QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

OECD (Q)SAR Toolbox v.4.4.1

An example illustrating RAAF Scenario 5 and related assessment elements

Outlook

- **Background**
- Keywords
- Objectives
- Specific Aims
- Read Across Assessment Framework (RAAF)
- The exercise
- Workflow

Background

- This is a step-by-step presentation designed to take the Toolbox user through the workflow of a data gap filling exercise and assessing of the outcome whether read across is scientifically acceptable or not

- The read-across prediction will be justified by fulfilling all information requirements according to the Read Across Assessment Framework (RAAF).

Outlook

- Background
- **Keywords**
- Objectives
- Specific Aims
- Read Across Assessment Framework (RAAF)
- The exercise
- Workflow

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
- Read Across Assessment Framework (RAAF)
- The exercise
- Workflow

Objectives

This presentation demonstrates a number of functionalities of the Toolbox:

- Define target endpoint;
- Relevancy of profiles and data availability;
- Searching of analogues accounting for metabolism;
- Category consistency check;
- Selection of RAAF scenario;
- Filling in the report sections related to each read-across assessment element.

Outlook

- Background
- Keywords
- Objectives
- **Specific Aims**
- Read Across Assessment Framework (RAAF)
- The exercise
- Workflow

Specific Aims

- To familiarize the user with the Read Across Assessment Framework (RAAF) and more specifically with Scenario 5;
- To explain to the user how to search for analogues producing common metabolite;
- To introduce to the user the read across assessment elements (AE) and to provide examples with possible content of them;
- To introduce to the user the report basket;
- To provide to the Toolbox user the rationale behind each step of the exercise.

Outlook

- Background
- Keywords
- Objectives
- Specific Aims
- **Read Across Assessment Framework (RAAF)**
- The exercise
- Workflow

Read Across Assessment Framework (RAAF)

Overview

- RAAF has been developed by ECHA as an internal tool providing a framework for a consistent and structured assessment of grouping and read across approaches under REACH.
- The outcome of the assessment is a conclusion on whether the read across is scientifically acceptable or not.
- The RAAF defines different scenarios for different read-across approaches.
- Each scenario is associated with particular aspects (assessment elements, AEs).
- Total six scenarios are available: two for analogue approach and four for category approach

Read Across Assessment Framework (RAAF)

Criteria for the different RAAF scenarios

SCENARIO	APPROACH	READ-ACROSS HYPOTHESIS BASED ON	QUANTITATIVE VARIATIONS
1	Analogue	(Bio)transformation to common compound(s)	Property of the target substance predicted to be quantitatively equal to those of the source substance or prediction based on a worst-case approach.
2	Analogue	Different compounds have qualitatively similar properties	Properties of the target substance predicted to be quantitatively equal to those of the source substance or prediction based on a worst-case approach.
3	Category	(Bio)transformation to common compound(s)	Variations in the properties observed among source substances. Prediction based on a regular pattern or on a worst-case approach.
4	Category	Different compounds have qualitatively similar properties	Variations in the properties observed among source substances. Prediction based on a regular pattern or on a worst-case approach.
5	Category	(Bio)transformation to common compound(s)	No relevant variations in properties observed among source substances and the same strength predicted for the target substance.
6	Category	Different compounds have qualitatively similar properties	No relevant variations in properties observed among source substances and the same strength predicted for the target substance

*Read-Across Assessment Framework (RAAF) available at https://echa.europa.eu/documents/10162/13628/raaf_en.pdf

Read Across Assessment Framework (RAAF)

Selection of RAAF scenario

1. Distinguish whether analogue or category approach is decided based on number (N) of analogues*:
   a) N of analogues ≤ 3 is Analogue approach (scenario 1-2)
   b) N of analogues > 3 is Category approach (scenario 3-6)
2. To identify the basis of the read across hypothesis
   a) (Bio)transformation to common compound(s) – the read across hypothesis is that different substances give rise to (the same) common compounds to which the organism is exposed
   b) Different compounds have the same type of effect(s) – the read across hypothesis is that the organism is not exposed to common compounds but rather, as a result of similarity, that different compounds have similar (eco)toxicological and fate properties. These compounds may be the **source** and **target substances themselves** or **one or more of their (bio)transformation products**.
3. For a category approach (scenario 3-6) there is a need to take further account whether or not quantitative variations in the properties are observed among the category members:
   a) There is quantitative variation in the (eco) toxicity when it is more than 1 log units** (scenario 3 and 4)
   b) Quantitative variation is not expected in the (eco) toxicity when it is less or equal to 1 log unit (scenario 5-6)

*** The threshold for number of analogues which distinguishes analogue from category approach is proposed by LMC**
****The quantitative variation in the (eco)toxicity of 1 log unit is proposed by LMC due to empirically observations.**

Read Across Assessment Framework (RAAF)
Selection of RAAF scenario

- Each scenario consists of a pre-defined set of assessment elements (AEs) that, when taken together, cover all of the essential scientific aspects that need to be addressed in the read-across approach for a particular scenario.*
- Each AE reflects a critical scientific aspect of a read-across.
- The AEs could be:
  - **common** for all scenario within one approach - common AEs for Scenario 1 and 2 (analogue approach) and common AEs for Scenario 3, 4, 5 and 6 (category approach)
  - **specific** – addressing specific scenario.

*Read-Across Assessment Framework (RAAF) available at https://echa.europa.eu/documents/10162/13628/raaf_en.pdf

Outlook

- Background
- Keywords
- Objectives
- Specific Aims
- Read Across Assessment Framework (RAAF)
- **The exercise**
- Workflow

The Exercise

- In this exercise we will predict *Repeated dose toxicity* of *N*-(2-Hydroxyethyl)ethylenediamine [CAS# 111-41-1], which will be the "target" chemical;
- The target endpoint will be preliminary defined;
- The category will be defined based on analogues having common metabolite produced after *in vivo* Rat liver metabolism;
- A read-across approach will be used for the prediction. The prediction will be based on category approach relying on common metabolite generated for the source and target substances;
- Read across assessment elements will be included to the report.
- Examples for the possible content of each of AEs will be provided.

Outlook

- Background
- Keywords
- Objectives
- Specific Aims
- Read-Across Assessment Framework (RAAF)
- The example
- **Workflow**

Workflow

- **The Toolbox has six modules which are used in a sequential workflow:**
  - Input
  - Profiling
  - Data
  - Category Definition
  - Data Gap Filling
  - Report

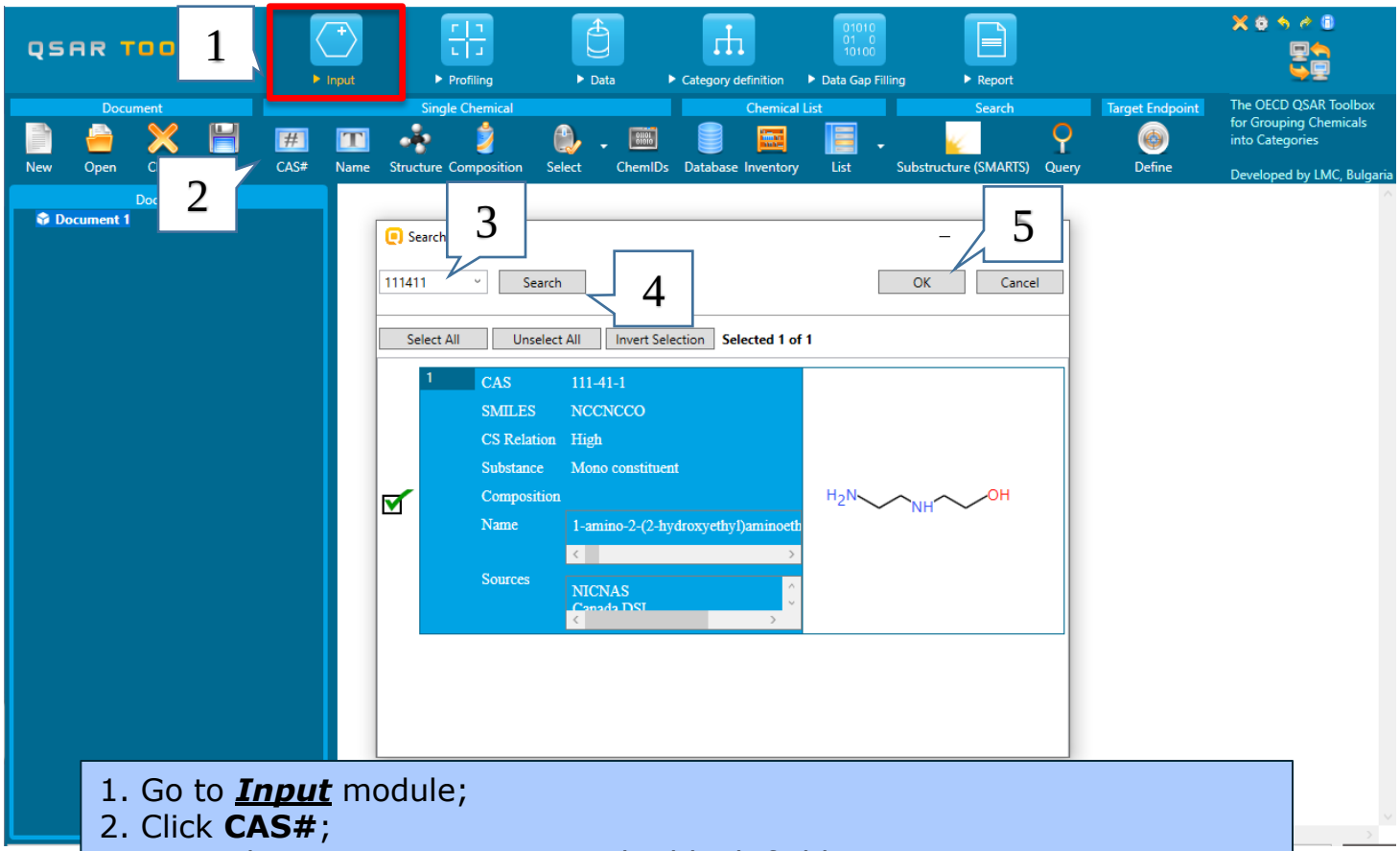
The modules will be presented in different sequence than the showed above.

Input
Overview

- This module provides the user with several means of entering the chemical of interest or the target chemical.
- Since all subsequent functions are based on chemical structure, the goal here is to make sure the molecular structure assigned to the target chemical is the correct one.

Input

Input target chemical by CAS#



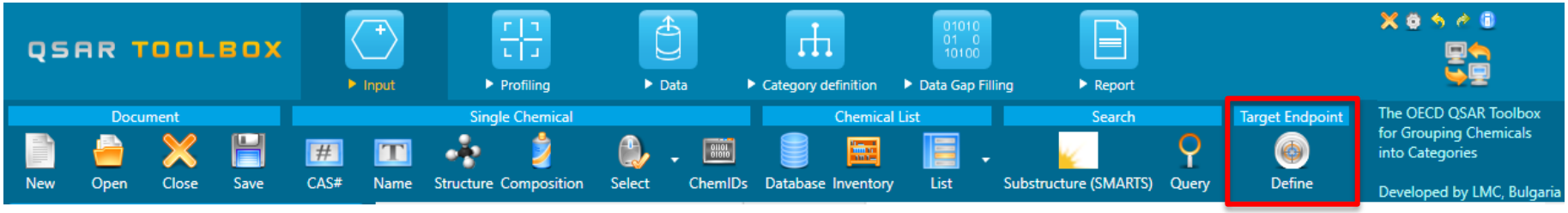
1. Go to ***Input*** module;
2. Click **CAS#**;
3. Enter the **CAS# 111-41-1** in the blank field;
4. Click ***Search;***

When the structure with the requested CAS# appears, click ***OK (5)***.

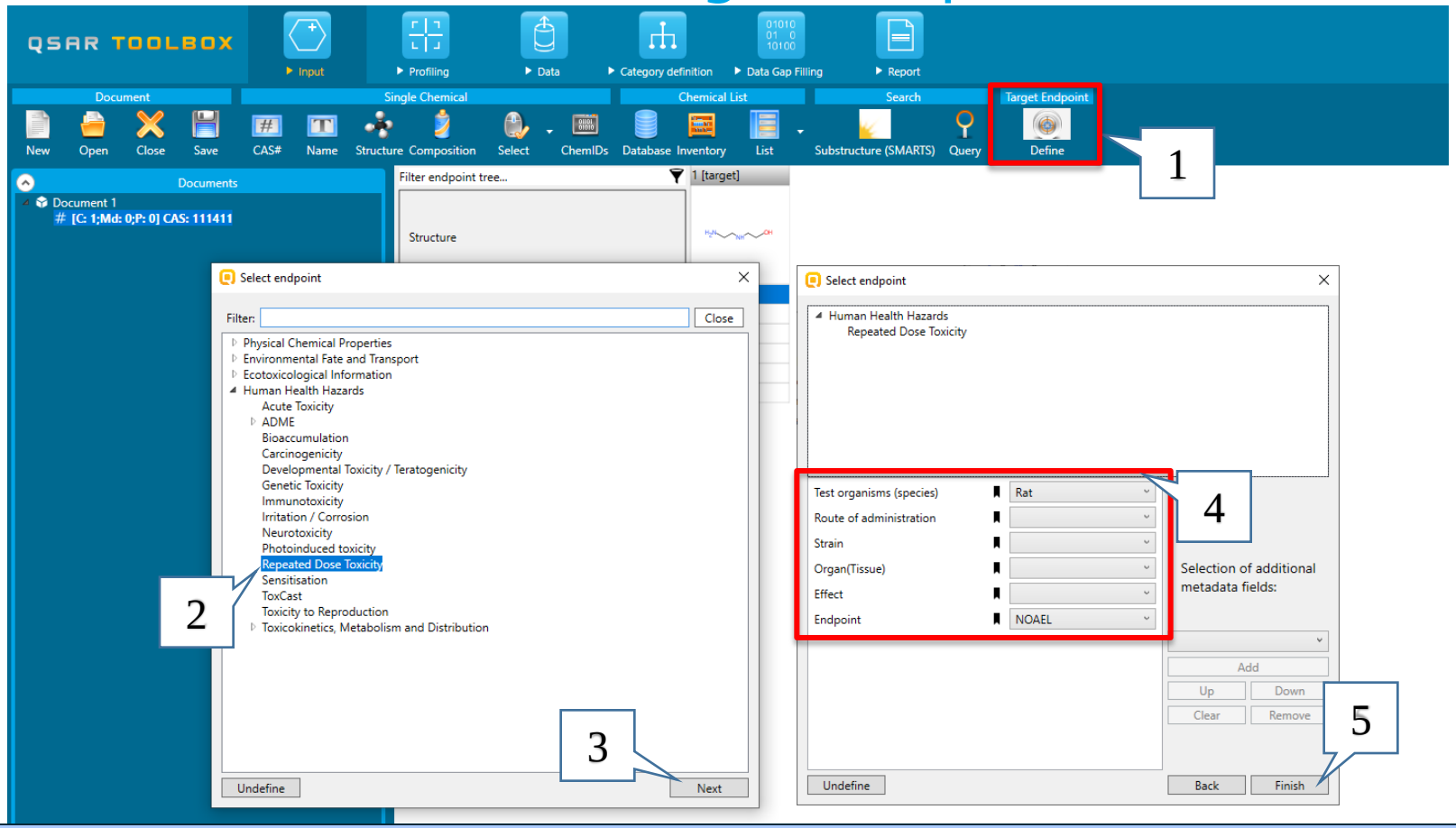
Input
Define target endpoint

Defining of the endpoint allows entering the endpoint of interest e.g. EC3, Chromosome aberration, LC50 etc., along with specific metadata information. Based on the metadata, different relevancy scores for profiles could be provided for same endpoint.

Calculation of alert performance (AP) is only possible if the target endpoint is preliminary defined.



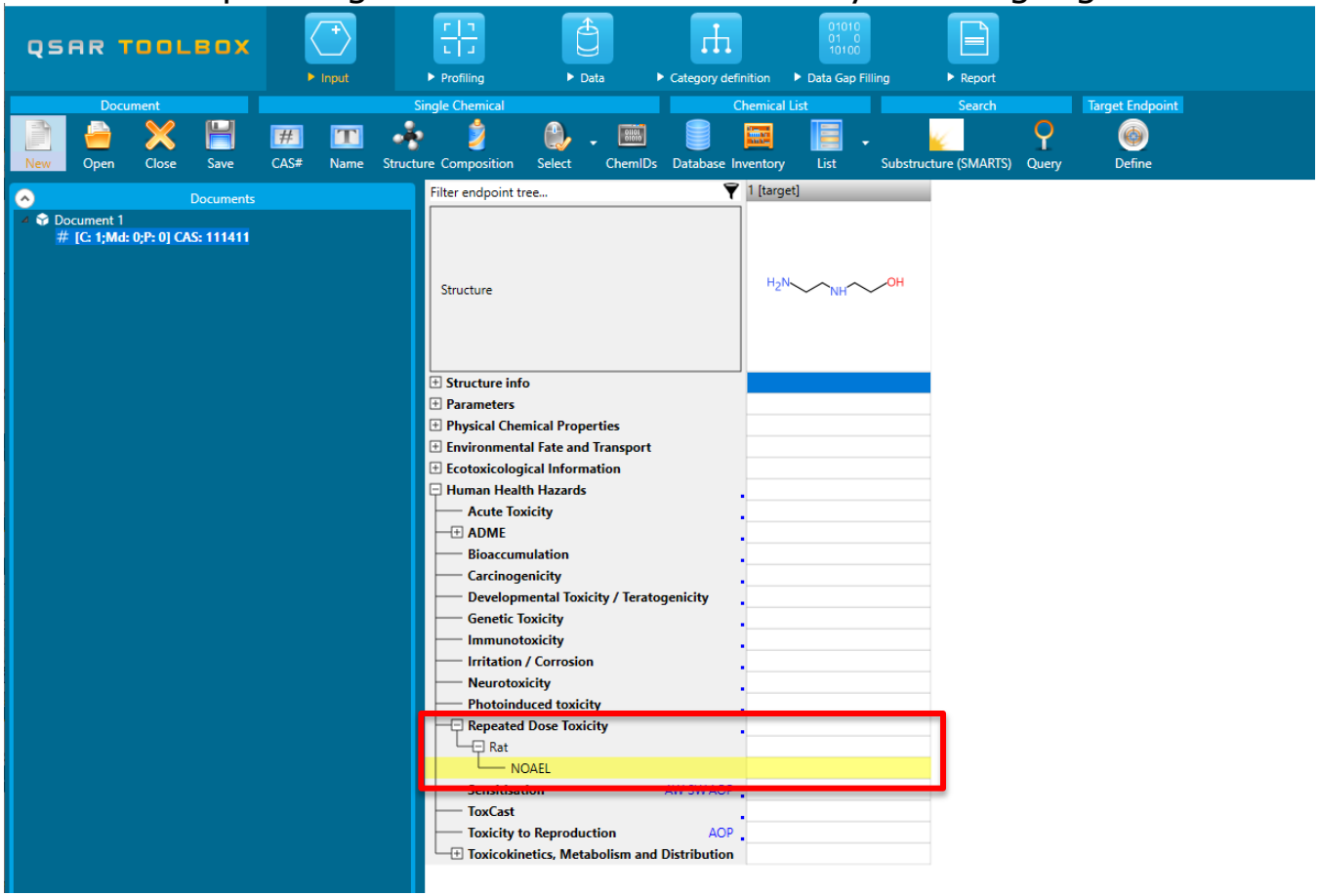
Input
Define target endpoint



When click on ***Define*** (1) you should select the target endpoint. Select **Repeated Dose Toxicity** in the *Human health Hazards* level (2) and click on ***Next*** (3). Select **NOAEL** endpoint (4) and **Rat** test organism (4) from the drop-down menu. Finally click on ***Finish*** (5).

Input
Define target endpoint

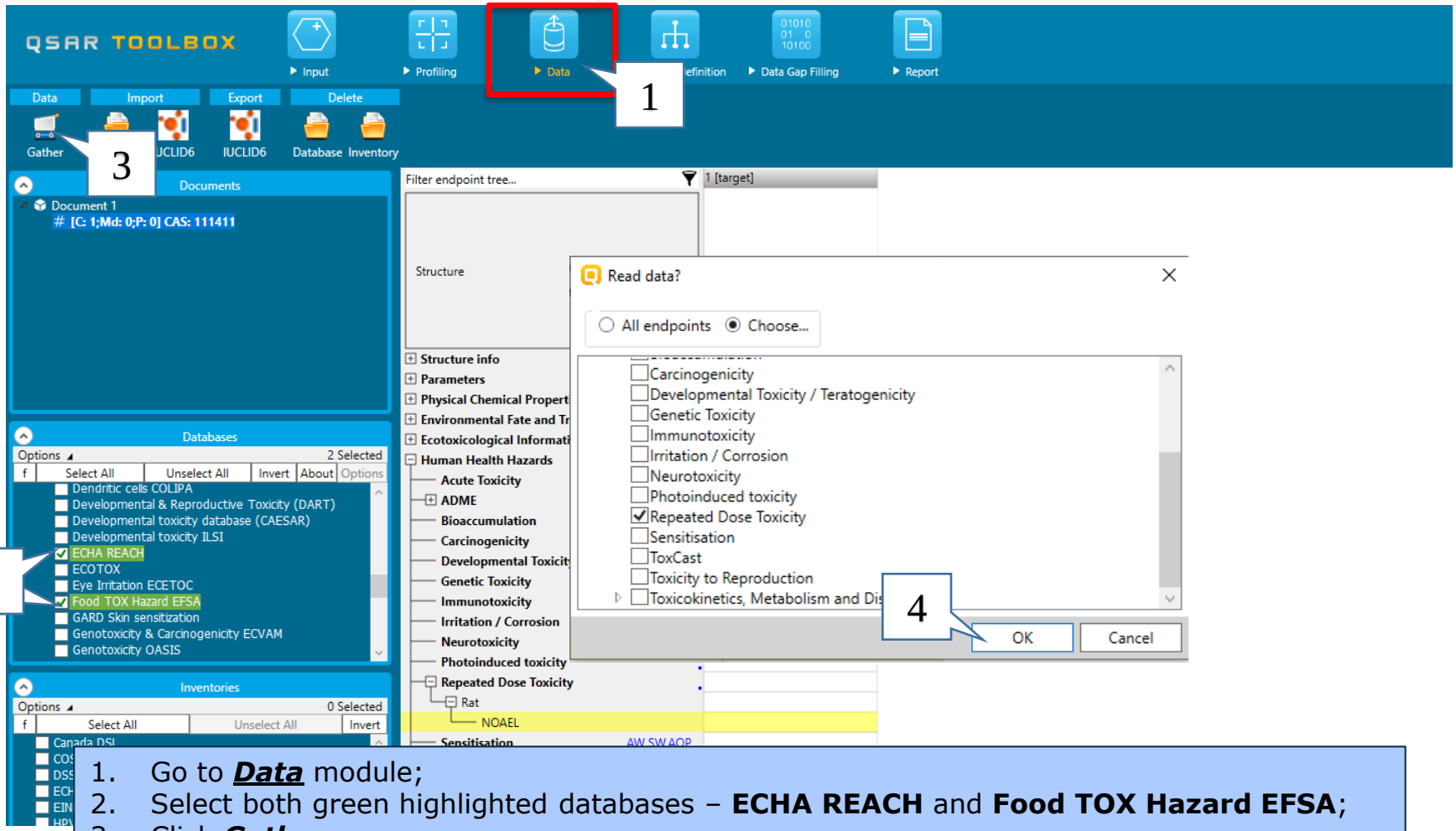
Once the endpoint is defined along with its metadata, they appear in the endpoint tree and the corresponding row of the data matrix is yellow highlighted.



Data
Overview

- "Data" refers to the electronic process of retrieving the environmental fate, eco-toxicity and toxicity data that are stored in the Toolbox.
- Data gathering can be executed in a global fashion (i.e., collecting all data for all endpoints) or on a more narrowly defined basis (e.g., collecting data for a single or limited number of endpoints).

Data
Collecting experimental data



1. Go to ***Data*** module;
2. Select both green highlighted databases – **ECHA REACH** and **Food TOX Hazard EFSA**;
3. Click ***Gather***;
4. Choose to collect repeated dose toxicity data, only and click ***OK.***

Data
Extracted data

- Toxicity information on the target chemical is electronically collected from the selected dataset(s).
- It should be kept in mind that the search for data and analogues is performed only among the chemicals which are listed in the selected databases. In this example ECHA REACH and *Food TOX Hazard EFSA* databases are selected.
- In this example, an insert window appears stating there are two experimental data points for the Repeated dose toxicity.
- Go to the *Profiling* module to check for the reason of the possible effect (to check for an alert identified in the target chemical).

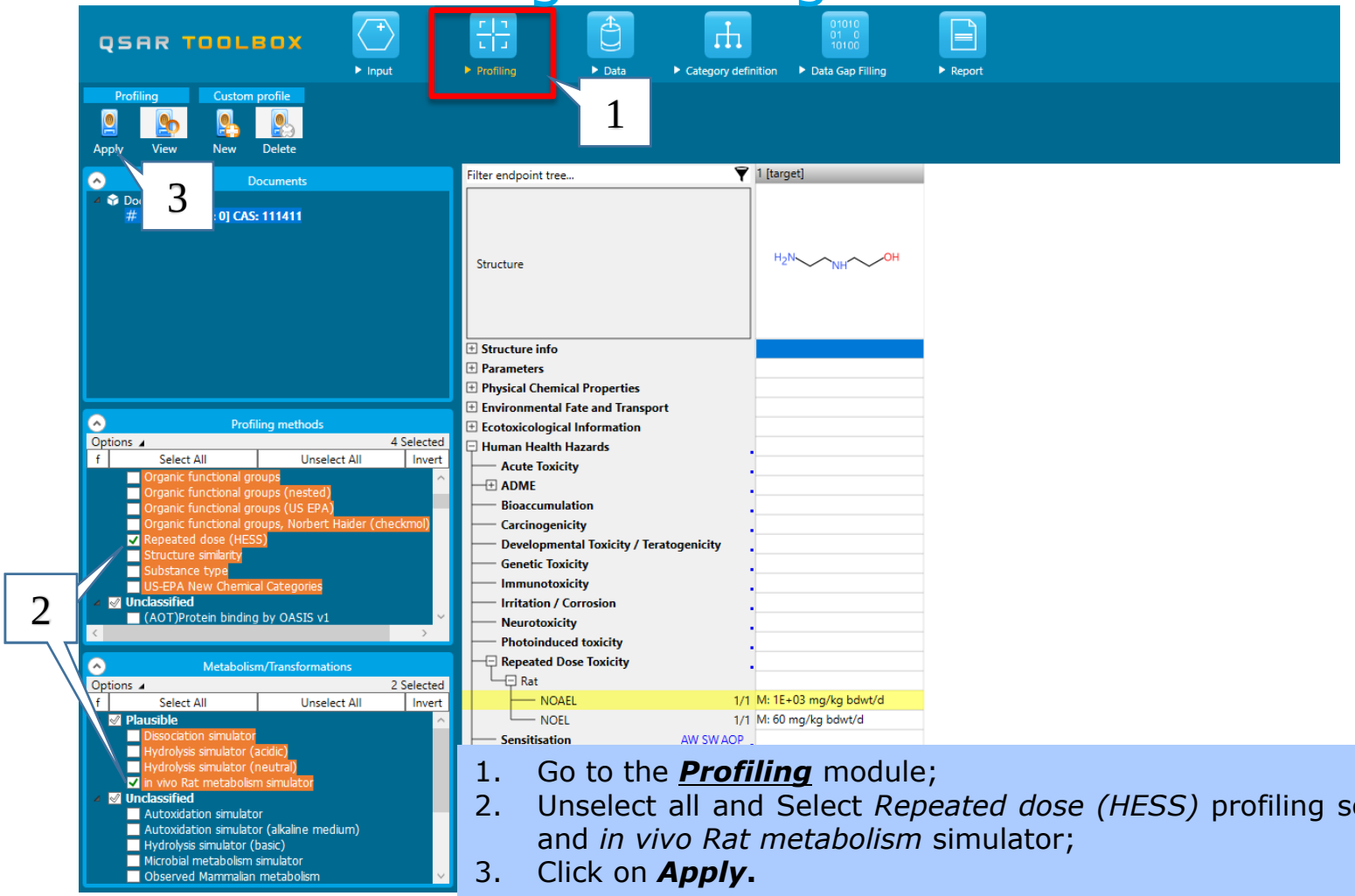
Profiling
Overview

- "Profiling" refers to the electronic process of retrieving relevant information for the target compound, other than its environmental fate, ecotoxicity and toxicity data, which are stored in the Toolbox databases.
- "Profiling" module contains all the knowledge in the system coded in profiling schemes (profilers);
- "Profilers" are a collection of empirical and mechanism knowledge (expertly derived) which could be used to analyse the structural properties of chemicals;
- The "profilers" identify the affiliation of the target chemical(s) to preliminary defined categories (functional groups/alerts);
- The "Profiling" module contains also observed and simulated metabolisms/transformations, which could be used in combination with the profiling schemes;
- The outcome of the profiling determines the most appropriate way to search for analogues, but they are also useful for preliminary screening or prioritization of substances;
- The "profilers" are not (Q)SARs, i.e. they are not prediction models themselves;
- Based on the "profilers' relevancy" the most suitable once are getting colour highlighted*
- For the purpose of this example suitable profilers in combination with simulators are used (see next slide).

*For more details regarding relavancy of profilers see ppt: *Example for predicting skin sensitization taking into account alert performance*

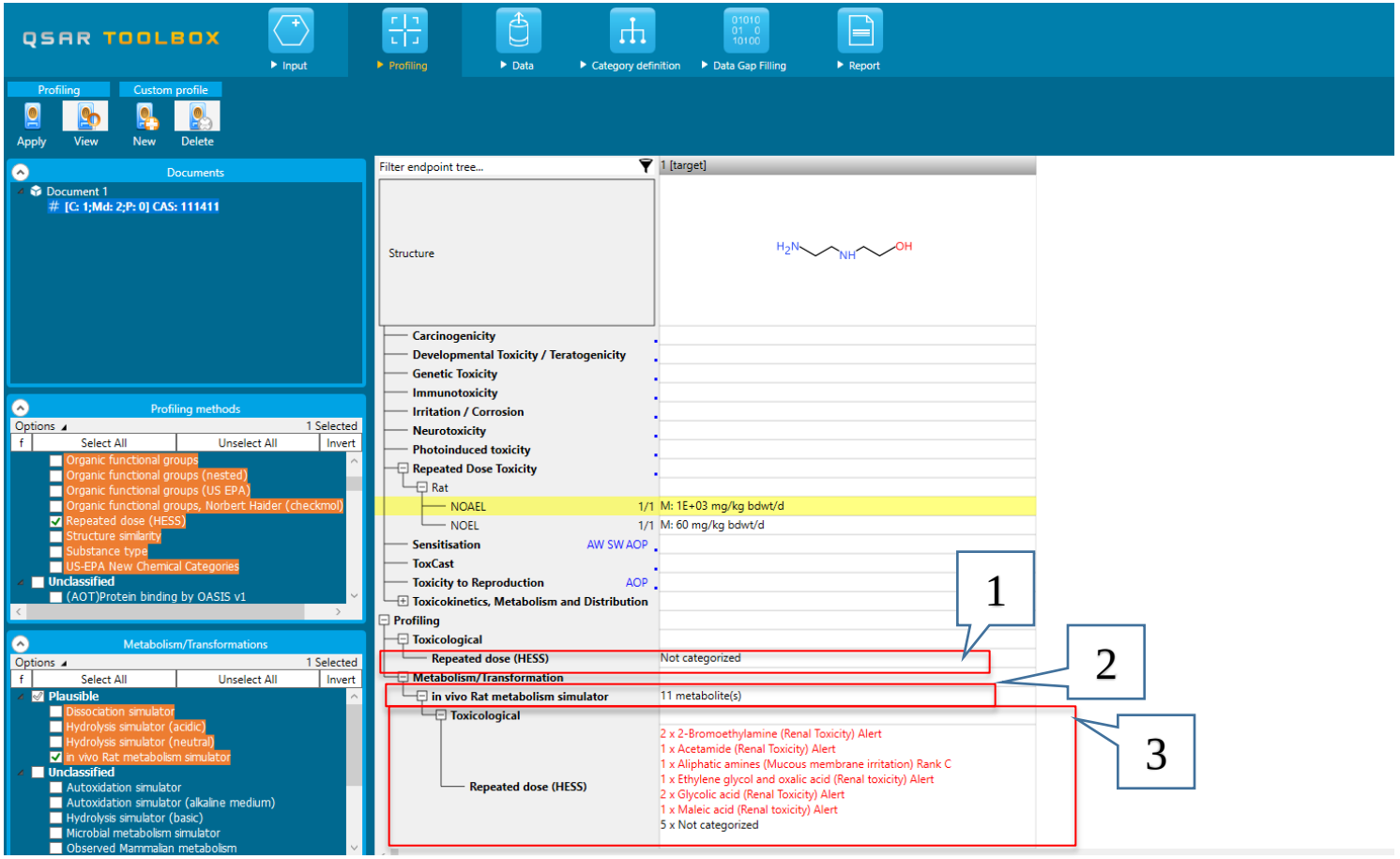
Profiling

Profiling the target chemical



1. Go to the ***Profiling*** module;
2. Unselect all and Select *Repeated dose (HESS)* profiling scheme and *in vivo Rat metabolism* simulator;
3. Click on ***Apply***.

Profiling
Profiling results

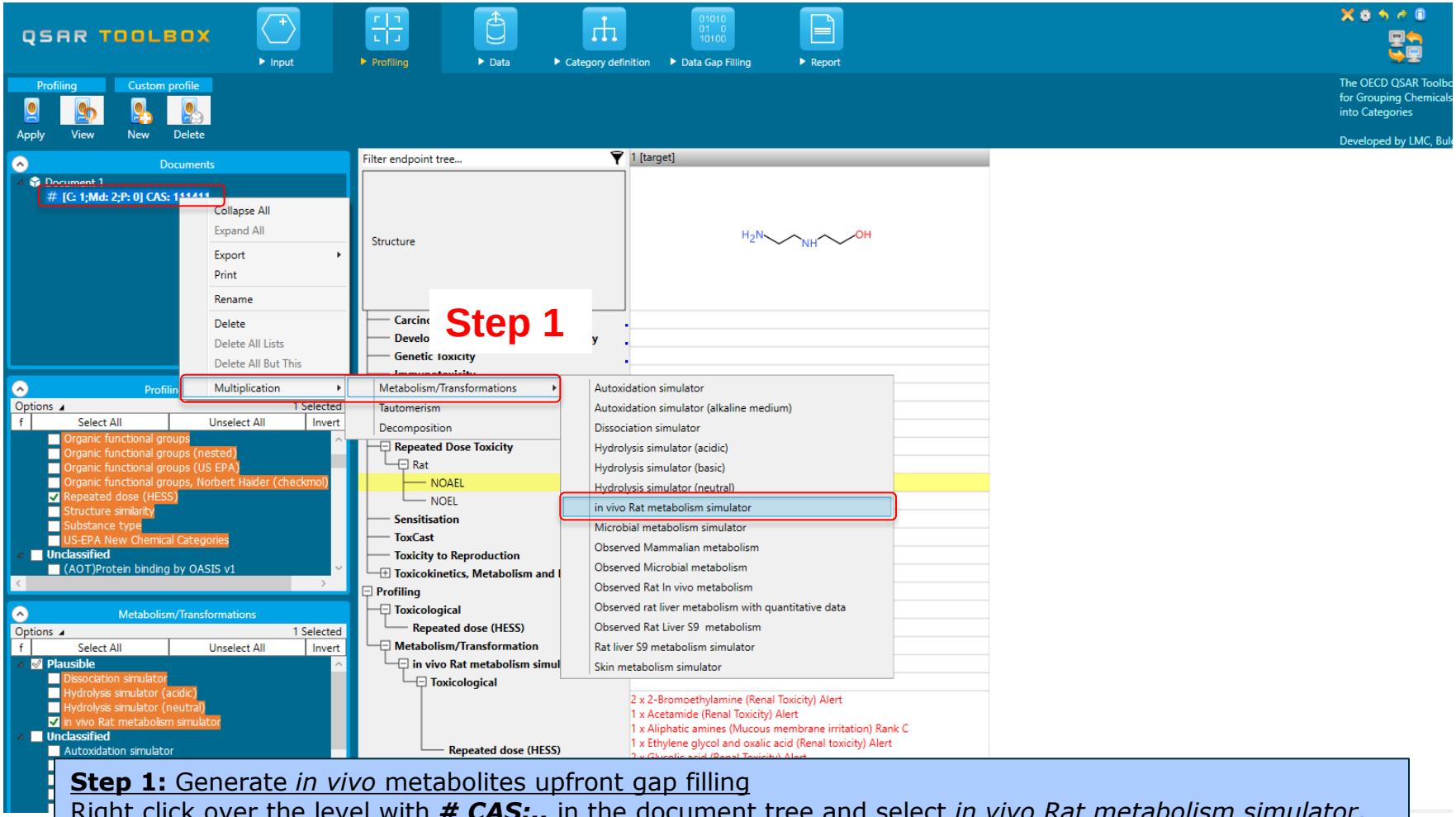


1) No alerts are identified in the target structure as a parent;
2) 11 metabolites are generated as a result of *in vivo Rat metabolism simulator*;
3) Alerts for repeated dose toxicity are identified in six of the generated metabolites.

Recap

- In the *Input* module, you entered the target chemical and defined the target endpoint.
- In the *Data* module, you saw the databases corresponding to the defined target endpoint and collect data for the target.
- In the *Profiling* module, you profiled the target chemical with profiling scheme and metabolic simulator related to the selected target endpoint.
- Alerts for repeated dose toxicity were identified for some of the metabolites produced after *in vivo* rat liver metabolic activation.
- Hence the next step of the workflow is to collect analogues accounting for an *in vivo* rat metabolism (pretending that experimental data for the target does not exist).
- Before collecting analogue let's analyse in more details the simulated *in vivo* metabolites (see next slides).

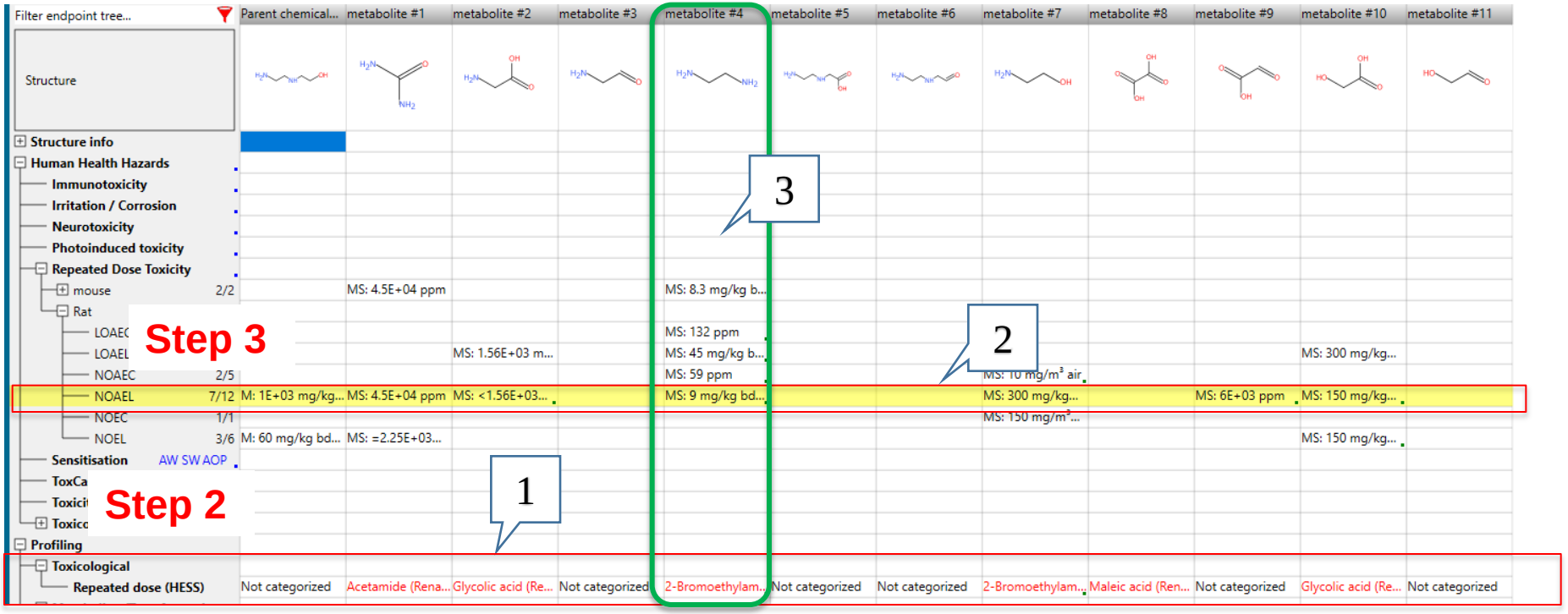
Handling of in vivo rat liver metabolism



Step 1: Generate *in vivo* metabolites upfront gap filling

Right click over the level with ***# CAS:..*** in the document tree and select *in vivo Rat metabolism simulator*. 11 metabolites are produced. The metabolites appeared next to the parent (see next slide).

Handling of in vivo rat liver metabolism



Step 2: Profile the package: parent and metabolites according to *Repeated dose (HESS)* profiler (RDT) only (uncheck the metabolic simulator).
Alerts are identified in six out of 11 generated metabolites (1).

Step 3: Gather data for package: parent and metabolites from the selected databases (gather only repeated dose toxicity data). Experimental data for the defined target endpoint is found for six of the metabolites (2).

The metabolite having an alert according to RDT profiler and have experimental data falling in the category 1 (GHS) will be used for searching of analogues (3) (see next slide).

Category Definition
Overview

- This module provides the user with several means of grouping chemicals into a toxicologically meaningful category that includes the target molecule.
- This is the critical step in the workflow.
- Several options are available in the Toolbox to assist the user in refining the category definition.

Category Definition
Grouping methods

- The different grouping methods allow the user to group chemicals into chemical categories according to different measures of "similarity" so that within a category data gaps can be filled by read-across.

- For example, starting from a target chemical for which a specific protein binding mechanism is identified, analogues can be found which can bind by the same mechanism and for which experimental results are available.

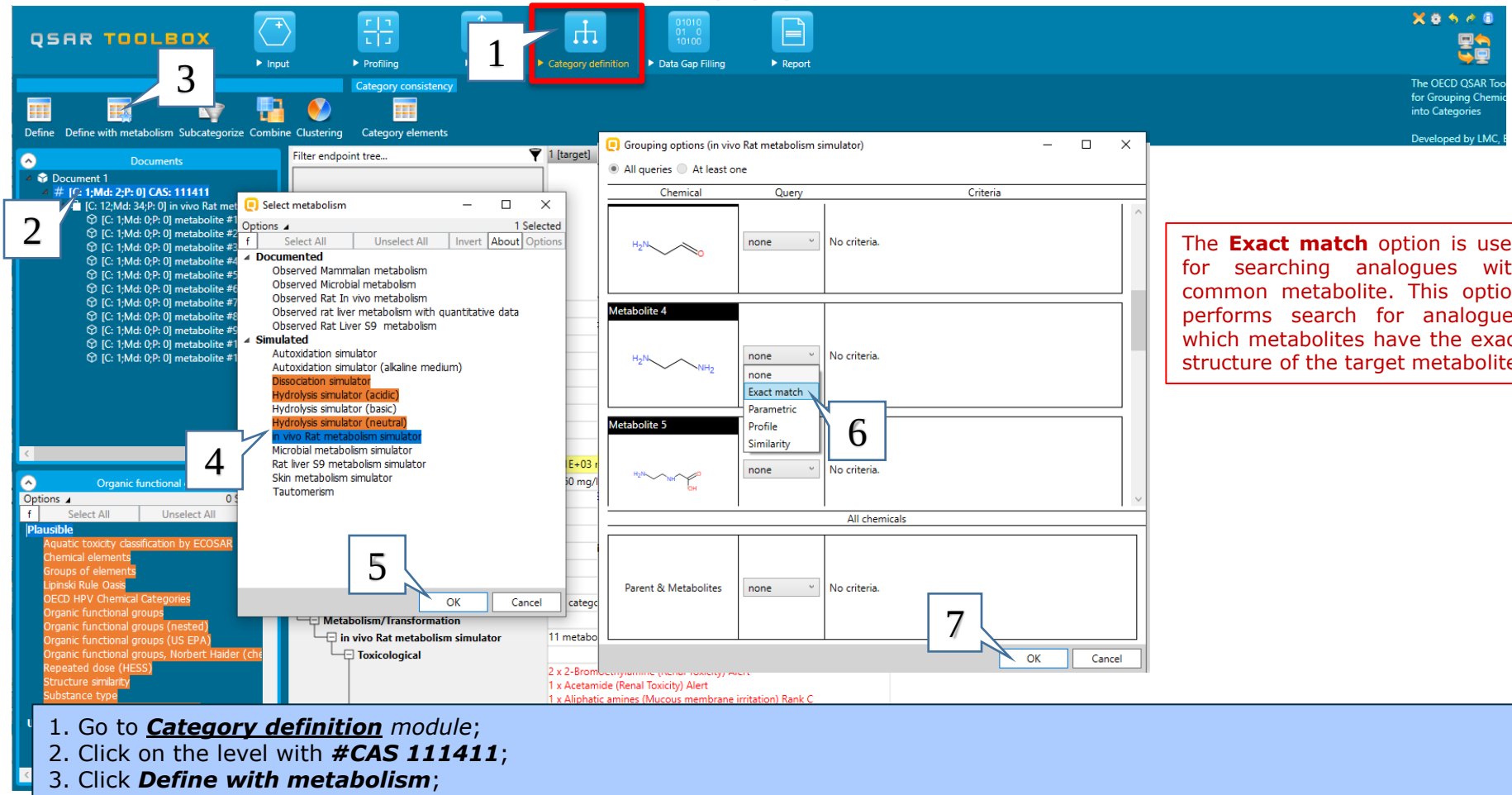
- If no alert is identified in the target structure, but is identified in its metabolites, analogues can be searched accounting for metabolism. In this way the target chemical and the identified analogues will have similar metabolic pattern.

- In our case we will use *Food TOX Hazard EFSA* database and ECHA REACH databases to search for suitable analogues.

- Searching for analogues will be based on a common metabolite (Ethylenediamine) generated as a result of *in vivo* Rat metabolism (see next slide)

Category Definition

Searching for analogues accounting for in vivo rat liver metabolism

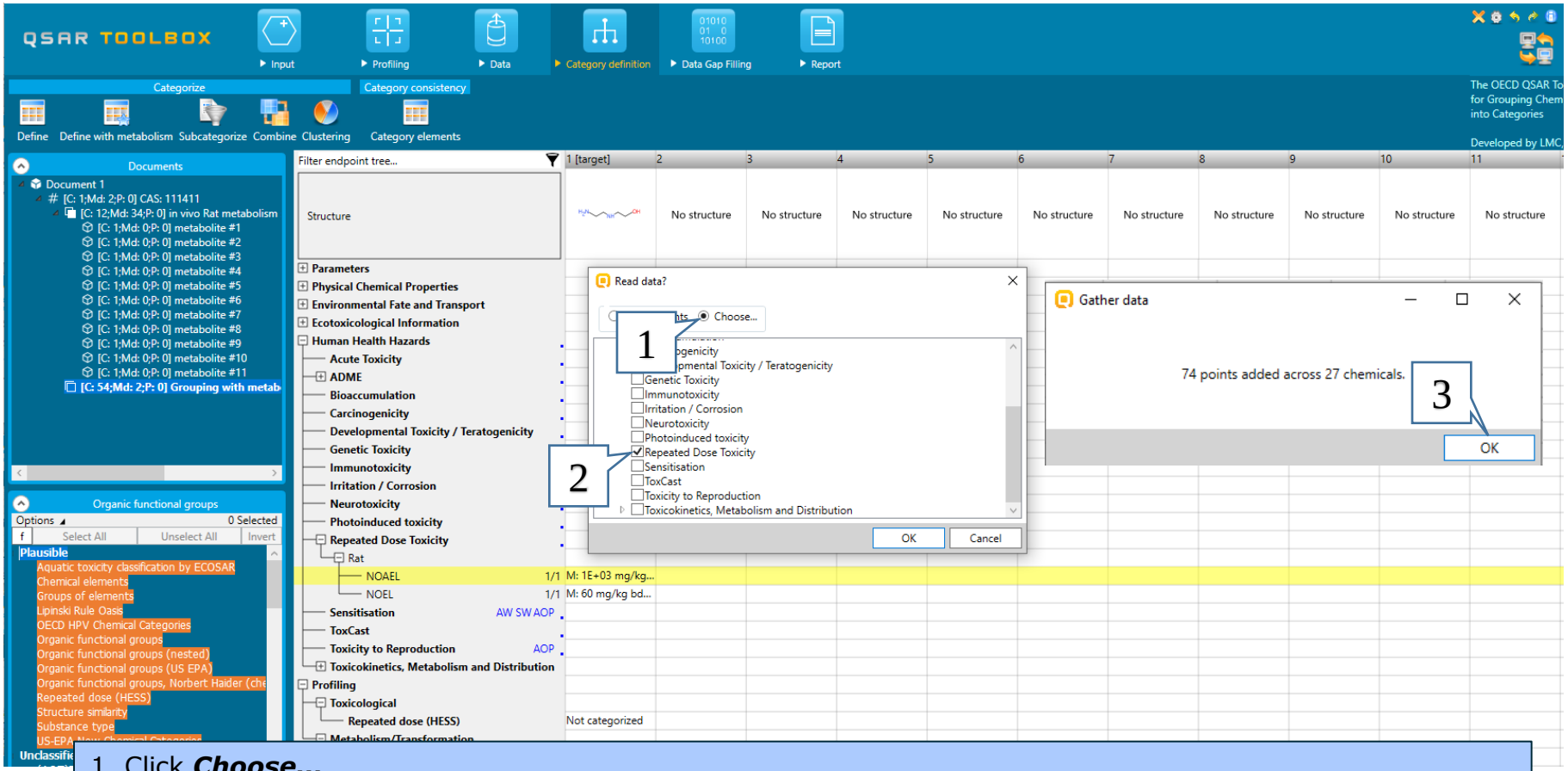


The **Exact match** option is used for searching analogues with common metabolite. This option performs search for analogues which metabolites have the exact structure of the target metabolite.

1. Go to ***Category definition*** *module*;
2. Click on the level with ***#CAS 111411***;
3. Click ***Define with metabolism***;
4. Select ***in vivo Rat metabolism simulator***; 5. Click ***OK***; 6. Target and all metabolites produced by the selected simulator appear. Find the Ethylenediamine structure (Metabolite #4) and specify "**Exact match**" query; 7. Execute the search by clicking ***OK***. The selected databases are not cached. Therefore, first running of this example will take a few minutes.

Category Definition

Searching for analogues accounting for in vivo rat liver metabolism



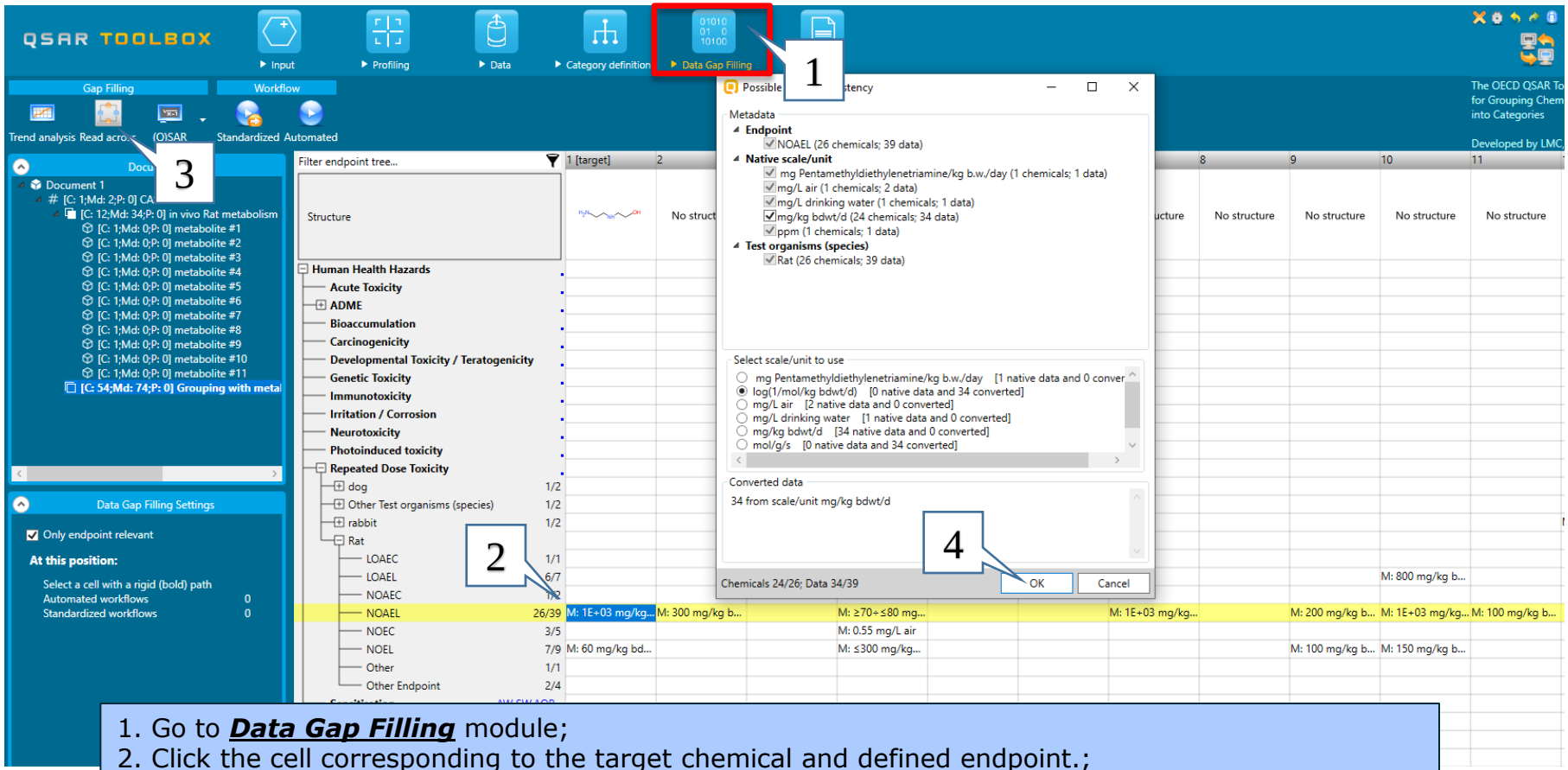
1. Click ***Choose…***
2. Select ***Repeated Dose Toxicity*** data to be collected only;
3. An information window appears informing about the number of collected, click ***OK***. 27 chemicals with 74 experimental data has been found related to the target endpoint.

Data Gap Filling
Overview

- "Data Gap Filling" module give access to five different data gap filling tools:
  - Read-across
  - Trend analysis
  - (Q)SAR models
  - Standardized workflow
  - Automated workflow
- Depending on the situation, the most relevant data gap mechanism should be chosen, taking into account the following considerations:
  - Read-across is the appropriate data-gap filling method for "qualitative" endpoints like skin sensitisation or mutagenicity for which a limited number of results are possible (e.g. positive, negative, equivocal). Furthermore read-across is recommended for "quantitative endpoints" (e.g., 96h-LC50 for fish) if only a low number of analogues with experimental results are identified.
  - Trend analysis is the appropriate data-gap filling method for "quantitative endpoints" (e.g., 96h-LC50 for fish) if a high number of analogues with experimental results are identified.
  - "(Q)SAR models" can be used to fill a data gap if no adequate analogues are found for a target chemical.

In this example we will use the read-across approach.

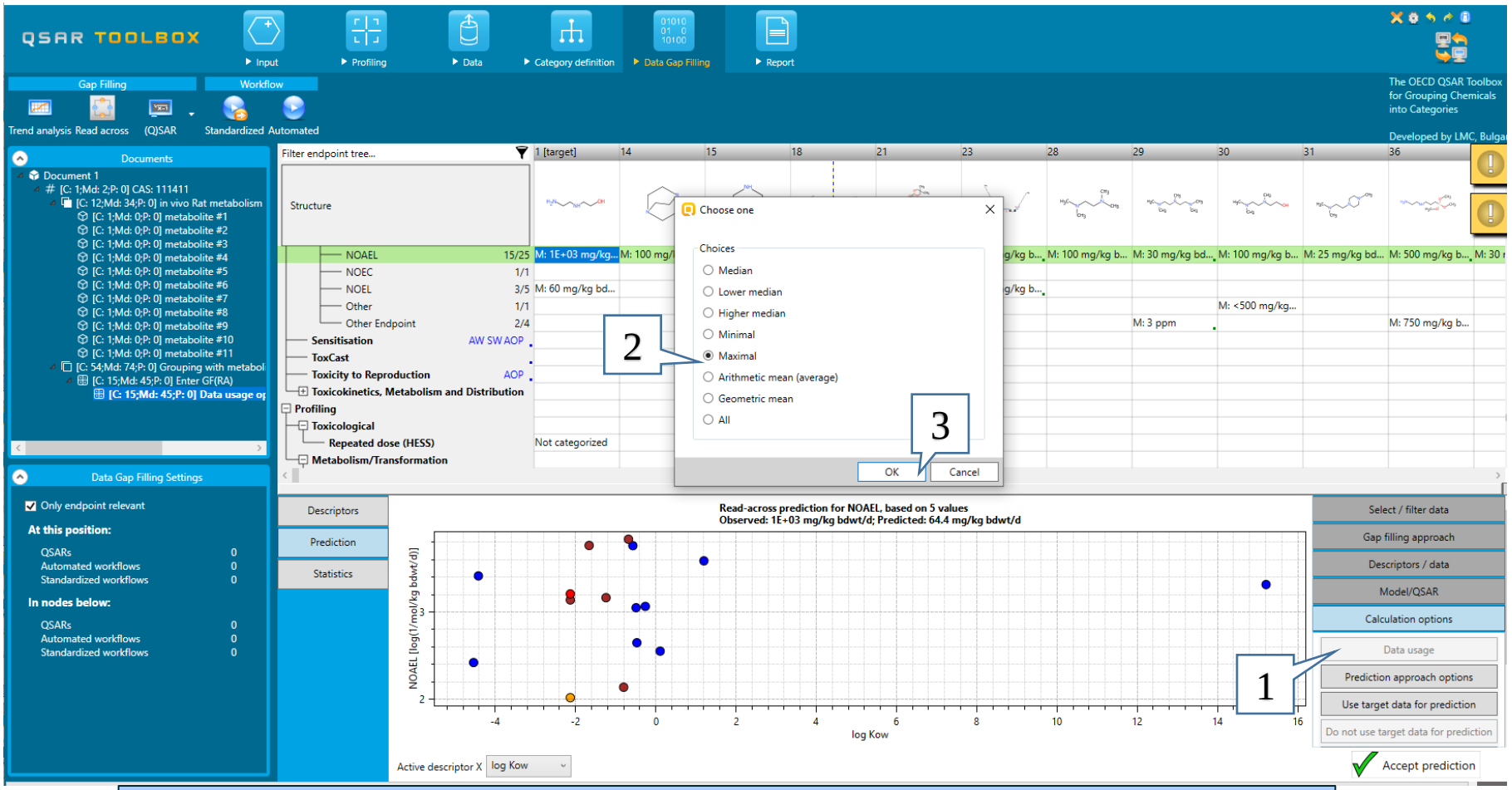
Data Gap Filling
Apply Read-across



1. Go to ***Data Gap Filling*** module;
2. Click the cell corresponding to the target chemical and defined endpoint.;
3. Apply ***Read across***;
4. A pop-up window informing about possible data inconsistency appears click ***OK.***

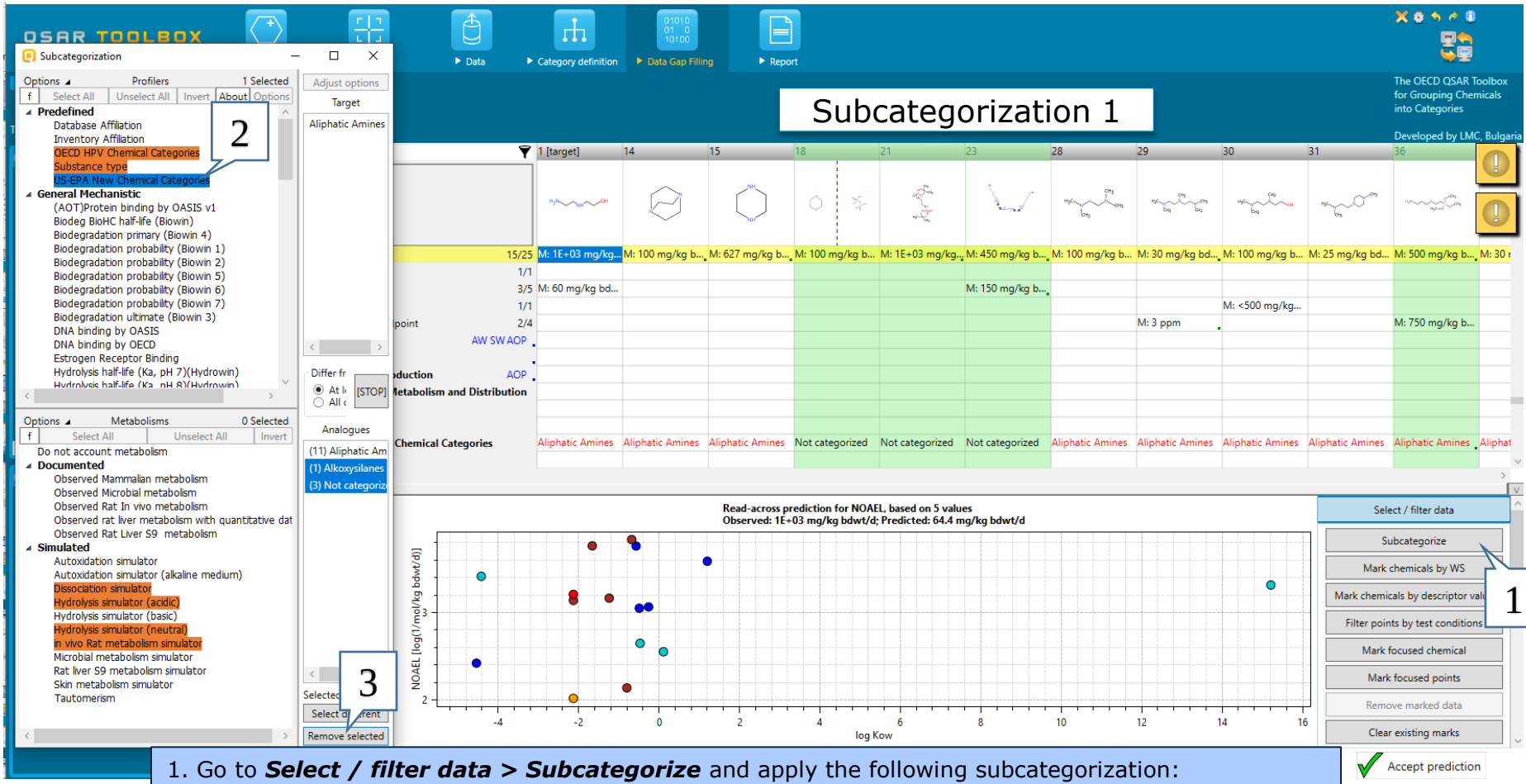
Data Gap Filling

Apply worst-case scenario



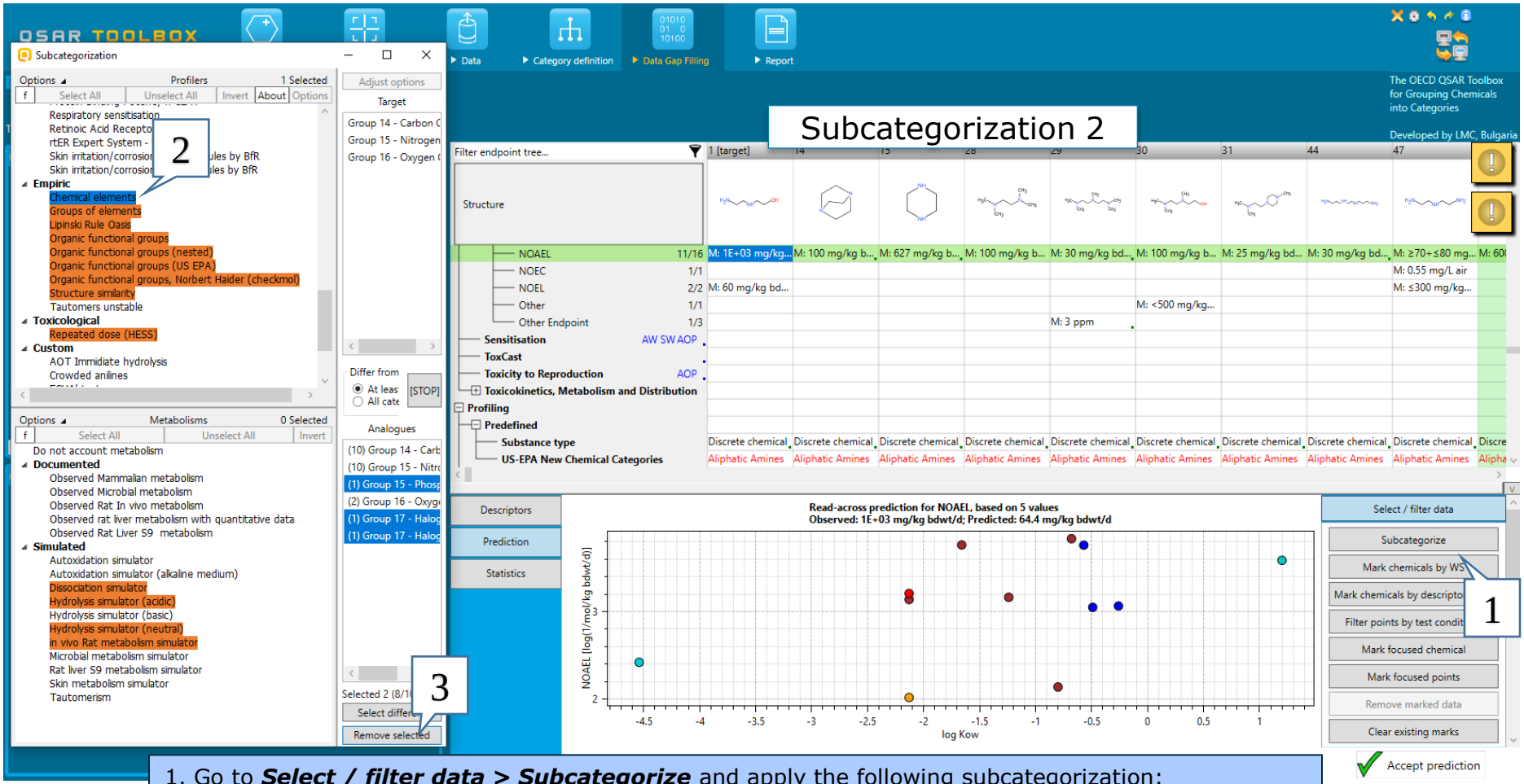
Apply the worst case scenario: 1) Go to ***Calculation options > Data usage***; 2) Click ***Maximal*** ratio button; 3. Confirm with ***OK***.

Data Gap Filling
Subcategorize



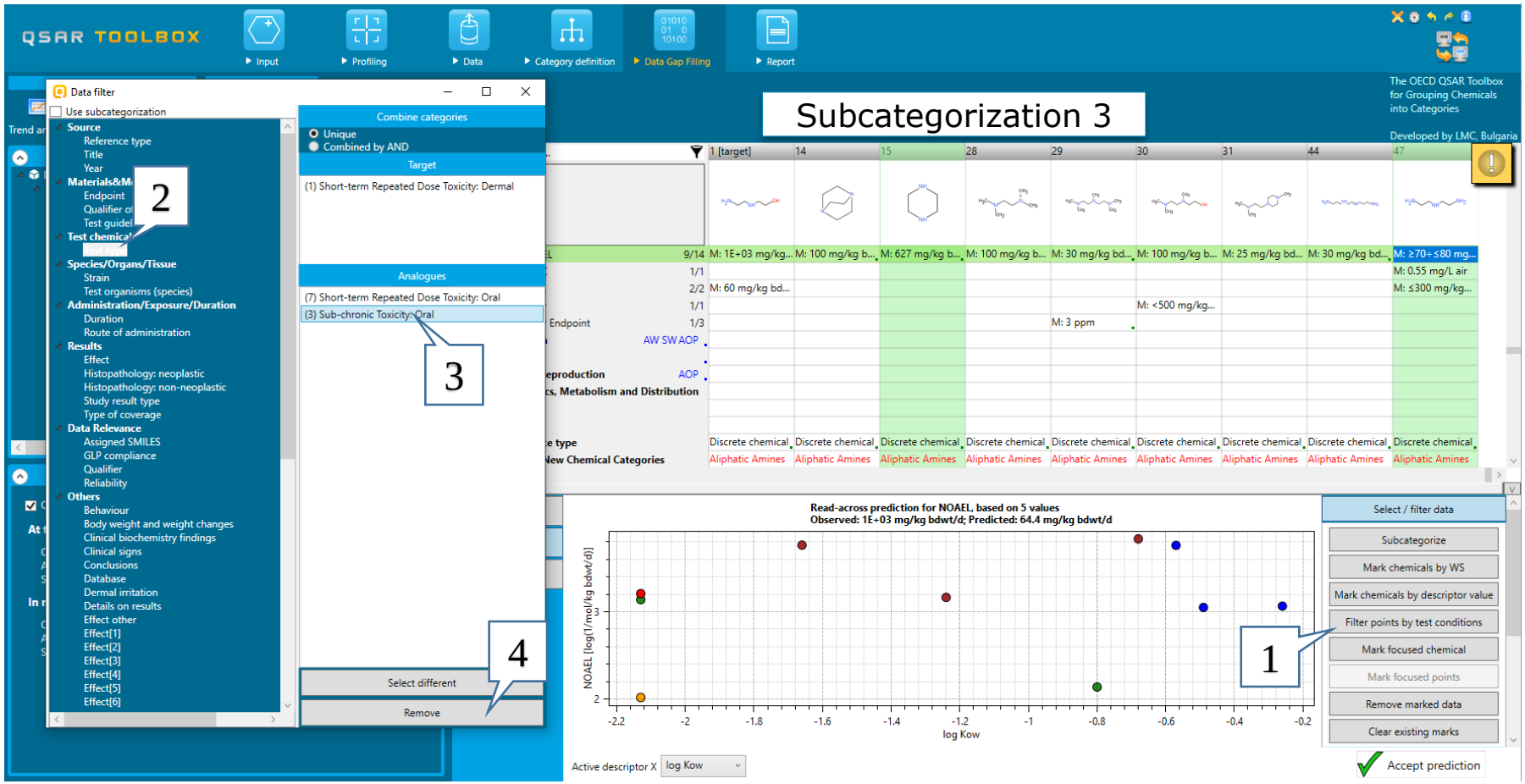
1. Go to ***Select / filter data > Subcategorize*** and apply the following subcategorization:
2. *US-EPA New Chemical Categories*.
3. Eliminate dissimilar chemicals after applied subcategorization using the ***Remove selected*** button.

Data Gap Filling
Subcategorize



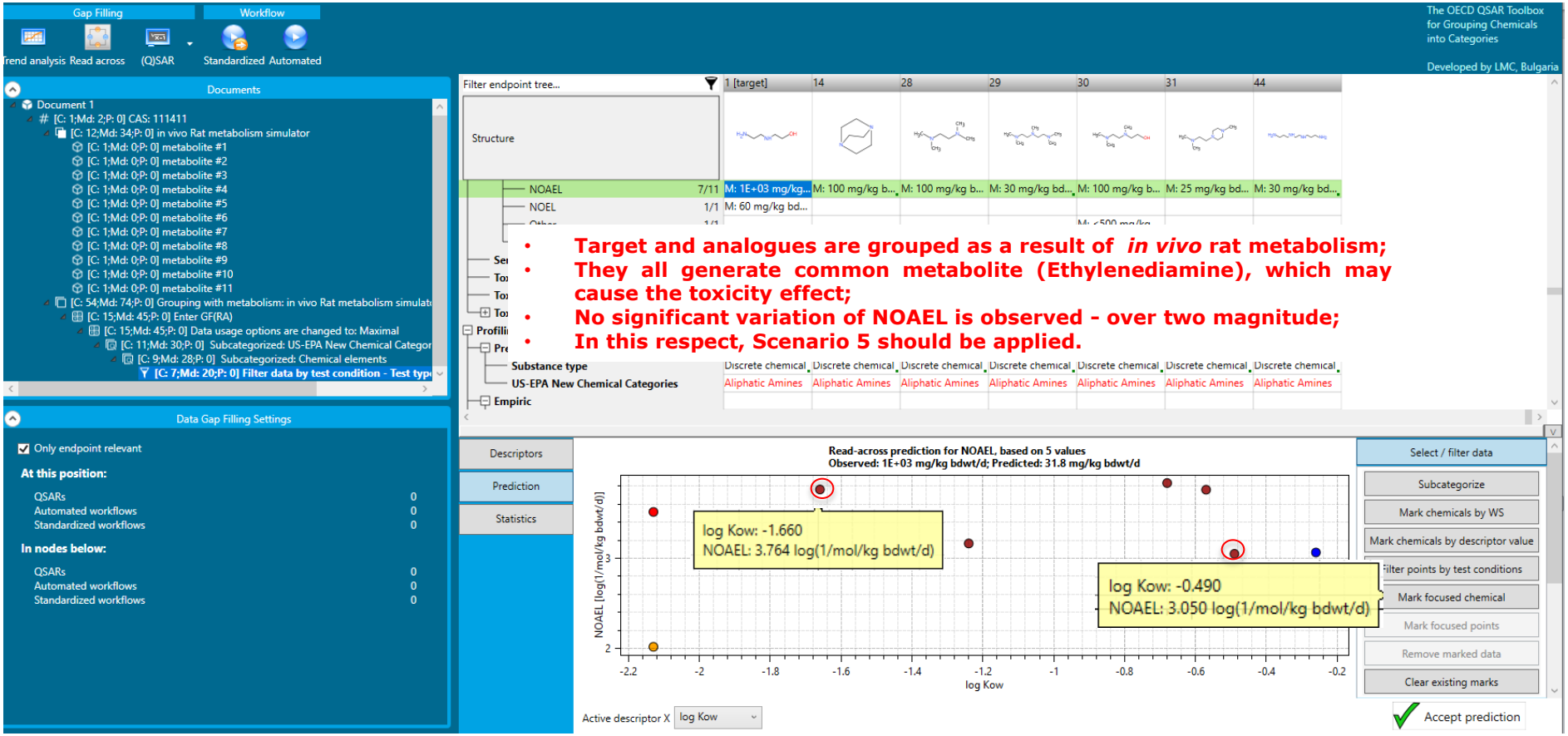
1. Go to ***Select / filter data > Subcategorize*** and apply the following subcategorization:
2. *Chemical elements*.
3. Eliminate dissimilar chemicals after applied subcategorization using the ***Remove selected*** button.

Data Gap Filling
Subcategorize



1. Go to ***Select / filter data > Filter points by test conditions***
2. *Select* ***Test chemical > Test type***
3. Eliminate *Sub-chronic Toxicity Oral* (select it manually) and click the ***Remove*** button (4).

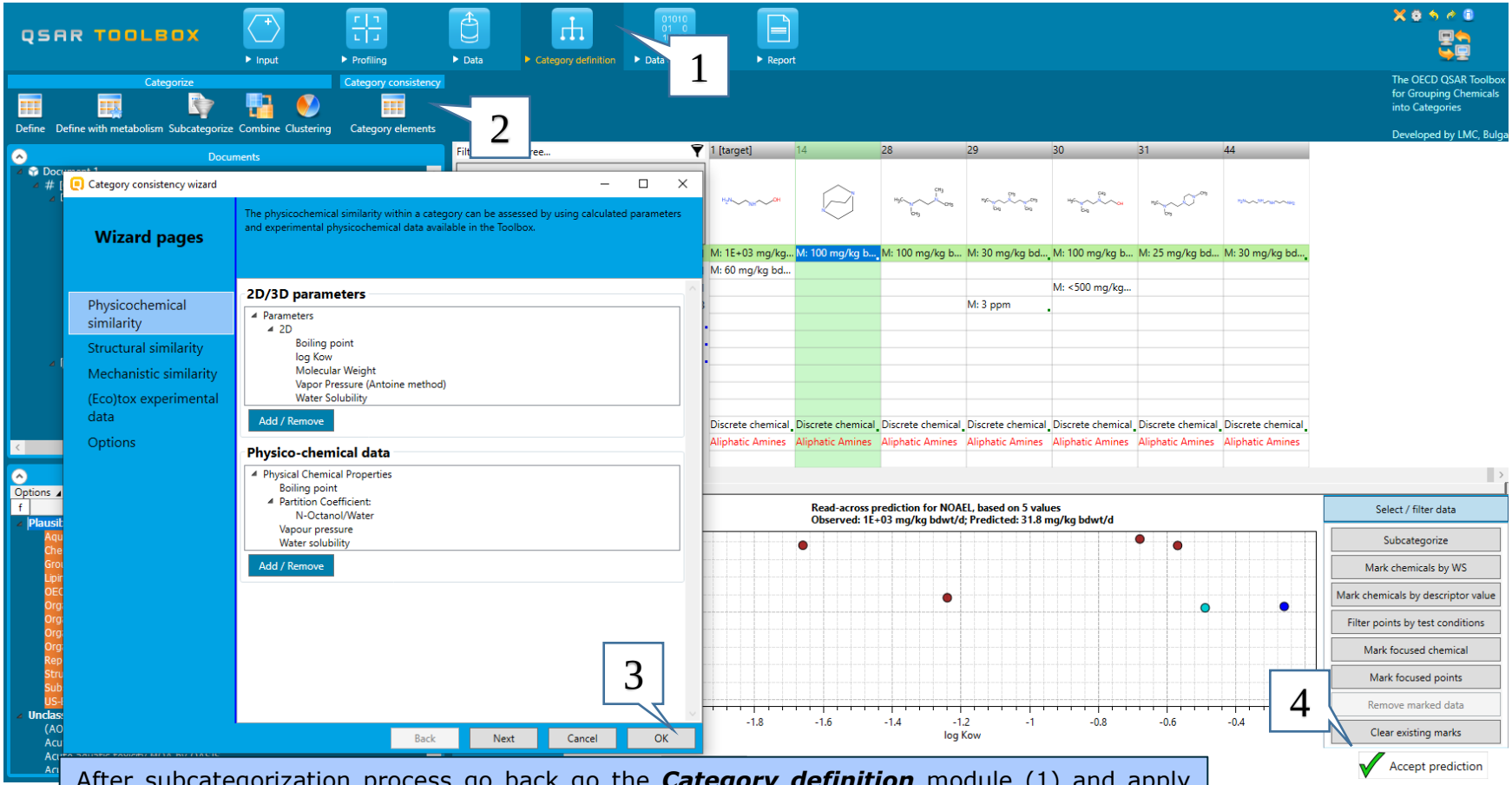
Data Gap Filling
Data variation



- **Target and analogues are grouped as a result of *in vivo* rat metabolism;**
- **They all generate common metabolite (Ethylenediamine), which may cause the toxicity effect;**
- **No significant variation of NOAEL is observed - over two magnitude;**
- **In this respect, Scenario 5 should be applied.**

NOAEL is in the range from 3.05 to 3.76 log (1/mol/kg bdwt/d) for the 5 analogues

Data Gap Filling
Category consistency check



After subcategorization process go back go the ***Category definition*** module (1) and apply ***Category elements**** (2). No different selection than the default is needed – click ***OK*** (3). Once the category elements are applied ***Accept prediction*** (4).

Recap

- In the *Category definition* module you found 27 chemicals having a common metabolite (Ethylenediamine) as a result of *in vivo* rat metabolism.
- All chemicals have data for the defined endpoint.
- In *Data gap filling* module you applied a read-across approach. As a result of subcategorizations the number of analogues was reduced to 5.
- No significant variation of NOAEL data was observed for the 5 closest analogues.
- Category consistency was checked by applying the category elements.
- You are now ready to complete the final module and to create the report.
- Click "Report" to proceed to the last module.

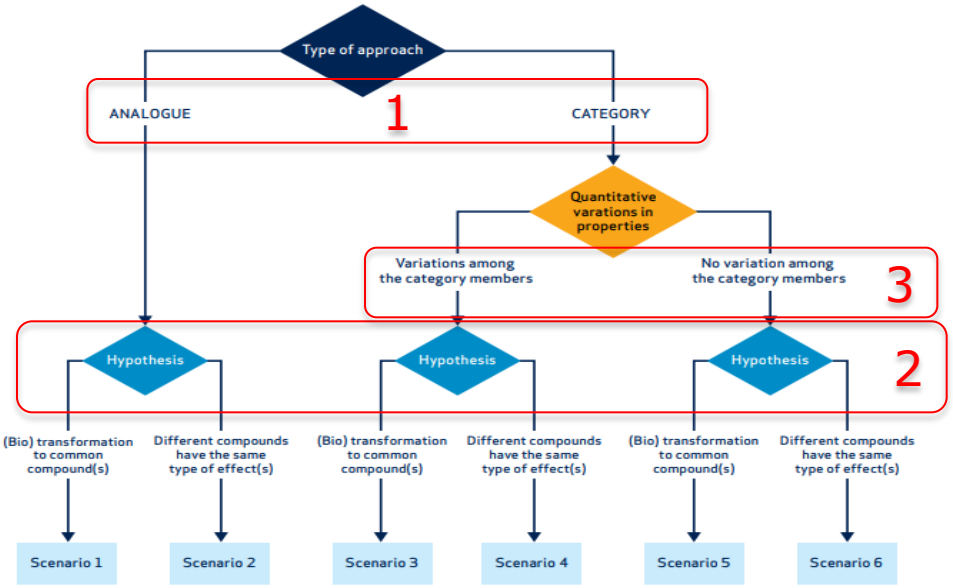
Report
Overview

- The report module allows generating a report for predictions performed within the Toolbox.
- The report module contains a predefined report template which users can customize.
- Additionally specific RAAF scenario could be chosen. Selection of one of the scenarios will append automatically the related assessment elements related to the corresponding report sections.

Report
Selection of RAAF scenario

To select the applicable RAAF scenario for assessment, the following aspect should be identified*:

1) the type of approach applied - analogue approach or category approach;
2) the read-across hypothesis;
3) For category approach - whether quantitative variations in the properties are observed among the category members must be considered.



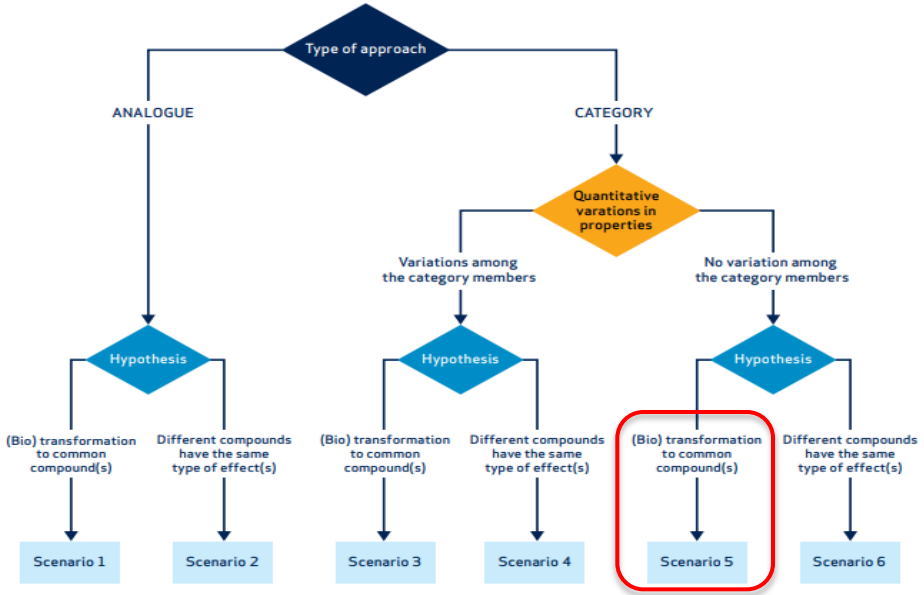
*Read-Across Assessment Framework (RAAF) available at https://echa.europa.eu/documents/10162/13628/raaf_en.pdf

Report
Selection of RAAF scenario

For this example the following criteria are met :

- the type of approach applied - category approach is used (threshold of > 3 analogues is proposed by LMC for the category approach);
- the read-across hypothesis – different compounds (bio)transformed to the common compound;
- There is no significant variation in the property under investigation (NOAEL) among the category members

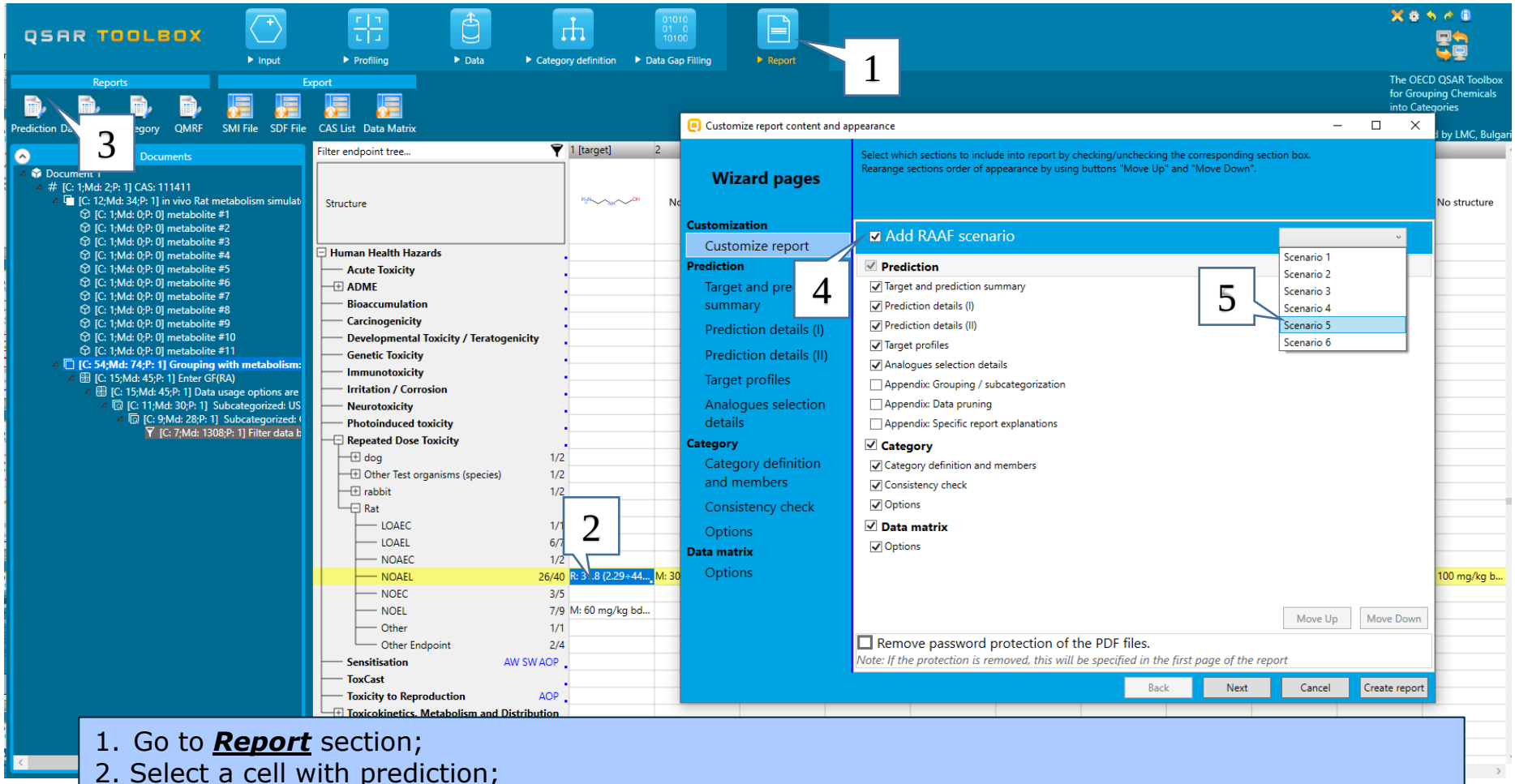
Based on that RAAF scenario 5 was identified as the most appropriate for the current example.



*Read-Across Assessment Framework (RAAF) available at https://echa.europa.eu/documents/10162/13628/raaf_en.pdf

Report

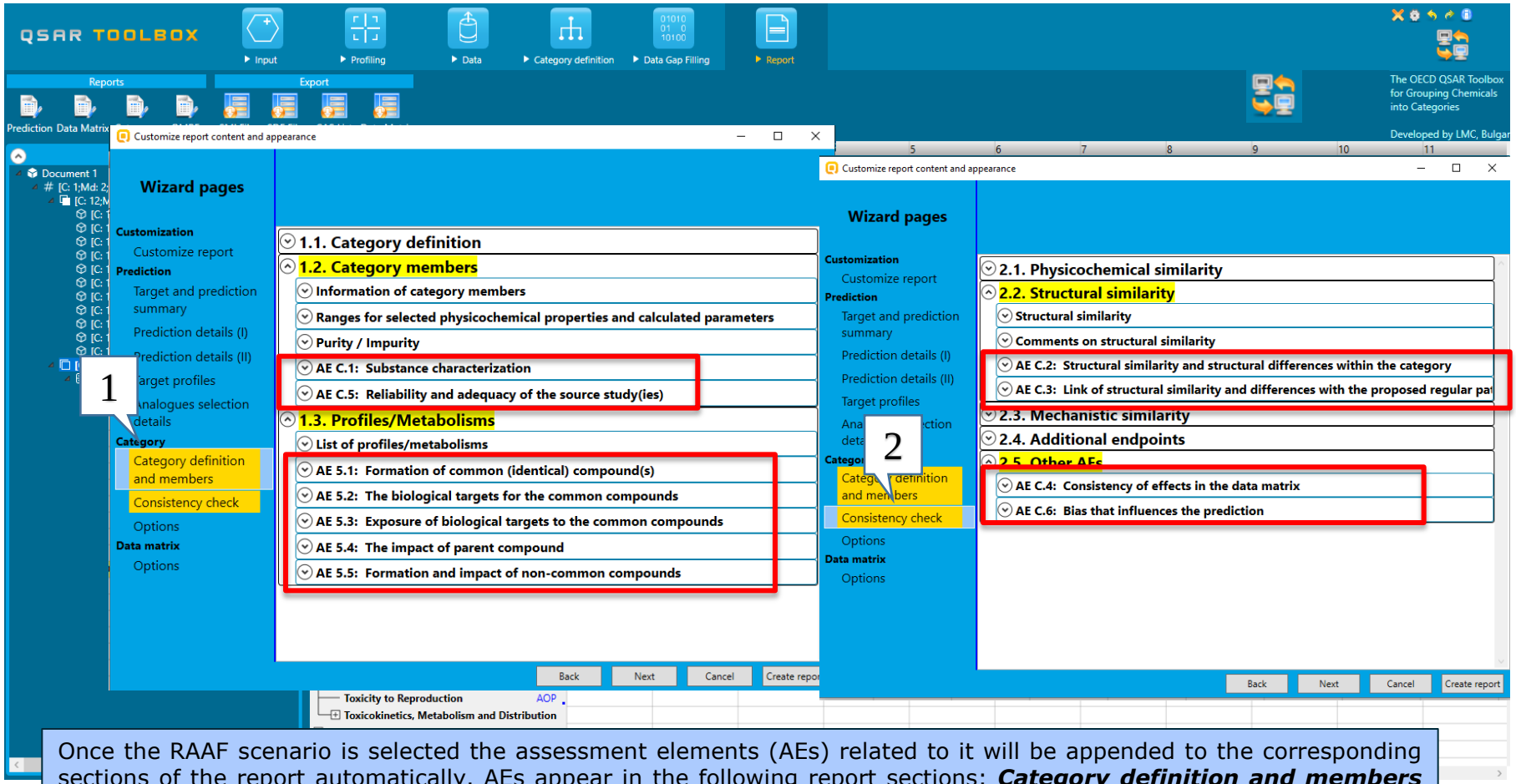
Generation report according to RAAF-Scenario 5



1. Go to ***Report*** section;
2. Select a cell with prediction;
3. Click ***Prediction***;
4. Check the box at the top to add RAAF scenario;
5. Select ***Scenario 5*** from the drop-down menu.

Report

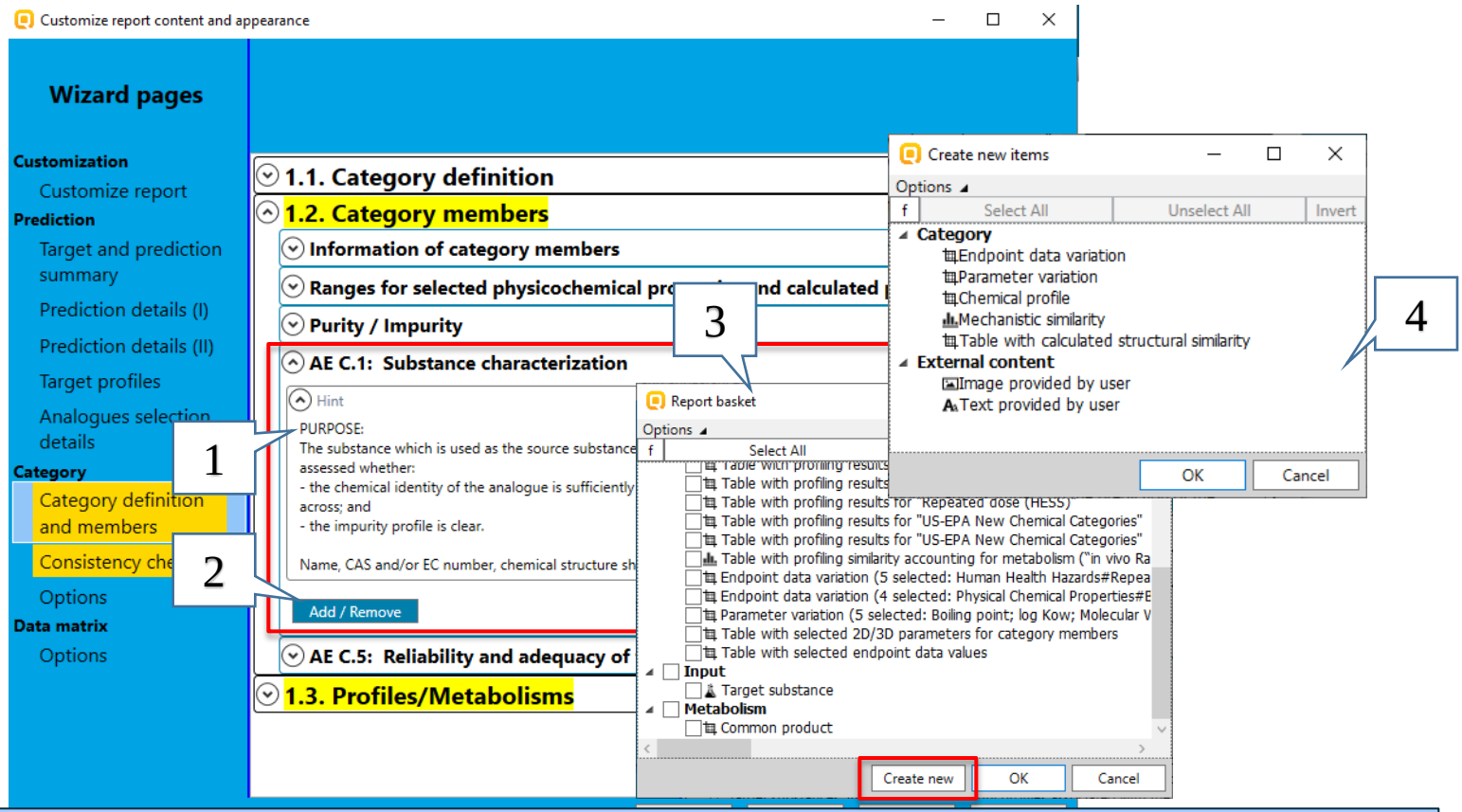
Generation report according to RAAF-Scenario 5



Once the RAAF scenario is selected the assessment elements (AEs) related to it will be appended to the corresponding sections of the report automatically. AEs appear in the following report sections: ***Category definition and members*** (1) and ***Consistency check*** (2).

Each of the AEs will be considered in the next slides.

Report
Assessment elements of Scenario 5



Hint for each of the assessment elements is available (1). Information can be included by the ***Add/Remove*** button (2) located below the corresponding AE. The ***Add/Remove*** button invokes so-called *Report basket* (3). The latter contains different items triggered by the actions of the user during the workflow (e.g. Alert performance calculation, applying of category elements, etc.).
Additionally, new items (including items with external content) can be created (4).

Items with external content (picture and text) will be added for some AEs.

Report
Assessment elements of Scenario 5

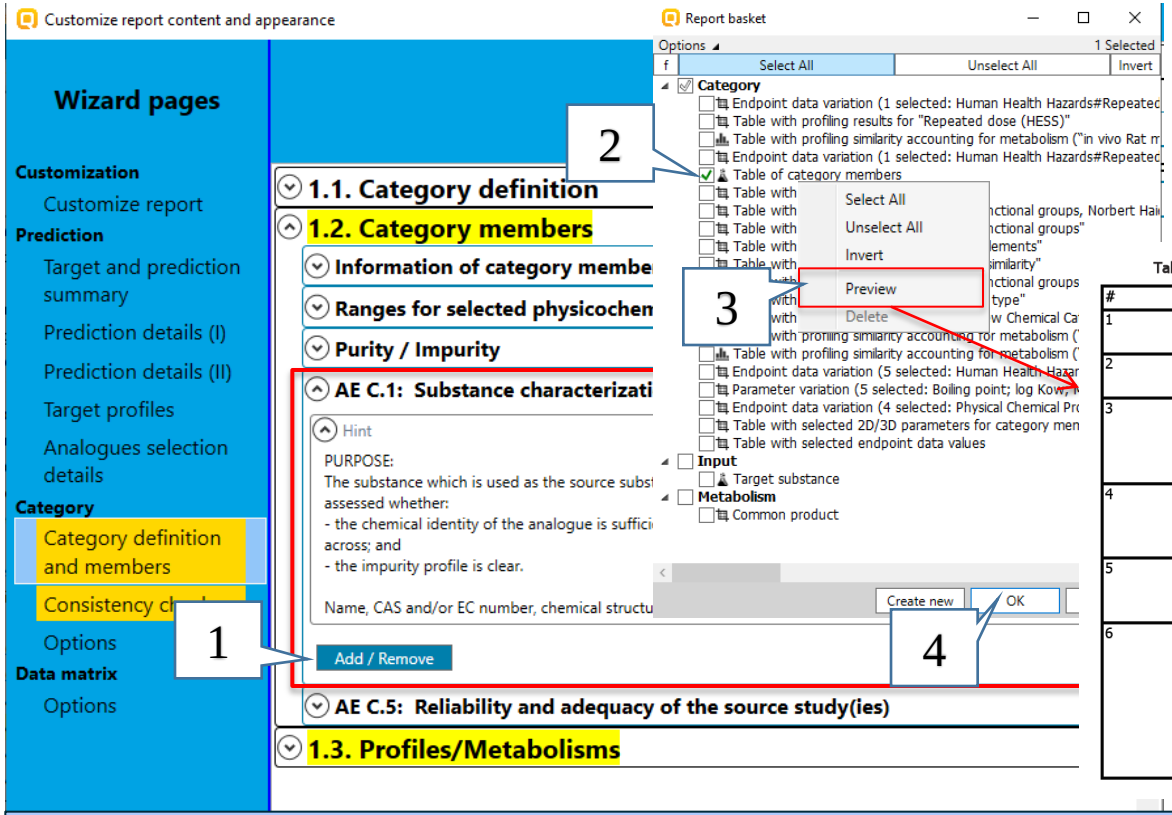
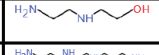
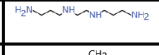
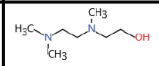
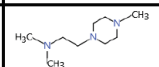
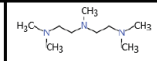
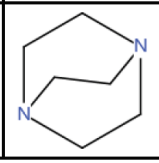


Table of category members

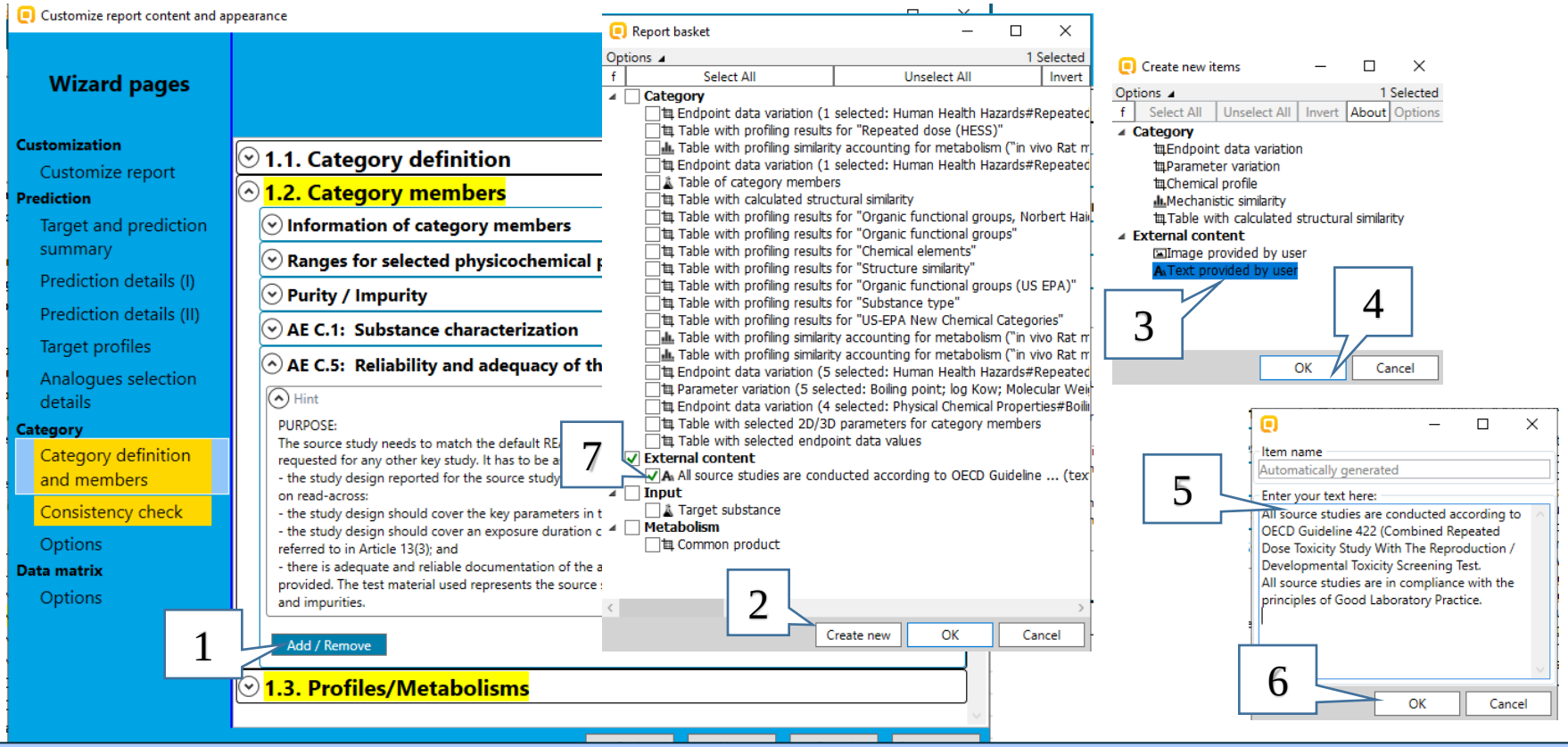
#	CAS	Name	SMILES	Structure
1	111-41-1	Aminoethyl ethanolamine	NCCNCCO	
2	10563-26-5	1,3-Propanediamine, N,N''-ethylenebis-	NCCCNCCNCCCN	
3	2212-32-0	2-[[2-(dimethylamino)ethyl]methylamino]ethanol	CN(C)CCN(C)CCO	
4	104-19-8	N,N,4-trimethylpiperazine-1-ethylamine	CN(C)CCN1CCN(C)CC1	
5	3030-47-5	bis(2-dimethylaminoethyl)(methyl)amine	CN(C)CCN(C)CCN(C)C	
6	280-57-9	TEDA	C1CN2CCN1CC2	

Five AE (AE C.1, 5.1, 5.2, 5.3 and 5.5) related to Scenario 5 are included in the *Category definition and members* section.

AE C.1: Substance characterization

Click ***Add/Remove*** button (1) in the second part of AE C.1. Check the box next to ***Table of category members*** item (2). Right click over the item and select ***Preview*** to see the content (3). Finally confirm by ***OK*** (4). By this way a list with category members is added to the report

Report
Assessment elements of Scenario 5



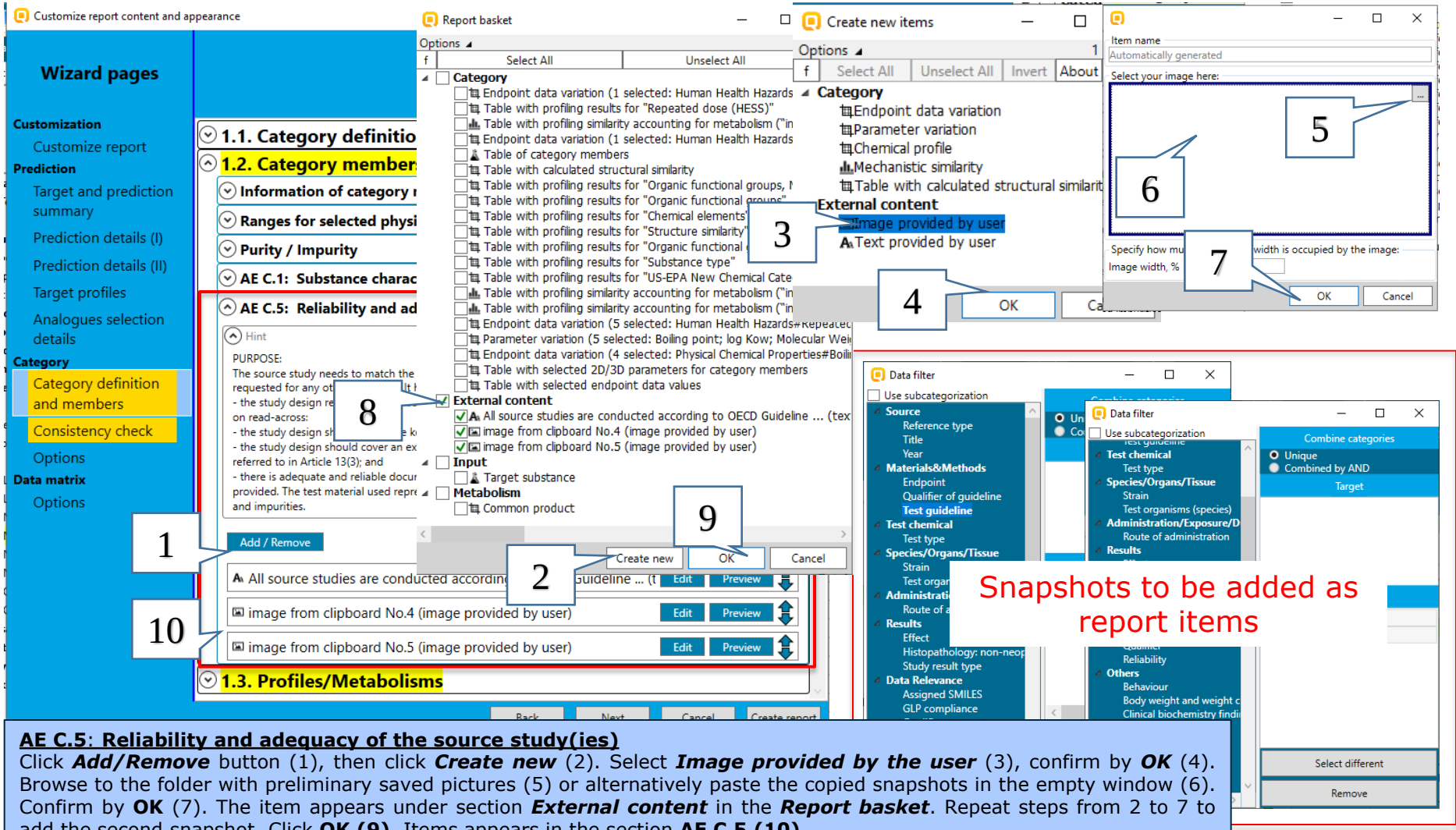
AE C.5: Reliability and adequacy of the source study(ies)

The example text can be added for AE C.5. Click Add/Remove button (2), then ***Create new*** (2), select ***Text provided by user*** (3) and click ***OK*** (4). Copy the following example text:

- *All source studies are conducted according to OECD Guideline 422 (Combined Repeated Dose Toxicity Study With The Reproduction / Developmental Toxicity Screening Test.*
- *All source studies are in compliance with the principles of Good Laboratory Practice.*

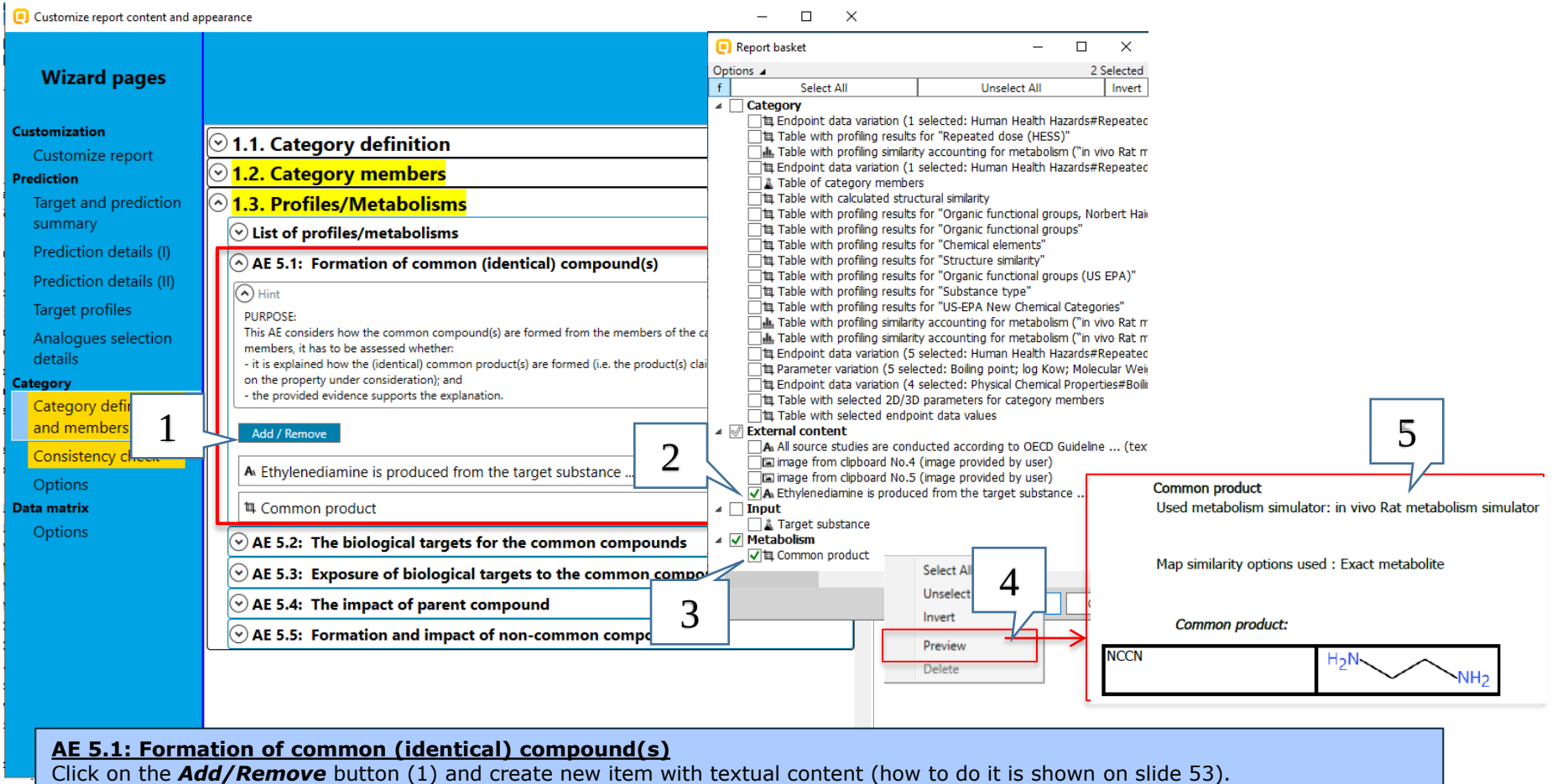
And paste it in the appeared windows (5). Confirm by **OK** (6). The item appears in the Report basket (7). Additionally snapshots of the "filter by test conditions" window could be added to confirm the consistency regarding the guideline and GLP compliance (see next slide).

Report
Assessment elements of Scenario 5



AE C.5: **Reliability and adequacy of the source study(ies)**
Click ***Add/Remove*** button (1), then click ***Create new*** (2). Select ***Image provided by the user*** (3), confirm by ***OK*** (4). Browse to the folder with preliminary saved pictures (5) or alternatively paste the copied snapshots in the empty window (6). Confirm by **OK** (7). The item appears under section ***External content*** in the ***Report basket***. Repeat steps from 2 to 7 to add the second snapshot. Click **OK (9)**. Items appears in the section **AE C.5 (10)**.

Report
Assessment elements of Scenario 5



AE 5.1: Formation of common (identical) compound(s)

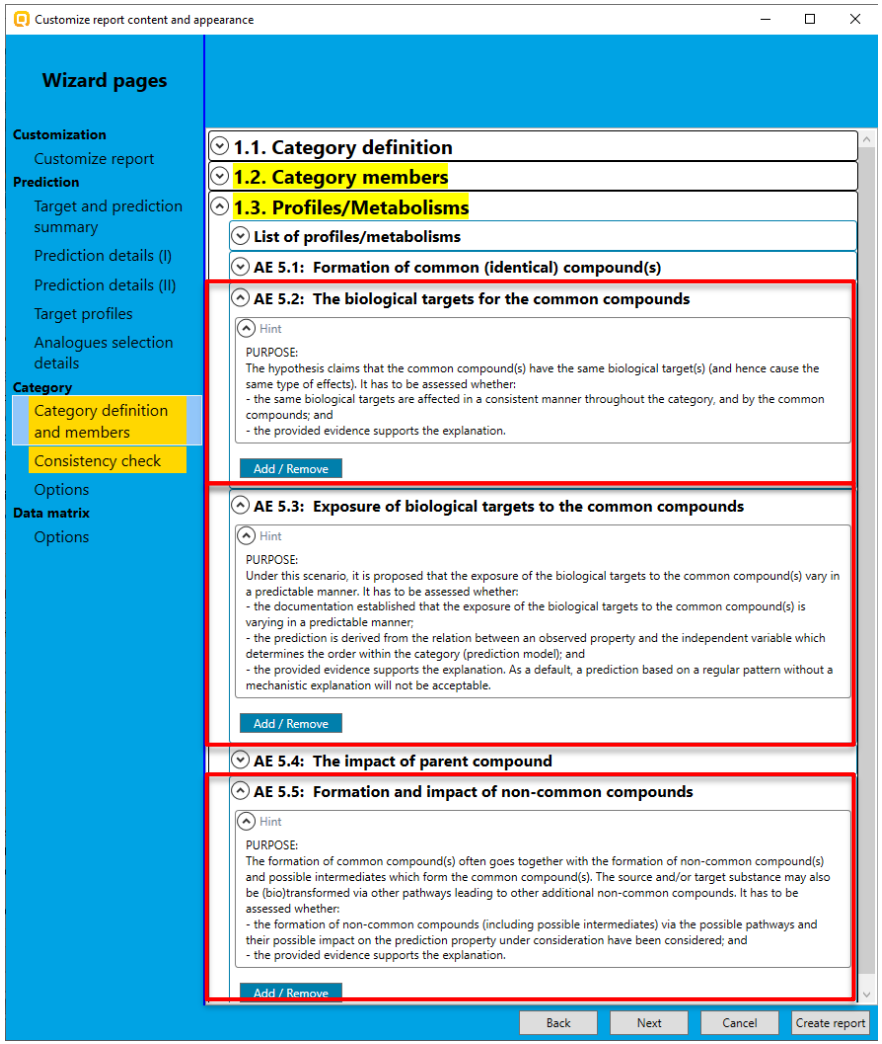
Click on the ***Add/Remove*** button (1) and create new item with textual content (how to do it is shown on slide 53).

In the text field you can paste the following example text:

- *Ethylenediamine is produced from the target substance A and the source substances B-E by in vivo rat metabolism simulator*
- *Alert for repeated dose toxicity and experimental data for the property under consideration are found for the common metabolite.*

The item appears in the Report basket (2). Once the text item is created, check the box next to the *Common product* item (3). Right click over the item and select ***Preview*** (4) to see the content (5). Finally confirm by ***OK***.

Report
Assessment elements of Scenario 5



Example text for **AE 5.2: The biological targets for the common compounds**

- *The target and source substances form a common metabolite: Ethylenediamine.*
- *No alerts are identified in the structures of the Target A.*
- *The common compound is supposed that may cause the toxic effect.*

Example text for **AE 5.3: Exposure of biological targets to the common compounds**

- *Target chemical A and source substances from B to F are metabolized to the common reactive product: Ethylenediamine.*
- *It is considered that for low molecular weight aliphatic amines the most important biological effect is the acute effect, especially the strong local irritation which they can evoke.*

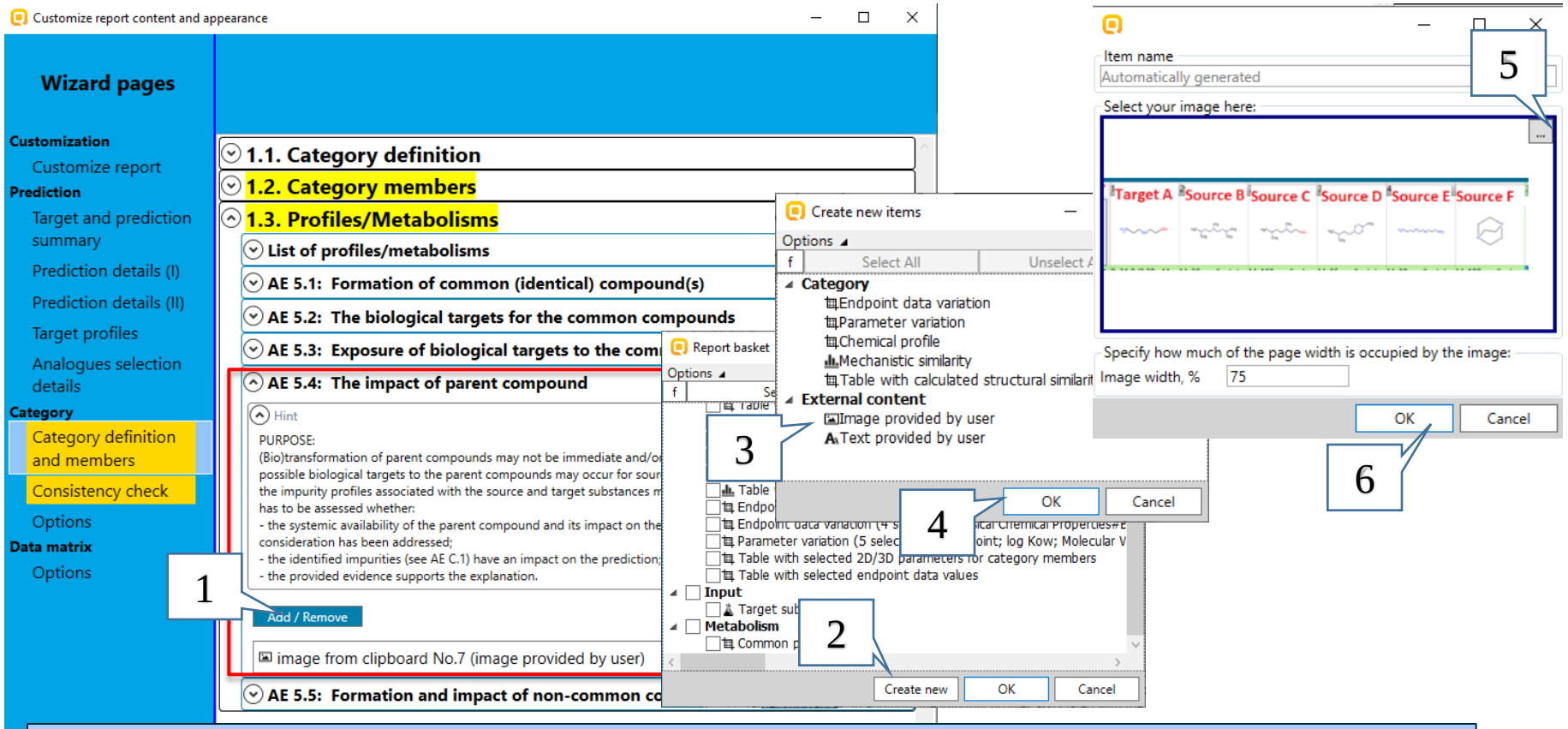
References:

1. F. Gagnaire et al. J Appl Toxicol. 1993; 13: 129.

Example text for **AE 5.5: Formation and impact of non-common compounds**

- *The target substance A and the five source substances (analogues) are metabolized to the common - Ethylenediamine*
- *The positive effect of Ethylenediamine is supported by experimental NOAEL data.*
- *Another alerts related to Mucous membrane irritation and Renal toxicity are identified in some of the produced non-common compounds.*
- *The lowest experimental NOAEL value was found for the common metabolite.*
- *The common compound is supposed to be responsible for the repeated dose toxicity effect.*

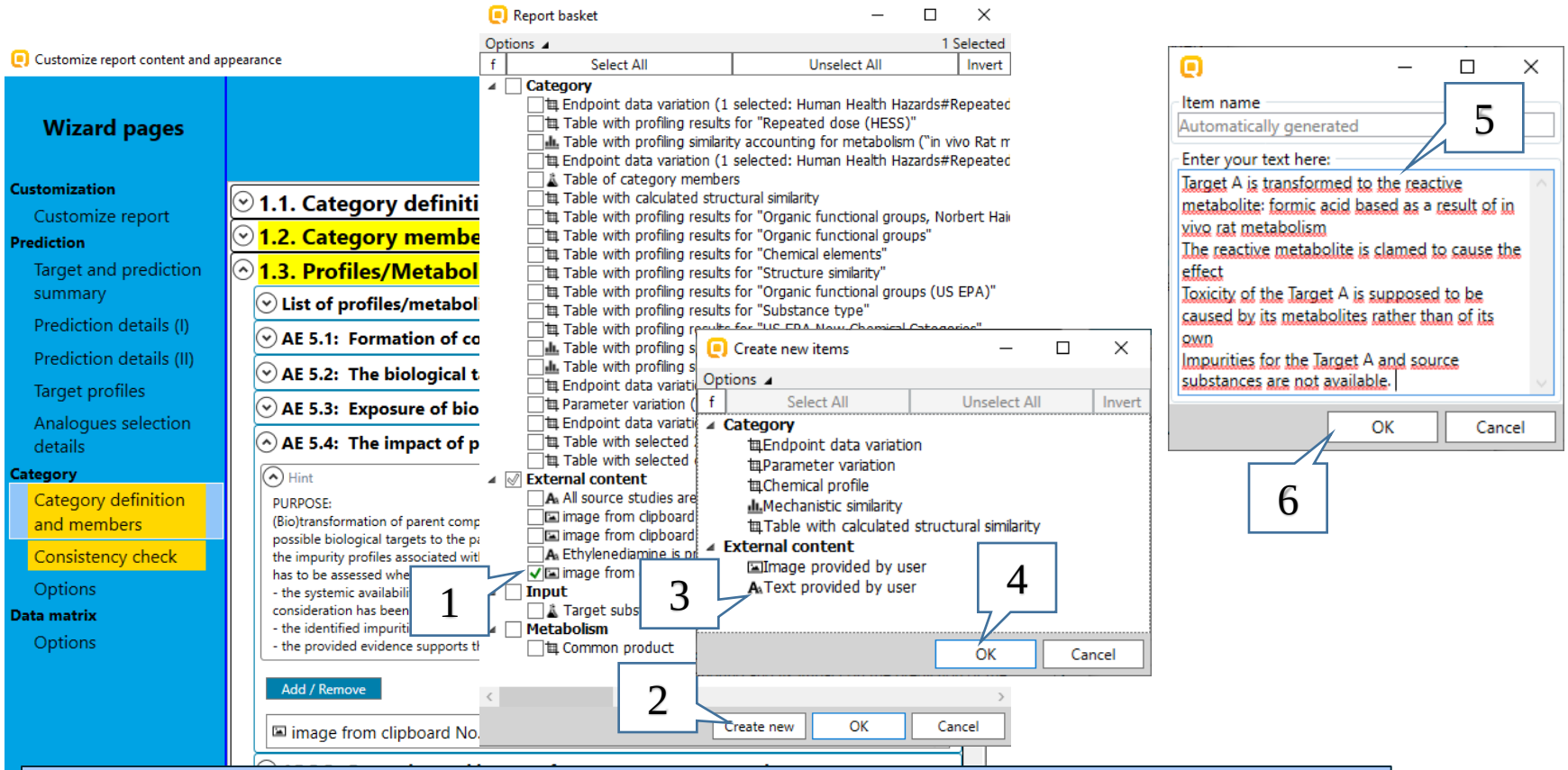
Report
Assessment elements of Scenario 5



Click on the ***Add/Remove*** button (1) and then ***Create new*** (2). Select to create item with external content – ***Image provided by user*** (3) and click ***OK*** (4). New window appears where you can add your custom picture by Copy/Paste or browsing (5) to the directory in your PC where the desired picture is saved*. Finally confirm by ***OK*** (6).

*In the current example a picture illustrating the target chemical marked as **Target A** and source chemicals marked as **Source B**, **C, D, E, F** was prepared in advance.

Report
Assessment elements of Scenario 5

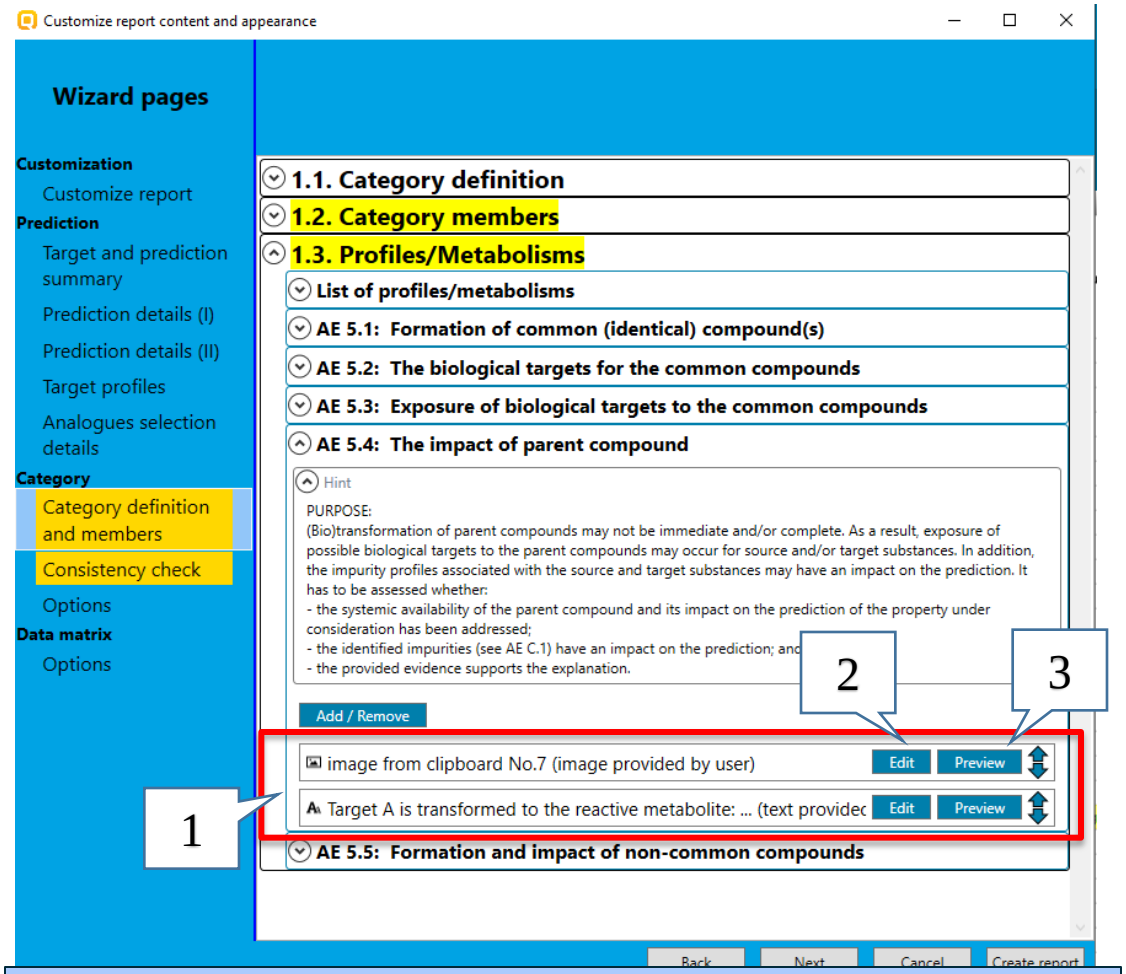


The newly created item appears in the *Report basket* (1).
Now text will be also included. Click ***Create new*** (2), select ***Text provided by user*** (3) and click ***OK*** (4).
Copy the following example text:

- *Target A is transformed to the reactive metabolite: Ethylenediamine based on the result of in vivo rat metabolism*
- *The reactive metabolite is clamed to cause the effect*
- *Toxicity of the Target A is supposed to be caused by its metabolites rather than of its own*
- *Impurities for the Target A and source substances are not available*

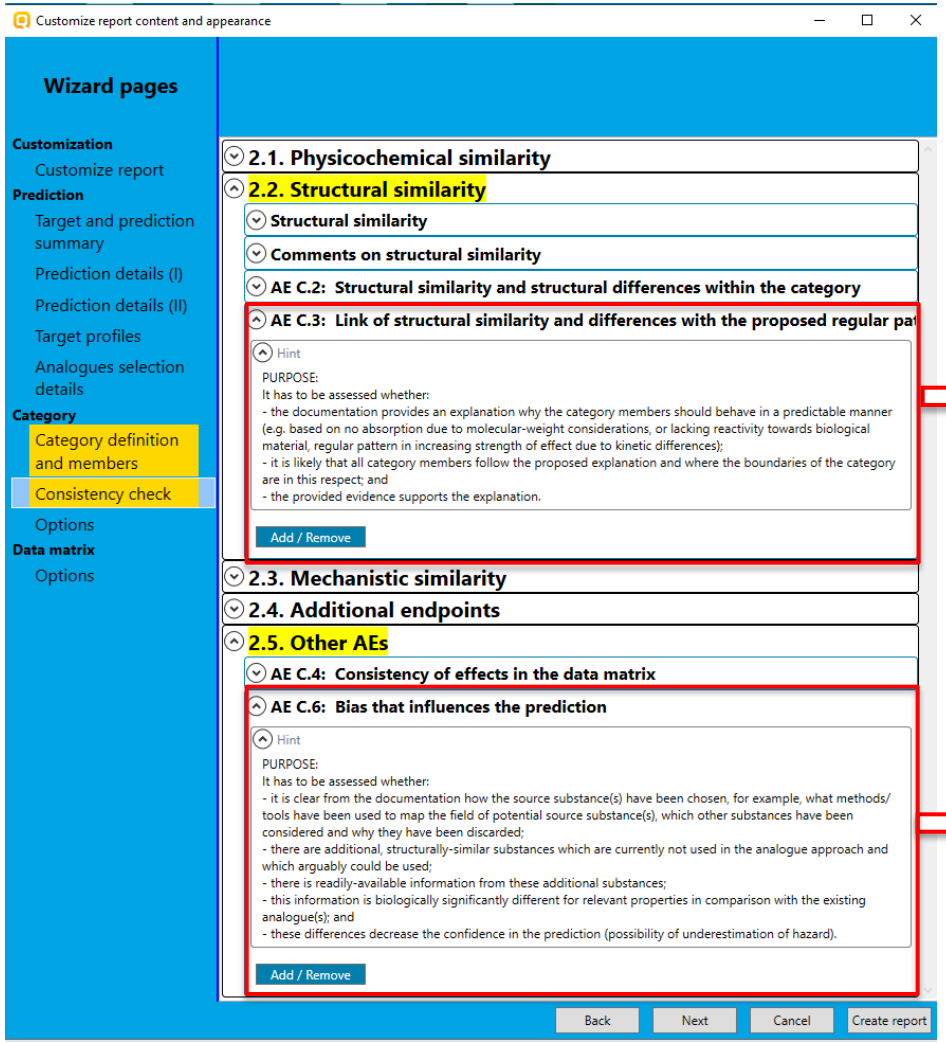
and paste it in the new window (5). Finally confirm by ***OK*** (6).

Report
Assessment elements of Scenario 5



Both newly created items appear under the **AE 5.4** (1). Each of the items can be ***edited*** (2) or just ***previewed*** (3) in a *.pdf* format.

Report
Assessment elements of Scenario 5



Four AEs are included to the *Consistency check* section. Example content for two of them (AE C.3 and AE C.6) is given below.

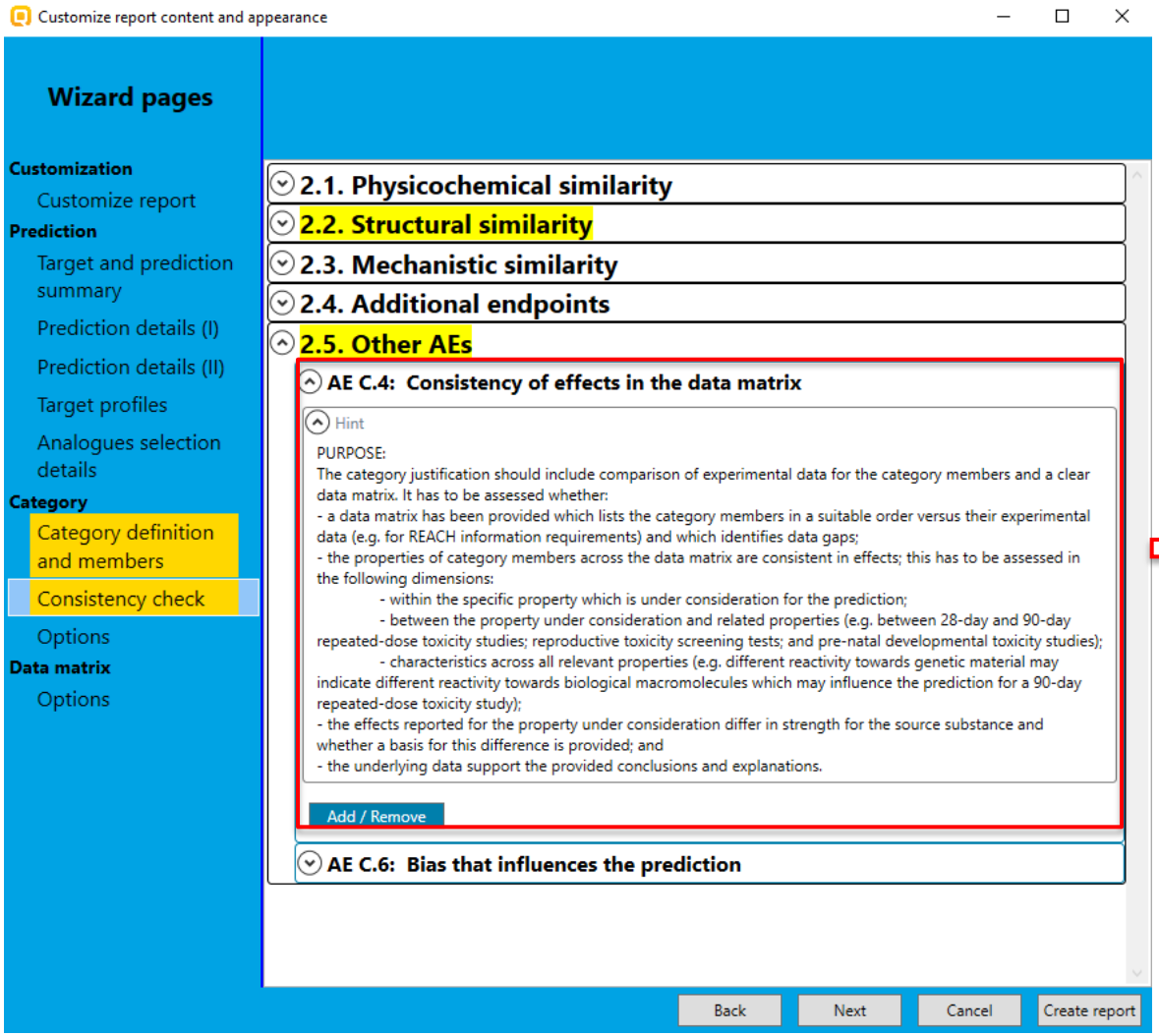
Example text for **AE C.3: Link of structural similarity and differences with the proposed regular pattern**

- *The category is structurally defined as target (A) and five source substances (B, C, D, E, F) all form a common product – Ethylenediamine*
- *They all consist of common reactivity pattern responsible for the formation of reactive metabolites*

Example text for **AE C.6: Bias that influences the prediction**

- *Source substances for the target chemical A have been searched based on formation of a specific metabolite as a result of in vivo rat metabolism;*
- *On the next level all analogues that differ from the Target A according to US-EPA New Chemical Categories profiling scheme have been removed.*
- *Five source substances with no significant variation in the property under consideration were used for the prediction.*

Report
Assessment elements of Scenario 5



AE C.4: Consistency of the effects in the data matrix

The following example text can be added for AE C.4 by analyzing the structural similarity items:

- *Physico-chemical properties, identified alerts and experimental data along with the characteristics of the studies (species, duration, test type, references, etc.) are provided in the generated Data matrix file.*

Report
Assessment elements of Scenario 5

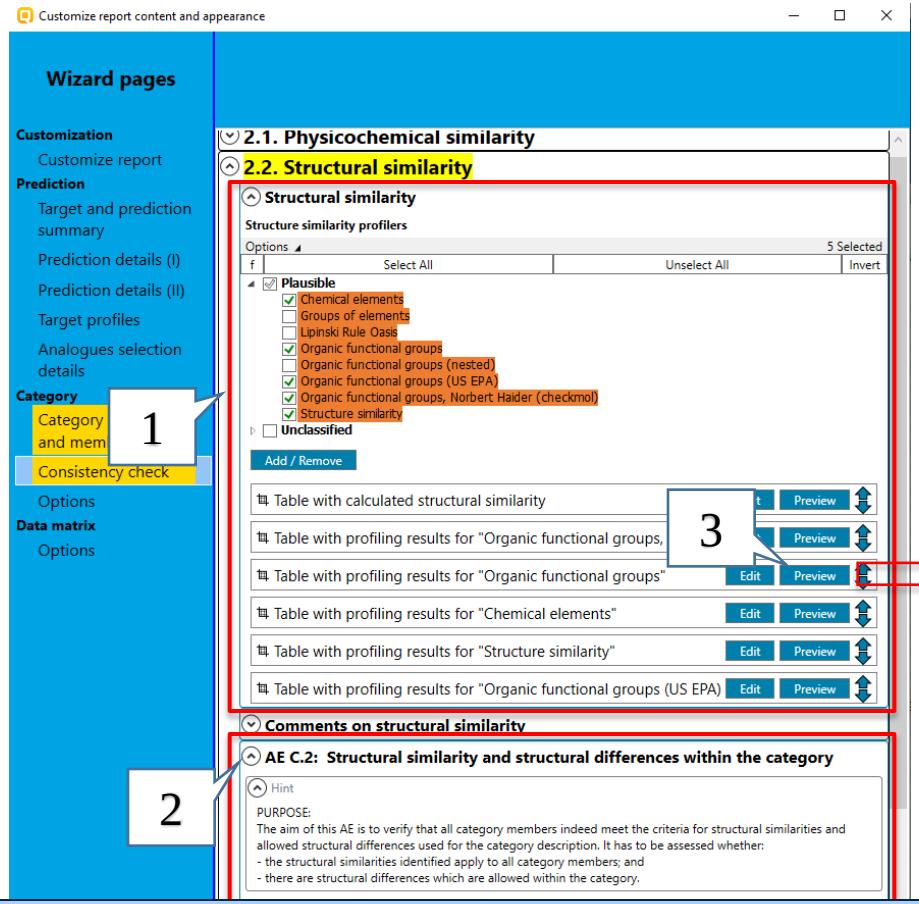
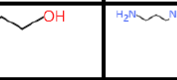
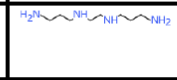
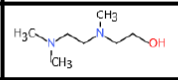
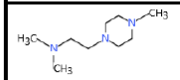
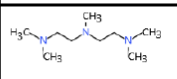
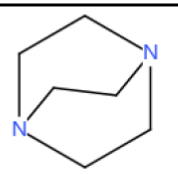


Table with profiling results for "Organic functional groups"

1 CAS# 111-41-1	2 CAS# 10563-26-5	3 CAS# 2212-32-0
		
Alcohol Amine, primary Amine, secondary Aliphatic amine, primary Aliphatic amine, secondary	Amine, primary Amine, secondary Aliphatic amine, primary Aliphatic amine, secondary	Alcohol Amine, tertiary Aliphatic amine, tertiary

4 CAS# 104-19-8	5 CAS# 3030-47-5	6 CAS# 280-57-9
		
Amine, tertiary Saturated heterocyclic fragment Piperazine Aliphatic amine, tertiary	Amine, tertiary Aliphatic amine, tertiary	Amine, tertiary Saturated heterocyclic fragment Piperazine Bridged-ring heterocycles Aliphatic amine, tertiary

AE C.2. Link of structural similarity and structural differences within the category is related to the structural similarity of the final category.
All items in the report basket related to the structural consistency of the category (1) are added automatically.
The following example text can be added for AE C.2. by analyzing the structural similarity items:

- *All category members are identified as Aliphatic amines.*

Preview of the item is possible once click on the Preview button (3).

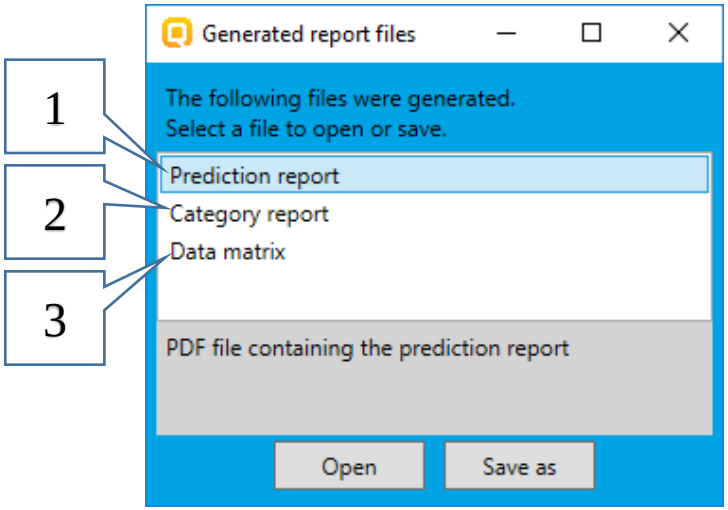
Report
Generation report

After clicking on the *Create report* button, the *Generated report files* window appears. It contains three types 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.

RAAF AEs are included in the second file.

All generated files should be provided when submitting a prediction.



Report
Generated report files

Prediction report

QSAR Toolbox prediction for single chemical

(in accordance with RAAF scenario 5)

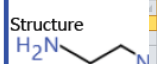
The selected RAAF scenario is specified in the first page

Date: 15 Apr 2020
Author(s):
Contact details:

Target information

Structural information	Numerical identifiers	Chemical names
SMILES: NCCNCCO	CAS#: 111-41-1 Other: EC Number:2038675	1-amino-2-(2-hydroxy ethyl)aminoethane

Structure



Predicted endpoint
Predicted value:
Unit/scale: mg/kg
Data gap filling m
Summary: manua
Not provided by t

Category report

QSAR Toolbox report for category

(in accordance with RAAF scenario 5)

1. Category definition

1.1. Category definition

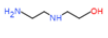
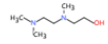
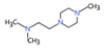
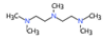
Category name *manually editable field*
Not provided by the user

Covered (target) endpoint(s)
- Human Health Hazards/Repeated Dose Toxicity: NOAEL, Rat

Category hypothesis *manually editable field*
Not provided by the user

1.2. Category members

Data matrix report

Substance identity		Target chemical	Neighbour #1	Neighbour #2	Neighbour #3	Neighbour #4	Neighbour #5
Structure							
CAS number		111-41-1	10563-26-5	2212-32-0	104-19-8	3030-47-5	280-57-9
Chemical name		Aminoethyl ethanolamine	1,3-Propanediamine, N,N''-ethylenebis-	2-[[2-(dimethylamino)ethyl]methylamino]et	N,N,4-trimethylpiperazine-1-ethylamine	bis(2-dimethylaminoethyl)(methyl)amine	TEDA
Other identifier							
SMILES		NCCNCCO	NCCCNCCNCCCN	CN(C)CCN(C)CCO	CN(C)CCN1CCN(C)CC1	CN(C)CCN(C)CCN(C)C	C1CN2CCN1CC2
Parameters	**unit**						
Boiling point	°C	211	293	219	226	204	152
log Kow		-2.13	-1.66	-1.24	-0.68	-0.57	-0.49
Molecular Weight	Da	104	174	146	171	173	112
Vapor Pressure (Antoine method)	mm Hg	0.00693	0.000569	0.025	0.0754	0.319	0.0596
Water Solubility	mg/L	1000000	1000000	1000000	1000000	1000000	1000000
Profilers							
Profiles used for grouping/subcategorization							
Using of "in vivo Rat metabolism simulator"		metabolite #1; Aliphatic Amines	metabolite #1; Aliphatic Amines	metabolite #1; Aliphatic Amines	metabolite #1; Aliphatic Amines	metabolite #1; Aliphatic Amines	metabolite #1; Aliphatic Amines
US-EPA New Chemical Categories		Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines
Predefined							
Substance type		Discrete chemical	Discrete chemical	Discrete chemical	Discrete chemical	Discrete chemical	Discrete chemical
Substance type, with in vivo Rat metabolism		11 x Discrete chemical	11 x Discrete chemical	24 x Discrete chemical	26 x Discrete chemical	21 x Discrete chemical	7 x Discrete chemical
US-EPA New Chemical Categories, with in vivo		4 x Aldehydes (Acute toxicity)	4 x Aldehydes (Acute toxicity)	7 x Aldehydes (Acute toxicity)	7 x Aldehydes (Acute toxicity)	5 x Aldehydes (Acute toxicity)	3 x Aldehydes (Acute toxicity)
Empiric							
Organic functional groups, Norbert Haider		Alcohol	Amine	Alcohol	Amine	Amine	Amine
Organic functional groups		Alcohol	Aliphatic amine, primary	Alcohol	Aliphatic amine, tertiary	Aliphatic amine, tertiary	Aliphatic amine, tertiary
Structure similarity		[90%,100%]	[50%,60%)	[20%,30%)	[0%,10%)	[0%,10%)	[0%,10%)
Organic functional groups (US EPA)		Aliphatic Carbon [CH]	Aliphatic Carbon [CH]	Aliphatic Carbon [CH]	Aliphatic Carbon [CH]	Aliphatic Carbon [CH]	Aliphatic Carbon [CH]
Toxicological							
Repeated dose (HESS)		Not categorized	Aliphatic amines (Mucous membrane	Not categorized	Aliphatic amines (Mucous membrane	Aliphatic amines (Mucous membrane	Aliphatic amines (Mucous membrane
Repeated dose (HESS), with in vivo Rat		2 x 2-Bromoethylamine (Renal Toxicity)	1 x 2-Bromoethylamine (Renal Toxicity)	2 x 2-Bromoethylamine (Renal Toxicity)	1 x 2-Bromoethylamine (Renal Toxicity)	1 x 2-Bromoethylamine (Renal Toxicity)	1 x 2-Bromoethylamine (Renal Toxicity)

Congratulations!

- You have now been introduced to the RAAF scenario;
- You have now been introduced to the *Report basket*.
- You have now been introduced to the AEs related to Scenario 5.
- Note, proficiency comes with practice!