OECD (Q)SAR Toolbox v.4.4

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).

```
Workflow components
1. Input
2. Profiling
3. Category building
4. Data gathering
5. Sub-categorization
6. Data gap filling
7. Reporting
```

![Diagram showing the workflow components and algorithm](Diagram.png)
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

<table>
<thead>
<tr>
<th>Unified skin sensitization scale in Toolbox</th>
<th>LLNA</th>
<th>GPMT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive</strong></td>
<td>Extreme, Strong, Moderate, Weak</td>
<td>Strong &amp; Moderate, Weak</td>
</tr>
<tr>
<td><strong>Negative</strong></td>
<td>Non</td>
<td>Non</td>
</tr>
</tbody>
</table>
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:
  o **Single chemical – Chlorpyrifos [CAS# 2921-88-2];**
    o 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-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 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.
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 (log Kow 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)
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).
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).
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

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

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.

![Generated report files window](image)
Execution of AW for Skin sensitization

Report: Generation of report

Prediction report

QSAR Toolbox prediction for single chemical

Date: 3 Sep 2019
Author(s):
Contact details:

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

Report: Category report

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

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:
  o Single chemical – Chlorpyrifos [CAS# 2921-88-2];
  o 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

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)

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)
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)

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.
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!