Automated and Standardized workflows for predicting Acute aquatic toxicity and Skin sensitization

Laboratory of Mathematical Chemistry,
University “Prof. Assen Zlatarov” - Bourgas, Bulgaria

February 2020

Outlook

- Endpoints
- Specificities
- Components
- Executing module
- Algorithm of Ecotoxicological workflow
- Predicting Ecotoxicity by using the workflows - examples
- Algorithm of Skin sensitization workflow
- Predicting Skin sensitization by using the workflows - examples

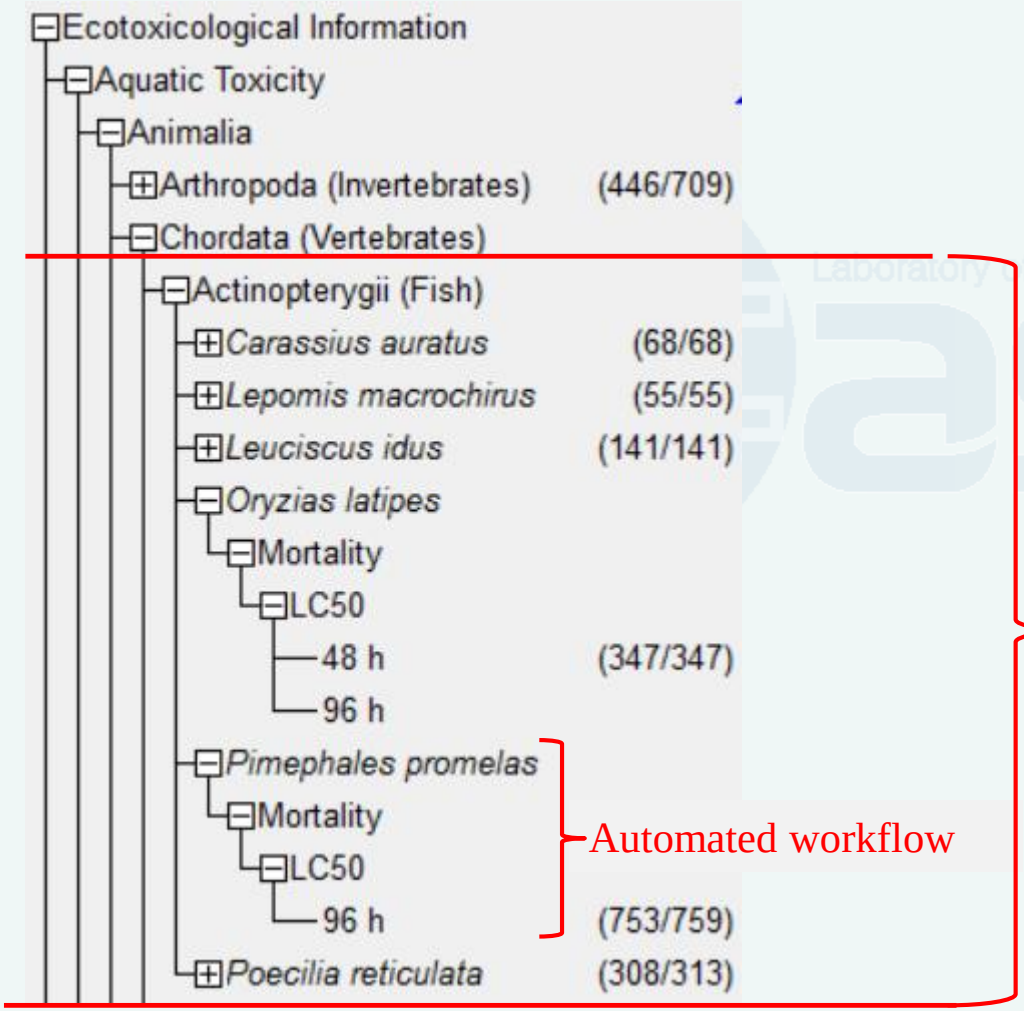
Outlook

- Endpoints
- Specificities
- Components
- Executing module
- Algorithm of Ecotoxicological workflow
- Predicting Ecotoxicity by using the workflows - examples
- Algorithm of Skin sensitization workflow
- Predicting Skin sensitization by using the workflows - examples

Automated and Standardized workflows in Toolbox

Endpoints

- **Acute aquatic toxicity – Domain of application**

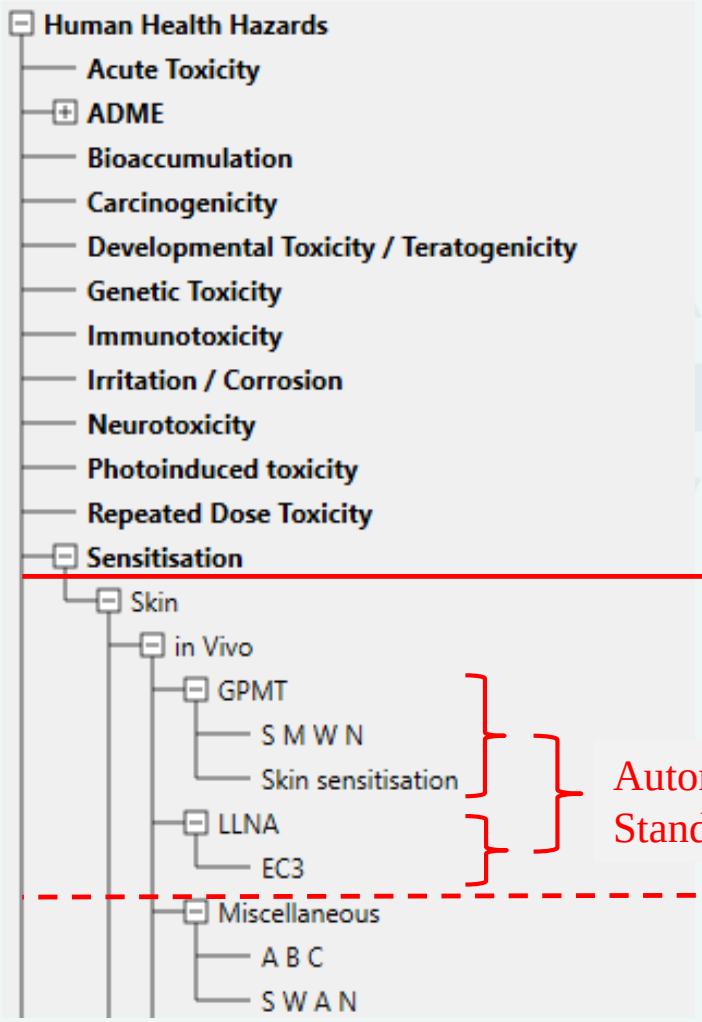


- Sub-hazard – Aquatic toxicity
- Effect – Mortality, Intoxication, Immobilization, Growth, Population
- Endpoint – LC50;EC50; IC50
- Duration: 48 - 96 h
- Species:
  - all fish*: P. promelas*, *P. reticulate*, etc.
  - Crustaceans
  - Green algae

Automated and Standardized workflows in Toolbox

Endpoints

- **Skin sensitization – Domain of application**



Automated workflow /
Standardized workflow

- Sub-hazard – Skin sensitization
- Method – *in vivo*
- Assays – LLNA + GPMT, LLNA only, GPMT only
- Endpoint – EC3 + SMWN converted in dichotomous scale

Outlook

- Endpoints
- Specificities
- Components
- Executing module
- Algorithm of Ecotoxicological workflow
- Predicting Ecotoxicity by using the workflows - examples
- Algorithm of Skin sensitization workflow
- Predicting Skin sensitization by using the workflows - examples

Automated and Standardized workflows in Toolbox

Specificities

Automated workflow (AW)

- On the fly data gap filling – the user is not able to control the workflow
- Input information – only target chemical ID
- Independency – application is not affected by the user activities (proceeding or subsequent)
- Batch mode execution

Standardized workflow (SW)

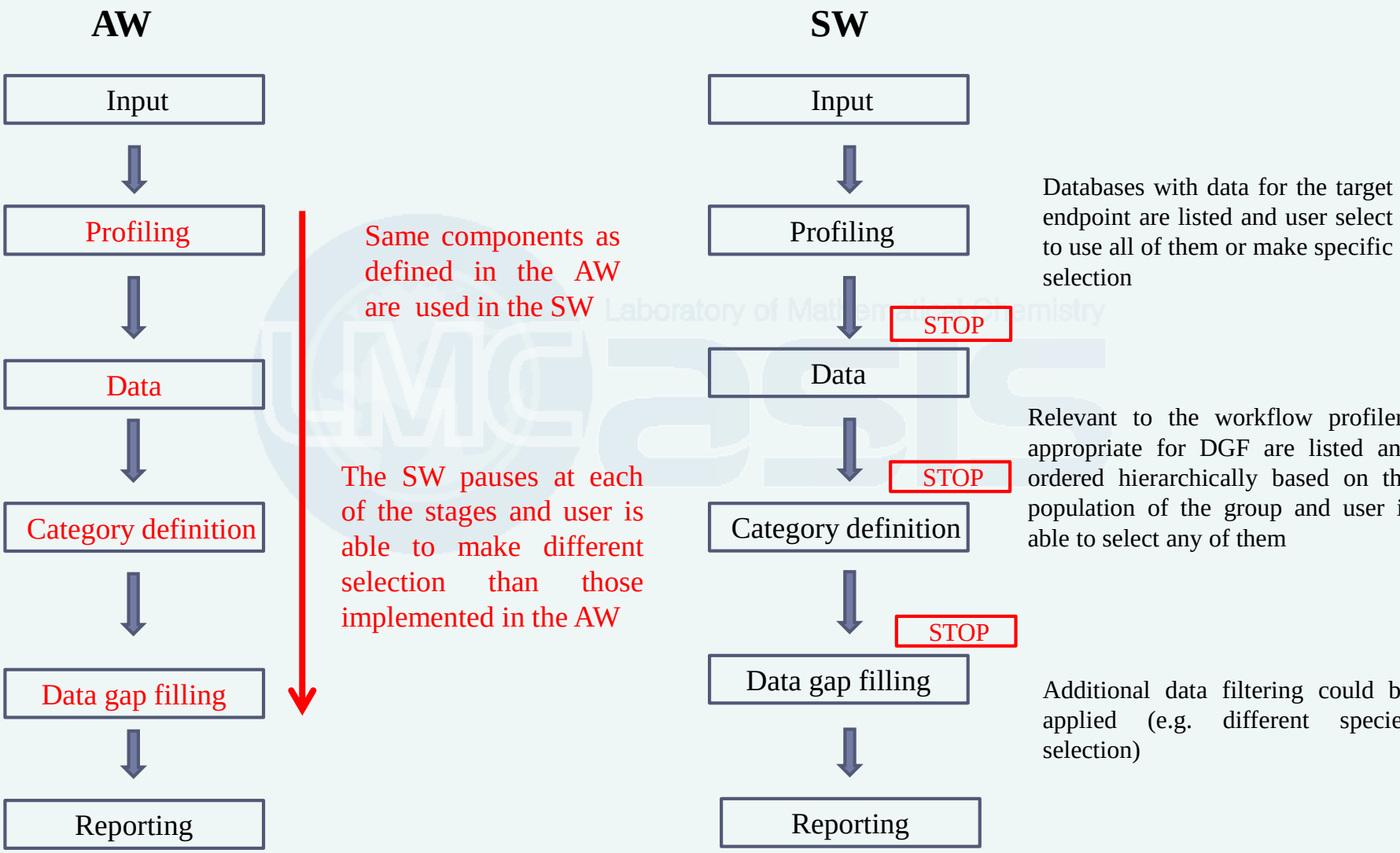
- Applicability domain of SW is expanded as compared to the AWs, including additional endpoints, effects, durations and species.
- The SWs follow same steps as in the AWs.
- The SWs stop at the each step of the workflows allowing the user to make different selection.

Outlook

- Endpoints
- Specificities
- Components
- Executing module
- Algorithm of Ecotoxicological workflow
- Predicting Ecotoxicity by using the workflows - examples
- Algorithm of Skin sensitization workflow
- Predicting Skin sensitization by using the workflows - examples

Automated and Standardized workflows in Toolbox

Components



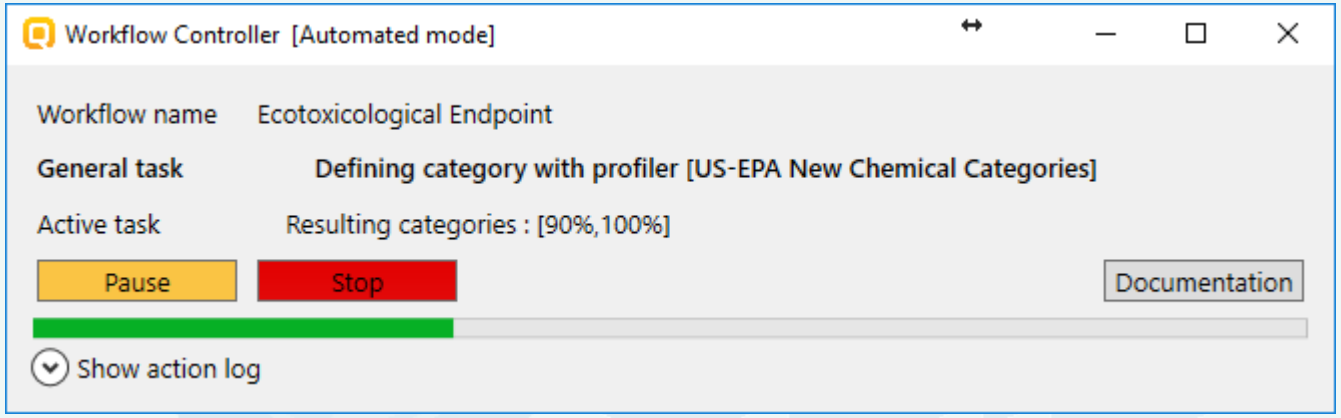
Outlook

- Endpoints
- Specificities
- Components
- Executing module
- Algorithm of Ecotoxicological workflow
- Predicting Ecotoxicity by using the workflows - examples
- Algorithm of Skin sensitization workflow
- Predicting Skin sensitization by using the workflows - examples

Automated and Standardized workflows in Toolbox

Executing module

- The AWs and SWs are controlled by the "*Workflow controller*"



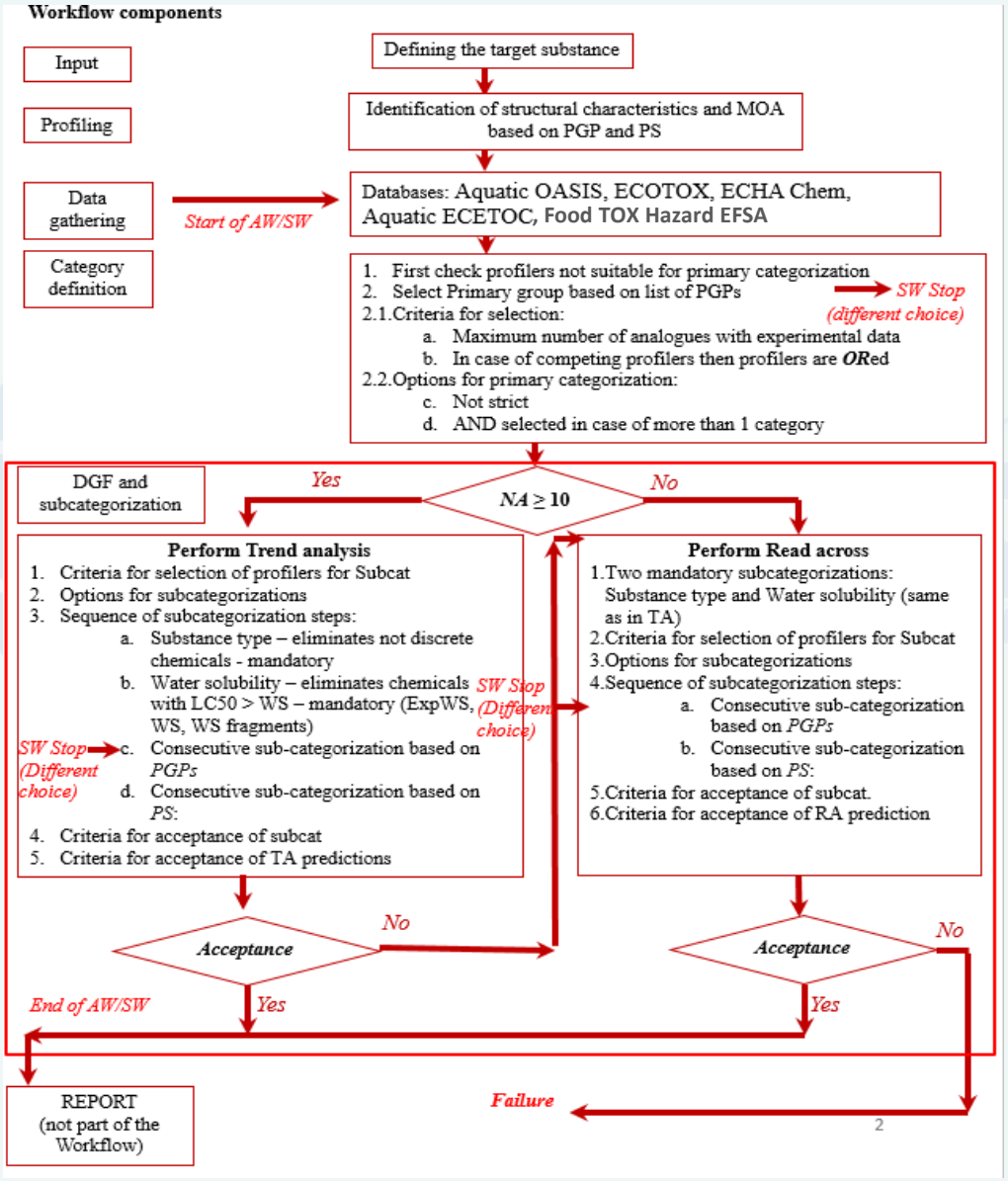
- The workflow is navigated by two main buttons:
  - *Continue/Pause* allowing to continue or pause the workflow
  - *Stop* which cancel (deactivate) the workflow.
- The algorithm of the current workflow could be seen by the *Documentation* button.
- Additionally, all actions that are done during the execution of the workflow are tracked down and could be seen in the "*Show action log*".

Outlook

- Endpoints
- Specificities
- Components
- Executing module
- Algorithm of Ecotoxicological workflow
- Predicting Ecotoxicity by using the workflows - examples
- Algorithm of Skin sensitization workflow
- Predicting Skin sensitization by using the workflows - examples

Automated and Standardized workflows in Toolbox

Algorithm of Ecotoxicological workflow



Automated and Standardized workflows in Toolbox

Algorithm of Ecotoxicological workflow

- **Databases**
  - Aquatic OASIS (used in AW)
  - Aquatic ECETOX (used in AW)
  - ECOTOX (used in AW)
  - ECHA CHEM (used in AW)
  - Aquatic Japan*

*Aquatic japan MoE is not used in the AW because there are no data for L(E)C50, 96h, Pimephales promelas, Mortality

Automated and Standardized workflows in Toolbox

Algorithm of Ecotoxicological workflow

- **Profilers used for primary categorization**:
  - US EPA New Chemical Categories
  - Acute aquatic toxicity classification by Verhaar (Modified)
  - Acute aquatic toxicity MOA by OASIS
  - Aquatic toxicity classification by ECOSAR
  - Organic functional groups (OFG)
  - Organic functional groups US-EPA,
  - Organic functional groups, Norbert Haider
- **Options for primary categorization**:
  - Maximum number of analogues with experimental data
  - In case of competing profilers then profilers are ***OR***ed

Automated and Standardized workflows in Toolbox

Algorithm of Ecotoxicological workflow

- **Data Gap Filling**
  - Trend analysis is the default approach
  - Read across is applied if:
    - Prediction by Trend analysis is not acceptable, or
    - The number of analogues is < 10
  - Gap filling and subcategorizations are sequence of logical operations (if, then), combined with criteria for acceptance.

Automated and Standardized workflows in Toolbox

Algorithm of Ecotoxicological workflow

- **Subcategorization**
  - The aim is to increase the similarity of analogues with the target
  - It is consecutive process of application of primary grouping profilers (PGPs) and profilers for subcategorization (PS)
  - Hierarchy of application of PGPs and PS depends on the number of analogues they have collected
  - Sub-categorization process is based on:
    - Sequence of subcategorization steps
    - Criteria for acceptance of subcategorization steps

PGPs – primary grouping profilers
PS – profilers for subcategorization

Automated and Standardized workflows in Toolbox

Algorithm of Ecotoxicological workflow

- **Subcategorization**

Sequence of subcategorization steps

1. *Substance type* – eliminates not discrete chemicals
2. *Water solubility* (WSKOWWIN + WATERNT)
   – eliminates chemicals with *LC50 > WS*
3. Consecutive sub-categorization based on *PGPs*:
   - US EPA,
   - Verhaar,
   - MOA,
   - ECOSAR,
   - OFG (without nested)

4. Consecutive sub-categorization based on *PS*:
   - Protein binding (OASIS + OECD),
   - Chemical elements,
   - Str. Similarity* (remove all constituents with similarity <50%)

*Default options: Dice, Atom centered fragments

Automated and Standardized workflows in Toolbox

Algorithm of Ecotoxicological workflow

- **Subcategorization**

  Criteria for acceptance of subcategorization step:

  - Depends on the specific statistical and structural criteria (e.g, experimental error, 95% of residuals, log K_{OW}, range of variation of the analogues etc.)
  - Criteria are different for RA and TA

Automated and Standardized workflows in Toolbox

Algorithm of Ecotoxicological workflow

Criteria for acceptance of prediction

Trend analysis

After sub-categorization by all PGPs and PS

IF ($R^2 \geq 0.7$ **and** $NA \geq 10$)
THEN accept the prediction and generate report,
ELSE switch to Read across

Read-across

After sub-categorization by all PGPs and PS

IF *Interpolation* **AND** *LC50* ≤ 2 log units **OR** log *Kow* ≤ 2 log units **AND** *NA* ≥ 5
THEN accept prediction and proceed with Report

In case, criteria are not met, then no prediction is obtained

Definitions

R^2 – *Correlation coefficient*
Interpolation: log K_{OW} of the target should be within the range of log K_{OW} of analogues
LC50 ≤ 2*:* for the 5 closest analogues the range of variation of *LC50* is ≤ 2 log units
log K_{OW} ≤ 2*:* for the 5 closest analogues the range of variation of log K_{OW} is ≤ 2 log units
NA – Number of analogues

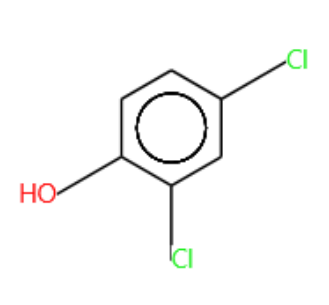
Outlook

- Endpoints
- Specificities
- Components
- Executing module
- Algorithm of Ecotoxicological workflow
- Predicting Ecotoxicity by using the workflows - examples
- Algorithm of Skin sensitization workflow
- Predicting Skin sensitization by using the workflows - examples

Automated and Standardized workflows in Toolbox

SW for Ecotoxicological endpoint - example

Target chemical: CAS 120-83-2

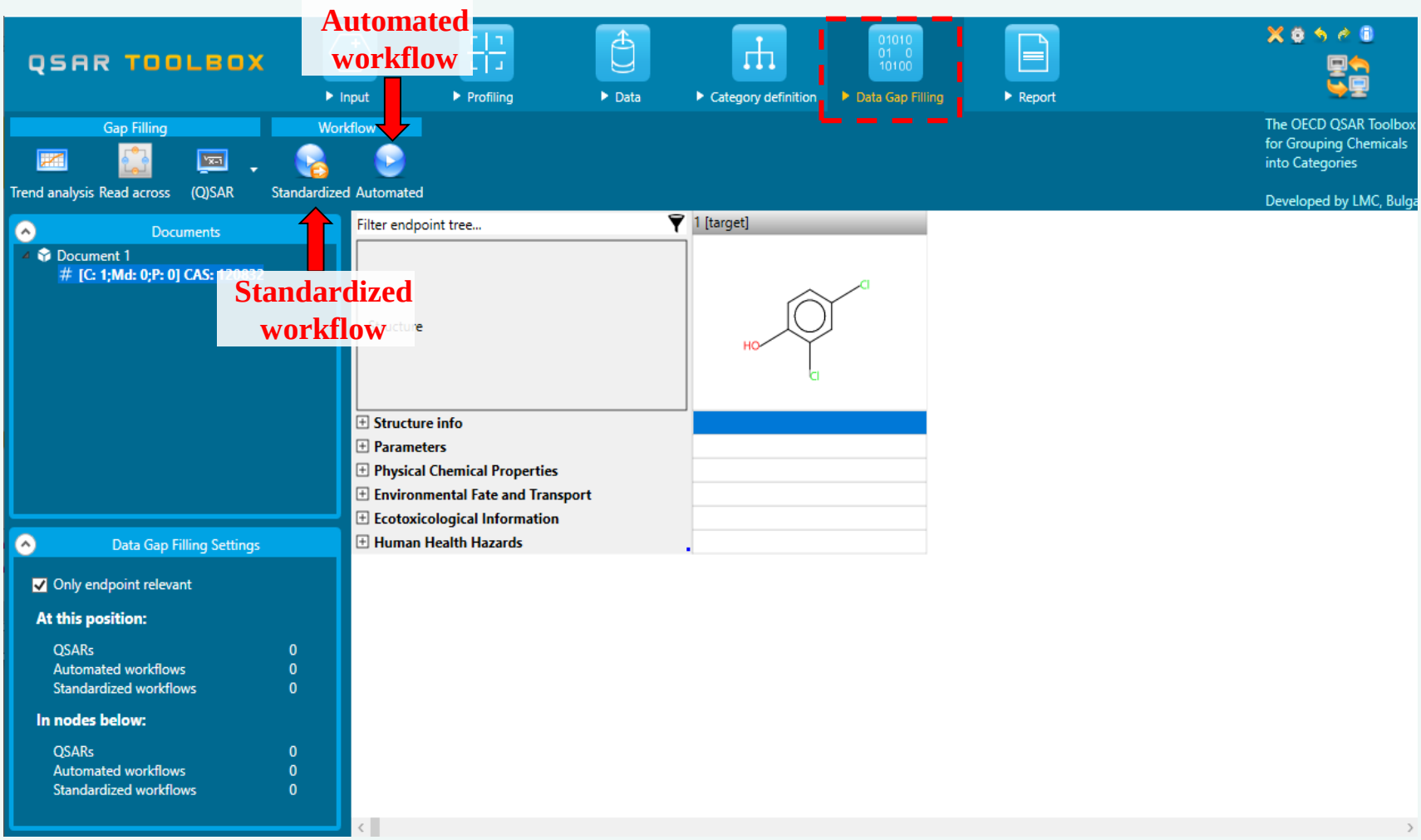


Endpoint: LC50 or EC50
Effect: Mortality
Species: Actinopergyii (all fish) – illustrated in SW
Species: Pimephales promelas – illustrated in AW
Duration: 96h

Automated and Standardized workflows in Toolbox

SW for Ecotoxicological endpoint - example

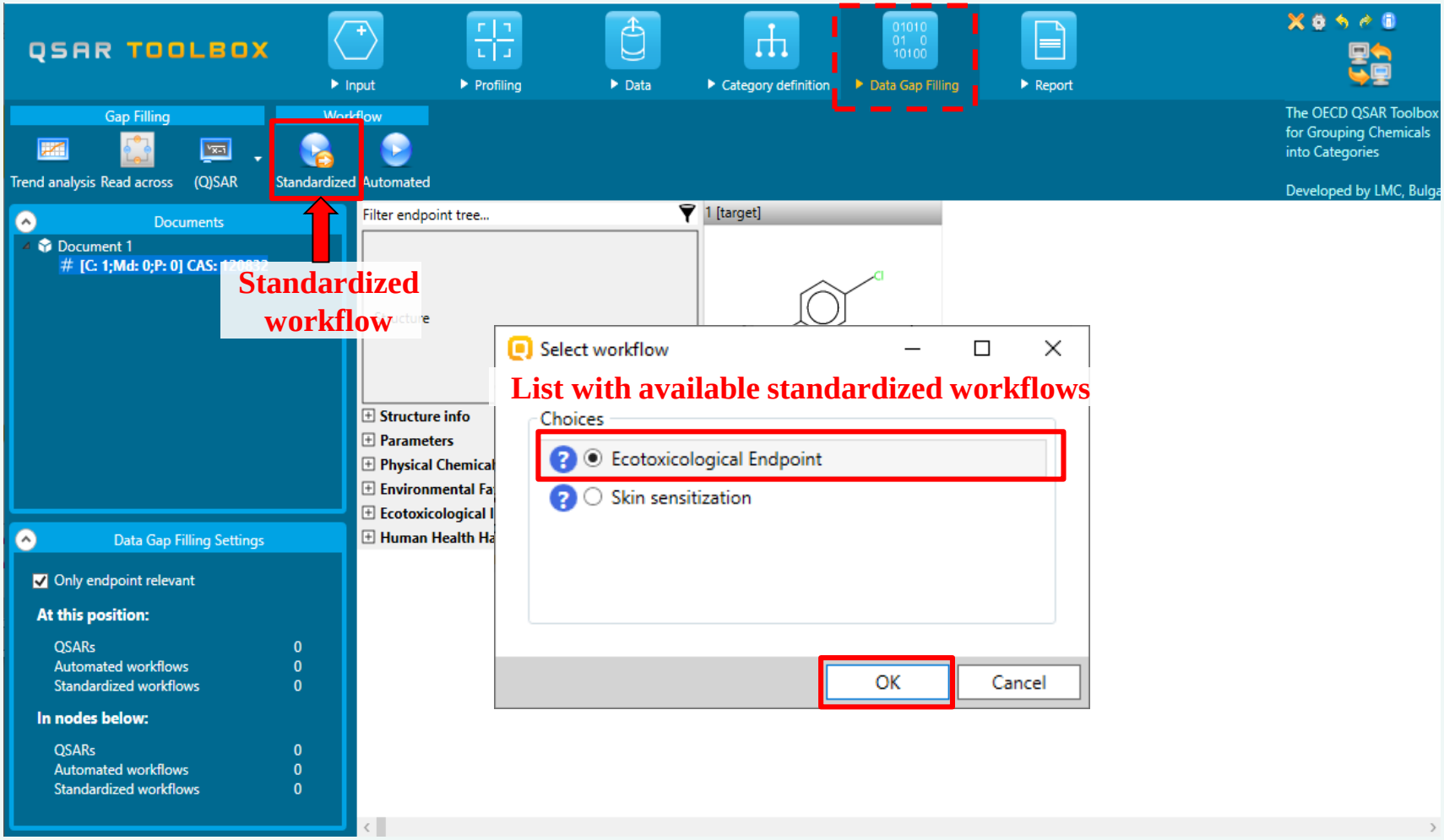
Location: The AWs and SWs are part of the DGF



Automated and Standardized workflows in Toolbox

SW for Ecotoxicological endpoint - example

Activation: The Standardized workflow is activated by clicking on the corresponding button

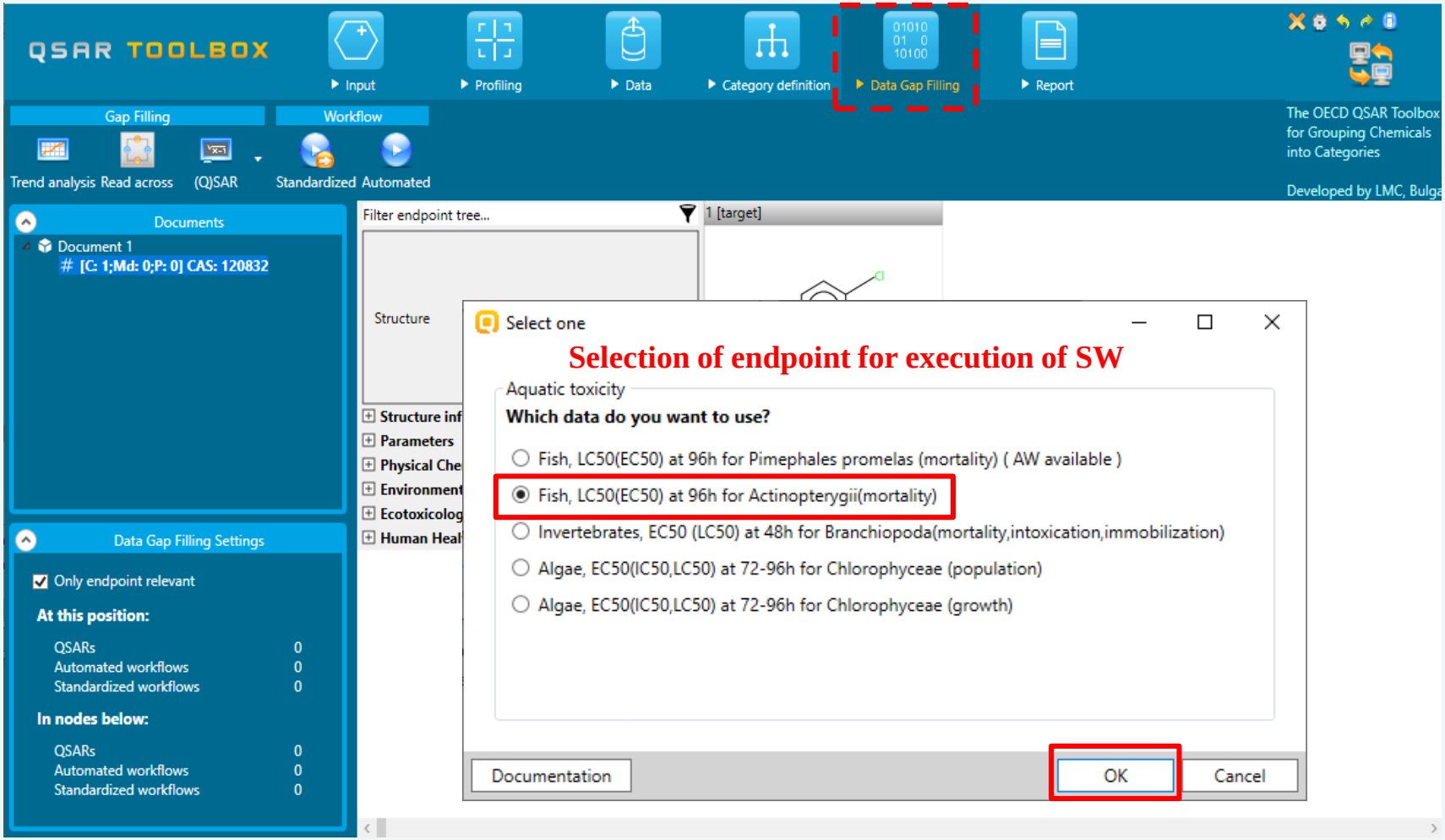


Automated and Standardized workflows in Toolbox

SW for Ecotoxicological endpoint - example

Selection of endpoint: The SW for Ecotoxicological endpoint needs the endpoint to be confirmed. There are 5 options for endpoint selection. In our case study we apply SW for:

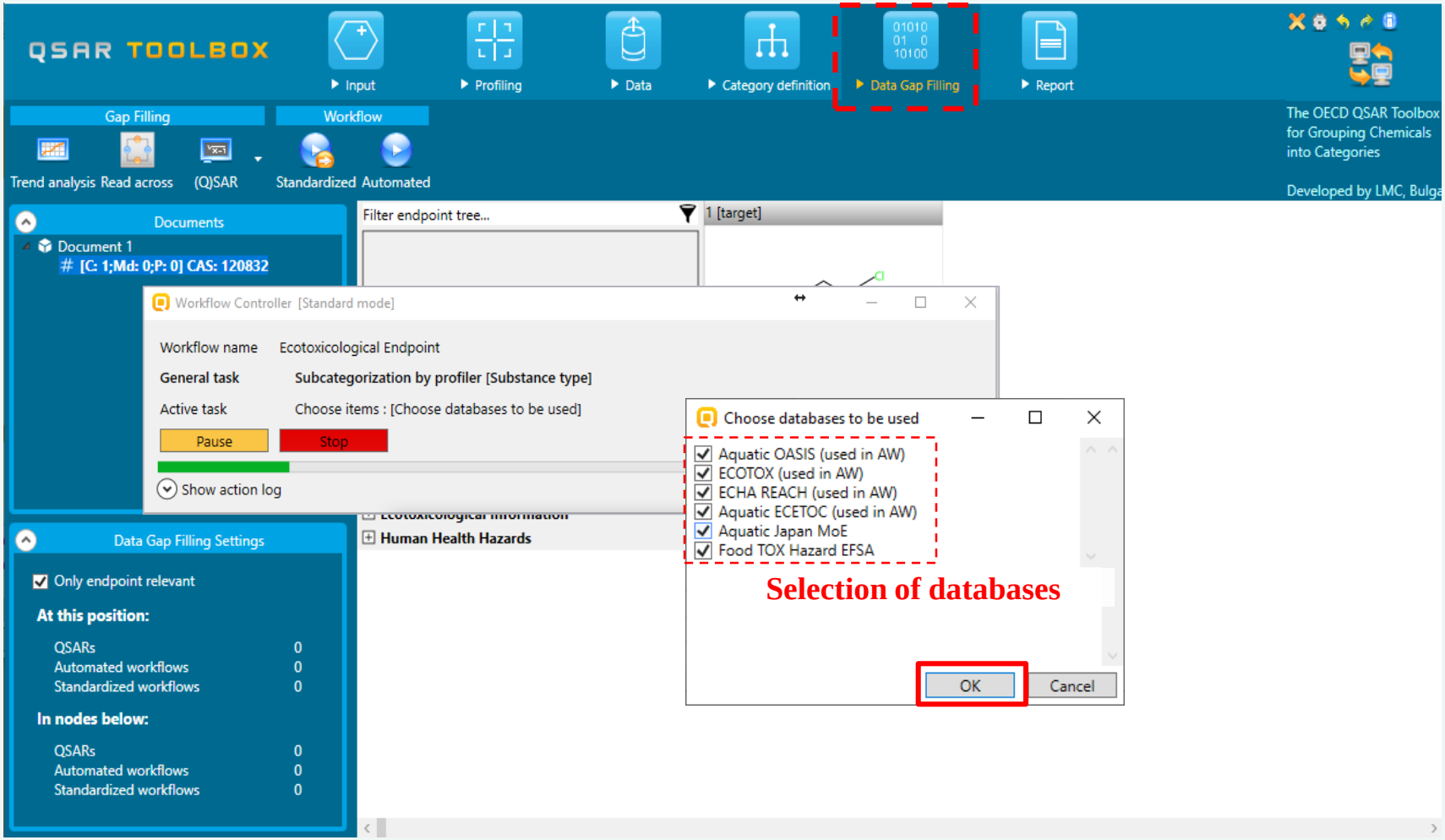
Endpoint: EC50 or LC50 for Fish, duration: 96h



Automated and Standardized workflows in Toolbox

SW for Ecotoxicological endpoint - example

Selection of databases: Six databases are available. All of them are selected in this example

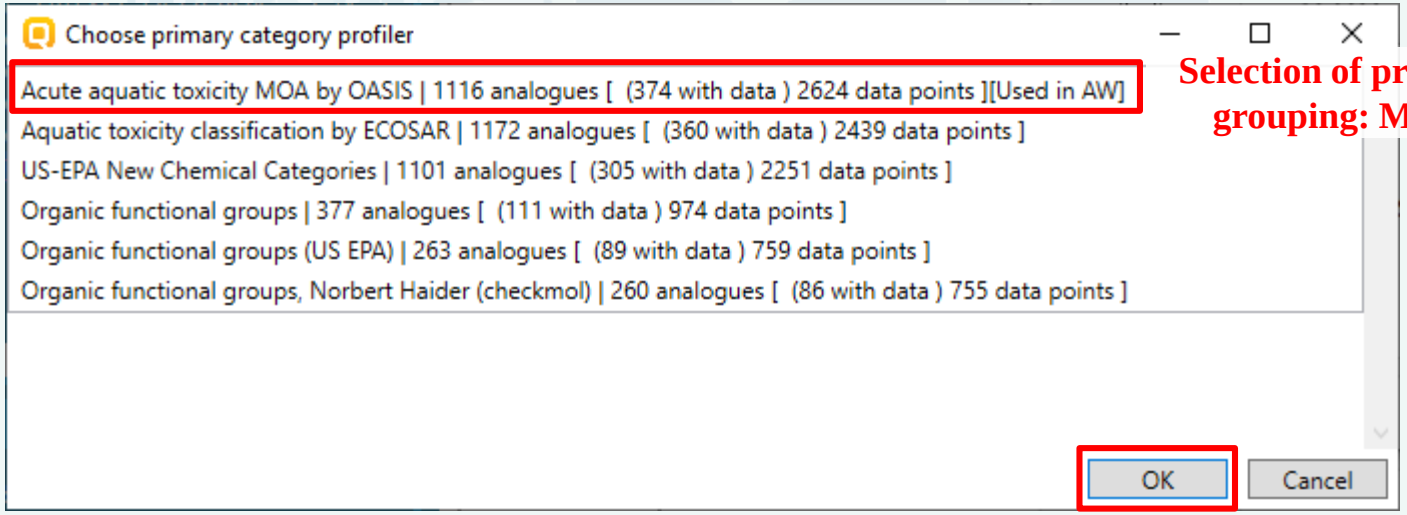


Automated and Standardized workflows in Toolbox

SW for Ecotoxicological endpoint - example

Category definition:

- Categories are hierarchically order depending on the number of analogues with data
- The most populated category is used in the AW
- In SW the user is able to make different selection
- In this example we apply Acute aquatic toxicity MOA

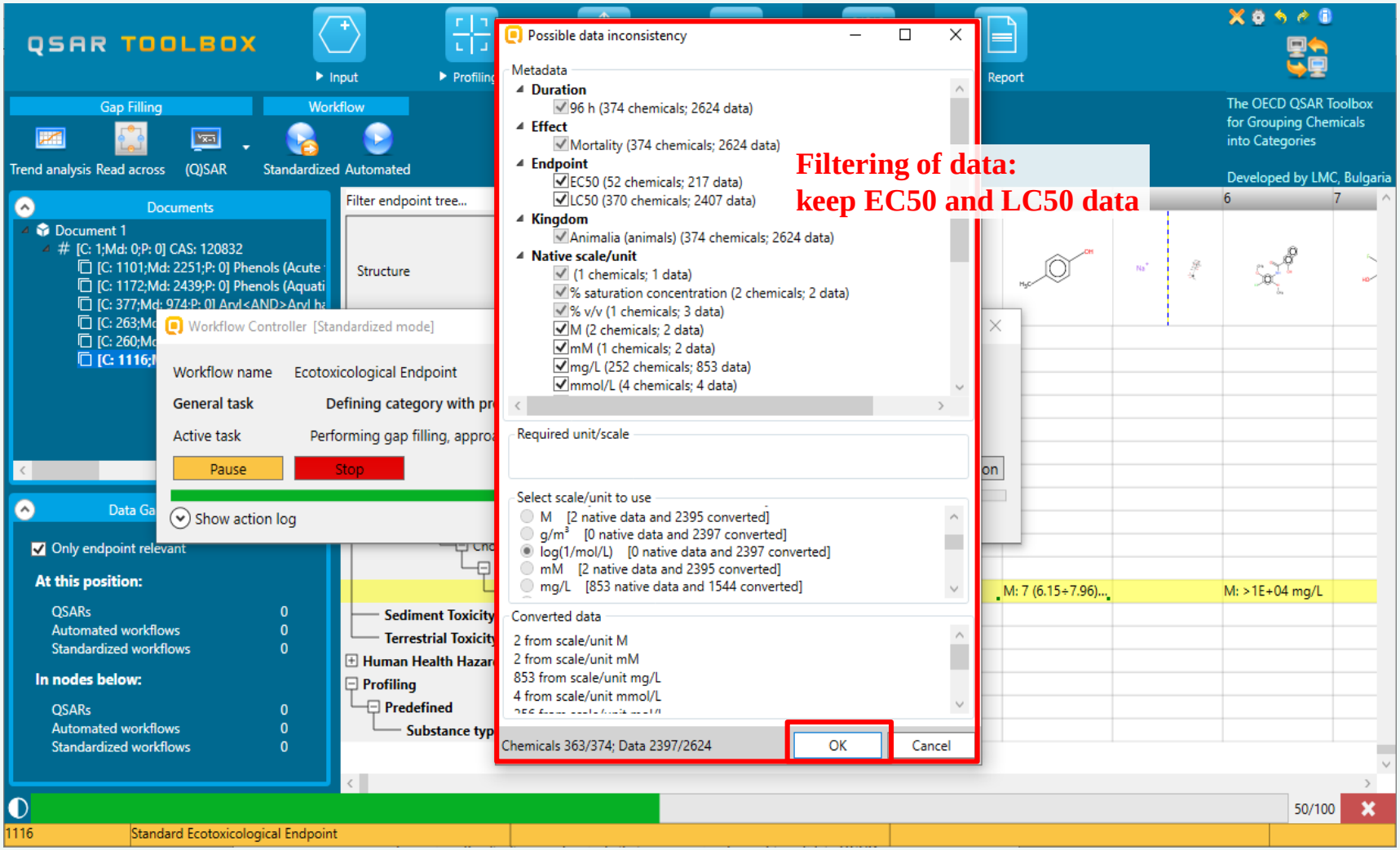


Selection of primary grouping: MOA

Automated and Standardized workflows in Toolbox

SW for Ecotoxicological endpoint - example

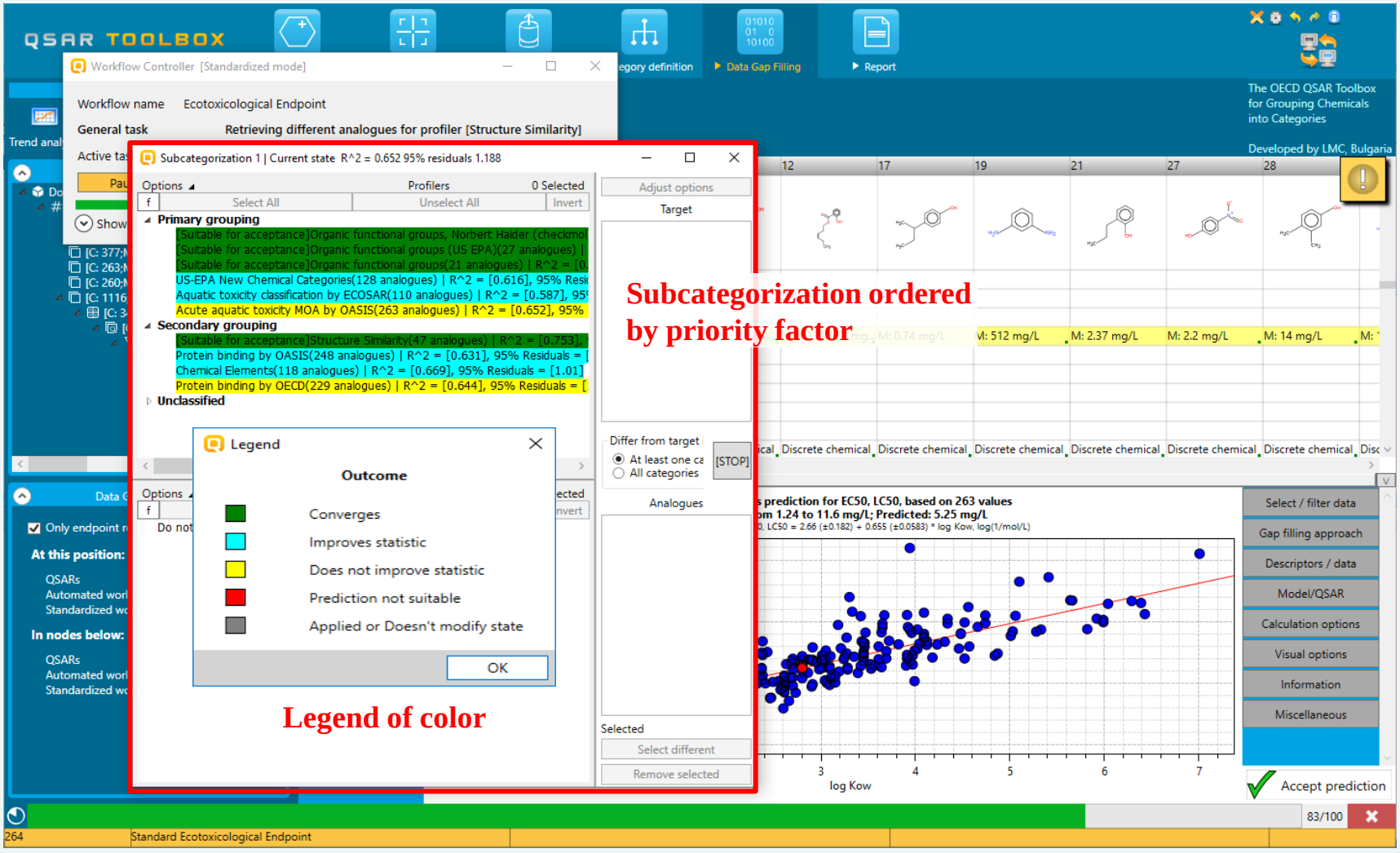
Data Gap Filling: Before entering to DGF, the workflow allows filtering of duration; endpoint and unit



Automated and Standardized workflows in Toolbox

SW for Ecotoxicological endpoint - example

Data Gap Filling: Subcategorization dialogue provides the list with relevant profilers highlighted in different color and order by priority factor



Automated and Standardized workflows in Toolbox

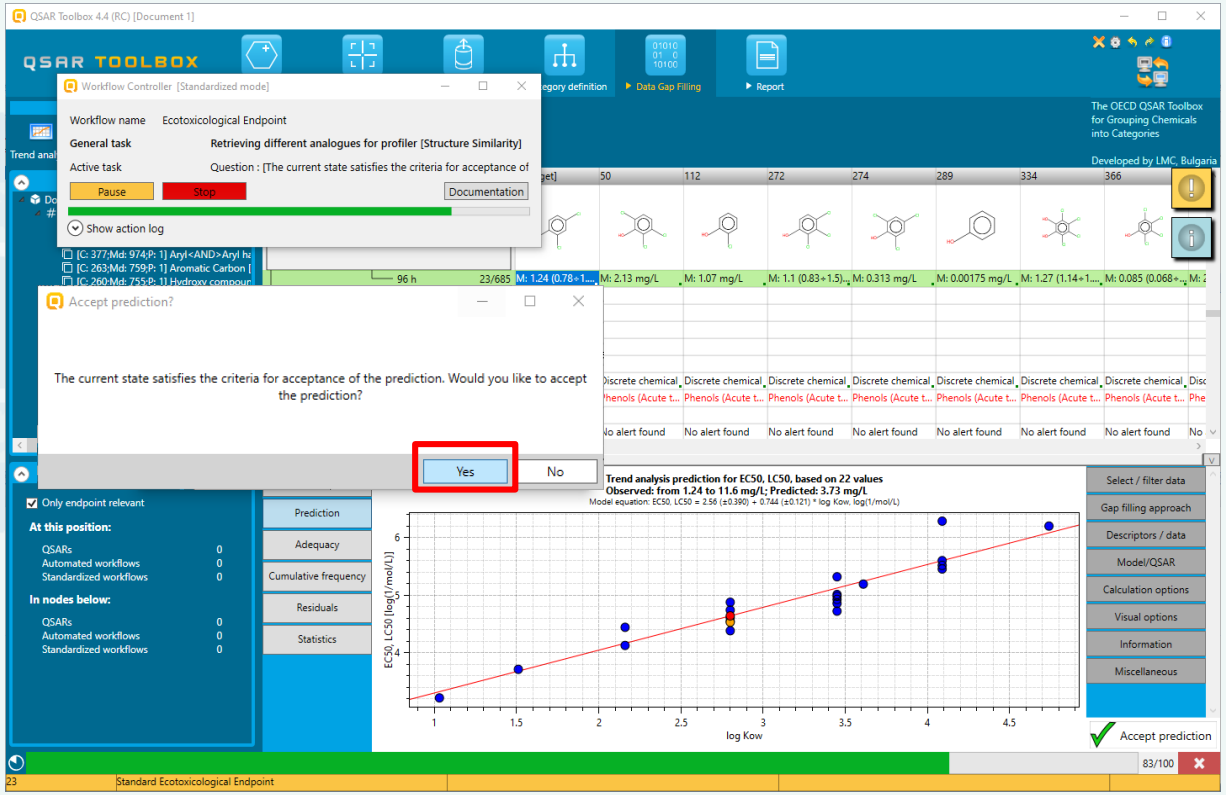
SW for Ecotoxicological endpoint - example

Subcategorization: The applied subcategorizations are as follows:

- ✓ Substance type (mandatory – automatically applied)
- ✓ Filter by WS (mandatory – automatically applied)
- ✓ OFG (USEPA)
- ✓ US-EPA Chemical categories

Note: After each subcategorization the workflow controller wait for activity by the user – "Continue" button needs to be clicked.

Once the prediction satisfy the acceptance criteria a message appears.



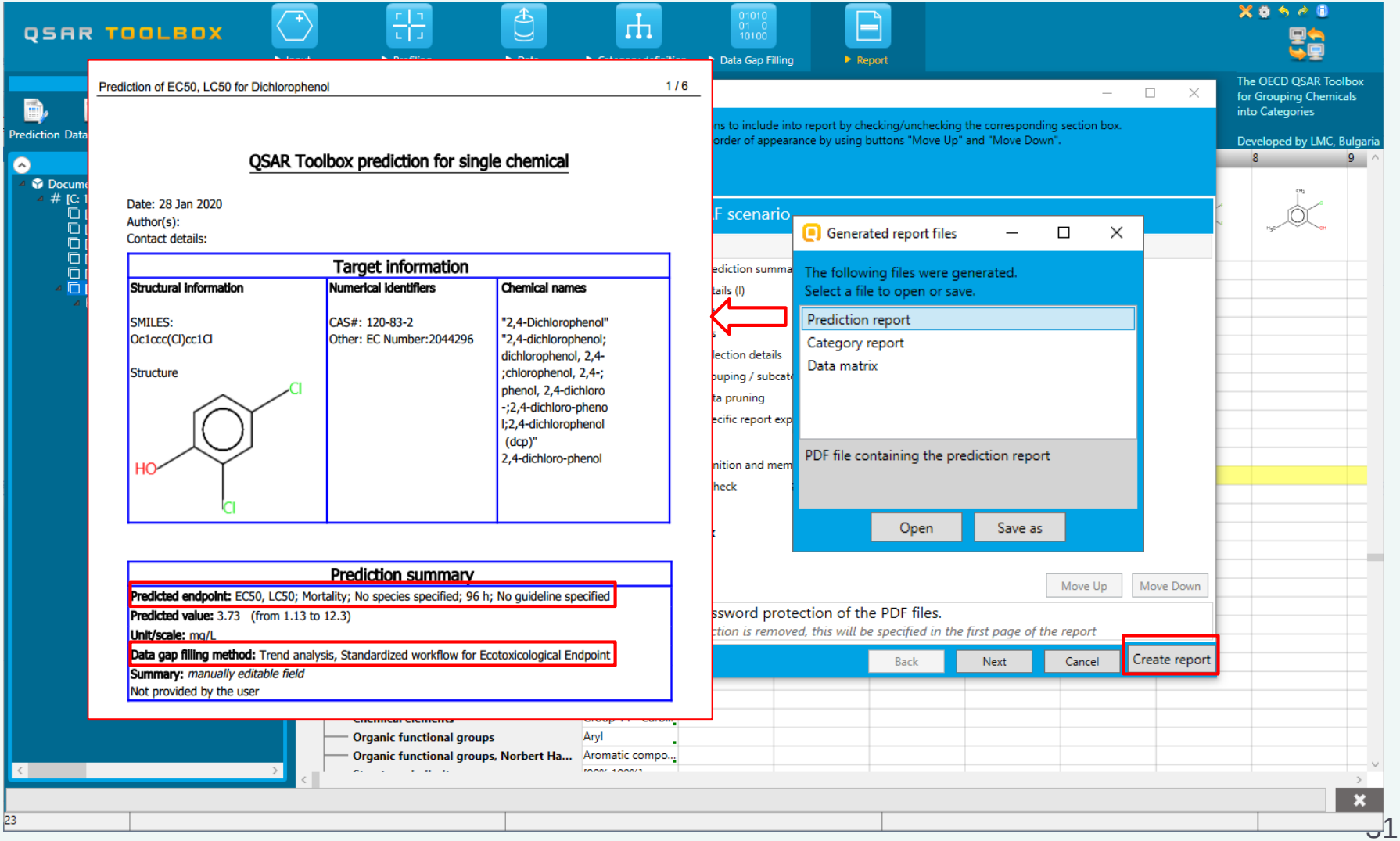
Predicted LC50(EC50) for all fish of 2,4-dichlorophenol is 3.73 mg/L (toxic) after application of SW for Ecotox

Automated and Standardized workflows in Toolbox

SW for Ecotoxicological endpoint - example

Report:

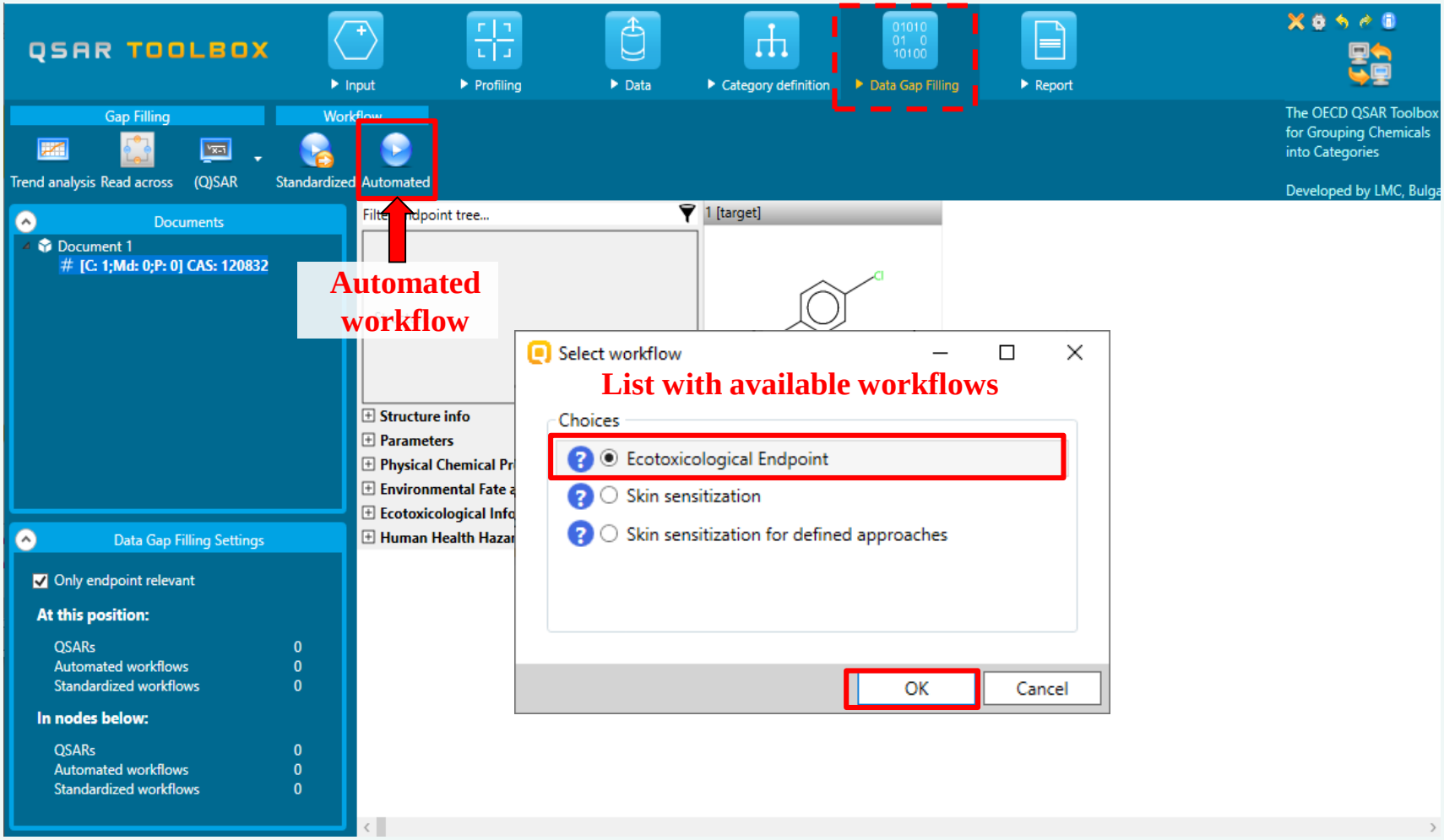
- The report section is not part of the workflows
- Report could be generated once the prediction is accepted



Automated and Standardized workflows in Toolbox

AW for Ecotoxicological endpoint - example

Activation: The Automated workflow is activated by clicking on the corresponding button

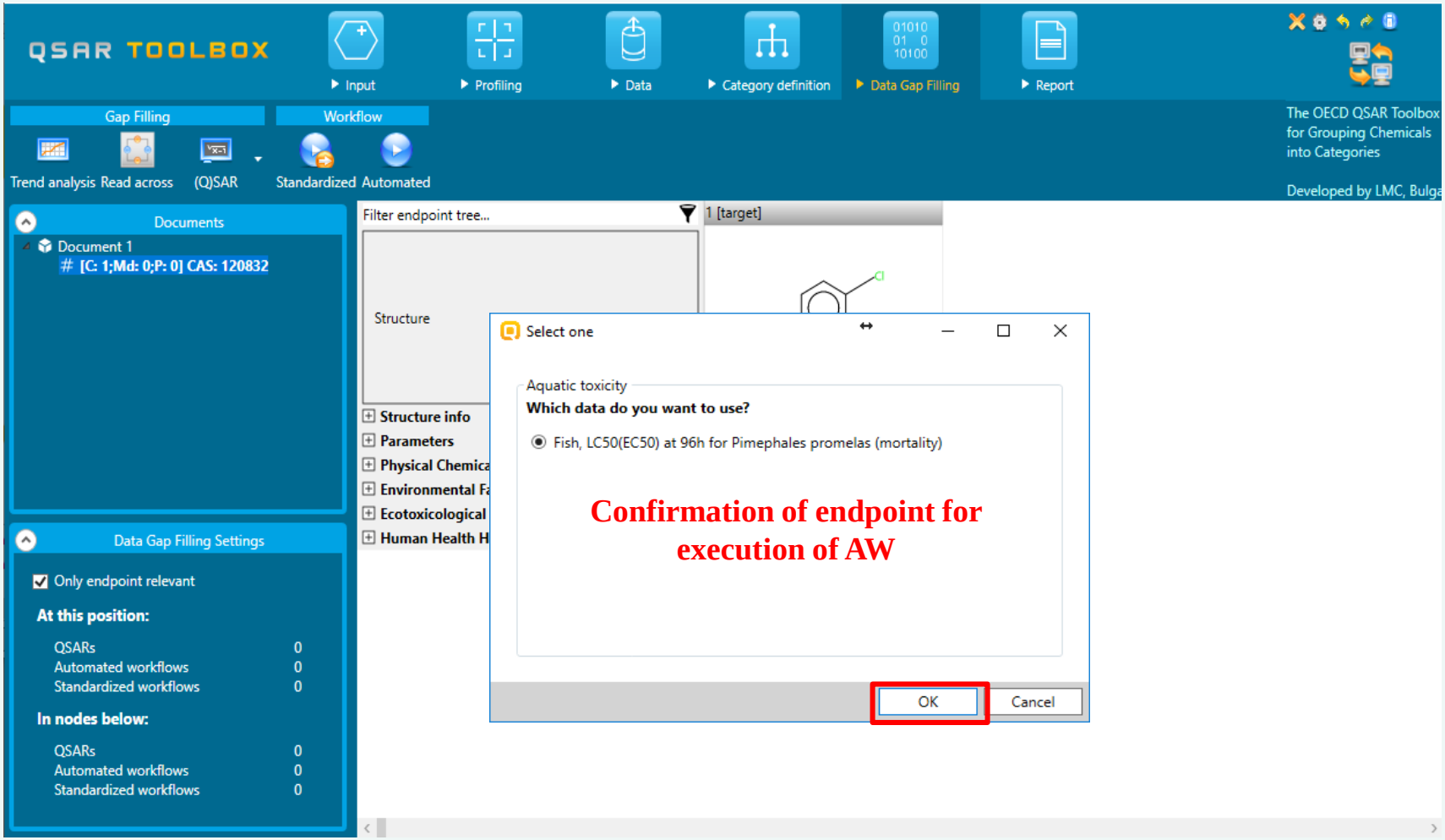


Automated and Standardized workflows in Toolbox

AW for Ecotoxicological endpoint - example

Selection of endpoint: The AW for Ecotoxicological endpoint needs the endpoint to be confirmed.

Endpoint: EC50 or LC50 for Pimephales promelas, duration: 96h

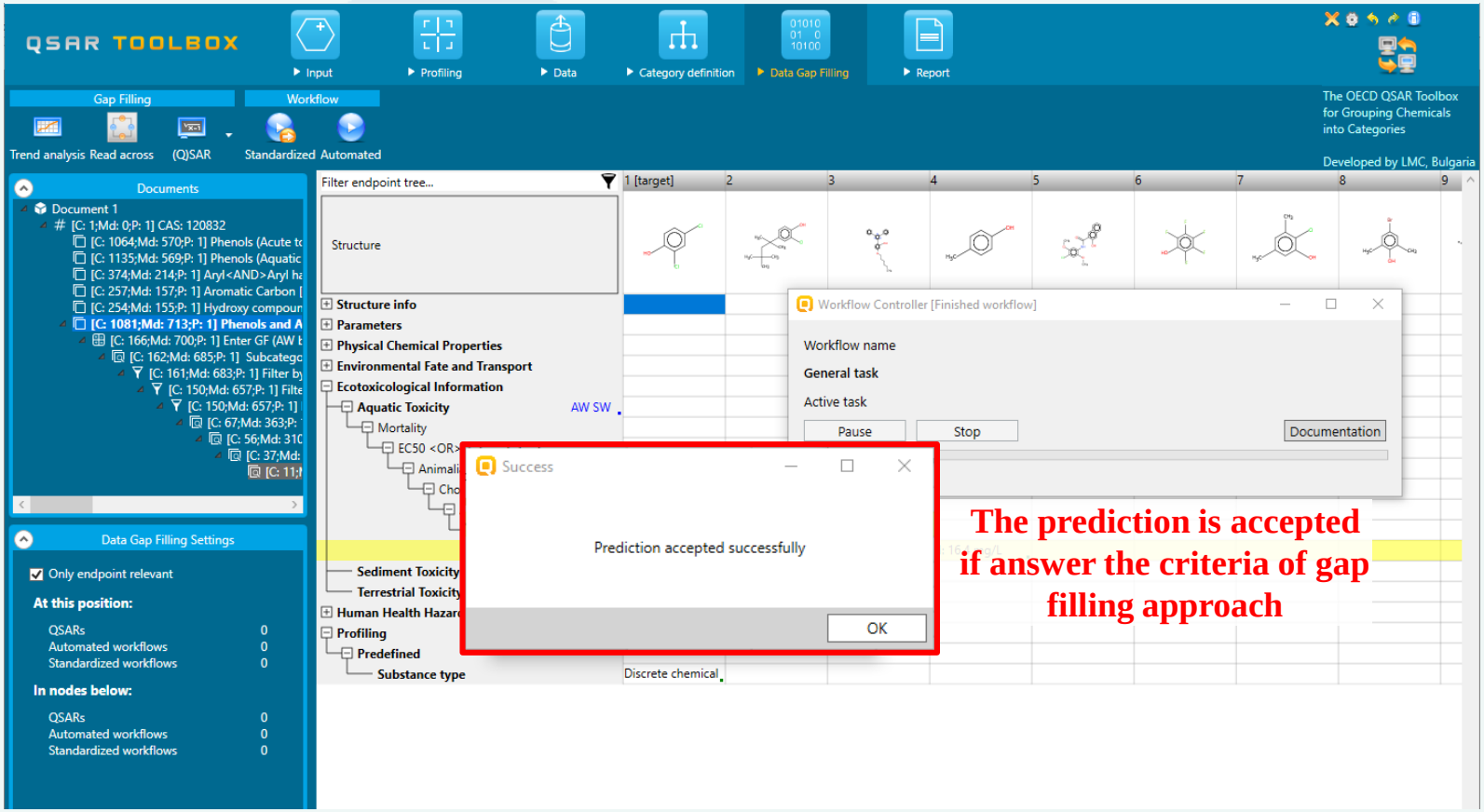


Automated and Standardized workflows in Toolbox

AW for Ecotoxicological endpoint - example

The following steps are applied automatically:

- Selection of databases
- Selection of primary group
- Sequence of subcategorizations
- Acceptance of the prediction



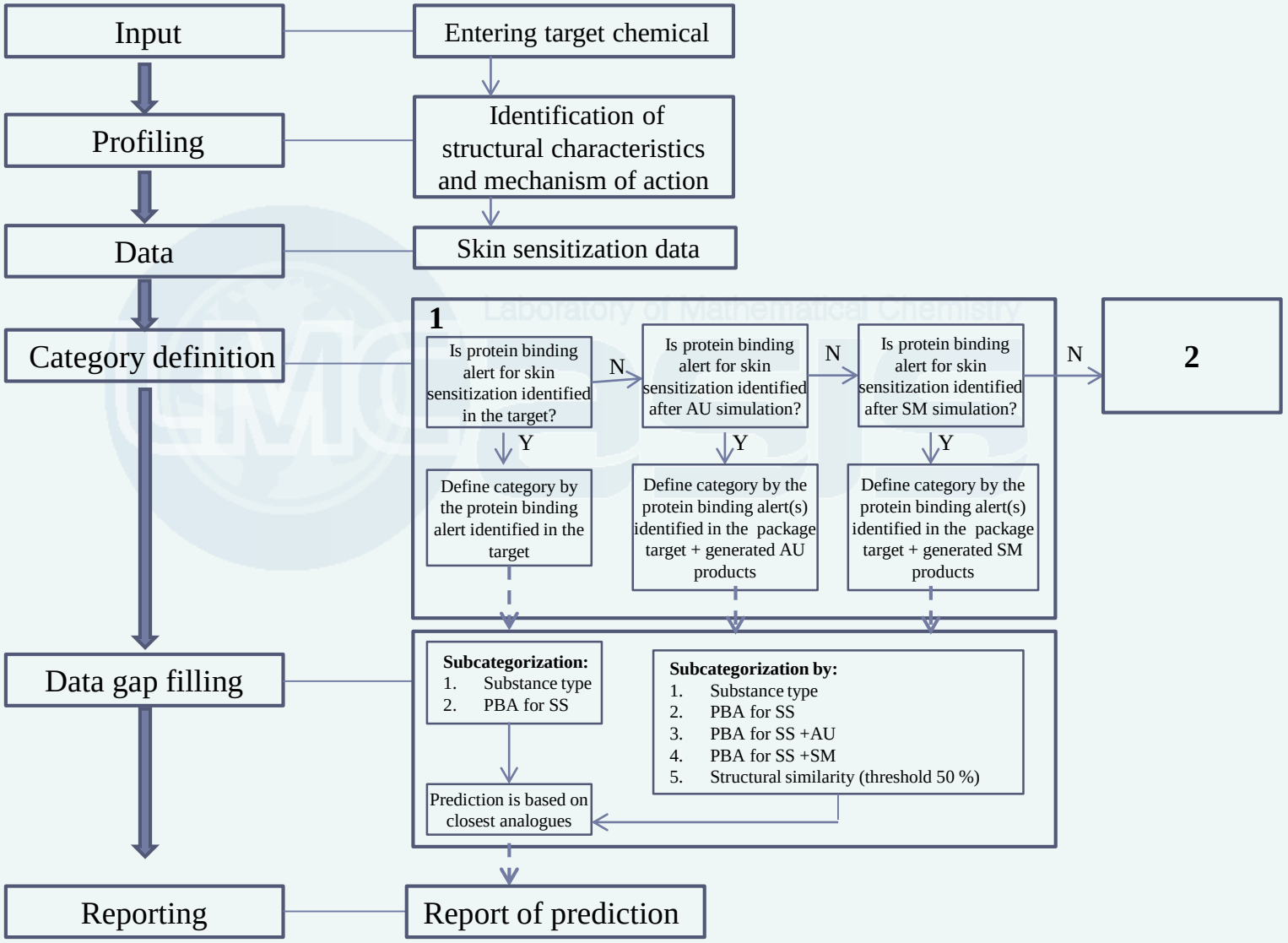
Outlook

- Endpoints
- Specificities
- Components
- Executing module
- Algorithm of Ecotoxicological workflow
- Predicting Ecotoxicity by using the workflows - examples
- Algorithm of Skin sensitization workflow
- Predicting Skin sensitization by using the workflows - examples

Automated and Standardized workflows in Toolbox

Algorithm of Skin sensitization workflow

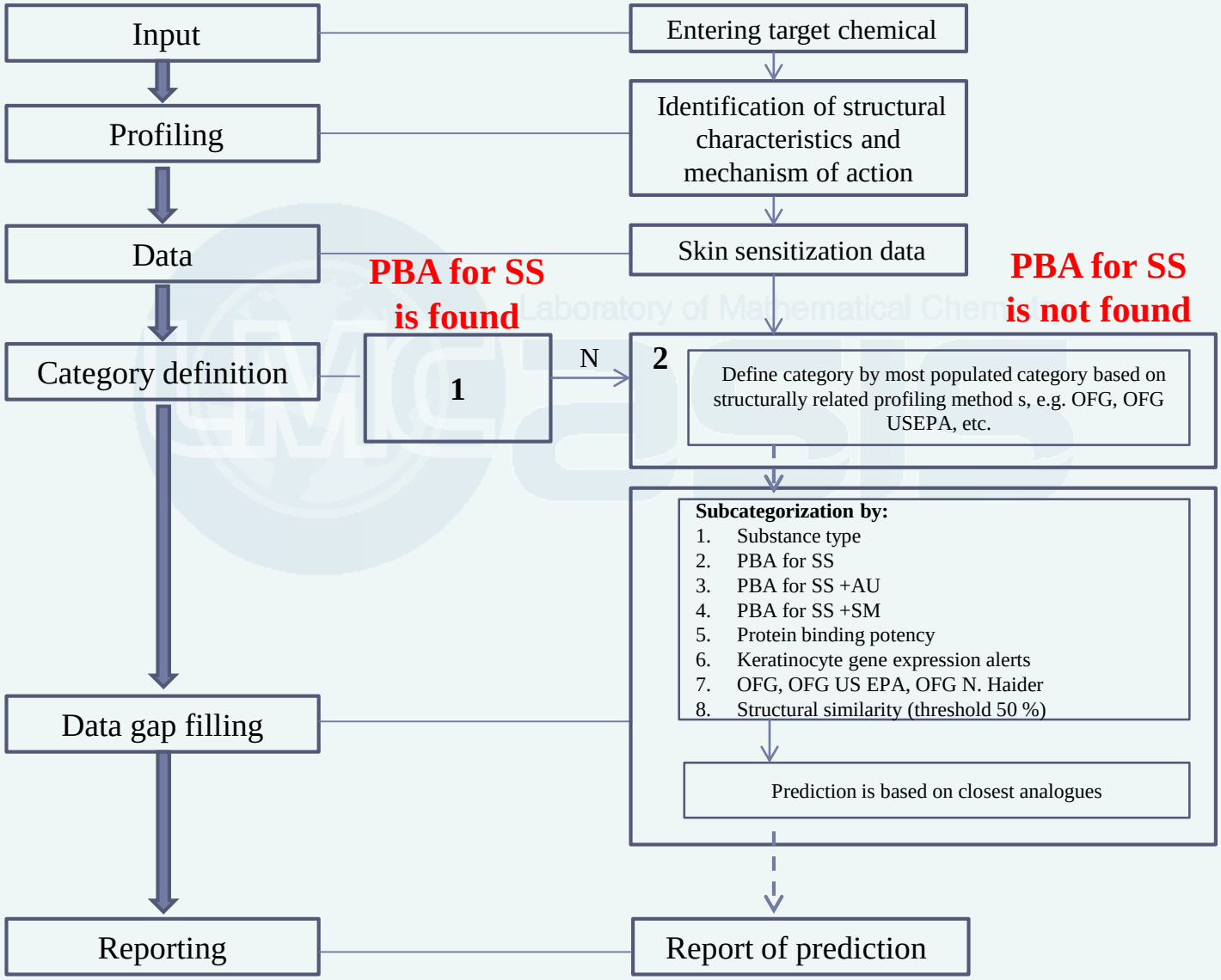
Part 1



Automated and Standardized workflows in Toolbox

Algorithm of Skin sensitization workflow

Part 2



Automated and Standardized workflows in Toolbox

Algorithm of Skin sensitization workflow

- **Databases**
  - Skin sensitization
  - Skin sensitization ECETOC
  - REACH Skin sensitisation database (normalized)
- SS predictions are based on LLNA and GPMT exp. data.
  - ✓ LLNA – EC3, %
  - ✓ GPMT – Positive; Negative
  - ✓ LLNA and GPMT – Positive; Negative

Automated and Standardized workflows in Toolbox

Algorithm of Skin sensitization workflow

- **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 Norbert Haider

- **Abiotic and biotic activation of chemical is accounted for by application of respective Autoxidation (AU) and Skin metabolism (SM) simulators**

Automated and Standardized workflows in Toolbox

Algorithm of Skin sensitization workflow

- **Category definition**
  - If protein binding alert for skin sensitization (**PBA for SS)** is identified in the target structure then the primary category is based on this alert
  - If **PBA for SS** is identified after AU or SM simulation then the primary category is defined accounting the metabolic simulation
  - If **more than one PBA for SS** 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 for SS** is identified in the parent structure and in the generated metabolites, then the primary category is defined on global molecular features by using:
    - Organic *Functional Groups* (OFG) **OR**
    - *OFG USEPA* **OR**
    - *OFG N. Haider* **OR**
    - *Acute aquatic classification by ECOSAR* **OR**
    - *US-EPA New Chemical categories* **OR**

  In this case, the primary category will be formed based on the profiler leading to **the largest group** of analogues.

Automated and Standardized workflows in Toolbox

Algorithm of Skin sensitization workflow

- **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)

If no analogues with data are found, then no prediction is obtained after application of the workflows

Outlook

- Endpoints
- Specificities
- Components
- Executing module
- Algorithm of Ecotoxicological workflow
- Predicting Ecotoxicity by using the workflows - examples
- Algorithm of Skin sensitization workflow
- Predicting Skin sensitization by using the workflows - examples

Automated and Standardized workflows in Toolbox

SW for Skin sensitization endpoint - example

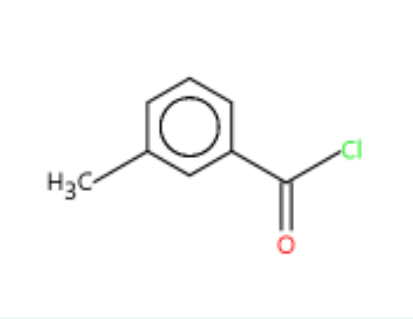
Example for the Standardized workflow:

Endpoint: EC3
Type of method: in vivo
Assay: LLNA
Organ: Skin

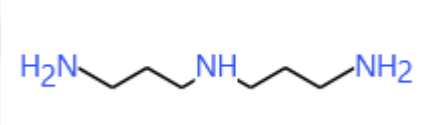
Example for the Automated workflow:

Endpoint: EC3 and Skin sensitization
Type of method: in vivo
Assay: LLNA and GPMT
Organ: Skin

Target chemical: CAS 1711-06-4



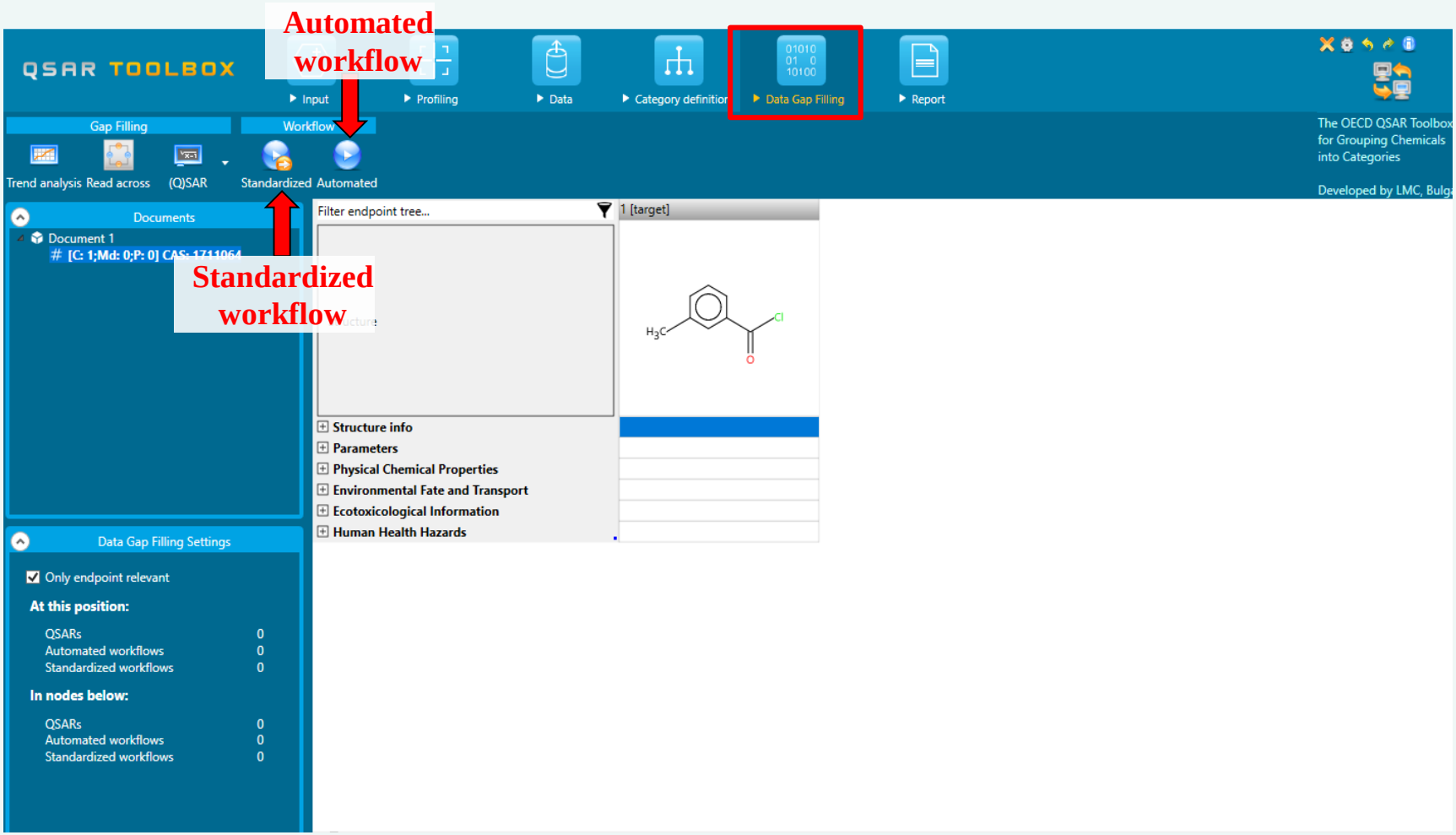
Target chemical: CAS 56-18-8



Automated and Standardized workflows in Toolbox

SW for Skin sensitization - examples

Location: The AWs and SWs are part of the DGF

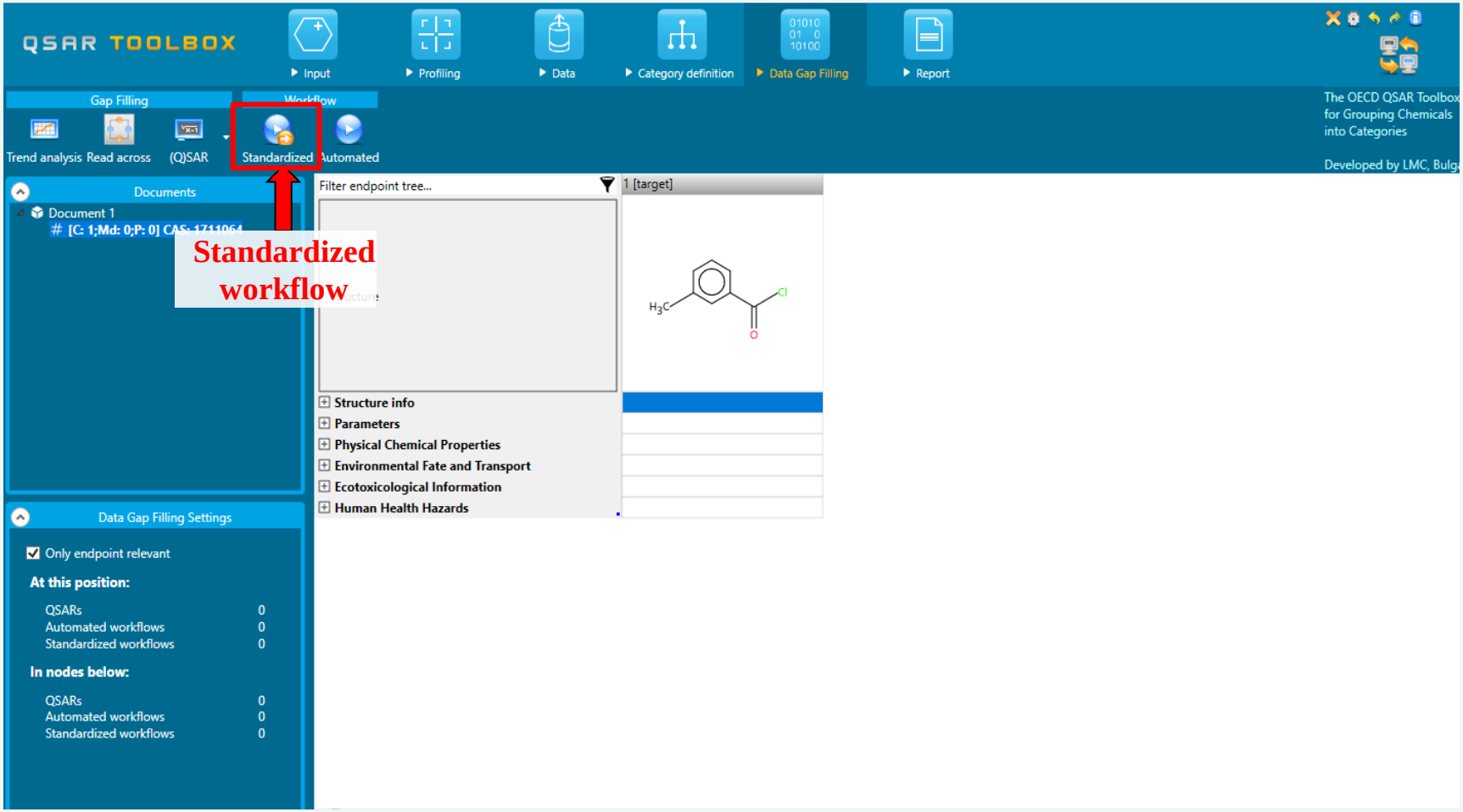


Automated and Standardized workflows in Toolbox

SW for Skin sensitization - examples

Activation: The SW is activated by clicking on the corresponding button

Example 1: Target chemical has protein binding alert

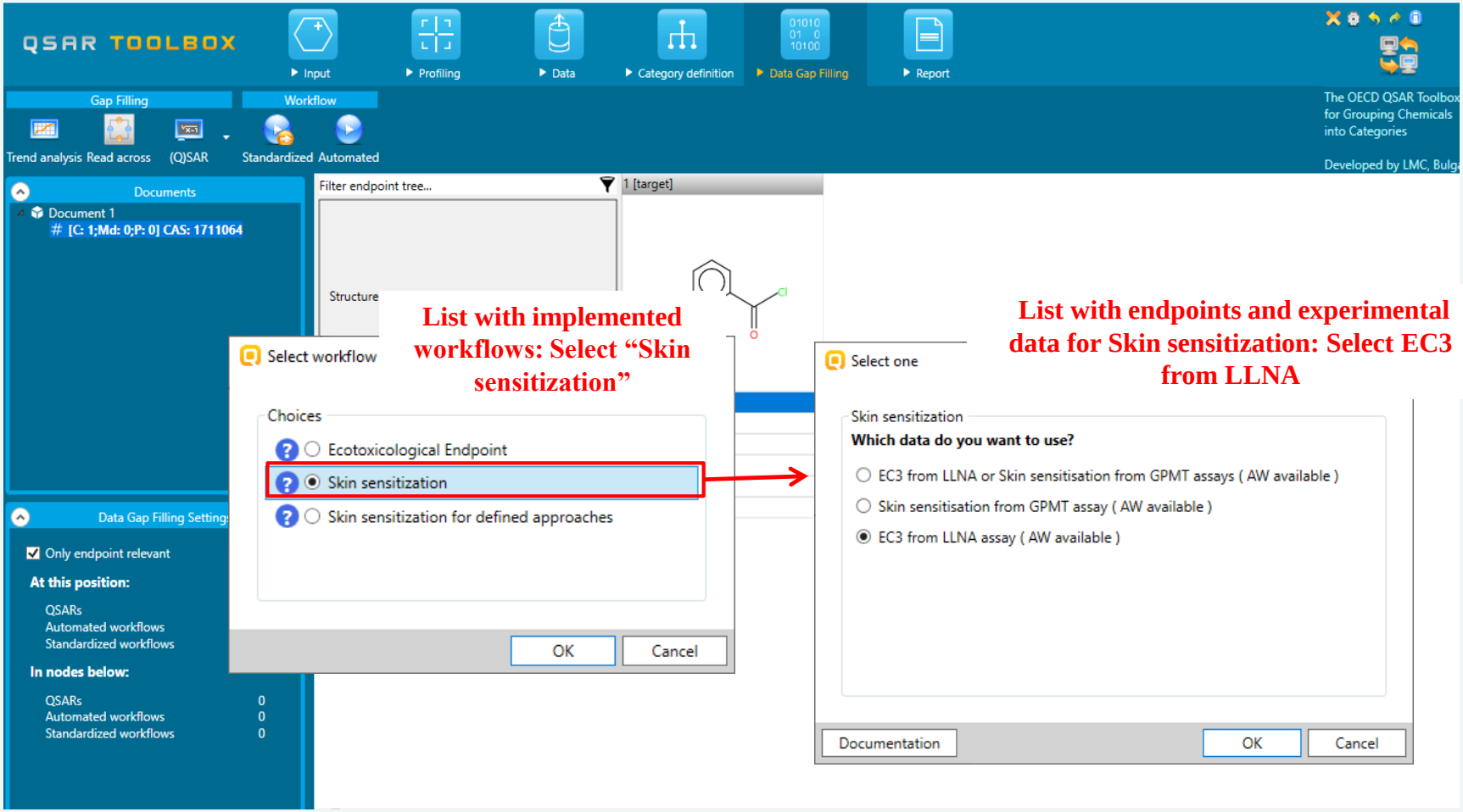


Automated and Standardized workflows in Toolbox

SW for Skin sensitization - examples

Selection of endpoint: Three options for data usage are provided: EC3, GPMT and EC3/GPMT

Example 1: Target chemical has protein binding alert

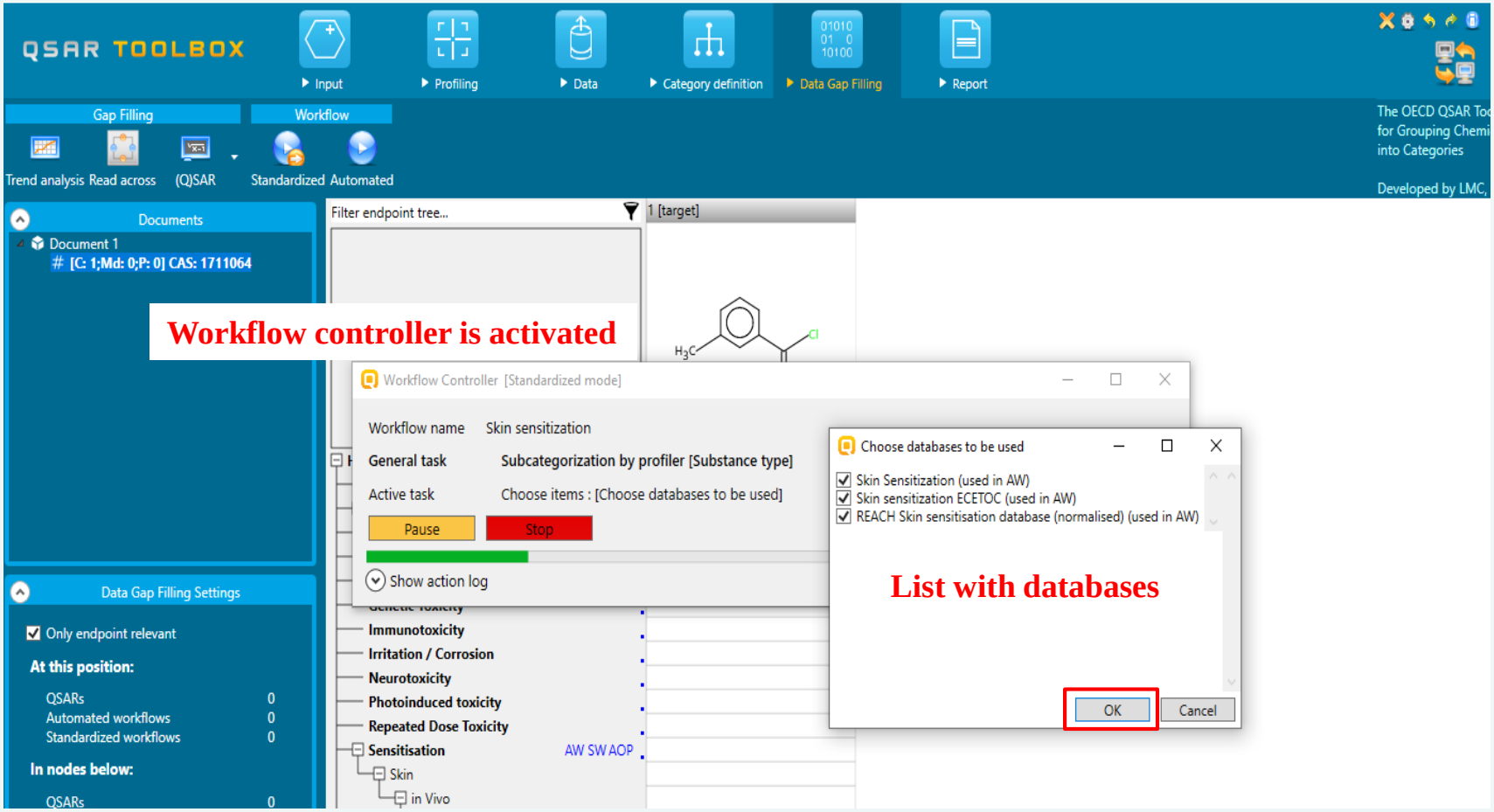


Automated and Standardized workflows in Toolbox

SW for Skin sensitization - examples

Selection of databases: Three databases are available

Example 1: Target chemical has protein binding alert

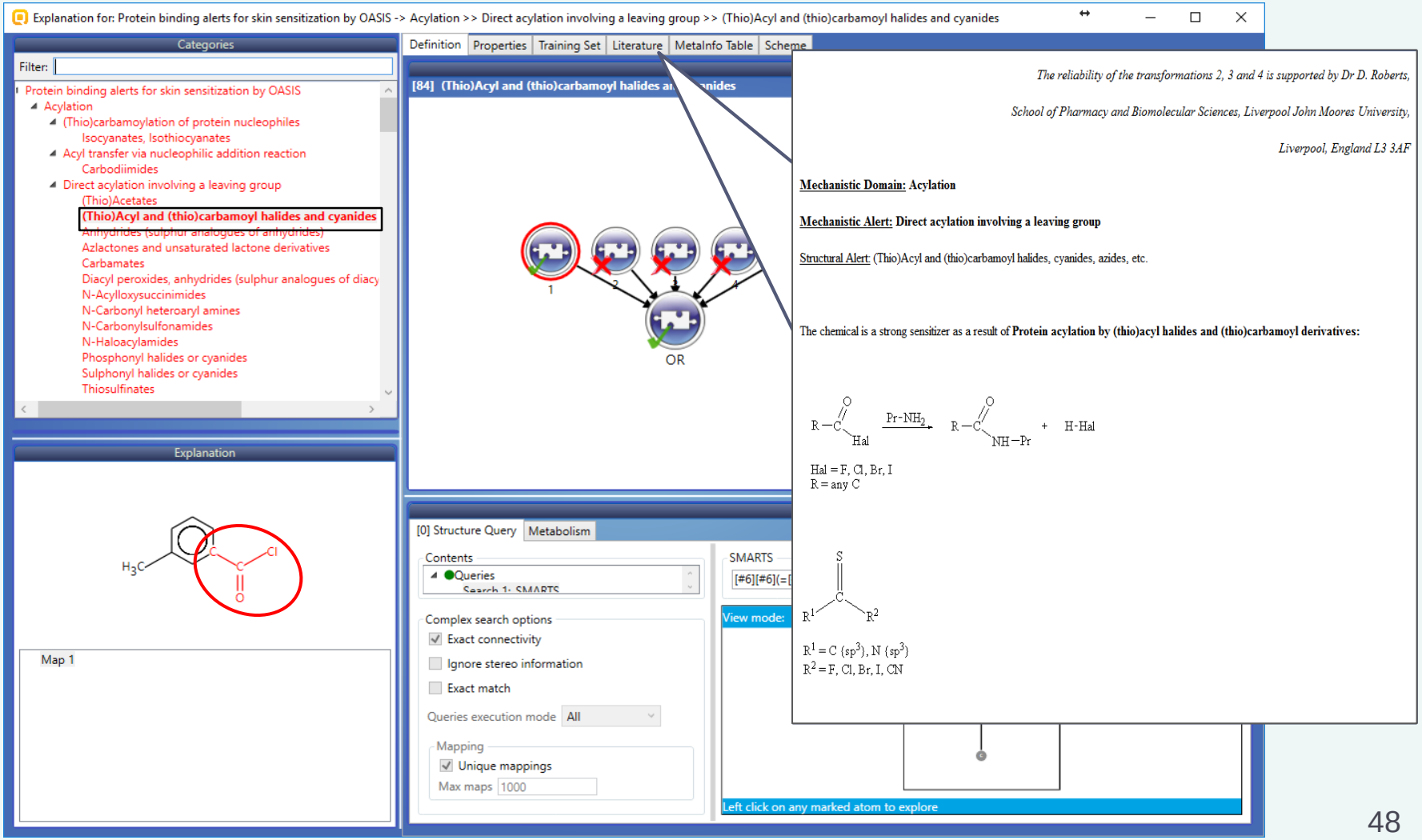


Automated and Standardized workflows in Toolbox

SW for Skin sensitization - examples

Category definition: Primary group is defined based on the alert found in the target

Example 1: Target chemical has protein binding alert

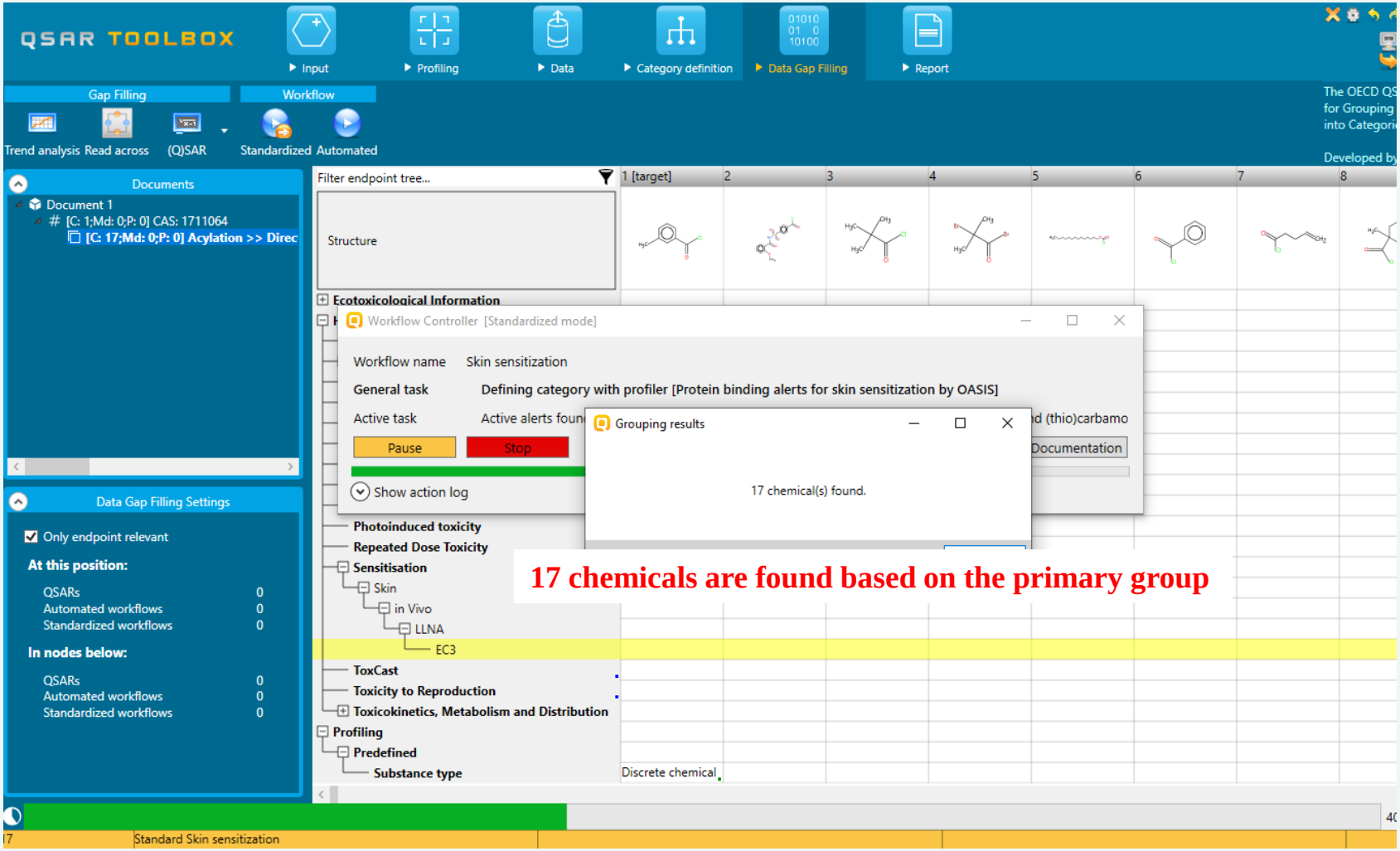


Automated and Standardized workflows in Toolbox

SW for Skin sensitization - examples

Category definition: Primary group is defined based on the alert found

Example 1: Target chemical has protein binding alert

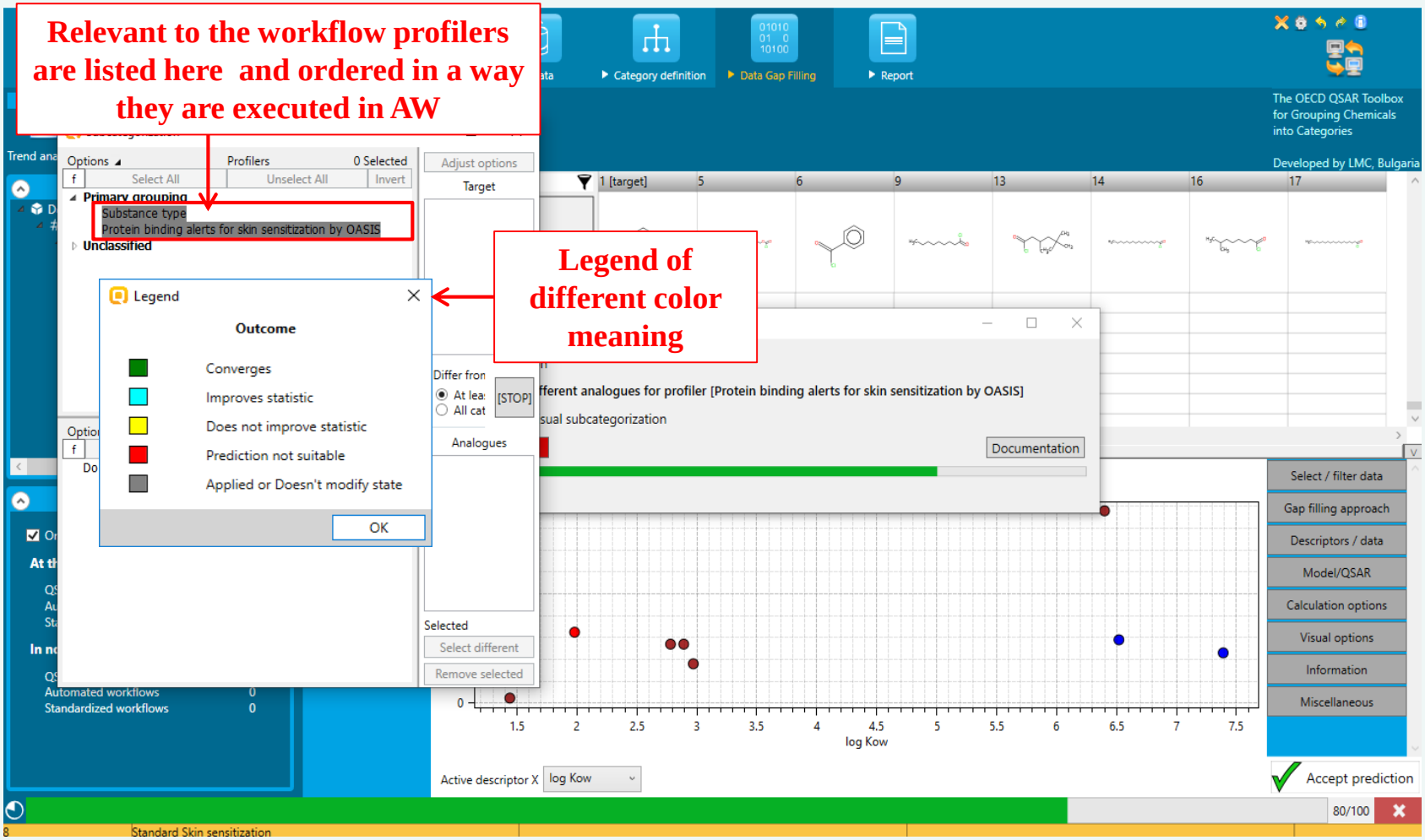


Automated and Standardized workflows in Toolbox

SW for Skin sensitization - examples

Data Gap Filling: Relevant to the workflow profilers are provided and colored

Example 1: Target chemical has protein binding alert

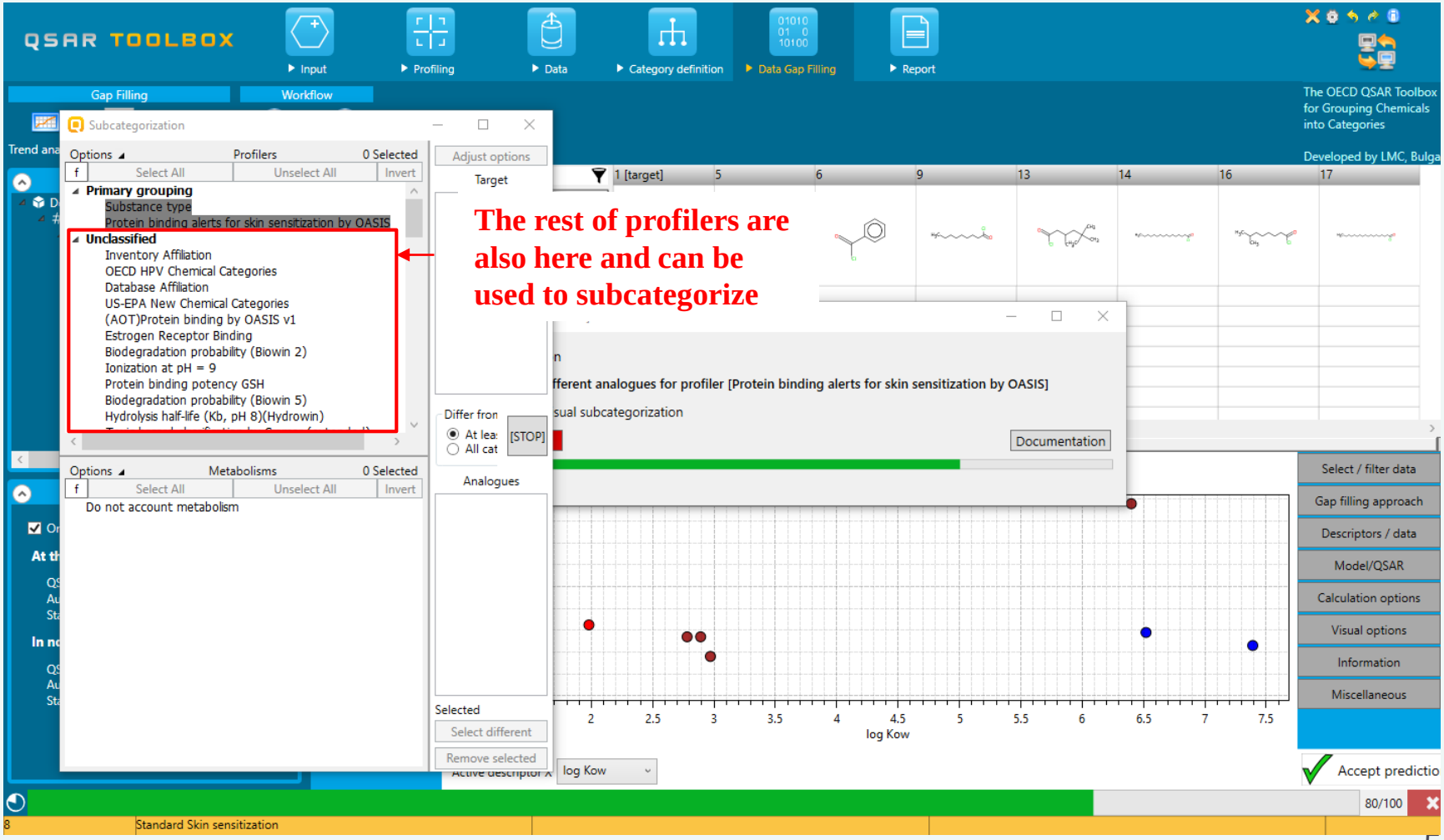


Automated and Standardized workflows in Toolbox

SW for Skin sensitization - examples

Data Gap Filling: Relevant to the workflow profilers are provided and colored

Example 1: Target chemical has protein binding alert

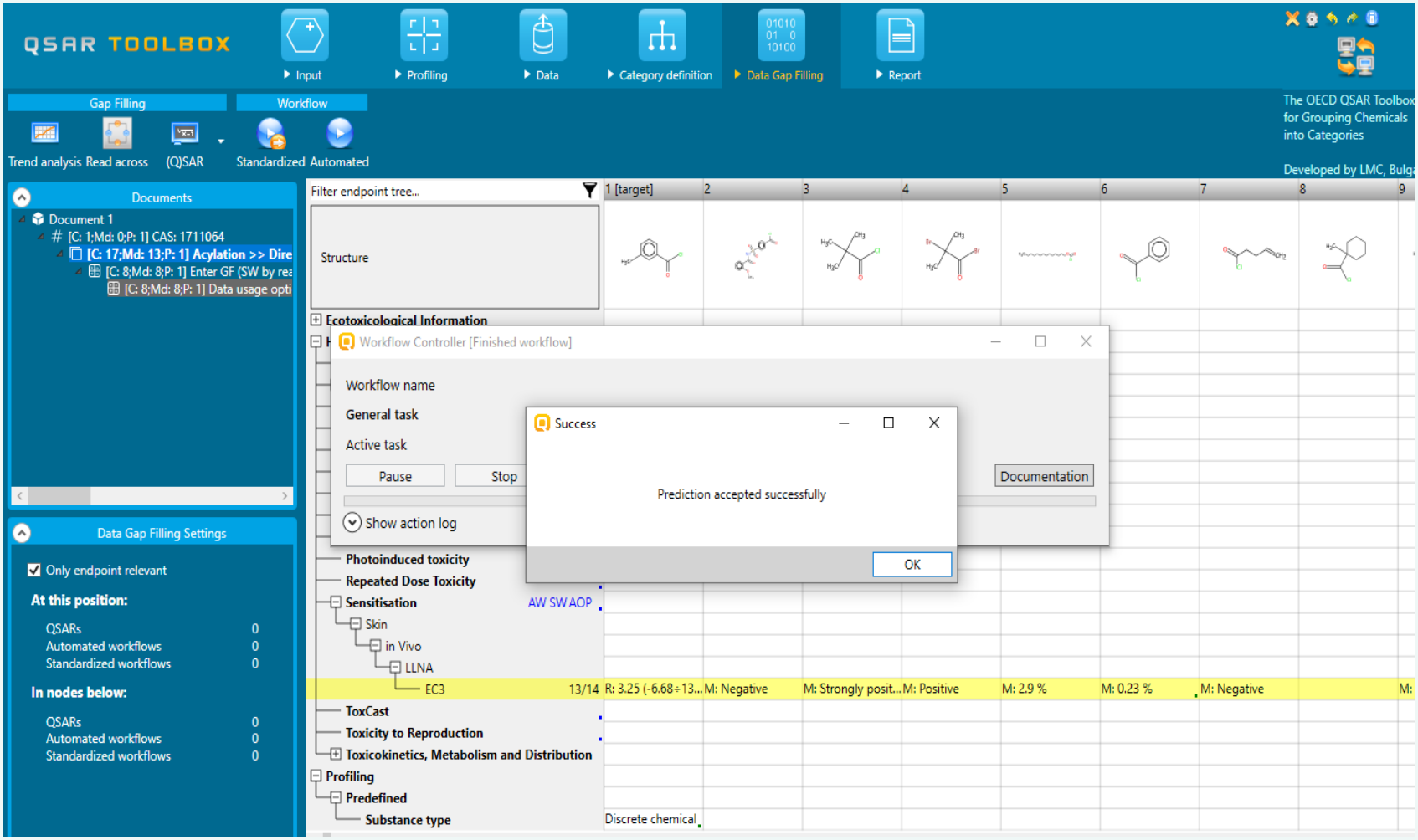


Automated and Standardized workflows in Toolbox

SW for Skin sensitization - examples

Data Gap Filling: Relevant to the workflow profilers are provided and colored

Example 1: Target chemical has protein binding alert

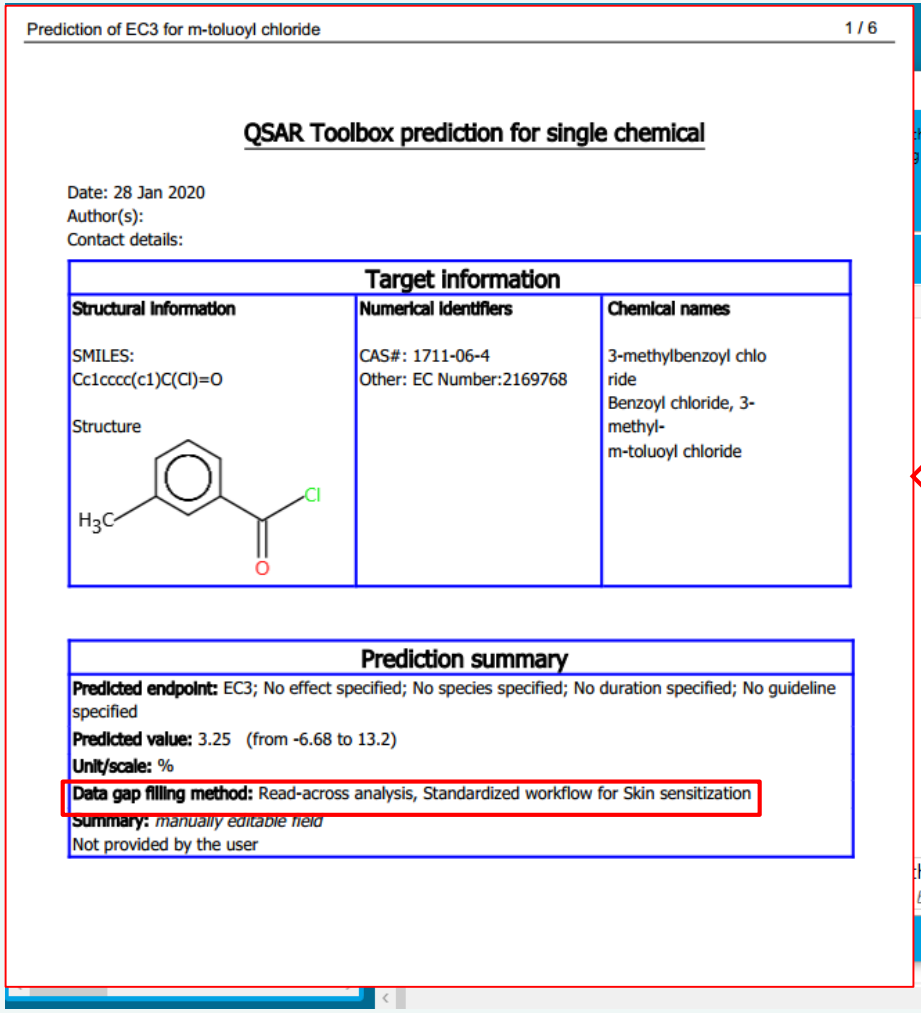


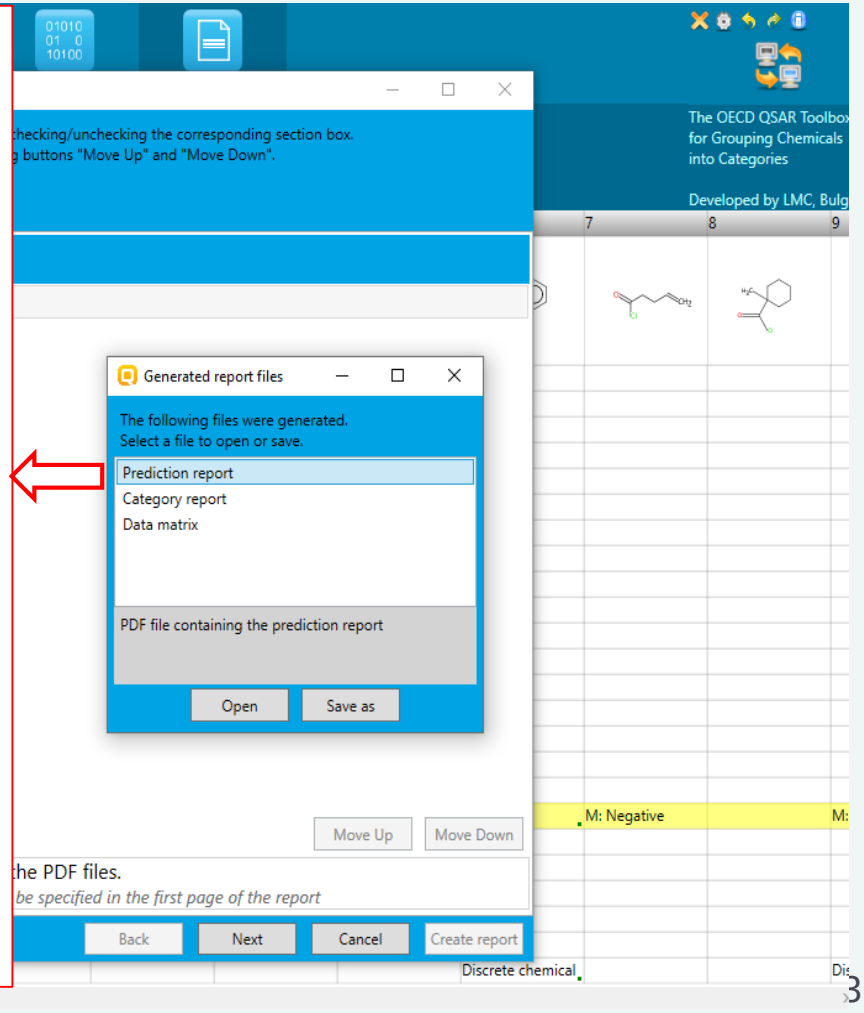
Automated and Standardized workflows in Toolbox

SW for Skin sensitization - examples

Report

- The report section is not part of the workflows
- Report could be generated once the prediction is accepted



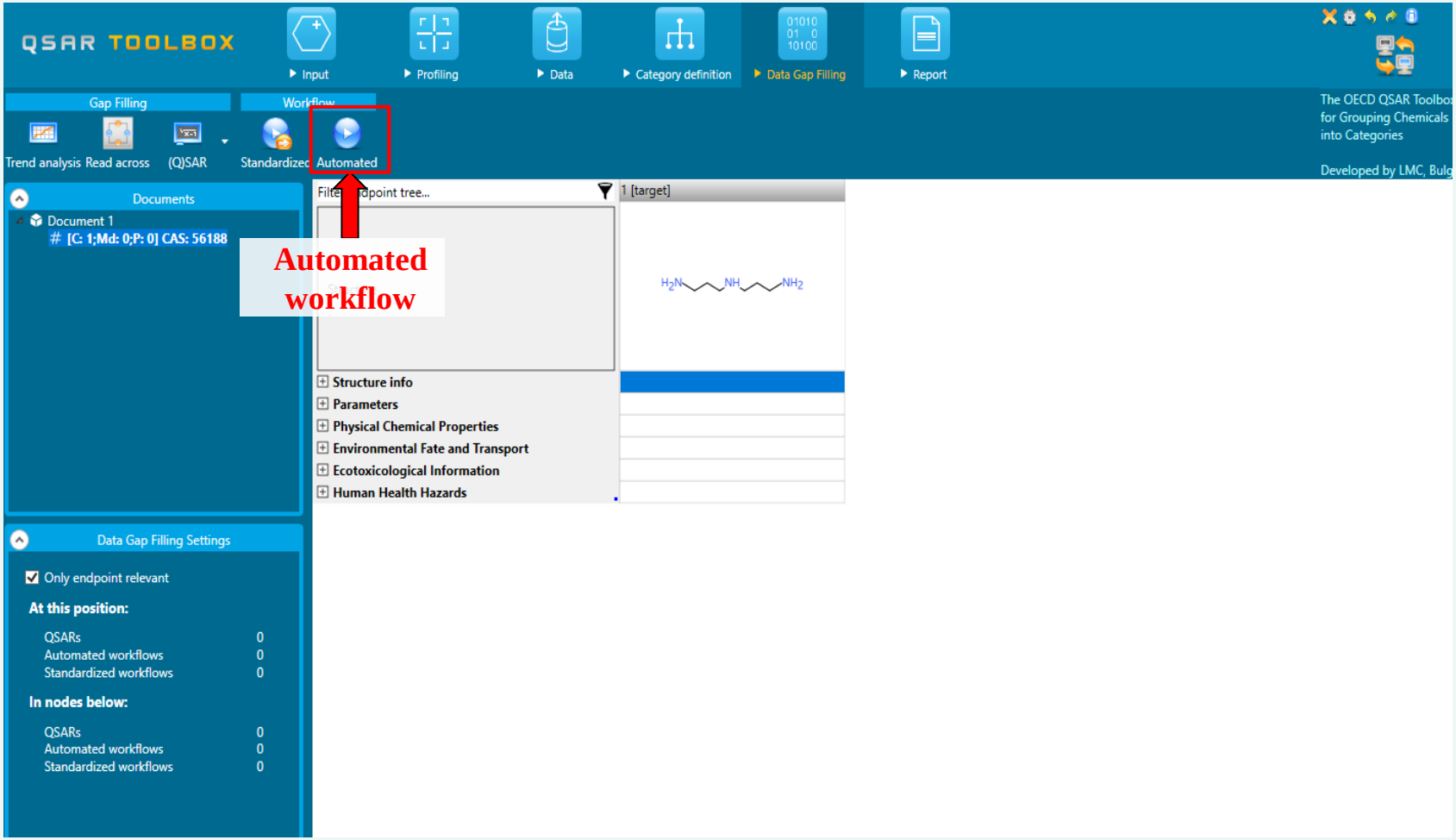


Automated and Standardized workflows in Toolbox

AW for Skin sensitization - examples

Activation: The Automated workflow is activated by clicking on the corresponding button

Example 2: Target chemical is activated metabolically

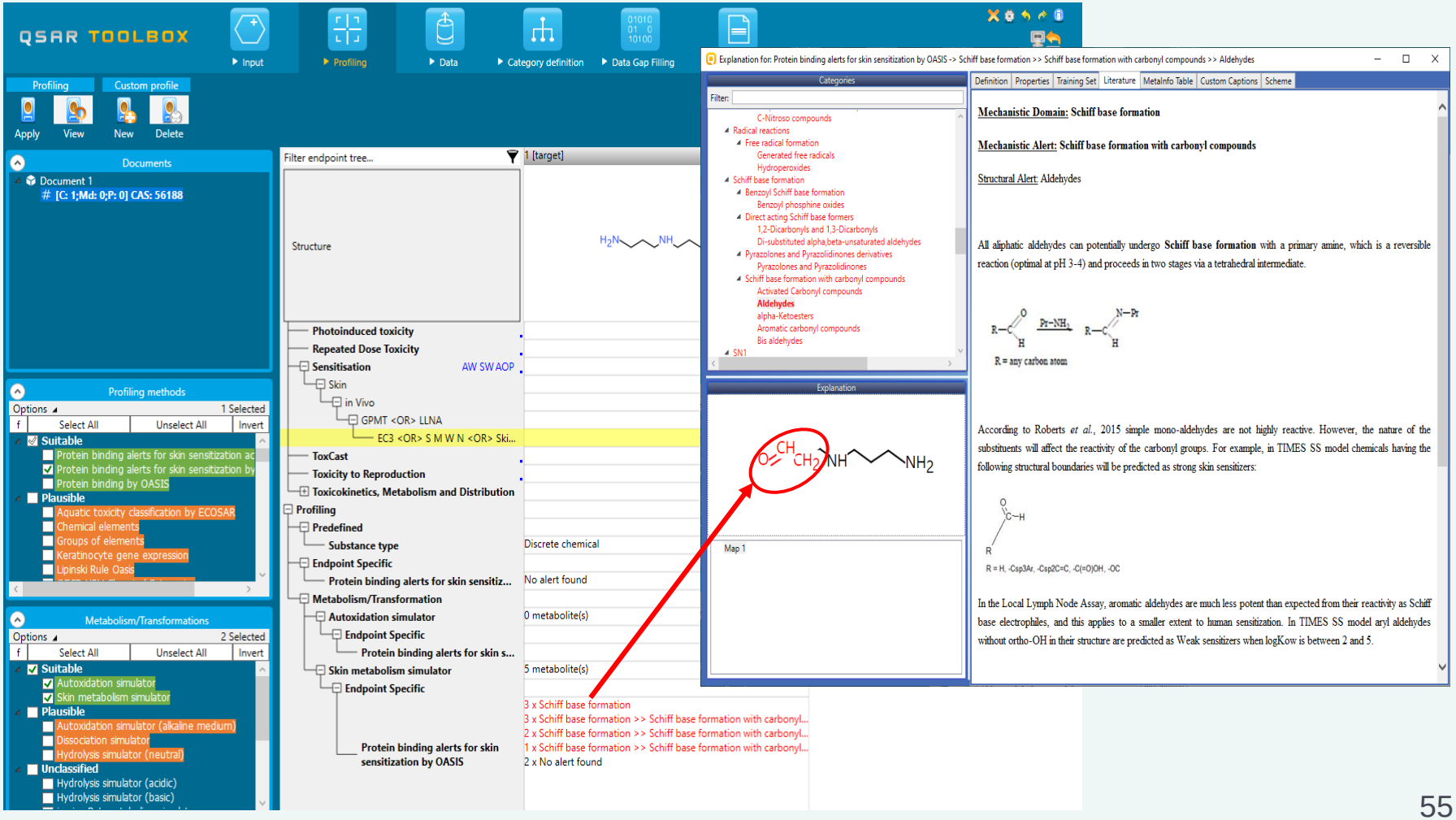


Automated and Standardized workflows in Toolbox

AW for Skin sensitization - examples

In this case the parent chemical is not active but is activated as a result of skin metabolism simulation

Example 2: Target chemical is activated metabolically. The primary group will be formed based on the alert(s) identified in the metabolite(s)

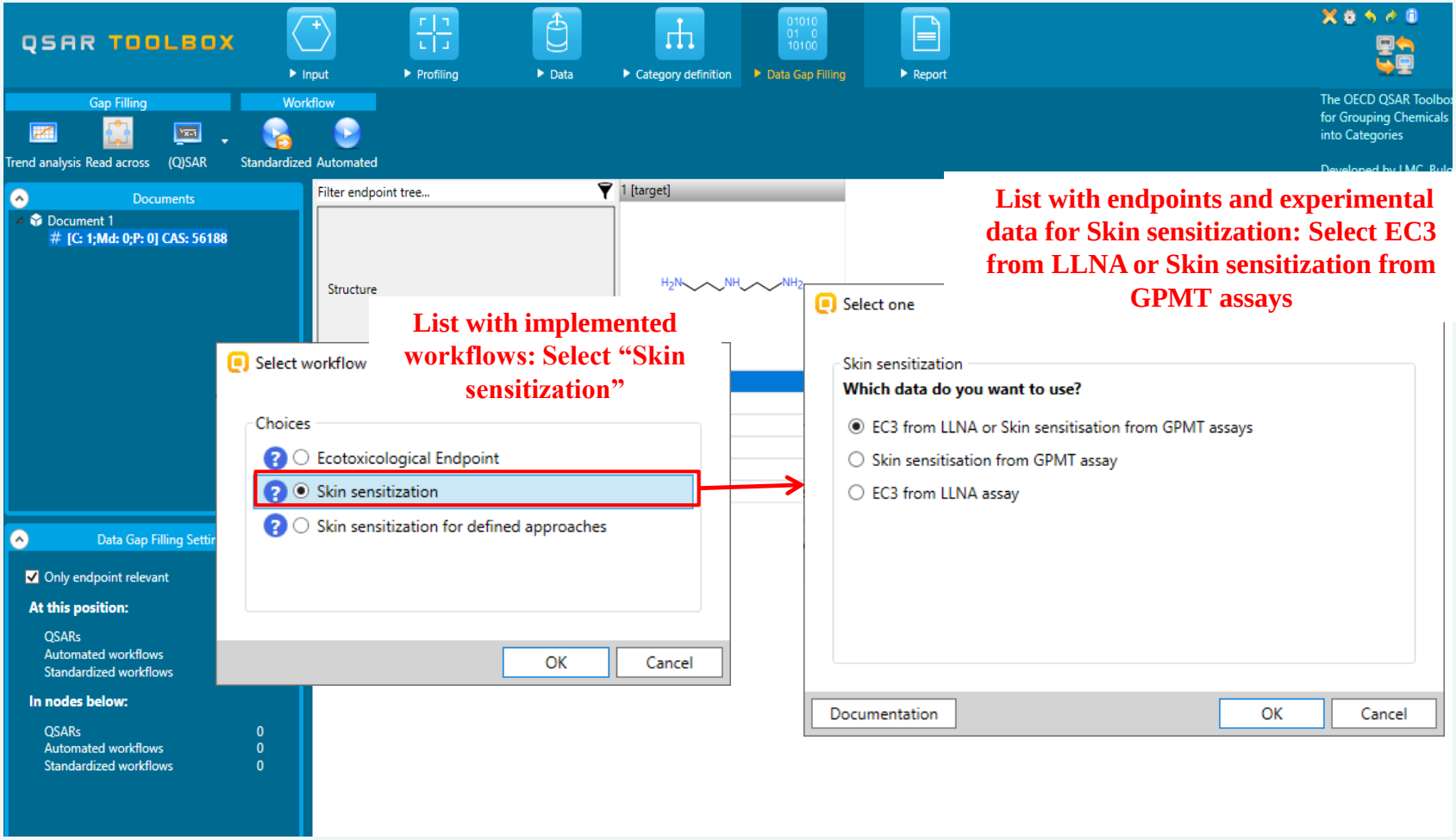


Automated and Standardized workflows in Toolbox

AW for Skin sensitization - examples

Activation: The Automated workflow is activated by clicking on the corresponding button

Example 2: Target chemical is activated metabolically

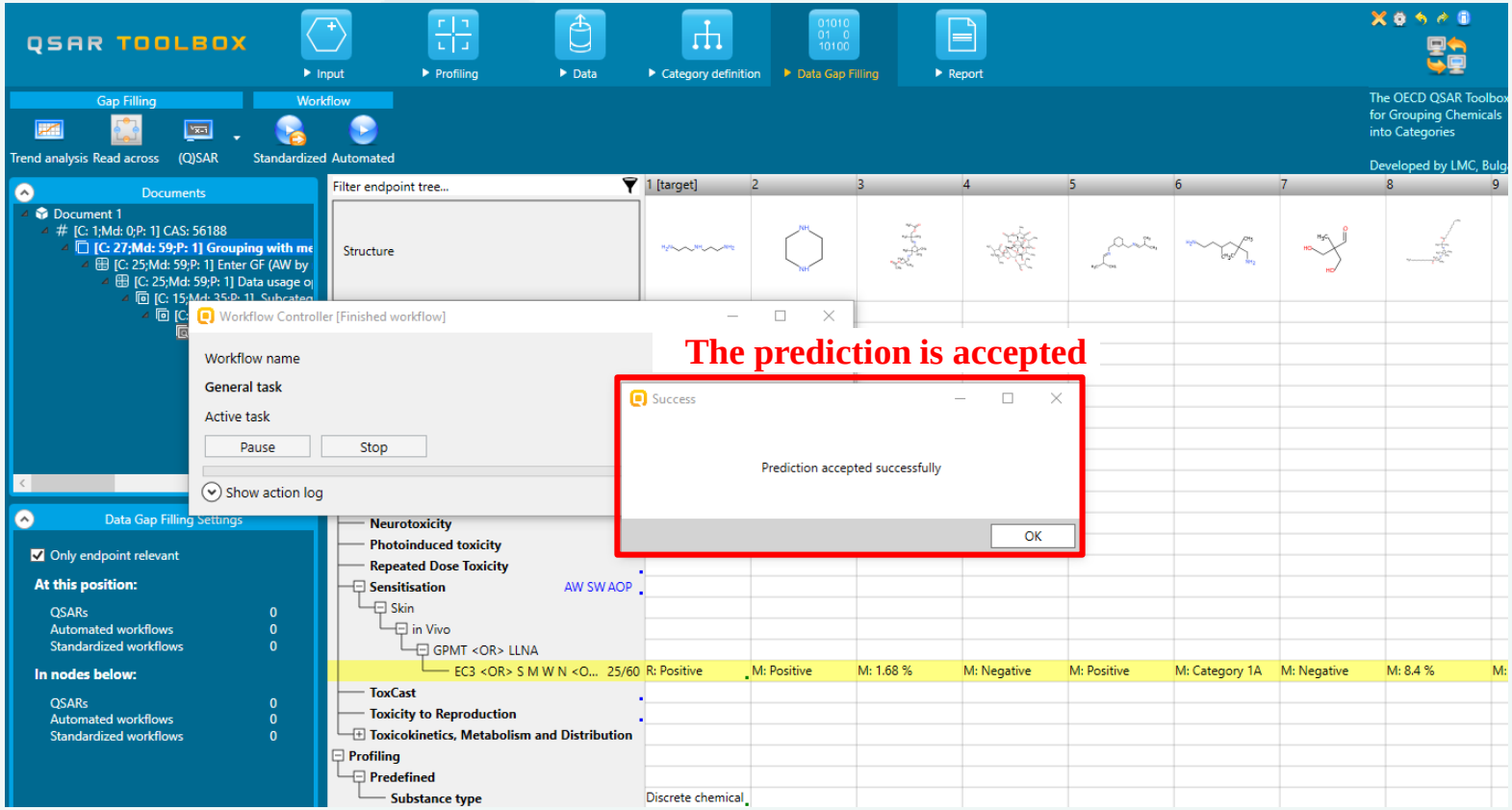


Automated and Standardized workflows in Toolbox

AW for Skin sensitization - examples

The following steps steps are applied automatically:

- Selection of databases
- Selection of primary group
- Sequence of subcategorizations
- Acceptance of the prediction



More information about the automated and standardized workflows could be found in:
Yordanova, D., Schultz, T.W., Kuseva, C., Tankova, K., Ivanova, H., Dermen, I., Pavlov, T., Temelkov, S., Chapkanov, A., Georgiev, M., Gissi, A., Sobanski, T., Mekenyan, O.G. 2019. Automated and standardized workflows in the OECD QSAR Toolbox. *Comput. Toxicol*. 10, pp. 89-104.

Additional materials related to Automated and standardized workflows are available at Toolbox website:
https://qsartoolbox.org

List with tutorials:

- Tutorial on how to predict skin sensitization potential by automated workflow
- Tutorial on how to predict skin sensitization potential by standardized workflow
- Tutorial of how to use Automated workflow for ecotoxicological prediction
- Tutorial of how to use Standardized workflow for ecotoxicological prediction