

OECD (Q)SAR Toolbox v.4.4.1

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

Keywords

TARGET CHEMICAL - chemical of interest

MODULE – a Toolbox module is a section dedicated to specific actions and options (e.g. Profiling)

WORKFLOW – the use, in combination, of the different modules (e.g. prediction workflow: from input to report)

PROFILER - algorithm (rule set) for the identification of specific features of the chemicals. Several types of profilers are available, such as structural (e.g. Organic functional groups), mechanistic (e.g. Protein binding by OECD) and endpoint-specific (e.g. in vitro in vitro mutagenicity (Ames test) alerts by ISS) profilers.

ALERT - the profilers consist of sets of rules or alerts. Each of the rules consists of a set of queries. The queries could be related to the chemical structure, physicochemical properties, experimental data, comparison with the target or list with substances and external queries from other predefined profilers (reference queries).

CATEGORY – “group” of substances sharing same characteristics (e.g. the same functional groups or mode of action). In a typical Toolbox workflow, it consists of the target chemical and its analogues gathered according to the selected profilers

ENDPOINT TREE – Endpoints are structured in a branched scheme, from a broader level (Phys-Chem properties, Environmental Fate and transport, Ecotoxicology, Human health hazard) to a more detailed one (e.g. EC3 in LLNA test under Human health hazard-Skin sensitization)

DATA MATRIX – Table reporting the chemical(s) and data (experimental results, profilers outcomes, predictions). Each chemical is in a different column and each data in a different row

Outlook

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

Automated workflow for skin sensitization

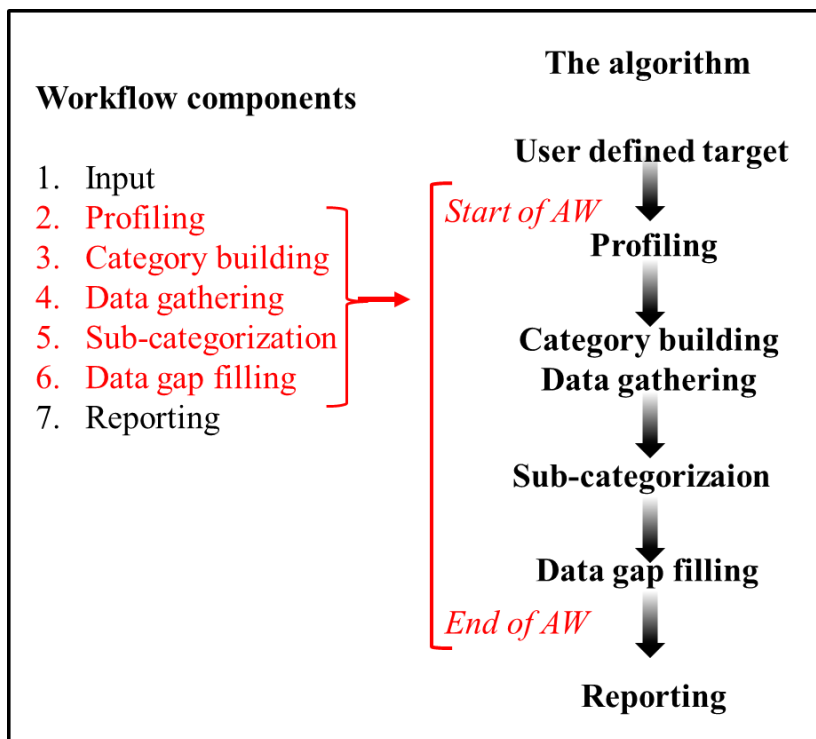
Mechanistic understanding

- Skin sensitization effect is a complex endpoint resulting in a multifactorial sequence of events.
- The initiating event of the 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).

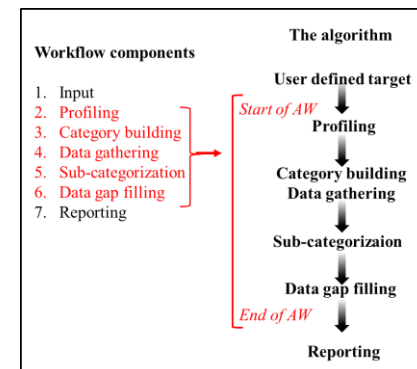


Automated workflow for skin sensitization

Workflow components

- **Profiling**

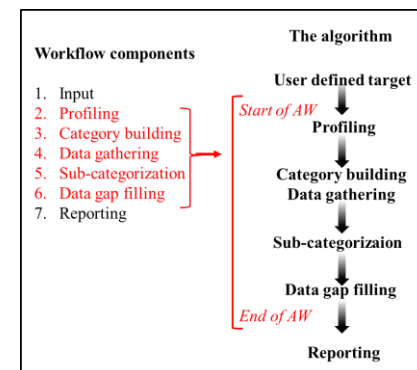
- ✓ Profilers for primary grouping
 - US EPA New Chemical Categories
 - Aquatic toxicity classification by ECOSAR
 - Protein binding alerts for Skin sensitization effect
 - Organic Functional Groups
 - Organic Functional Groups by US EPA
 - Organic Functional Groups by N. Haider
- ✓ Supporting profilers 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



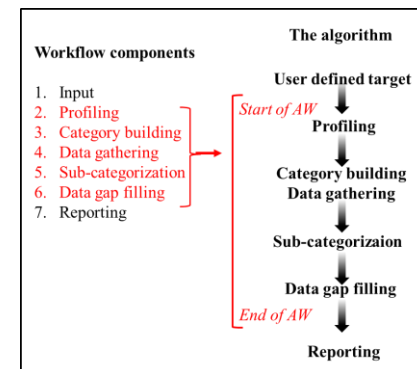
Unified skin sensitization scale in Toolbox	LLNA	GPMT
Positive	Extreme, Strong, Moderate, Weak	Strong & Moderate, Weak
Negative	Non	Non

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.



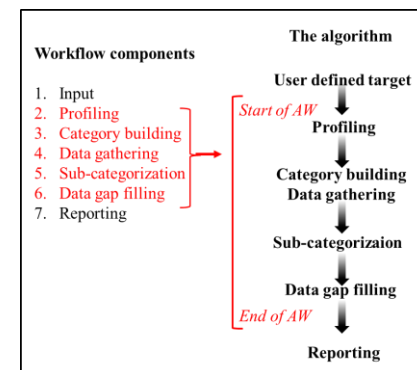
- ✓ If **more than one PBA** are identified in the parent structure or in the generated metabolites, then:
 - the category is defined based on all available PBA as presented in the target structure
 - In case no analogues found, the selection of alert is based on the criteria for **reliability** of alerts, 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

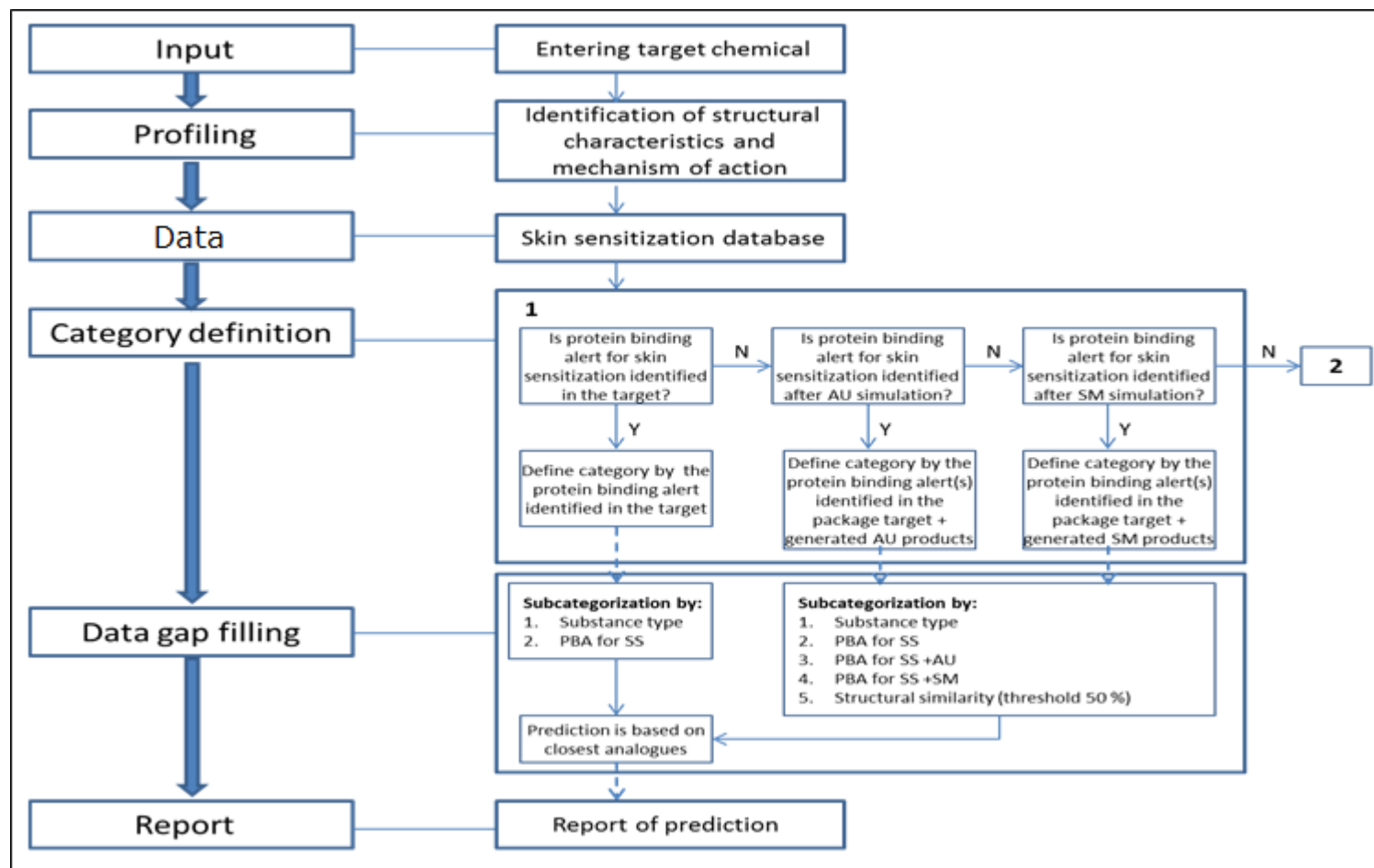
Workflow components

- **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)



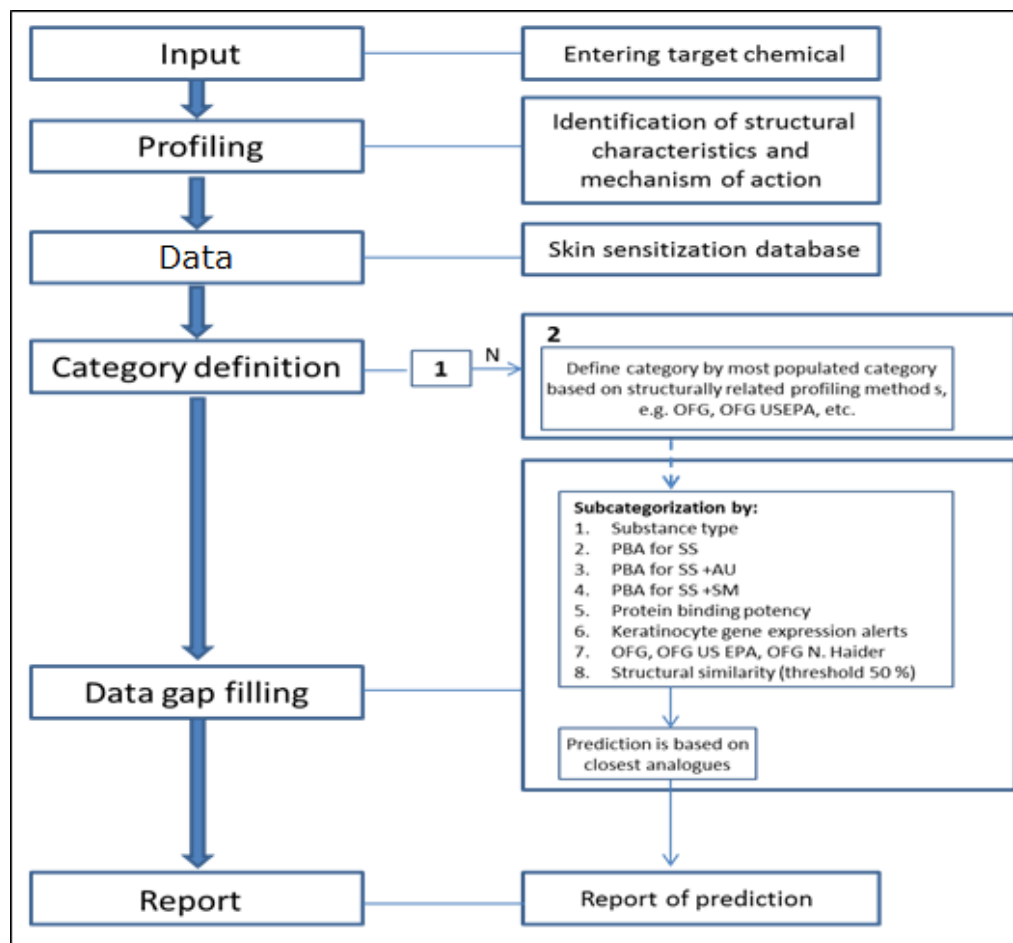
Automated workflow for skin sensitization

Algorithm of the workflow – Part 1



Automated workflow for skin sensitization

Algorithm of the workflow – Part 2



Automated workflow for skin sensitization

Summary

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

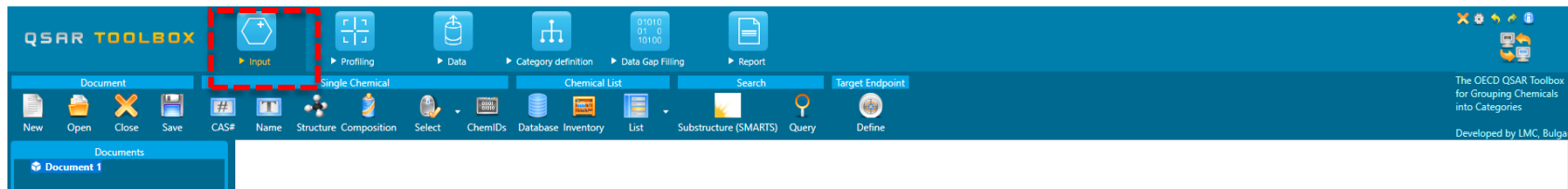
The Exercise

- 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 using 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

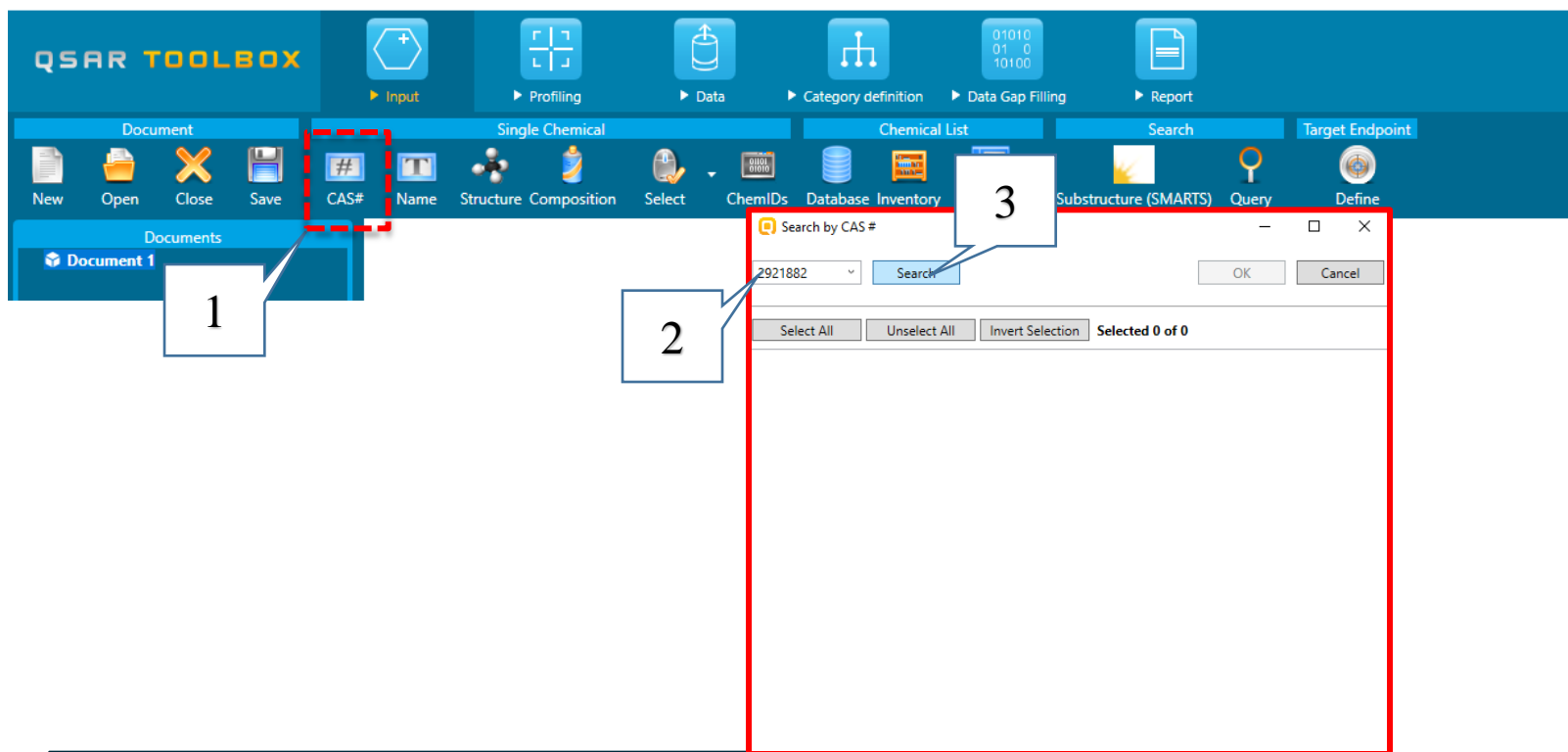


- Chemical Name
- Chemical Abstract Services (CAS) number (#)
- Drawing chemical structure with or without additional defining of the composition
- Select from User List/Inventory/Databases

Execution of AW for Skin sensitization

Single chemical, CAS # 2921-88-2

Input by CAS number



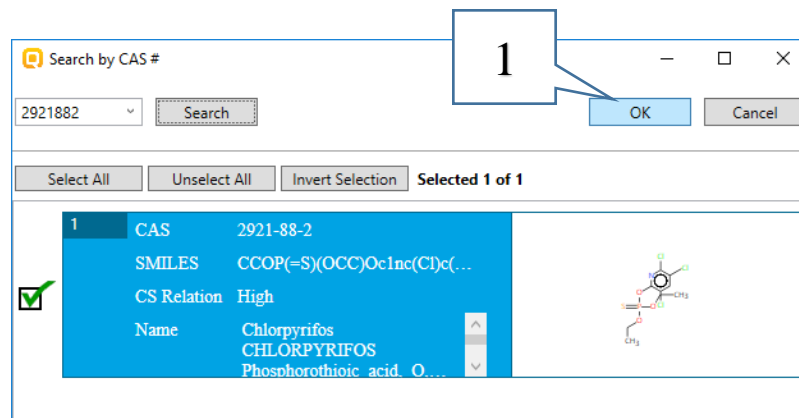
1. Click on **CAS#**;
2. Enter the CAS# of Chlorpyrifos (CAS **2921-88-2**);
3. Click on **Search**.

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-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 only one substance is identified based on the this CAS #.

Execution of AW for Skin sensitization

Location of the Automated workflow for skin sensitization

1. Go to **Data Gap Filling** module;

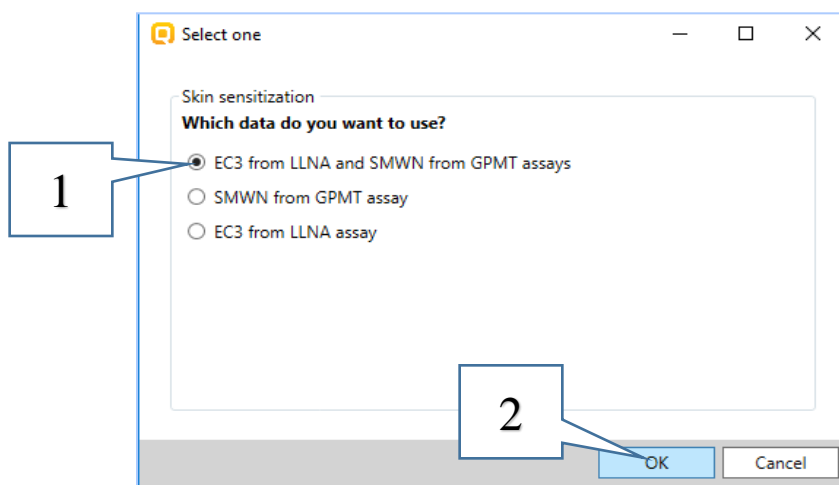
2. Click on **Automated** button;

3. Select **Skin sensitization**;

4. Confirm with **OK**

Execution of AW for Skin sensitization

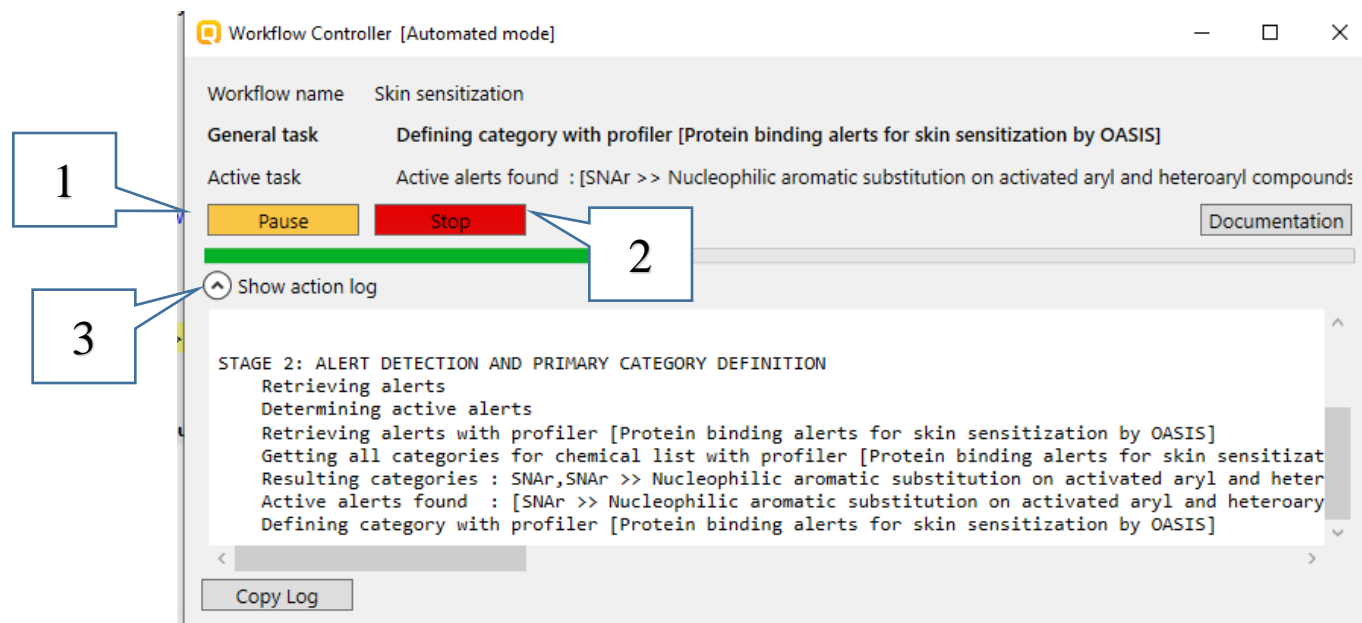
There are three options for endpoint selection and the user should select one of them.



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

The screenshot displays the QSAR Toolbox software interface during the execution of an Automated Workflow (AW) for Skin sensitization. The interface is divided into several sections:

- Top Menu Bar:** Contains icons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report.
- Left Sidebar:**
 - Documents:** Lists Document 1 with a tree structure showing chemical structures and their associated data.
 - Data Gap Filling Settings:** Includes checkboxes for "Only endpoint relevant" and "At this position:" (QSARs, Automated workflows, Standardized workflows).
 - In nodes below:** Lists QSARs, Automated workflows, and Standardized workflows.
- Central Workspace:**
 - Filter endpoint tree:** Shows a tree structure with endpoints like Immunotoxicity, Irritation / Corrosion, Neurotoxicity, Photoinduced toxicity, Repeated Dose Toxicity, Sensitisation, and Skin.
 - Chemical Structure Grid:** Displays a grid of chemical structures, with a callout box labeled "1" pointing to a specific structure.
 - Workflow Controller:** A window titled "Workflow Controller [Automated mode]" showing the workflow name "Skin sensitization", general task "Retrieving different analogues for profiler (Substance type)", and active task "Pause".
- Bottom Section:**
 - Read-across prediction for EC3, Skin sensitisation, based on 6 values:** Observed: Positive; Predicted: Positive.
 - Scatter Plot:** A plot of log Kow (x-axis) vs. EC3, Skin sensitisation (y-axis, Positive/Negative).
 - Active descriptor X:** log Kow.
 - Accept prediction:** A green checkmark indicating the prediction is accepted.

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

A yellow rectangular box with a black border. On the right side, there is a circular icon containing a black exclamation mark. The text "The current gap filling state contains data with qualifiers" is written in black font on the left side.

The current gap filling state contains data with qualifiers

- for analogue(s) with composition

A yellow rectangular box with a black border. On the right side, there is a circular icon containing a black exclamation mark. The text "The current gap filling state contains chemicals with composition" is written in black font on the left side.

The current gap filling state contains chemicals 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 (log Kow values)

A light blue rectangular box with a black border. On the right side, there is a circular icon containing a black lowercase letter 'i'. The text "Currently read-across uses data from 6 neighbours for prediction" is written in black font on the left side.

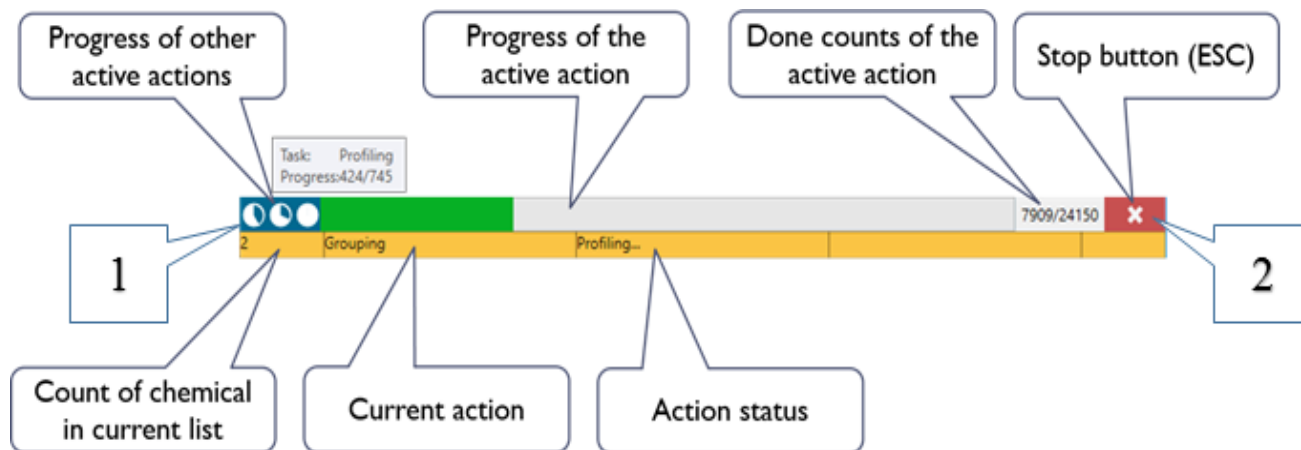
Currently read-across uses data from 6 neighbours for prediction

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 displays the QSAR Toolbox software interface during the execution of an automated workflow (AW) for skin sensitization. The interface is divided into several panels:

- Top Panel:** Contains the QSAR Toolbox logo and a series of icons representing different workflow stages: Input, Profiling, Data, Category definition, Data Gap Filling, and Report.
- Left Panel:** Includes a 'Documents' section showing a list of documents and a 'Data Gap Filling Settings' section with checkboxes for 'Only endpoint relevant' and 'At this position:'. Below these are sections for 'In nodes below:' and 'In nodes below:'.
- Central Panel:** Features a 'Filter endpoint tree...' on the left and a large matrix on the right. The matrix displays chemical structures and their corresponding predicted (R:) and experimental (M:) results. A red box labeled '1' highlights a specific row in the matrix where the predicted result (R:) matches the experimental data (M:).
- Right Panel:** Contains a 'Workflow Controller' window showing the workflow steps and a 'Success' dialog box indicating 'Prediction accepted successfully'. A red box labeled '2' highlights the 'OK' button in the 'Success' dialog.

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 shows the QSAR Toolbox interface. The 'Data points' window is open, displaying a table of data points. The table has the following columns: Datapoints, Value, Original value, Assay, Assigned SMILE, Author, Comments, Database, Endpoint, Identity, Institution and, Organ, and Prediction approach. The table contains two rows of data. The first row is highlighted with a red border. The second row is highlighted with a blue border. A red dashed box highlights the cell containing 'R: Positive 47/93 M: 6.91 %' in the 'Data points' table. A red box highlights the 'Prediction approach' column, and a blue box highlights the 'Value' column. A red box highlights the 'Assay' column, and a blue box highlights the 'Endpoint' column. A red box highlights the 'Identity' column, and a blue box highlights the 'Institution and' column. A red box highlights the 'Organ' column, and a blue box highlights the 'Prediction approach' column.

Datapoints	Value	Original value	Assay	Assigned SMILE	Author	Comments	Database	Endpoint	Identity	Institution and	Organ	Prediction approach
Human Health Hazards;Sensitisation	R: Positive (Skin sensitisation II (ECETOC))	Positive (Skin sensitisation II (ECETOC))	GPMT <OR> LLNA					EC3 <OR> Skin sensitisation			Skin	Read-across, executed by AW 'Skin sensitization'
Human Health Hazards;Sensitisation	M: 6.91 % (Skin sensitization EC3 (ratio))	6.91 % (Skin sensitization EC3 (ratio))	LLNA	False	Submitted by Dow AgroScience	sensuser = Category B (solid-based indication for contact allergenic		EC3	1032	LMC,BUL	Skin	

The 'Data Gap Filling Settings' panel shows 'Only endpoint relevant' checked and 'At this position:' set to 'Sensitisation'. The 'In nodes below:' section shows 'QSARs', 'Automated workflows', and 'Standardized workflows' all set to 0.

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" (3).

Execution of AW for Skin sensitization

The screenshot shows the QSAR Toolbox interface with the Workflow Controller window open. The window displays the workflow steps for skin sensitization. A callout '1' points to the 'Show action log' button, and a callout '2' points to the close button (X) of the Workflow Controller window.

Workflow Controller [Finished workflow]

Workflow name:
General task: Passing control to the workflow controller

Active task:
Pause Stop Documentation

Show action log

STAGE 1: INITIAL DATA COLLECTION AND CHECKS

- Retrieving target
- Subcategorization by profiler [Substance type]
- Select databases
- Setting selected databases to [Skin Sensitization, Skin sensitization]

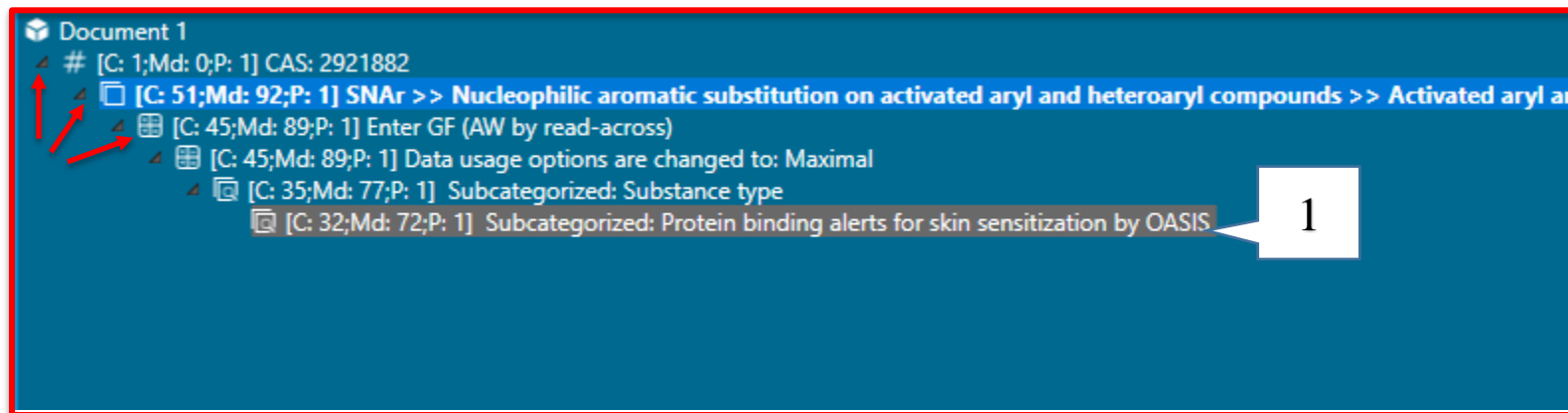
STAGE 2: ALERT DETECTION AND PRIMARY CATEGORY DEFINITION

- Retrieving alerts
- Determining active alerts
- Retrieving alerts with profiler [Protein binding alerts for skin sensitization]
- Getting all categories for chemical list with profiler [Protein binding alerts for skin sensitization]
- Resulting categories : SNAR, SNAR >> Nucleophilic aromatic substitution
- Active alerts found : [SNAR >> Nucleophilic aromatic substitution]
- Defining category with profiler [Protein binding alerts for skin sensitization]
- Computing alert performance for profiler [Protein binding alerts for skin sensitization]
- Alerts:
 - SNAR >> Nucleophilic aromatic substitution on activated aryl
 - SN2 >> Nucleophilic substitution at sp³ carbon atom >> (Thio)

Copy Log

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. The grey highlighting of the last leaf of the tree indicates that a prediction has been accepted at this level.

Execution of AW for Skin sensitization

Report: Generation of report

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes options like Input, Profiling, Data, Category definition, Data Gap Filling, and Report. The Report module is highlighted with a red dashed box and labeled '1'. On the left, the Reports panel shows various report types, with 'Prediction Data' highlighted by a red box and labeled '3'. The main workspace shows a table of results for Skin sensitization, with a yellow row highlighted and labeled '2'. The table includes columns for chemical structure, prediction results, and various metrics.

1. Go to the **Report** module;

2. Click on the cell with the prediction;

3. Click on **Prediction** button

Execution of AW for Skin sensitization

Report: Generation of report

Customize report content and appearance

Select which sections to include into report by checking/unchecking the corresponding section box. Rearrange the order of appearance by using buttons "Move Up" and "Move Down".

Wizard pages

Customization

Customize report

Prediction

Target and prediction summary

Prediction details (I)

Prediction details (II)

Target profiles

Analogues selection details

Category

Category definition and members

Consistency check

Options

Data matrix

Options

Add RAAF scenario

☒ **Prediction**

☒ Target and prediction summary

☒ Prediction details (I)

☒ Prediction details (II)

☒ Target profiles

☒ Analogues selection details

☐ Appendix: Grouping / subcategorization

☐ Appendix: Data pruning

☐ Appendix: Specific report explanations

☒ **Category**

☒ Category definition and members

☒ Consistency check

☒ Options

☒ **Data matrix**

☒ Options

☐ Remove password protection of the PDF files.
Note: If the protection is removed, this will be specified in the first page of the report

Move Up Move Down

Back Next Cancel **Create report**

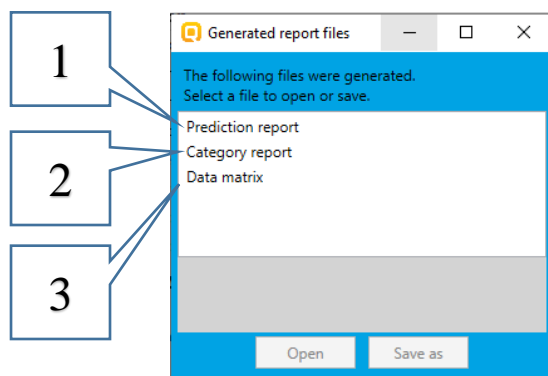
The user can customize the report content (1) and appearance (2). Generation of the reports happens by click on the **Create report** button (3).

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) Category report** - a PDF file containing information for the consistency of the final category (target plus used analogues).
- 3) 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 report

Prediction of EC3, Skin sensitisation for LORSBAN

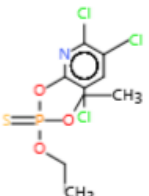
1 / 7

QSAR Toolbox prediction for single chemical

Date: 14 Apr 2020

Author(s):

Contact details:

Target information		
Structural information SMILES: <chem>CCOP(=S)(OCC)Oc1nc(C)cc(Cl)c1</chem>  Structure	Numerical identifiers CAS#: 2921-88-2 Other: EC Number: 2208644	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, Skin sensitisation; 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, Automated workflow for Skin sensitization Summary: manually editable field Not provided by the user

Execution of AW for Skin sensitization

Report: Category report

Category report

Chemicals category

QSAR Toolbox report for category

1. Category definition

1.1. Category definition

Category name

Not provided by the user

Covered (target) endpoint(s)

- Human Health Hazards/Sensitisation: EC3 <OR> S M W N <OR> Skin sensitisation, GPMT <OR> LLNA, In Vivo, Skin

Category hypothesis

Not provided by the user

1.2. Category members

Information of category members

Table of category members

#	CAS	Name	SMILES	Structure
1	2921-88-2	LORSBAN	<chem>CCOP(=S)(OCC)Oc1cc(Cl)c(Cl)cc1Cl</chem>	
2	877399-00-3	(r)-5-bromo-3-(1-(2,6-dichloro-3-fluorophenyl)ethoxy)pyridin-2-amine	<chem>CC(Oc1cc(Br)cn1N)c1c(Cl)c(Cl)cc1Cl</chem>	
3	42874-03-3	Oxyfluorfen	<chem>CCOCc1cc(Oc2cc(cc2C)C(F)(F)F)ccc1N+([O-])=O</chem>	

Chemicals category

2 / 7

4	81406-37-3	Fluroxypyr-methyl	<chem>CCCCC(C)OC(=O)CCc1nc(F)c(Cl)c(N)c1Cl</chem>	
5	130841-23-5	1,4-dichloro-2-((1,1,2,3,3,3-hexafluoropropoxy)-5-nitrobenzene	<chem>[O-][N+](=O)c1cc(Cl)c(Cl)c(OC(F)(F)C(F)(F)F)cc1Cl</chem>	
6	64700-56-7	TRICLOPYR ESTER	<chem>CCCCCOC(=O)CCc1nc(Cl)c(Cl)cc1Cl</chem>	
7	64470-88-8	Formulation 16	<chem>CCCCCOC(=O)CCc1nc(Cl)c(Cl)cc1Cl</chem>	

Ranges for selected physicochemical properties and calculated parameters

Not provided by user

Purity / Impurity

Not provided by the user

1.3. Profiles/Metabolisms

List of profiles/metabolisms

Profiles used for grouping/subcategorization:

- SHAr >> Nucleophilic aromatic substitution on activated aryl and heteroaryl compounds >> Activated aryl and heteroaryl compounds (Protein binding alerts for skin sensitization by QASIS) (primary grouping)
- Substance type (subcategorization)
- Protein binding alerts for skin sensitization by QASIS (subcategorization)

2. Consistency check

2.1. Physicochemical similarity

Information for the members of the category obtained as a result of AW application is included in the Category report.

Execution of AW for Skin sensitization

Report: Generation of report

[illegible]

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
- activating the automated workflow
- specification of the endpoint

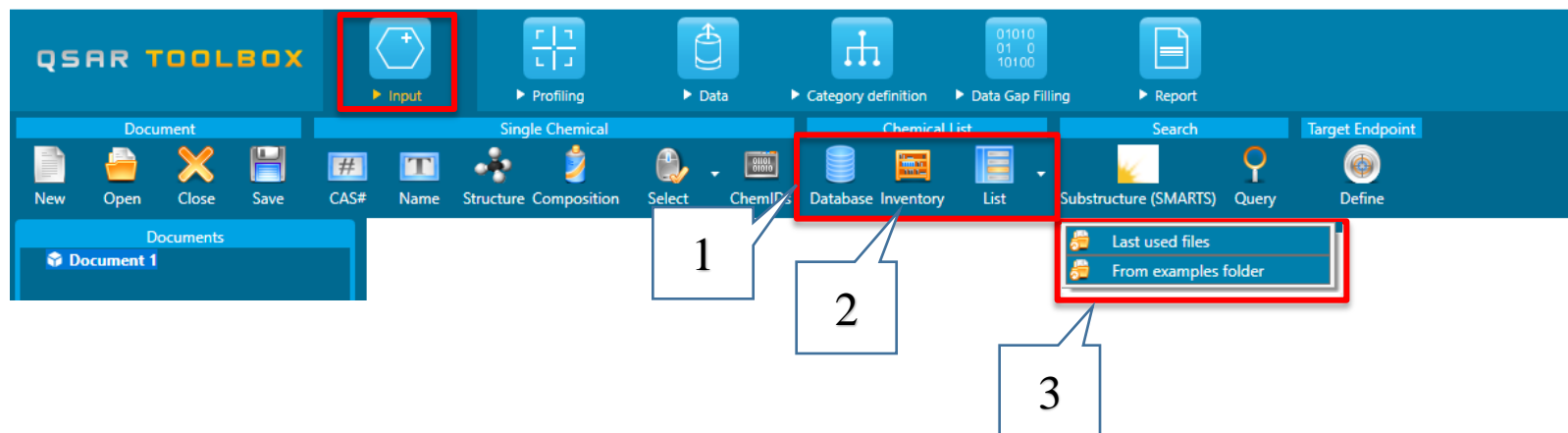
The Exercise

- In this exercise we will predict the skin sensitization effect for:
 - Single chemical – Chlorpyrifos [CAS# 2921-88-2];
 - **Batch of chemicals – all 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

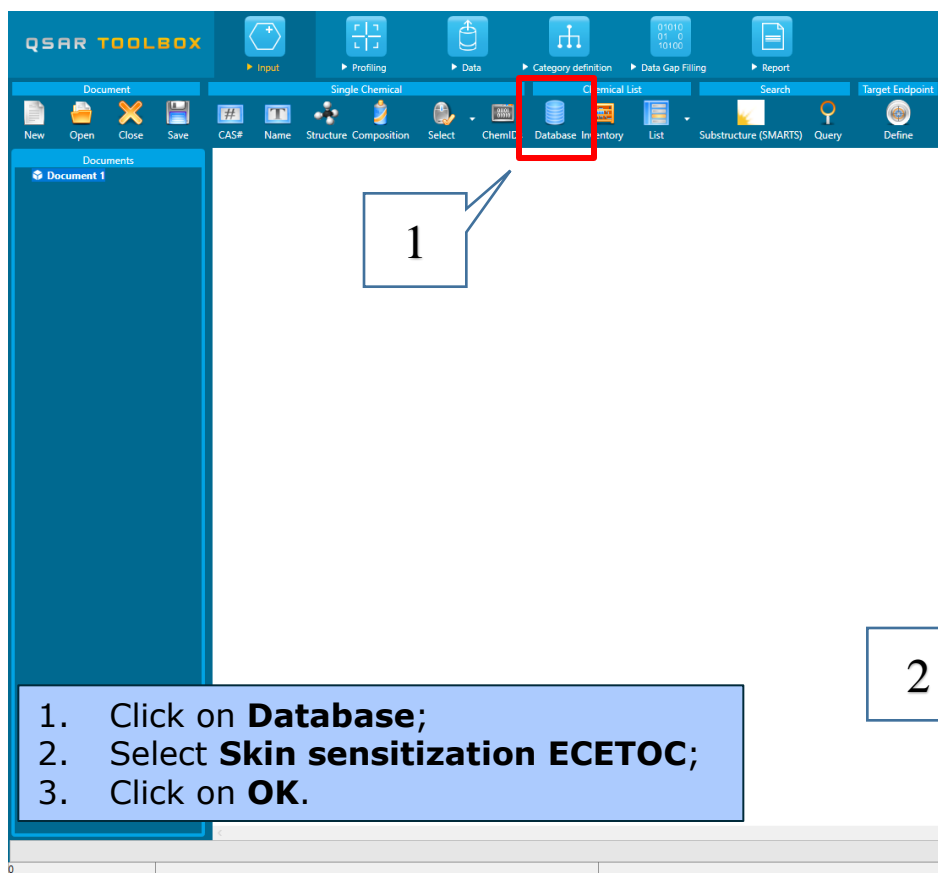


- 1) Database
- 2) Inventory
- 3) List: Last used files/ From examples folder

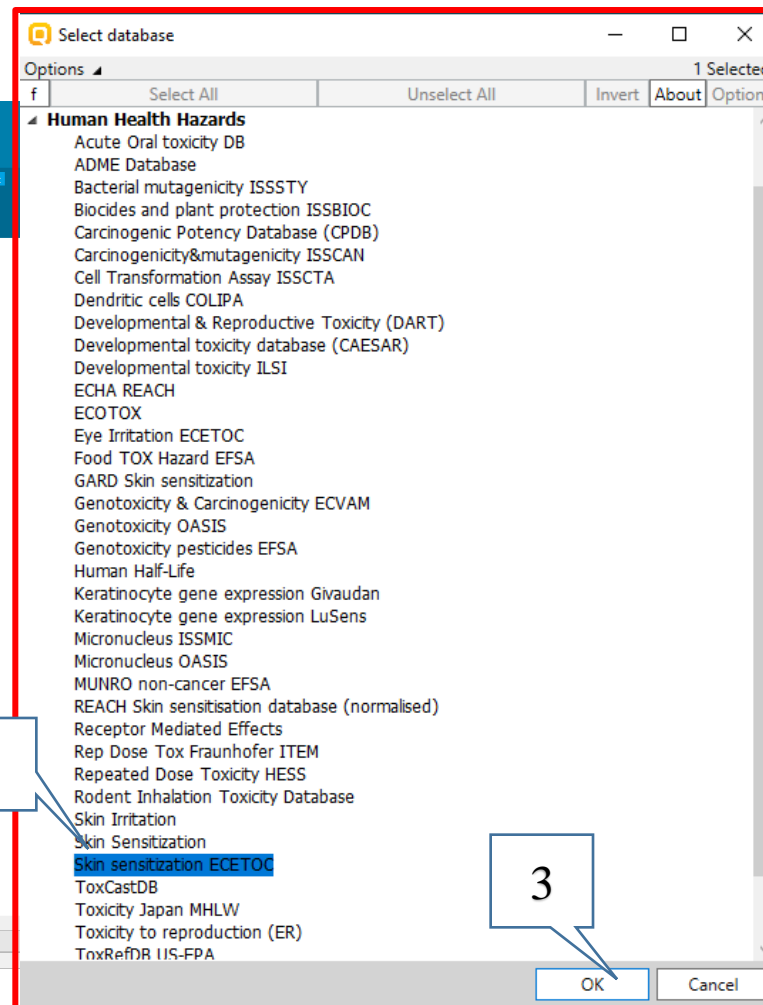
Execution of AW for Skin sensitization

List with chemicals (batch work)

Input: Load a database



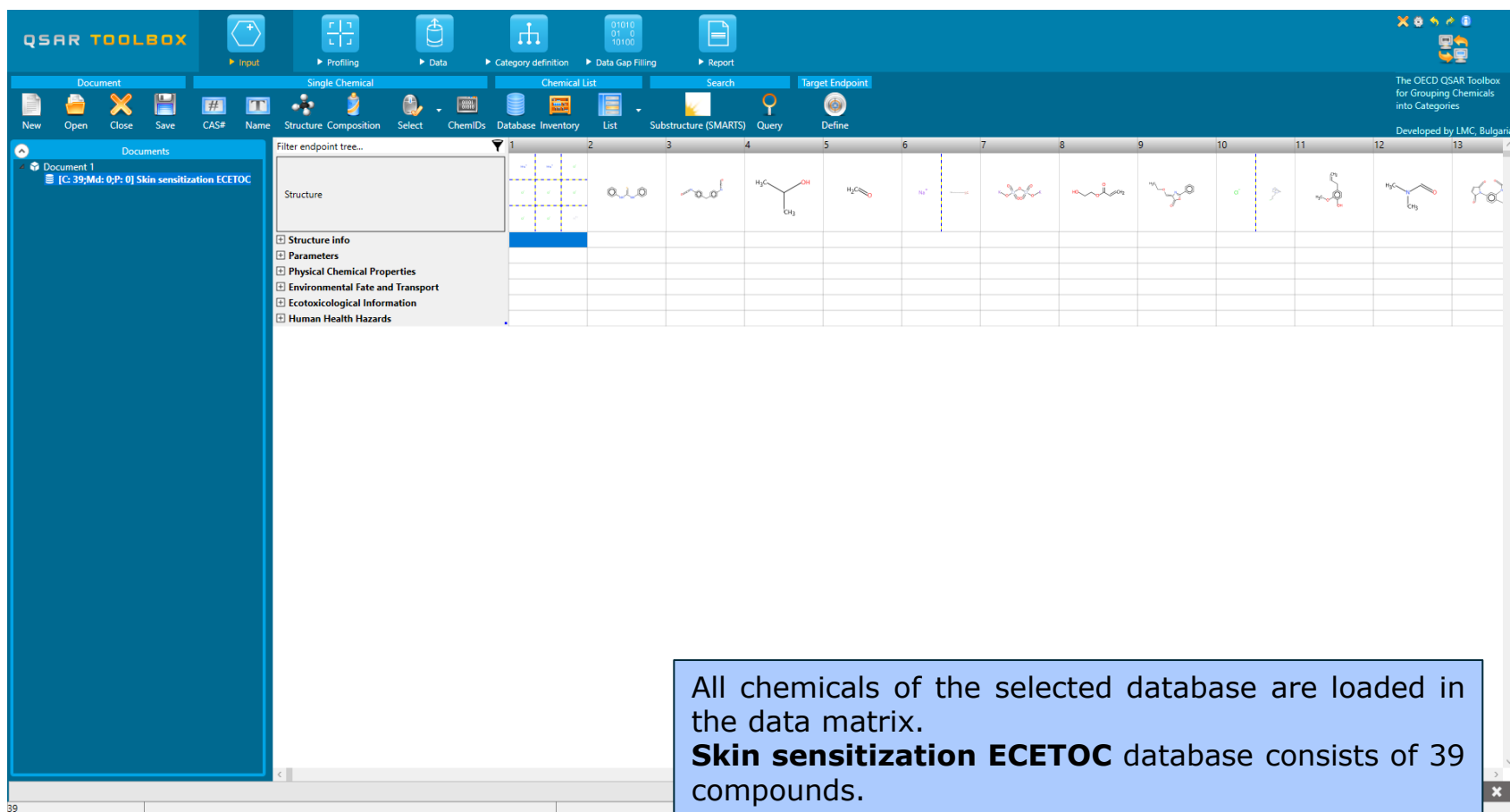
1. Click on **Database**;
2. Select **Skin sensitization ECETOC**;
3. Click on **OK**.



Execution of AW for Skin sensitization

List with chemicals (batch work)

Input: Load a database



The screenshot displays the QSAR Toolbox software interface. The top menu bar includes options like Document, Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below this, a toolbar contains icons for New, Open, Close, Save, CAS#, Name, Structure, Composition, Select, ChemIDs, Database, Inventory, List, Substructure (SMARTS), Query, and Define. The main window is divided into several panes. On the left, the 'Documents' pane shows a list of documents, including 'Document 1' and '[C: 39; Md: 0; P: 0] Skin sensitization ECETOC'. The 'Filter endpoint tree...' pane on the right lists various endpoints: Structure, Structure Info, Parameters, Physical Chemical Properties, Environmental Fate and Transport, Ecotoxicological Information, and Human Health Hazards. The central data matrix displays a grid of chemical structures, with columns numbered 1 through 13. The first row of the matrix shows chemical structures for each column, including benzene, ethanol, isopropanol, acetic acid, and others. A text box in the bottom right corner provides additional information about the database used.

All chemicals of the selected database are loaded in the data matrix.
Skin sensitization ECETOC database consists of 39 compounds.

Execution of AW for Skin sensitization

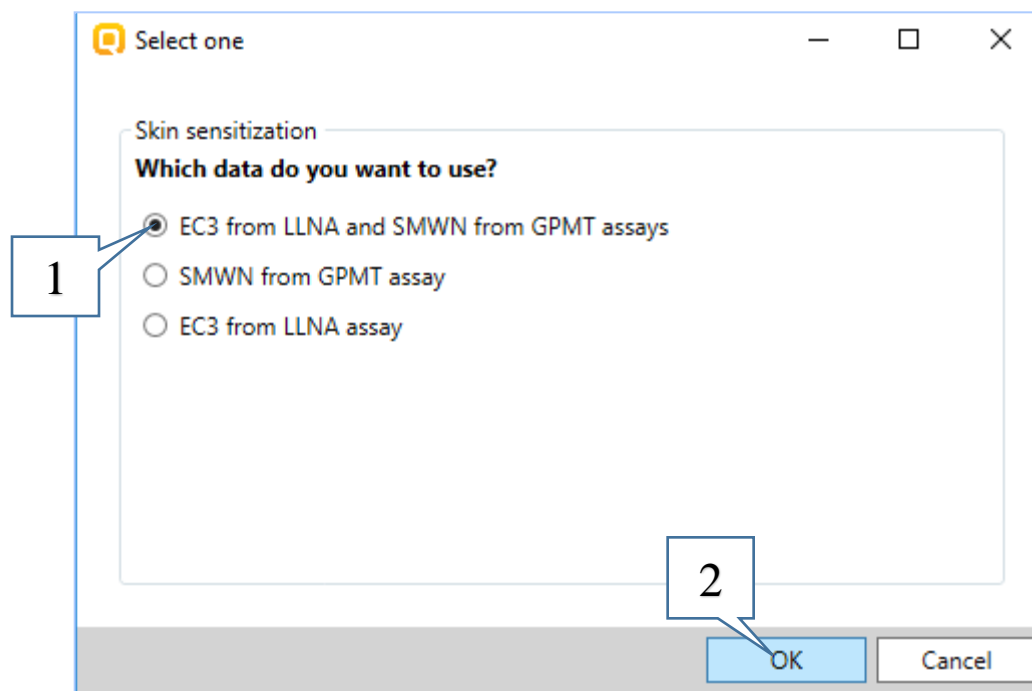
List with chemicals (batch work)

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The 'Data Gap Filling' module is highlighted with a red dashed box and labeled '1'. Below the menu bar, the 'Workflow' section shows 'Automated' selected, labeled '2'. A 'Select workflow' dialog box is open, showing 'Skin sensitization' selected under 'Choices', labeled '3'. The 'OK' button in the dialog is labeled '4'. The left sidebar shows 'Documents' with 'Document 1' and 'Skin sensitization ECETOC'. The 'Data Gap Filling Settings' panel is also visible, showing 'Only endpoint relevant' checked. The main workspace displays a list of chemicals with their structures and names.

1. Go to **Data Gap Filling** module;
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 whole range is set by default.

Execution of AW for Skin sensitization

List with chemicals (batch work)

The screenshot shows the QSAR Toolbox software interface. The top menu bar includes options like 'Gap Filling', 'Workflow', 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The left sidebar contains a 'Documents' panel with 'Document 1' and '[C: 39; Md: 0; P: 0] Skin sensitization ECETOC'. Below this is the 'Data Gap Filling Settings' panel, which has a 'Only endpoint relevant' checkbox checked and a table showing counts for 'At this position' and 'In nodes below' for 'QSARs', 'Automated workflows', and 'Standardized workflows'. The main workspace is divided into a 'Filter endpoint tree' on the left and a grid of chemical structures on the right. The 'Filter endpoint tree' lists various endpoints such as 'Acute Toxicity', 'ADME', 'Bioaccumulation', 'Carcinogenicity', 'Developmental Toxicity / Teratogenicity', 'Genetic Toxicity', 'Immunotoxicity', 'Irritation / Corrosion', 'Neurotoxicity', 'Photoinduced toxicity', 'Repeated Dose Toxicity', 'Sensitisation', 'Skin', 'In Vivo', 'GPMT <OR> LLNA', 'EC3 <OR> S M W N <OR> Ski...', 'ToxCast', 'Toxicity to Reproduction', and 'Toxicokinetics, Metabolism and Distribution'. The 'Sensitisation' endpoint is expanded, showing 'Skin' and 'In Vivo'. The 'In Vivo' endpoint is further expanded, showing 'GPMT <OR> LLNA' and 'EC3 <OR> S M W N <OR> Ski...'. The 'EC3 <OR> S M W N <OR> Ski...' endpoint is highlighted in yellow. A 'Workflow Controller' dialog box is open in the center, with a red box highlighting the 'Select the range on which to execute the ...' prompt. The dialog shows 'Workflow name: Skin sensitization', 'General task: Active task', and input fields for '≥ 1' and '≤ 39'. It also has 'Pause', 'Stop', and 'Documentation' buttons.

Execution of AW for Skin sensitization

List with chemicals (batch work)

- The workflow in a batch mode finishes when the automated workflow is executed for the 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 displays the QSAR Toolbox software interface. The top menu bar includes options like Input, Profiling, Data, Category definition, Data Gap Filling, and Report. The left sidebar shows a 'Documents' panel with a selected document '[C: 39;Md: 0;P: 25] AW "Skin sensitization"' and a 'Data Gap Filling Settings' panel with checkboxes for 'Only endpoint relevant' and 'At this position:'. The central workspace shows a grid of chemical structures. A 'Workflow Controller' dialog box is open, displaying the workflow name, general task, active task, and a log of actions. A red box highlights a row in the grid, and a callout points to it with the text 'There is no prediction'.

Workflow Controller [Finished workflow]

Workflow name

General task

Active task

Pause Stop Documentation

Show action log

STAGE 1: INITIAL DATA COLLECTION AND CHECKS

Retrieving target

Subcategorization by profiler [Substance type]

Copy Log

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. Another alternative is to execute Skin sensitization workflow for defined approaches.

The screenshot displays the QSAR Toolbox interface. On the left, the 'Filter endpoint tree...' panel shows a selection of endpoints including 'EC3 <OR> S M W N'. The main workspace shows a chemical structure of 2-propanone (acetone) and a table with one entry: (1/2) M: Negative. A '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 an error message box that appears: 'Message: Couldn't make a prediction using automated workflow, please try using standardized workflow instead'. Below the error message, a scatter plot titled 'Read-across prediction for EC3 <OR> S M W N, Observed: not converted (x2); Predicted: N/A' is shown. The plot has 'log Kow' on the x-axis (ranging from -1.8 to 0) and 'EC3 <OR> S M W N' on the y-axis (Positive/Negative). A single data point is plotted at approximately (-1.0, Negative). The right sidebar contains a 'Select / filter data' menu with options like 'Gap filling approach', 'Descriptors / data', 'Model/QSAR', 'Calculation options', 'Visual options', 'Information', and 'Miscellaneous'.

Execution of AW for Skin sensitization

List with chemicals (batch work)

Report: Generation of report

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

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

Congratulations!

- 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!