17	**Measured and predicted data**							
18	***Data used for prediction***							

Row	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, tes	value	unit	reference, species, duration, test type, type of	value	unit	reference, species, duration, tes
20	Sensitisation	EC3				6.15	%	Submitted by Dow AgroSciences Mouse	61.4	%	Submitted by Dow AgroSciences Mouse	6.59	%	Submitted by Dow AgroSciences	0.004	%	Dermatitis, 21(1): 8-32 Kern PS,	41.4	%	Submitted by ECPA Mouse
21	Sensitisation	EC3							7.31	%	Submitted by Dow AgroSciences Mouse							31.3	%	Submitted by ECPA Mouse

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

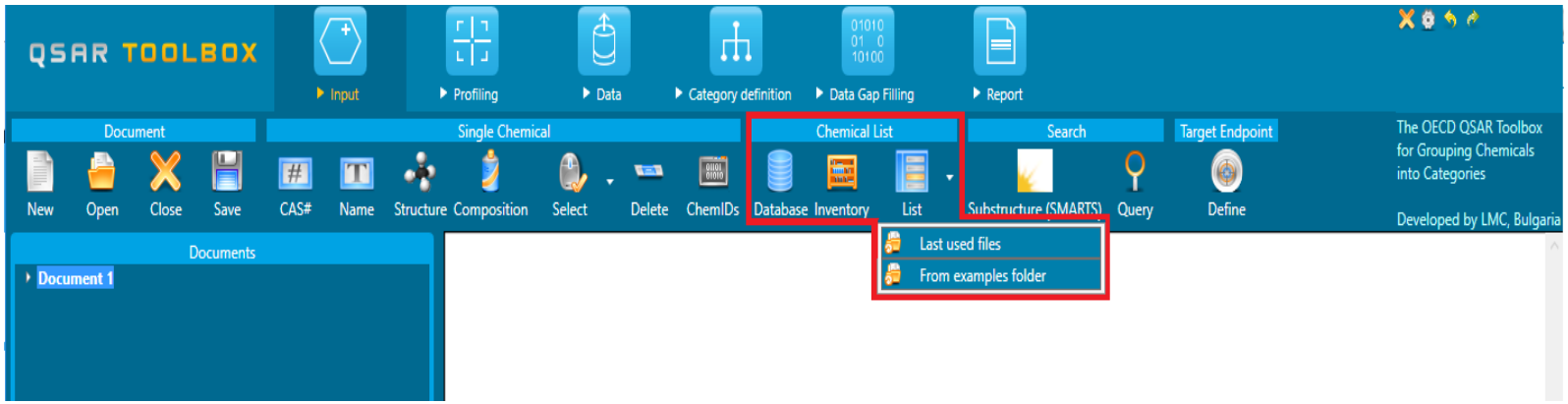
- In this exercise we will predict the skin sensitization effect for:
  - Single chemical – Chlorpyrifos [CAS# 2921-88-2];
  - **Batch of chemicals - the first five substances of the Skin sensitization ECETOC database.**

- 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

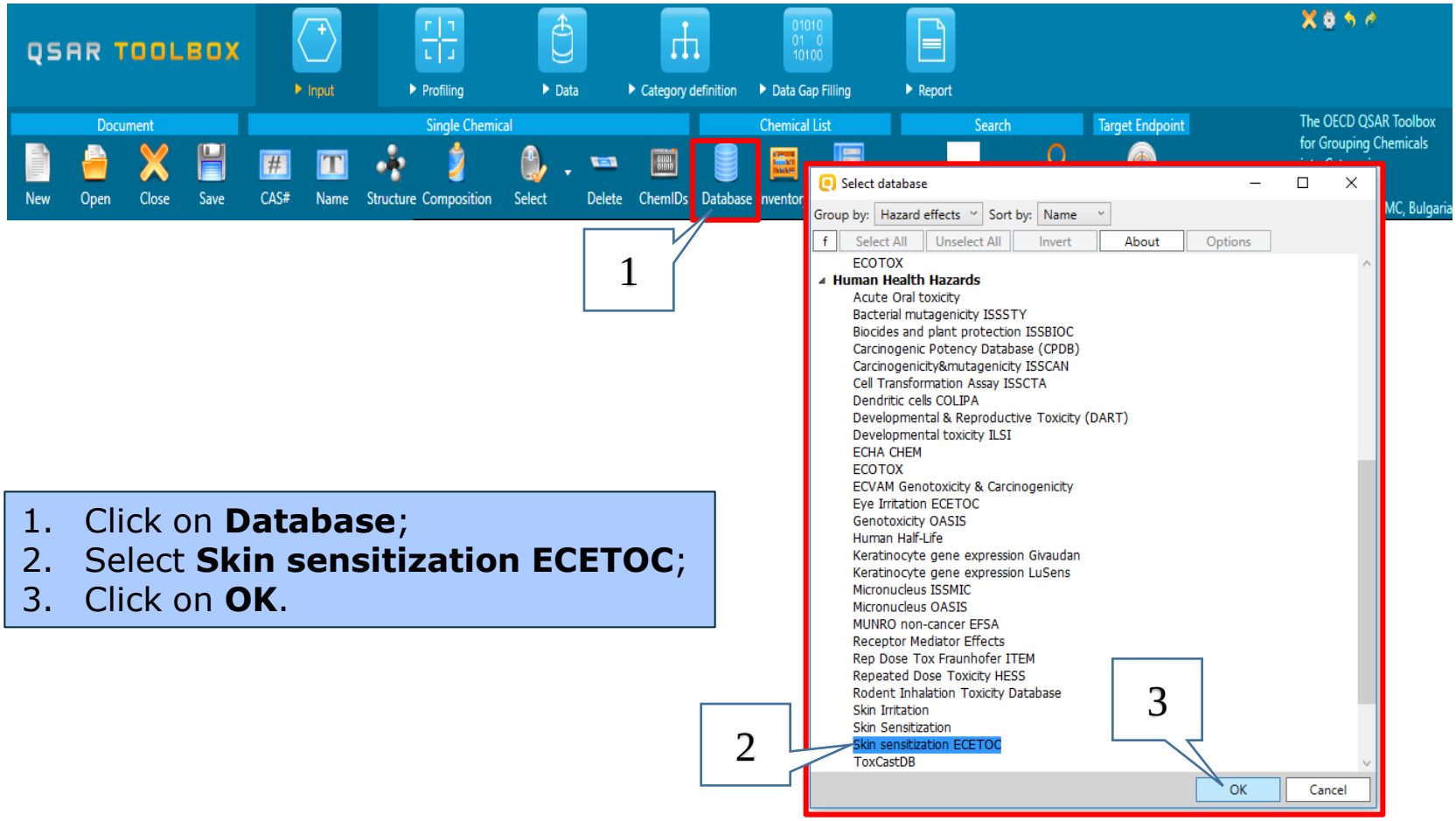


- Database
- Inventory
- List: Last used files/ From examples folder

Execution of AW for Skin sensitization

List with chemicals (batch work)

Input: Load a database

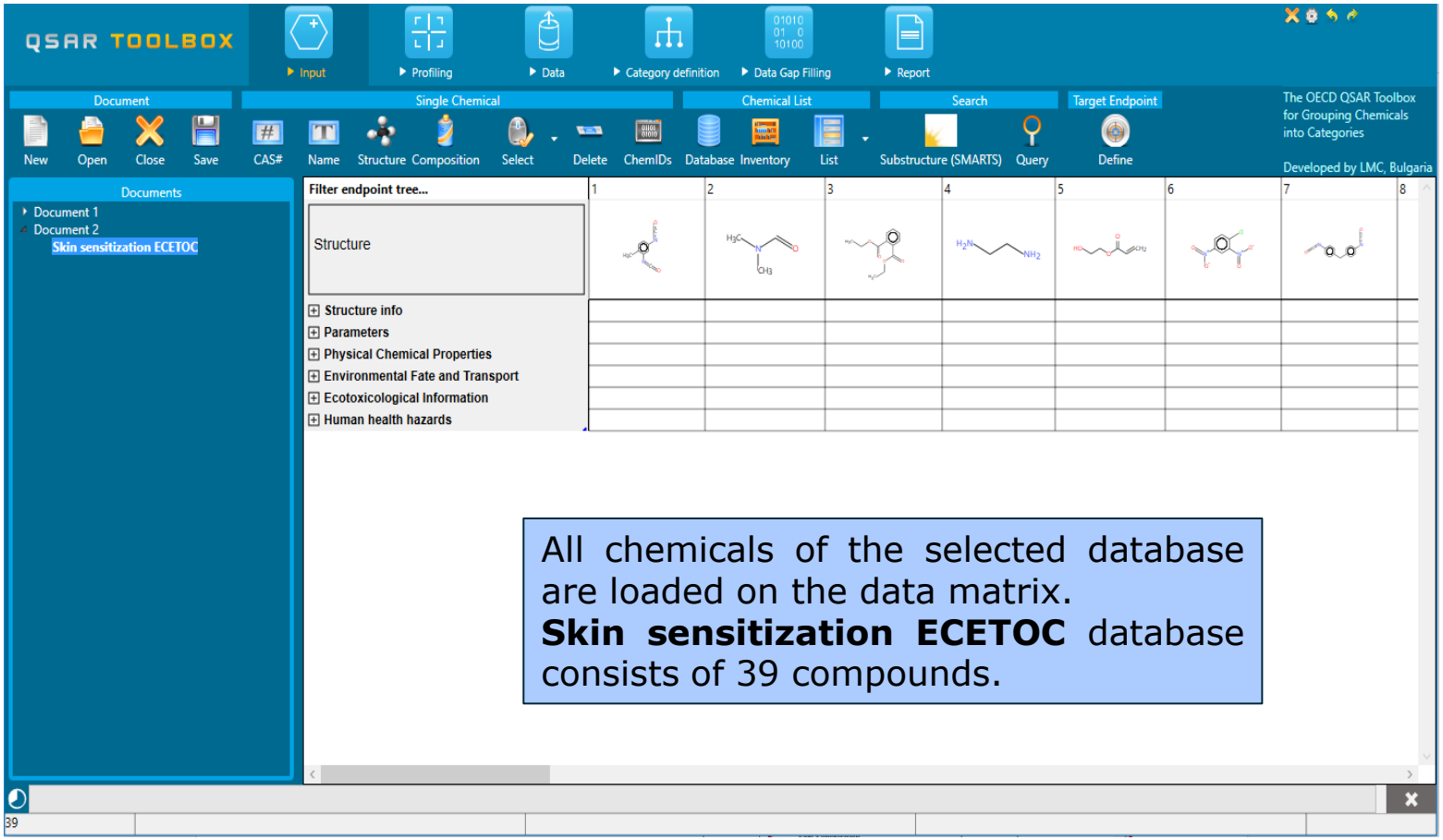


1. Click on **Database**;
2. Select **Skin sensitization ECETOC**;
3. Click on **OK**.

Execution of AW for Skin sensitization

List with chemicals (batch work)

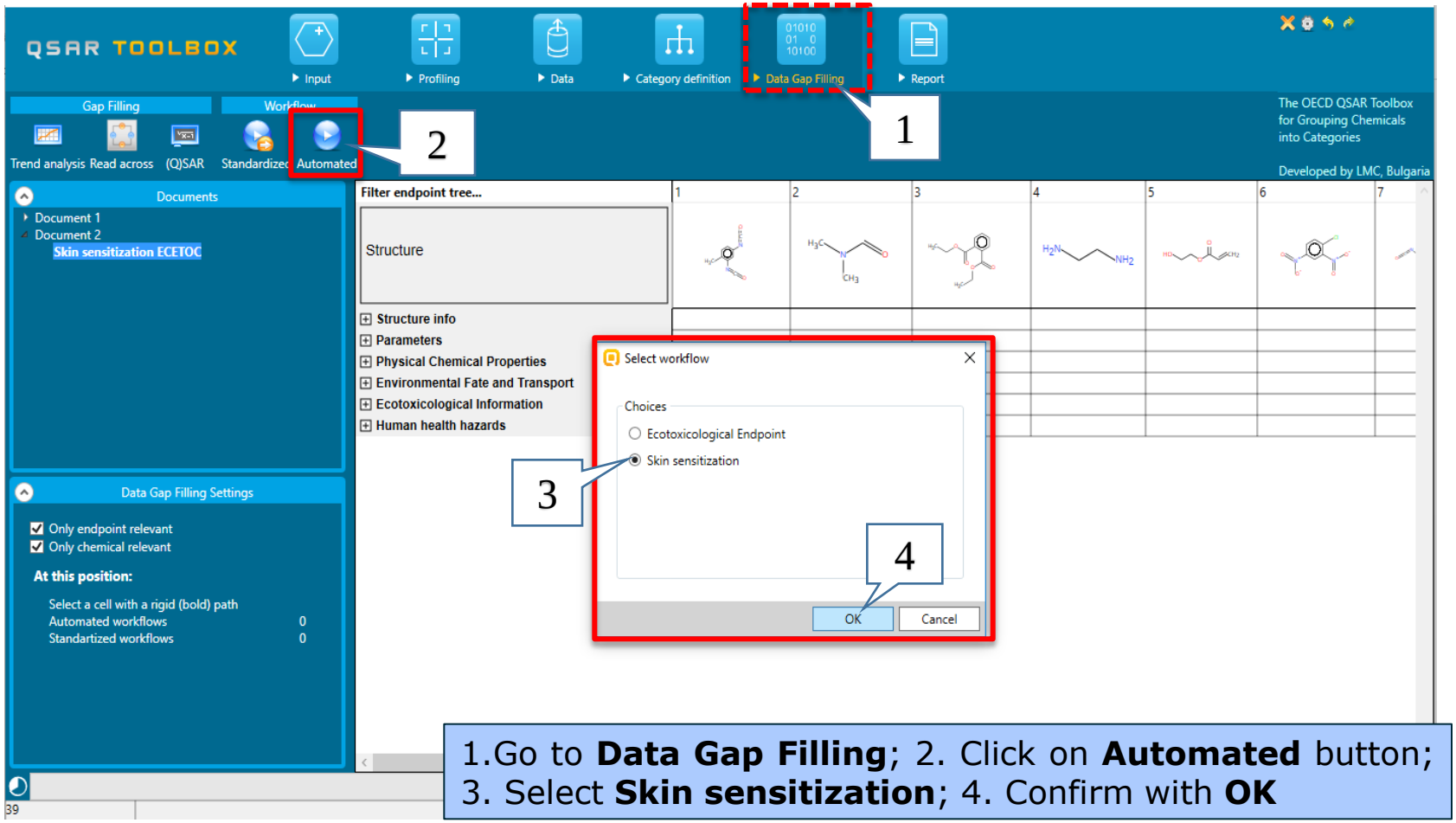
Input: Load a database



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

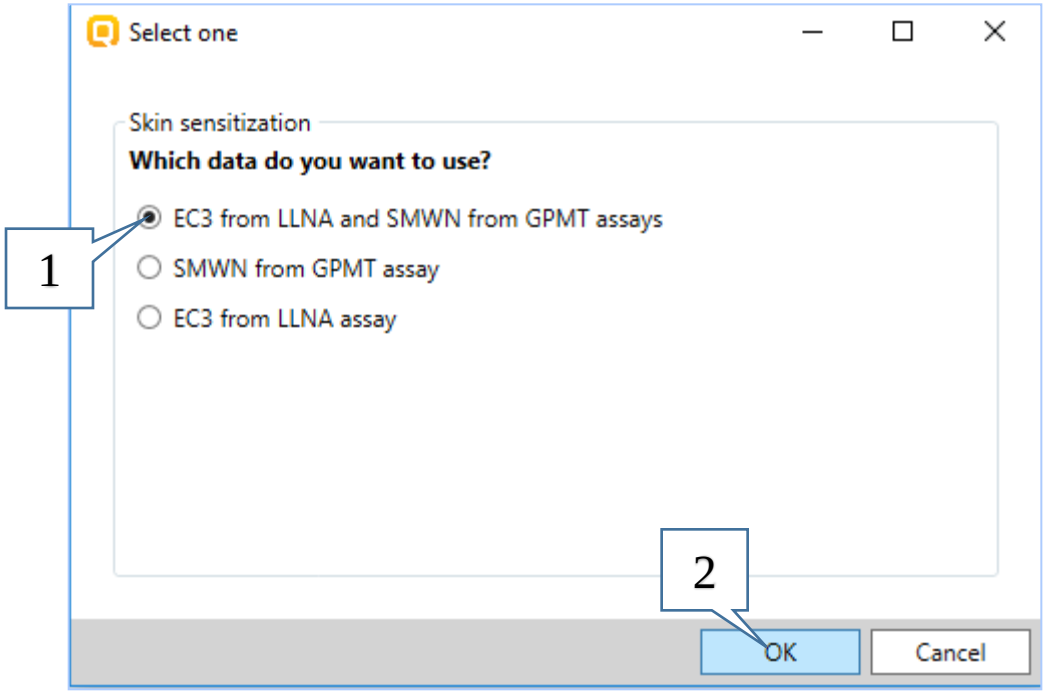
List with chemicals (batch work)



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)

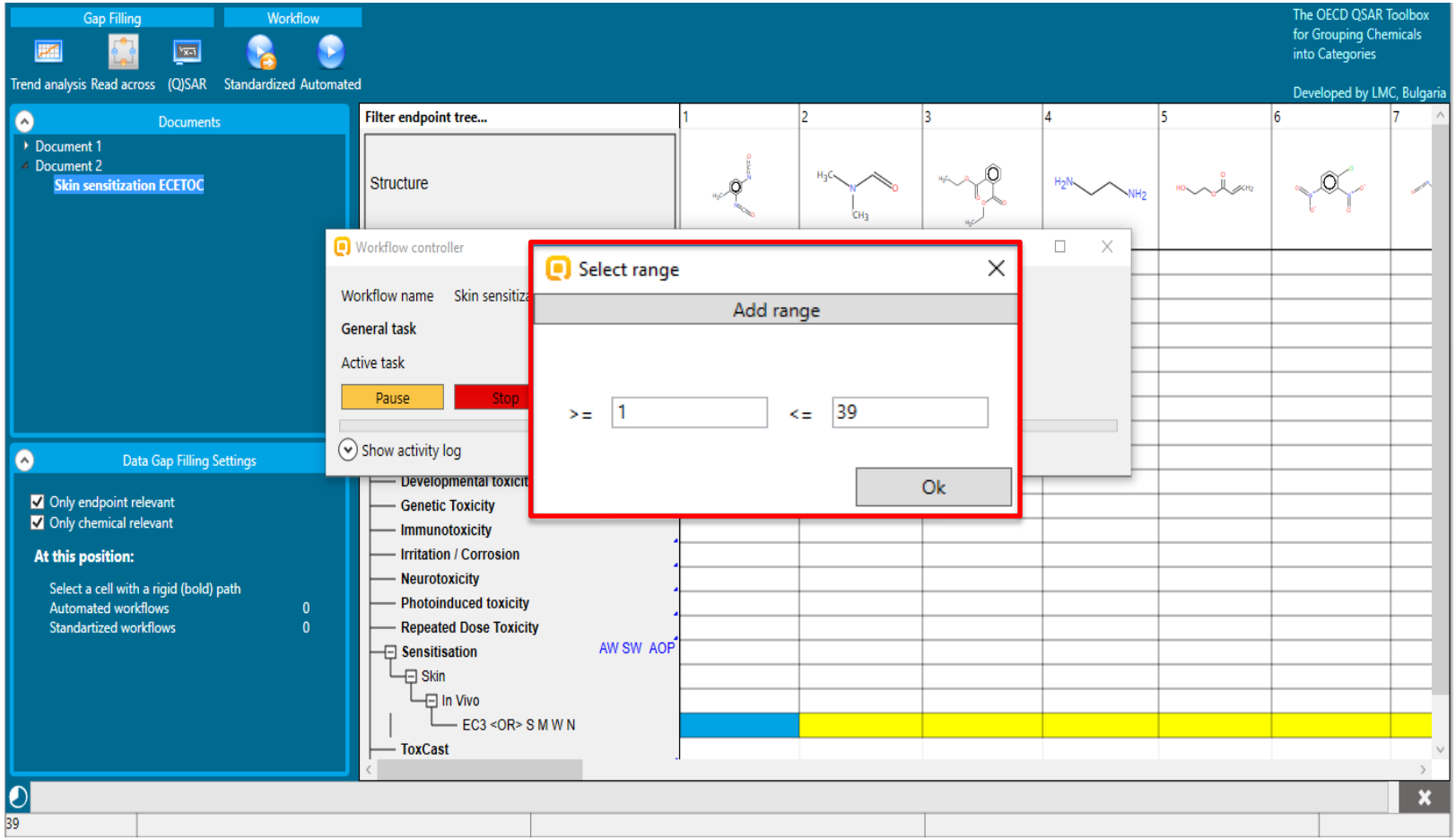
When the targeted endpoint is selected, a new dialog related to the application range of the workflow appears.

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.

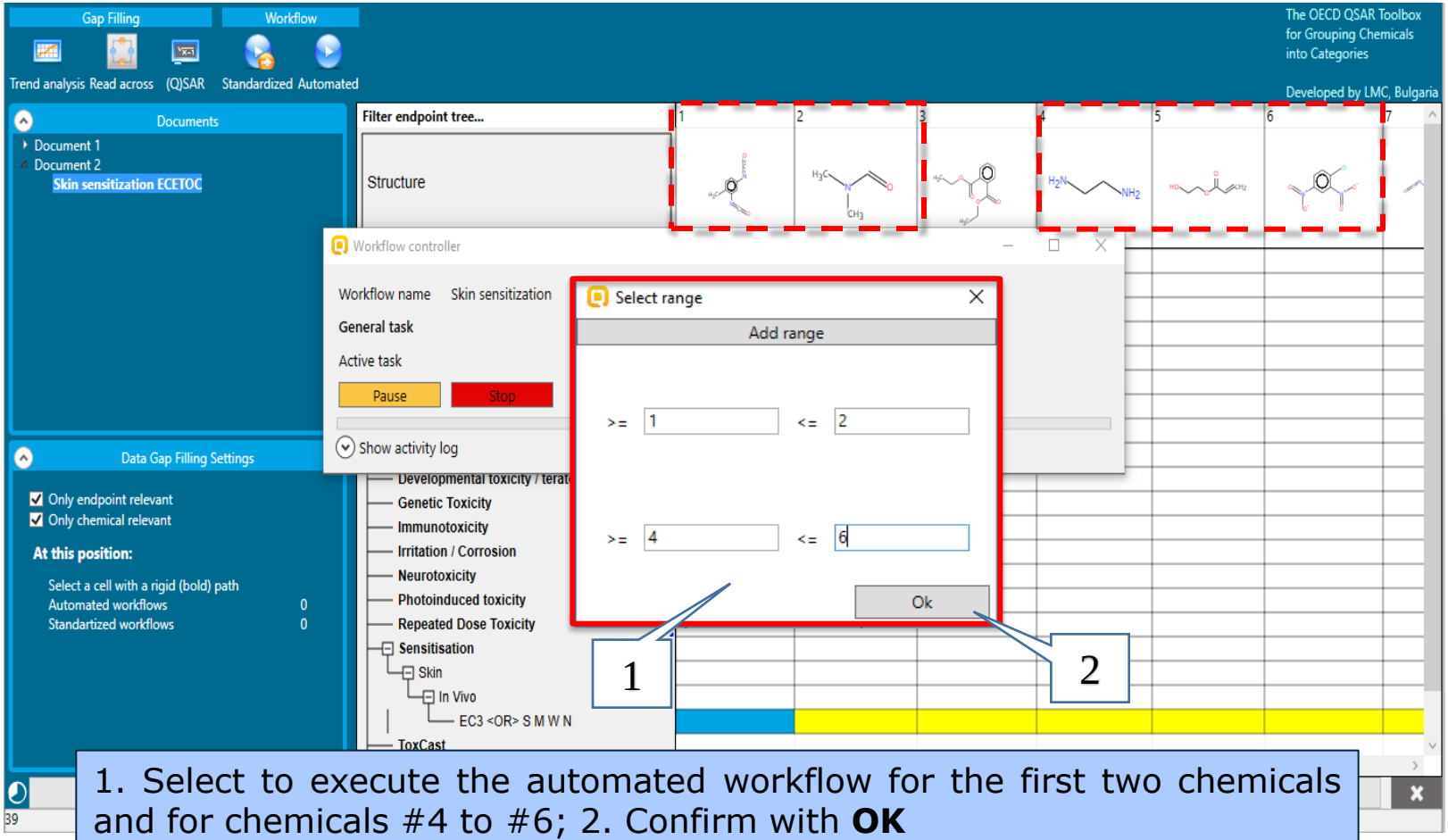
Execution of AW for Skin sensitization

List with chemicals (batch work)



Execution of AW for Skin sensitization

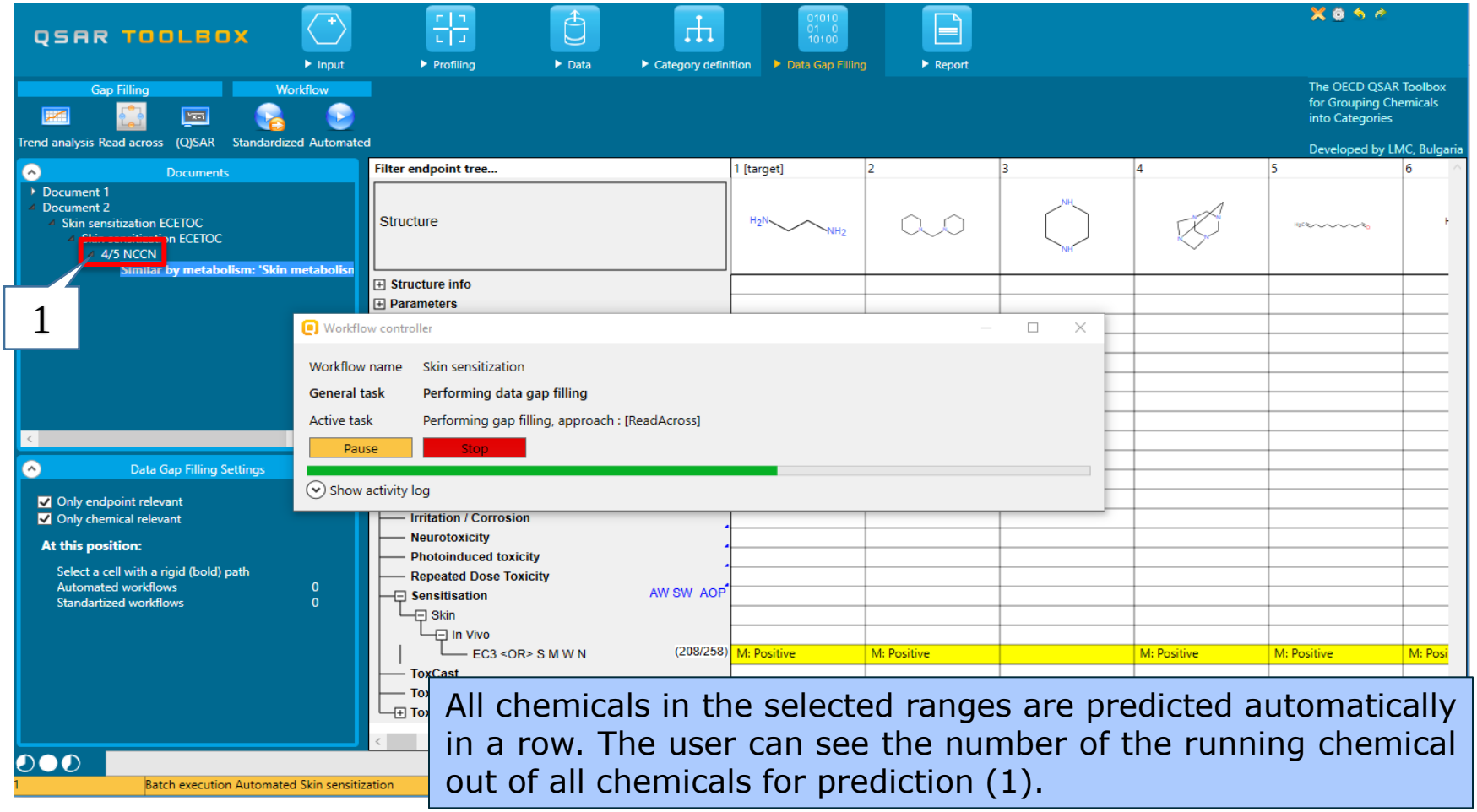
List with chemicals (batch work)



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)



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).

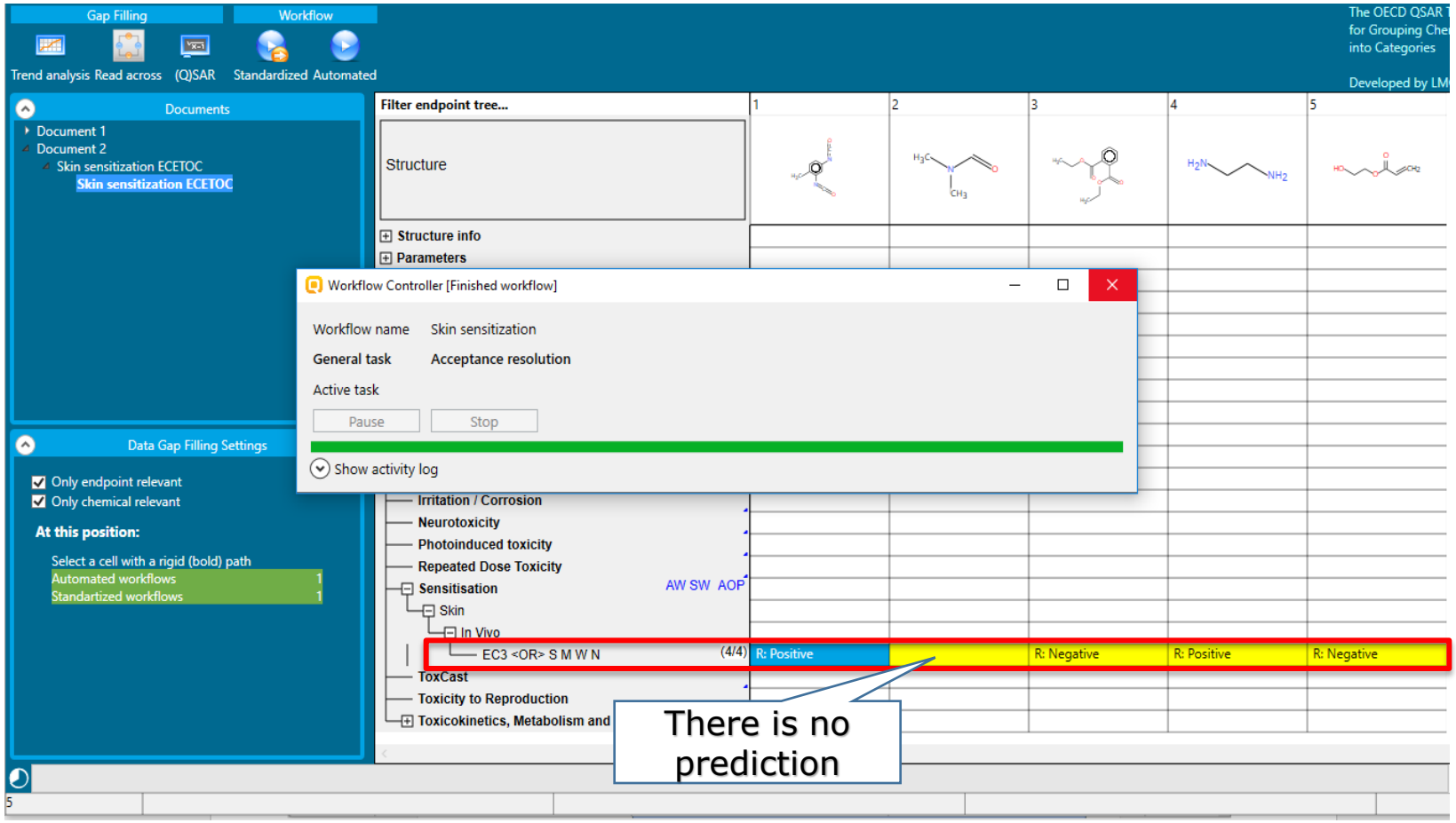
Execution of AW for Skin sensitization

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

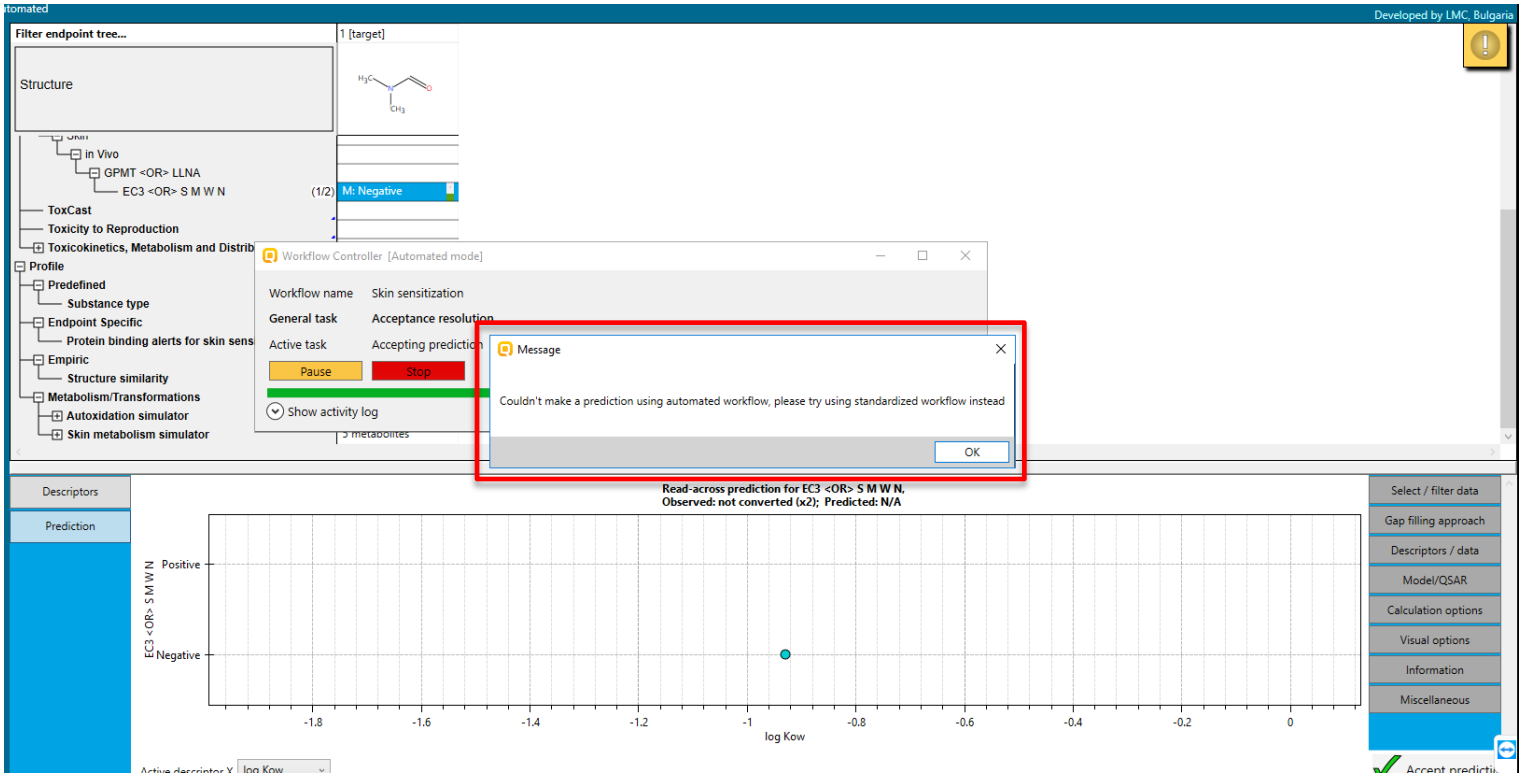
List with chemicals (batch work)



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.



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!