

## 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](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  $\leq 3$  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](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**;

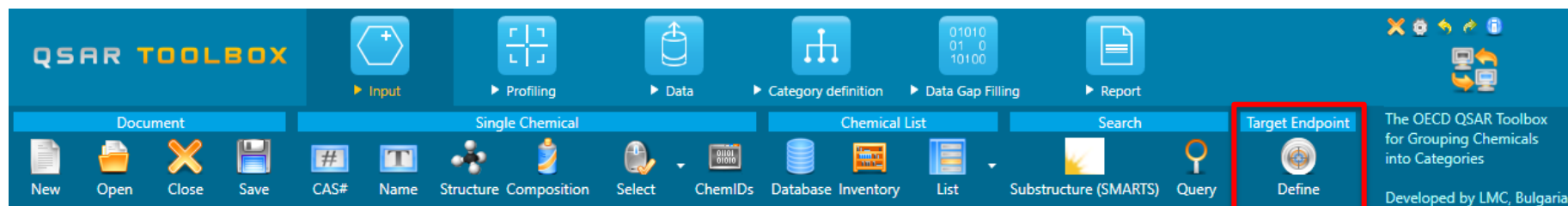
5. When the structure with the requested CAS# appears, click **OK**.

# 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

The screenshot shows the QSAR Toolbox software interface. The top toolbar has icons for Document, Single Chemical, Chemical List, Search, and Target Endpoint. The 'Target Endpoint' icon is highlighted with a red box and labeled '1'. Below the toolbar, the 'Select endpoint' dialog box is open, showing a tree view of endpoints. The 'Human Health Hazards' category is expanded, and 'Repeated Dose Toxicity' is selected, labeled '2'. The 'Next' button is labeled '3'. The 'Select endpoint' dialog box is also open, showing the 'Test organisms (species)' dropdown menu set to 'Rat', labeled '4'. The 'Endpoint' dropdown menu is set to 'NOAEL', labeled '5'. The 'Finish' button is labeled '5'.

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.

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes 'Document', 'Single Chemical', 'Chemical List', 'Search', and 'Target Endpoint'. The 'Target Endpoint' tab is active, showing a chemical structure (H<sub>2</sub>N-CH<sub>2</sub>-NH-CH<sub>2</sub>-OH) and a list of endpoints. The 'Filter endpoint tree...' dialog is open, showing a tree structure of endpoints. The 'Repeated Dose Toxicity' endpoint is highlighted in yellow, and its sub-entries, 'Rat' and 'NOEL', are also highlighted. The 'NOEL' entry is further highlighted with a red box. The 'Documents' panel on the left shows 'Document 1' with CAS# 111411.

Documents

- Document 1  
# [C: 1; Md: 0; P: 0] CAS: 111411

Filter endpoint tree... 1 [target]

Structure

H<sub>2</sub>N-CH<sub>2</sub>-NH-CH<sub>2</sub>-OH

- Structure info
- Parameters
- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
- Human Health Hazards
  - Acute Toxicity
  - ADME
    - Bioaccumulation
    - Carcinogenicity
    - Developmental Toxicity / Teratogenicity
    - Genetic Toxicity
    - Immunotoxicity
    - Irritation / Corrosion
    - Neurotoxicity
    - Photoinduced toxicity
    - Repeated Dose Toxicity
      - Rat
      - NOEL
    - Skin irritation
    - ToxCast
    - Toxicity to Reproduction
    - Toxicokinetics, Metabolism and Distribution

# 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

The screenshot displays the QSAR Toolbox software interface. The top toolbar features icons for 'Input', 'Profiling', 'Data' (highlighted with a red box and callout 1), 'Definition', 'Data Gap Filling', and 'Report'. Below the toolbar, the 'Databases' section on the left lists various databases, with 'ECHA REACH' and 'Food TOX Hazard EFSA' highlighted in green and callout 2. The 'Inventories' section below it shows a list of chemical inventories. The 'Read data?' dialog box is open in the center, with 'Choose...' selected under 'Filter endpoint tree...'. The dialog lists various endpoints, with 'Repeated Dose Toxicity' checked and callout 4 pointing to the 'OK' button. The 'Structure' panel on the right shows a tree view of chemical structure information.

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 ones 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 relevancy of profilers see ppt: [Example for predicting skin sensitization taking into account alert performance](#)*

# Profiling

## Profiling the target chemical

1

2

3

Documents

01 CAS: 111411

Profiling methods

Options Select All Unselect All Invert

- ☐ Organic functional groups
- ☐ Organic functional groups (nested)
- ☐ Organic functional groups (US EPA)
- ☐ Organic functional groups, Norbert Haider (checkmol)
- ☒ Repeated dose (HESS)
- ☐ Structure similarity
- ☐ Substance type
- ☐ US-EPA New Chemical Categories
- ☒ Unclassified
- ☐ (AOT)Protein binding by OASIS v1

Metabolism/Transformations

Options Select All Unselect All Invert

- ☒ Plausible
- ☐ Dissociation simulator
- ☐ Hydrolysis simulator (acidic)
- ☐ Hydrolysis simulator (neutral)
- ☒ in vivo Rat metabolism simulator
- ☒ Unclassified
- ☐ Autooxidation simulator
- ☐ Autooxidation simulator (alkaline medium)
- ☐ Hydrolysis simulator (basic)
- ☐ Microbial metabolism simulator
- ☐ Observed Mammalian metabolism

Filter endpoint tree... 1 [target]

Structure

Structure info

Parameters

Physical Chemical Properties

Environmental Fate and Transport

Ecotoxicological Information

Human Health Hazards

- Acute Toxicity
- ADME
- Bioaccumulation
- Carcinogenicity
- Developmental Toxicity / Teratogenicity
- Genetic Toxicity
- Immunotoxicity
- Irritation / Corrosion
- Neurotoxicity
- Photoinduced toxicity
- Repeated Dose Toxicity
- Sensitisation

Rat

Endpoint	Value	Unit
NOAEL	1/1	M: 1E+03 mg/kg bdwt/d
NOEL	1/1	M: 60 mg/kg bdwt/d

AW SW AOP

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

# Profiling

## Profiling results

The screenshot shows the QSAR TOOLBOX software interface. The top menu bar includes options: Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below the menu is a toolbar with icons for Apply, View, New, and Delete. The main window is divided into several panels. On the left, there's a 'Documents' panel showing 'Document 1' with chemical identifiers. Below it are 'Profiling methods' and 'Metabolism/Transformations' panels with various checkboxes. The central 'Filter endpoint tree...' panel shows a hierarchical list of endpoints. The rightmost panel displays the chemical structure of the target and a table of results. Three callout boxes with numbers 1, 2, and 3 point to specific results in the table.

Endpoint	Results
NOAEL	1/1 M: 1E+03 mg/kg bdwt/d
NOEL	1/1 M: 60 mg/kg bdwt/d
Repeated dose (HESS)	Not categorized
Metabolism/Transformation	11 metabolite(s)
in vivo Rat metabolism simulator	11 metabolite(s)
Toxicological	2 x 2-Bromoethylamine (Renal Toxicity) Alert 1 x Acetamide (Renal Toxicity) Alert 1 x Aliphatic amines (Mucous membrane irritation) Rank C 1 x Ethylene glycol and oxalic acid (Renal toxicity) Alert 2 x Glycolic acid (Renal Toxicity) Alert 1 x Maleic acid (Renal toxicity) Alert 5 x Not categorized

- 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

The screenshot displays the QSAR Toolbox interface. At the top, a navigation bar includes icons for Input, Profiling, Data, Category definition, Data Gap Filling, and Report. Below this, a 'Documents' panel on the left shows a tree structure with a document containing CAS: 111111. A right-click context menu is open over this document, with the 'Multiplication' option selected. This opens a sub-menu where 'Metabolism/Transformations' is chosen, leading to a list of simulation options. In this list, 'in vivo Rat metabolism simulator' is highlighted. The main workspace shows the chemical structure of 2-aminoethanol (NCCO) and a 'Filter endpoint tree' on the right. A red box labeled 'Step 1' points to the 'Multiplication' menu option.













**Step 1**

2 x 2-Bromoethylamine (Renal Toxicity) Alert  
 1 x Acetamide (Renal Toxicity) Alert  
 1 x Aliphatic amines (Mucous membrane irritation) Rank C  
 1 x Ethylene glycol and oxalic acid (Renal toxicity) Alert  
 2 x Glycolic acid (Renal Toxicity) Alert

**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

Filter endpoint tree...	Parent chemical...	metabolite #1	metabolite #2	metabolite #3	metabolite #4	metabolite #5	metabolite #6	metabolite #7	metabolite #8	metabolite #9	metabolite #10	metabolite #11
Structure												
Structure info												
Human Health Hazards												
Immunotoxicity												
Irritation / Corrosion												
Neurotoxicity												
Photoinduced toxicity												
Repeated Dose Toxicity												
mouse	2/2	MS: 4.5E+04 ppm			MS: 8.3 mg/kg b...							
Rat												
LOAEC					MS: 132 ppm							
LOAEL			MS: 1.56E+03 m...		MS: 45 mg/kg b...						MS: 300 mg/kg...	
NOAEC	2/5				MS: 59 ppm							
NOAEL	7/12	M: 1E+03 mg/kg...	MS: 4.5E+04 ppm	MS: <1.56E+03...	MS: 9 mg/kg bd...			MS: 300 mg/kg...		MS: 6E+03 ppm	MS: 150 mg/kg...	
NOEC	1/1							MS: 150 mg/m...				
NOEL	3/6	M: 60 mg/kg bd...	MS: =2.25E+03...								MS: 150 mg/kg...	
Sensitisation	AW SW AOP											
ToxCa												
Toxicity												
Toxicity												
Profiling												
Toxicological												
Repeated dose (HESS)	Not categorized	Acetamide (Rena...	Glycolic acid (Re...	Not categorized	2-Bromoethylam...	Not categorized	Not categorized	2-Bromoethylam...	Maleic acid (Ren...	Not categorized	Glycolic acid (Re...	Not categorized

**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 screenshot illustrates the steps for searching for analogues accounting for in vivo rat liver metabolism using the QSAR Toolbox. The interface shows the 'Category definition' module, the 'Define with metabolism' button, and the 'Grouping options' dialog box. The 'Grouping options' dialog box displays a list of chemicals and their metabolites, with Metabolite 4 (Ethylenediamine) highlighted. The 'Query' dropdown menu for Metabolite 4 is set to 'Exact match'. The 'Grouping options' dialog box also shows a list of chemicals and their metabolites, with Metabolite 4 (Ethylenediamine) highlighted. The 'Query' dropdown menu for Metabolite 4 is set to 'Exact match'. The 'Grouping options' dialog box also shows a list of chemicals and their metabolites, with Metabolite 4 (Ethylenediamine) highlighted. The 'Query' dropdown menu for Metabolite 4 is set to 'Exact match'.

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

The screenshot displays the QSAR Toolbox software interface. The main window shows a 'Filter endpoint tree' on the left, a central data table, and a right sidebar with various tool icons. The 'Filter endpoint tree' is expanded to show 'Human Health Hazards' and 'Acute Toxicity'. The 'Read data?' dialog box is open, showing a list of endpoints with 'Repeated Dose Toxicity' selected. The 'Gather data' dialog box is also open, displaying '74 points added across 27 chemicals.' and an 'OK' button.

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

The screenshot displays the QSAR Toolbox software interface during a Data Gap Filling workflow. The top menu bar includes options like Input, Profiling, Data, Category definition, Data Gap Filling, and Report. The left sidebar shows a 'Documents' panel with a list of metabolites and a 'Data Gap Filling Settings' panel with checkboxes for 'Only endpoint relevant' and 'At this position:'. The central workspace features a 'Filter endpoint tree...' on the left, a 'Structure' panel with chemical structures, and a 'Data' table with columns for endpoints like NOAEL, NOEC, and NOEL. A 'Choose one' dialog box is open in the center, showing a list of choices: Median, Lower median, Higher median, Minimal, **Maximal** (selected), Arithmetic mean (average), Geometric mean, and All. The bottom panel shows a scatter plot titled 'Read-across prediction for NOAEL, based on 5 values' with the text 'Observed: 1E+03 mg/kg bdwt/d; Predicted: 64.4 mg/kg bdwt/d'. The plot shows NOAEL [log (1/mol/kg bdwt/d)] on the y-axis and log Kow on the x-axis. A 'Calculation options' panel on the right includes 'Data usage' and 'Prediction approach options' sections. A green checkmark and 'Accept prediction' button are visible at the bottom right.

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

**Subcategorization 1**

The OECD QSAR Toolbox for Grouping Chemicals into Categories  
Developed by LMC, Bulgaria

**Subcategorization**

Options: 1 Selected  
f Select All Unselect All Invert About Options

**Predefined**

- Database Affiliation
- Inventory Affiliation
- OECD HPV Chemical Categories
- Substance type
- US-EPA New Chemical Categories

**General Mechanistic**

- (AOT)Protein binding by OASIS v1
- Biodeg BioHC half-life (Biowin)
- Biodegradation primary (Biowin 4)
- Biodegradation probability (Biowin 1)
- Biodegradation probability (Biowin 2)
- Biodegradation probability (Biowin 6)
- Biodegradation probability (Biowin 7)
- Biodegradation ultimate (Biowin 3)
- DNA binding by OASIS
- DNA binding by OECD
- Estrogen Receptor Binding
- Hydrolysis half-life (Ka, pH 7)(Hydrowin)
- Hydrolysis half-life (Ka, pH 7)(Hydrowin)

**Metabolisms**

Options: 0 Selected  
f Select All Unselect All Invert

Do not account metabolism

**Documented**

- Observed Mammalian metabolism
- Observed Microbial metabolism
- Observed Rat in vivo metabolism
- Observed rat liver metabolism with quantitative data
- Observed Rat Liver S9 metabolism

**Simulated**

- Autoxidation simulator
- Autoxidation simulator (alkaline medium)
- Dissociation simulator
- Hydrolysis simulator (acidic)
- Hydrolysis simulator (basic)
- Hydrolysis simulator (neutral)
- In vivo Rat metabolism simulator
- Microbial metabolism simulator
- Rat liver S9 metabolism simulator
- Skin metabolism simulator
- Tautomerism

**Chemical Categories**

Aliphatic Amines

Read-across prediction for NOEL, based on 5 values  
Observed: 1E+03 mg/kg bw/d; Predicted: 64.4 mg/kg bw/d

NOEL [log(1/mg/kg bw/d)]

log Kow

**Select / filter data**

- Subcategorize
- Mark chemicals by WS
- Mark chemicals by descriptor value
- Filter points by test conditions
- Mark focused chemical
- Mark focused points
- Remove marked data
- Clear existing marks

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.

Accept prediction



# Data Gap Filling Subcategorize

**Subcategorization 2**

The OECD QSAR Toolbox for Grouping Chemicals into Categories

Developed by LMC, Bulgaria

Options: Select All, Unselect All, Invert, About, Options

1 Selected

Adjust options

Target

Group 14 - Carbon C

Group 15 - Nitrogen

Group 16 - Oxygen

Filter endpoint tree...

Structure

NOAEL 11/16 M: 1E+03 mg/kg...

NOEC 1/1 M: 100 mg/kg b...

NOEL 2/2 M: 627 mg/kg b...

Other 1/1 M: 100 mg/kg b...

Other Endpoint 1/3 M: 30 mg/kg bd...

Sensitisation AW SW AOP

ToxCast AOP

Toxicity to Reproduction AOP

Toxicokinetics, Metabolism and Distribution

Profiling

Predefined

Substance type

US-EPA New Chemical Categories

Discrete chemical Discrete chemical Discrete chemical Discrete chemical Discrete chemical Discrete chemical Discrete chemical Discrete chemical Discrete chemical Discrete chemical

Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines

Read-across prediction for NOAEL, based on 5 values

Observed: 1E+03 mg/kg bdwt/d; Predicted: 64.4 mg/kg bdwt/d

NOAEL [log(1/mol/kg bdwt/d)]

log Kow

Selected 2 (8/1)

Remove selected

1

2

3

Select / filter data

Subcategorize

Mark chemicals by WS

Mark chemicals by descriptor

Filter points by test condition

Mark focused chemical

Mark focused points

Remove marked data

Clear existing marks

Accept prediction

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

**Subcategorization 3**

The screenshot displays the QSAR Toolbox interface during the 'Subcategorization' process. The top menu bar includes 'Input', 'Profiling', 'Data', 'Category definition', 'Data Gap Filling', and 'Report'. The left sidebar shows the 'Data filter' panel with various criteria like 'Source', 'Materials & Methods', 'Test chemical', 'Species/Organs/Tissue', 'Administration/Exposure/Duration', 'Results', 'Data Relevance', and 'Others'. The central area shows a table of chemical data with columns for chemical structures, endpoints, and categories. The bottom plot area shows a scatter plot of 'NOAEL [log(1/mol/kg bw/d)]' vs 'log Kow' with a 'Read-across prediction for NOAEL' line. The right sidebar contains buttons for 'Select / filter data', 'Subcategorize', 'Mark chemicals by WS', 'Mark chemicals by descriptor value', 'Filter points by test conditions', 'Mark focused chemical', 'Mark focused points', 'Remove marked data', 'Clear existing marks', and 'Accept prediction'.

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

The OECD QSAR Toolbox for Grouping Chemicals into Categories  
Developed by LMC, Bulgaria

Gap Filling Workflow  
Trend analysis Read across (Q)SAR Standardized Automated

Documents  
Document 1  
# [C: 1;Md: 2;P: 0] CAS: 111411  
[C: 12;Md: 34;P: 0] in vivo Rat metabolism simulator  
[C: 1;Md: 0;P: 0] metabolite #1  
[C: 1;Md: 0;P: 0] metabolite #2  
[C: 1;Md: 0;P: 0] metabolite #3  
[C: 1;Md: 0;P: 0] metabolite #4  
[C: 1;Md: 0;P: 0] metabolite #5  
[C: 1;Md: 0;P: 0] metabolite #6  
[C: 1;Md: 0;P: 0] metabolite #7  
[C: 1;Md: 0;P: 0] metabolite #8  
[C: 1;Md: 0;P: 0] metabolite #9  
[C: 1;Md: 0;P: 0] metabolite #10  
[C: 1;Md: 0;P: 0] metabolite #11  
[C: 54;Md: 74;P: 0] Grouping with metabolism: in vivo Rat metabolism simulator  
[C: 15;Md: 45;P: 0] Enter GF(RA)  
[C: 15;Md: 45;P: 0] Data usage options are changed to: Maximal  
[C: 11;Md: 30;P: 0] Subcategorized: US-EPA New Chemical Categories  
[C: 9;Md: 28;P: 0] Subcategorized: Chemical elements  
[C: 7;Md: 20;P: 0] Filter data by test condition - Test type

Data Gap Filling Settings  
Only endpoint relevant  
At this position:  
QSARs 0  
Automated workflows 0  
Standardized workflows 0  
In nodes below:  
QSARs 0  
Automated workflows 0  
Standardized workflows 0

Filter endpoint tree...  
Structure  
NOAEL 7/11 M: 1E+03 mg/kg... M: 100 mg/kg b... M: 100 mg/kg b... M: 30 mg/kg bd... M: 100 mg/kg b... M: 25 mg/kg bd... M: 30 mg/kg bd...  
NOEL 1/1 M: 60 mg/kg bd...  
Sei  
Toi  
Toi  
Profile  
Substance type Discrete chemical Discrete chemical Discrete chemical Discrete chemical Discrete chemical Discrete chemical  
US-EPA New Chemical Categories Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines Aliphatic Amines  
Empiric

Read-across prediction for NOAEL based on 5 values  
Observed: 1E+03 mg/kg bdwt/d; Predicted: 31.8 mg/kg bdwt/d

NOAEL [log(1/mg/kg bdwt/d)]  
log Kow: -1.660  
NOAEL: 3.764 log(1/mol/kg bdwt/d)  
log Kow: -0.490  
NOAEL: 3.050 log(1/mol/kg bdwt/d)

Active descriptor X log Kow

Select / filter data  
Subcategorize  
Mark chemicals by WS  
Mark chemicals by descriptor value  
Filter points by test conditions  
Mark focused chemical  
Mark focused points  
Remove marked data  
Clear existing marks  
Accept prediction

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

The screenshot displays the QSAR Toolbox software interface. The top menu bar includes options like Input, Profiling, Data, Category definition, Data, and Report. The main window is titled 'Category consistency wizard' and contains a 'Wizard pages' sidebar with options for Physicochemical similarity, Structural similarity, Mechanistic similarity, (Eco)tox experimental data, and Options. The 'Physicochemical similarity' page is active, showing '2D/3D parameters' and 'Physico-chemical data' sections. A 'Read-across prediction for NOAEL' plot is visible, showing observed vs. predicted values. A table of chemical data is also shown, with columns for chemical structures and various parameters. The interface is annotated with four numbered boxes: 1 points to the 'Data' menu, 2 points to the 'Category consistency' button, 3 points to the 'OK' button in the wizard, and 4 points to the 'Accept prediction' button in the bottom right corner.

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

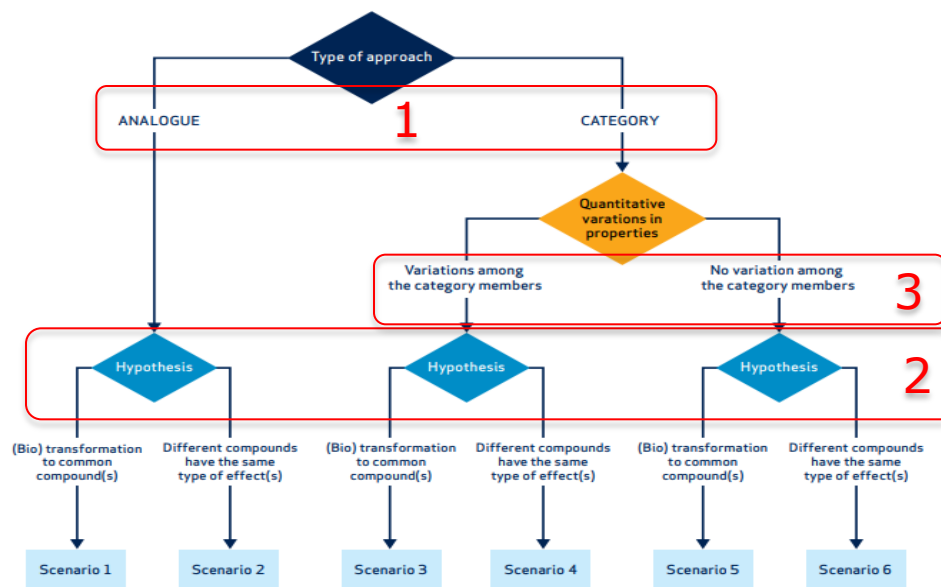
- The report module allows generating a report for predictions performed within the Toolbox.
- The report module contains a predefined report template which users can customize.
- Additionally specific RAAF scenario could be chosen. Selection of one of the scenarios will append automatically the related 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](https://echa.europa.eu/documents/10162/13628/raaf_en.pdf)

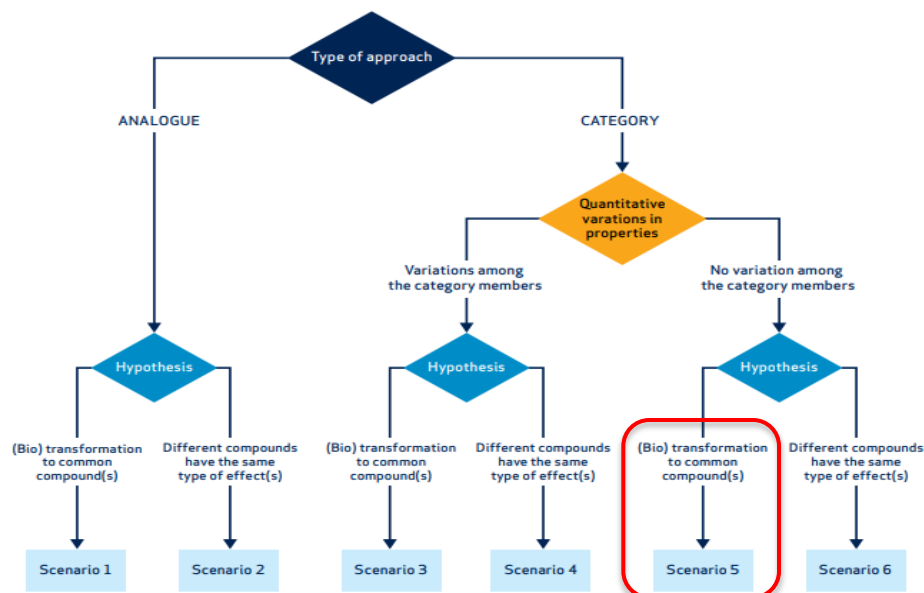
# Report

## 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](https://echa.europa.eu/documents/10162/13628/raaf_en.pdf)



# Report

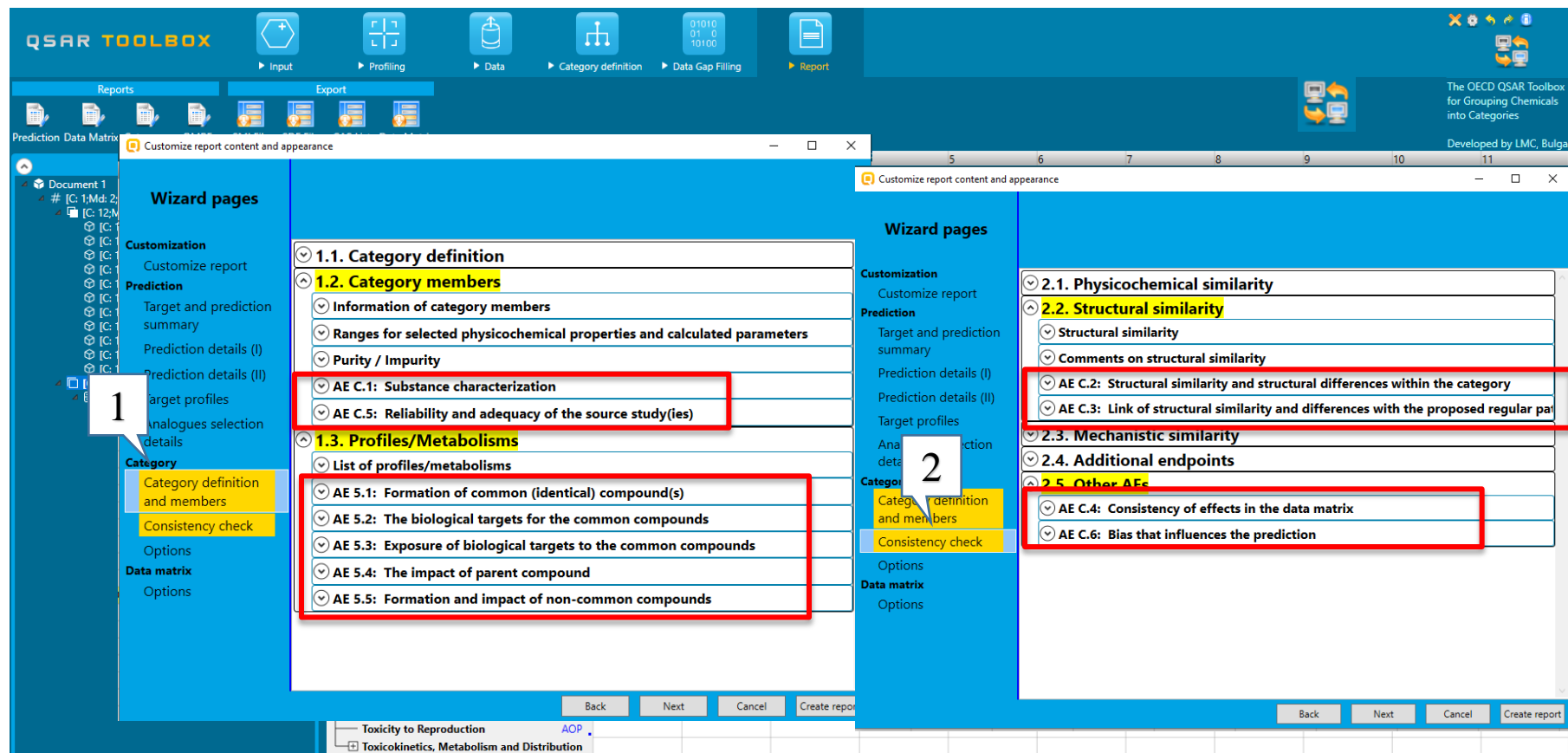
## Generation report according to RAAF-Scenario 5

The screenshot illustrates the steps to generate a report in the QSAR Toolbox. The 'Report' menu is selected (1). In the 'Documents' list, a specific chemical entry is highlighted (3). The 'Filter endpoint tree...' panel shows the selection of the 'NOAEL' endpoint (2). The 'Customize report content and appearance' dialog box is open, where the 'Add RAAF scenario' option is checked (4), and 'Scenario 5' is chosen from the drop-down menu (5). The dialog also shows other report sections like Prediction, Category, and Data matrix, with various sub-options checked. The 'Create report' button is visible at the bottom right of the dialog.

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

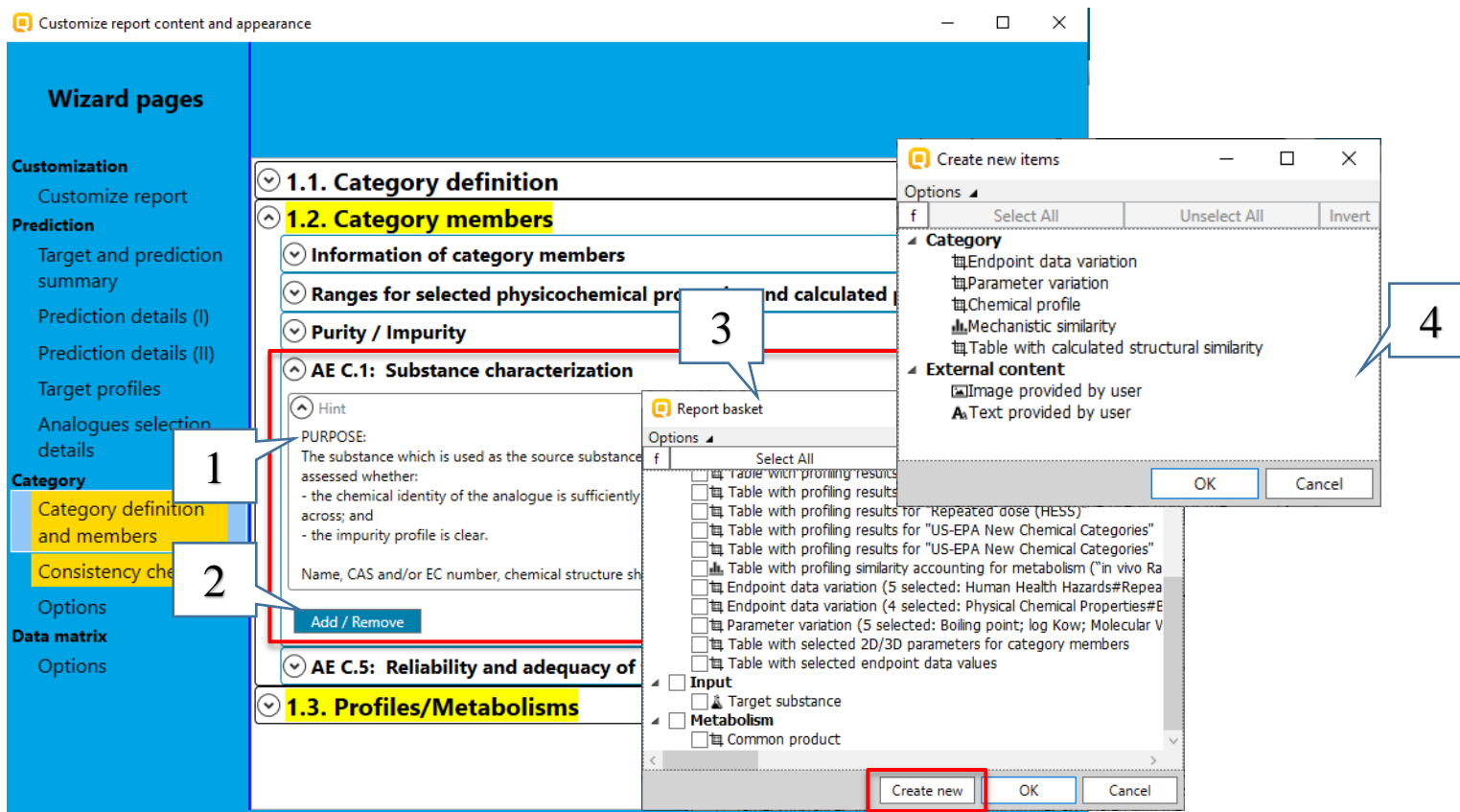


Once the RAAF scenario is selected the assessment elements (AEs) related to it will be appended to the corresponding sections of the report automatically. AEs appear in the following report sections: **Category definition and members** (1) and **Consistency check** (2).

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

The screenshot shows the 'Customize report content and appearance' window. The left sidebar has a 'Category' section with 'Category definition and members' highlighted. The main area shows a tree view of report sections. Annotations are as follows:

- 1**: Points to the 'Add / Remove' button at the bottom of the 'AE C.1: Substance characterization' section.
- 2**: Points to the '1.2. Category members' section header.
- 3**: Points to the 'Preview' button in the context menu for 'Table of category members'.
- 4**: Points to the 'OK' button at the bottom right of the window.

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

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

The image shows the QSAR Toolbox software interface with three main windows and numbered callouts (1-7) indicating the steps for adding external content to the report:

- 1:** In the 'Customize report content and appearance' wizard, under the 'Category' section, '1.2. Category members' is selected.
- 2:** In the 'Report basket' window, the 'Create new' button is clicked.
- 3:** In the 'Create new items' dialog, 'Text provided by user' is selected under the 'External content' category.
- 4:** The 'OK' button is clicked in the 'Create new items' dialog.
- 5:** A text entry window appears with the item name 'Automatically generated' and a text area containing the 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.'
- 6:** The 'OK' button is clicked in the text entry window.
- 7:** The 'Add / Remove' button is clicked in the '1.2. Category members' section of the wizard.

### 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

[illegible]

### **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

Customize report content and appearance

**Wizard pages**

- Customization
  - Customize report
- Prediction
  - Target and prediction summary
  - Prediction details (I)
  - Prediction details (II)
  - Target profiles
  - Analogues selection details
- Category
  - Category definition and members
  - Consistency check
  - Options
- Data matrix
  - Options

**1.1. Category definition**

**1.2. Category members**

**1.3. Profiles/Metabolisms**

**List of profiles/metabolisms**

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

Hint

PURPOSE:  
This AE considers how the common compound(s) are formed from the members of the category. It has to be assessed whether:  
- it is explained how the (identical) common product(s) are formed (i.e. the product(s) claim on the property under consideration); and  
- the provided evidence supports the explanation.

Add / Remove

A Ethylenediamine is produced from the target substance ...

Common product

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

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

**AE 5.4: The impact of parent compound**

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

**Report basket**

Options 2 Selected

Select All Unselect All Invert

**Category**

- ☐ Endpoint data variation (1 selected: Human Health Hazards#Repeated)
- ☐ Table with profiling results for "Repeated dose (HESS)"
- ☐ Table with profiling similarity accounting for metabolism ("in vivo Rat metabolism")
- ☐ Endpoint data variation (1 selected: Human Health Hazards#Repeated)
- ☐ Table of category members
- ☐ Table with calculated structural similarity
- ☐ Table with profiling results for "Organic functional groups, Norbert Hain"
- ☐ Table with profiling results for "Organic functional groups"
- ☐ Table with profiling results for "Chemical elements"
- ☐ Table with profiling results for "Structure similarity"
- ☐ Table with profiling results for "Organic functional groups (US EPA)"
- ☐ Table with profiling results for "Substance type"
- ☐ Table with profiling results for "US-EPA New Chemical Categories"
- ☐ Table with profiling similarity accounting for metabolism ("in vivo Rat metabolism")
- ☐ Table with profiling similarity accounting for metabolism ("in vivo Rat metabolism")
- ☐ Endpoint data variation (5 selected: Human Health Hazards#Repeated)
- ☐ Parameter variation (5 selected: Boiling point; log Kow; Molecular Weight)
- ☐ Endpoint data variation (4 selected: Physical Chemical Properties#Boiling point)
- ☐ Table with selected 2D/3D parameters for category members
- ☐ Table with selected endpoint data values

**External content**

- ☒ All source studies are conducted according to OECD Guideline ... (text)
- ☐ Image from clipboard No.4 (image provided by user)
- ☐ Image from clipboard No.5 (image provided by user)
- ☒ Ethylenediamine is produced from the target substance ...

**Input**

- ☐ Target substance

**Metabolism**

- ☒ Common product

Select All Unselect All Invert

Preview Delete

**Common product**

Used metabolism simulator: in vivo Rat metabolism simulator

Map similarity options used : Exact metabolite

**Common product:**

NCCN NCCN

### 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

Customize report content and appearance

**Wizard pages**

**Customization**

Customize report

**Prediction**

Target and prediction summary

Prediction details (I)

Prediction details (II)

Target profiles

Analogues selection details

**Category**

Category definition and members

Consistency check

Options

**Data matrix**

Options

**1.1. Category definition**

**1.2. Category members**

**1.3. Profiles/Metabolisms**

**List of profiles/metabolisms**

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

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

**Hint**

**PURPOSE:**  
The hypothesis claims that the common compound(s) have the same biological target(s) (and hence cause the same type of effects). It has to be assessed whether:

- the same biological targets are affected in a consistent manner throughout the category, and by the common compounds; and
- the provided evidence supports the explanation.

Add / Remove

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

**Hint**

**PURPOSE:**  
Under this scenario, it is proposed that the exposure of the biological targets to the common compound(s) vary in a predictable manner. It has to be assessed whether:

- the documentation established that the exposure of the biological targets to the common compound(s) is varying in a predictable manner;
- the prediction is derived from the relation between an observed property and the independent variable which determines the order within the category (prediction model); and
- the provided evidence supports the explanation. As a default, a prediction based on a regular pattern without a mechanistic explanation will not be acceptable.

Add / Remove

**AE 5.4: The impact of parent compound**

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

**Hint**

**PURPOSE:**  
The formation of common compound(s) often goes together with the formation of non-common compound(s) and possible intermediates which form the common compound(s). The source and/or target substance may also be (bio)transformed via other pathways leading to other additional non-common compounds. It has to be assessed whether:

- the formation of non-common compounds (including possible intermediates) via the possible pathways and their possible impact on the prediction property under consideration have been considered; and
- the provided evidence supports the explanation.

Add / Remove

Back Next Cancel Create report

### 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

The screenshot shows the 'Customize report content and appearance' window. On the left, the 'Wizard pages' sidebar includes 'Customization', 'Prediction', and 'Category'. The 'Category' section is expanded, showing 'Category definition and members' and 'Consistency check'. The 'Add / Remove' button is highlighted with a red box and labeled '1'. The 'Create new items' dialog is open, showing 'Options' and 'External content' sections. The 'Image provided by user' option is selected and labeled '3'. The 'OK' button is labeled '4'. The 'Create new items' dialog also shows a 'Select your image here' section with a grid of chemical structures labeled 'Target A', 'Source B', 'Source C', 'Source D', 'Source E', and 'Source F', labeled '5'. The 'Specify how much of the page width is occupied by the image' dialog is open, showing 'Image width, %' set to 75, and the 'OK' button is labeled '6'.

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

**Wizard pages**

- Customization**
  - Customize report
- Prediction**
  - Target and prediction summary
  - Prediction details (I)
  - Prediction details (II)
  - Target profiles
  - Analogues selection details
- Category**
  - Category definition and members
  - Consistency check
  - Options
- Data matrix**
  - Options

**Report basket**

Options: 1 Selected

**Category**

- ☐ Endpoint data variation (1 selected: Human Health Hazards#Repeated)
- ☐ Table with profiling results for "Repeated dose (HESS)"
- ☐ Table with profiling results accounting for metabolism ("in vivo Rat m
- ☐ Endpoint data variation (1 selected: Human Health Hazards#Repeated
- ☐ Table of category members
- ☐ Table with calculated structural similarity
- ☐ Table with profiling results for "Organic functional groups, Norbert Hai
- ☐ Table with profiling results for "Organic functional groups"
- ☐ Table with profiling results for "Chemical elements"
- ☐ Table with profiling results for "Structure similarity"
- ☐ Table with profiling results for "Organic functional groups (US EPA)"
- ☐ Table with profiling results for "Substance type"
- ☐ Table with profiling results for "US EPA New Chemical Categories"
- ☐ Table with profiling s
- ☐ Table with profiling s
- ☐ Endpoint data variati
- ☐ Parameter variation (
- ☐ Endpoint data variati
- ☐ Table with selected
- ☐ Table with selected

**External content**

- ☐ All source studies are
- ☐ Image from clipboard
- ☐ Image from clipboard
- ☐ Ethylenediamine is no
- ☒ Image from

**Input**

- ☐ Target subs

**Metabolism**

- ☐ Common product

**Create new items**

Options: 1 Selected

**Category**

- ☐ Endpoint data variation
- ☐ Parameter variation
- ☐ Chemical profile
- ☐ Mechanistic similarity
- ☐ Table with calculated structural similarity

**External content**

- ☐ Image provided by user
- ☒ Text provided by user

**Report basket**

Item name: Automatically generated

Enter your text here:

Target A is transformed to the reactive metabolite: formic acid based as a result of in vivo rat metabolism  
The reactive metabolite is claimed 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.

OK Cancel

1: Report basket

2: Create new

3: Text provided by user

4: OK

5: Item name

6: OK

The newly created item appears in the *Report basket* (1).

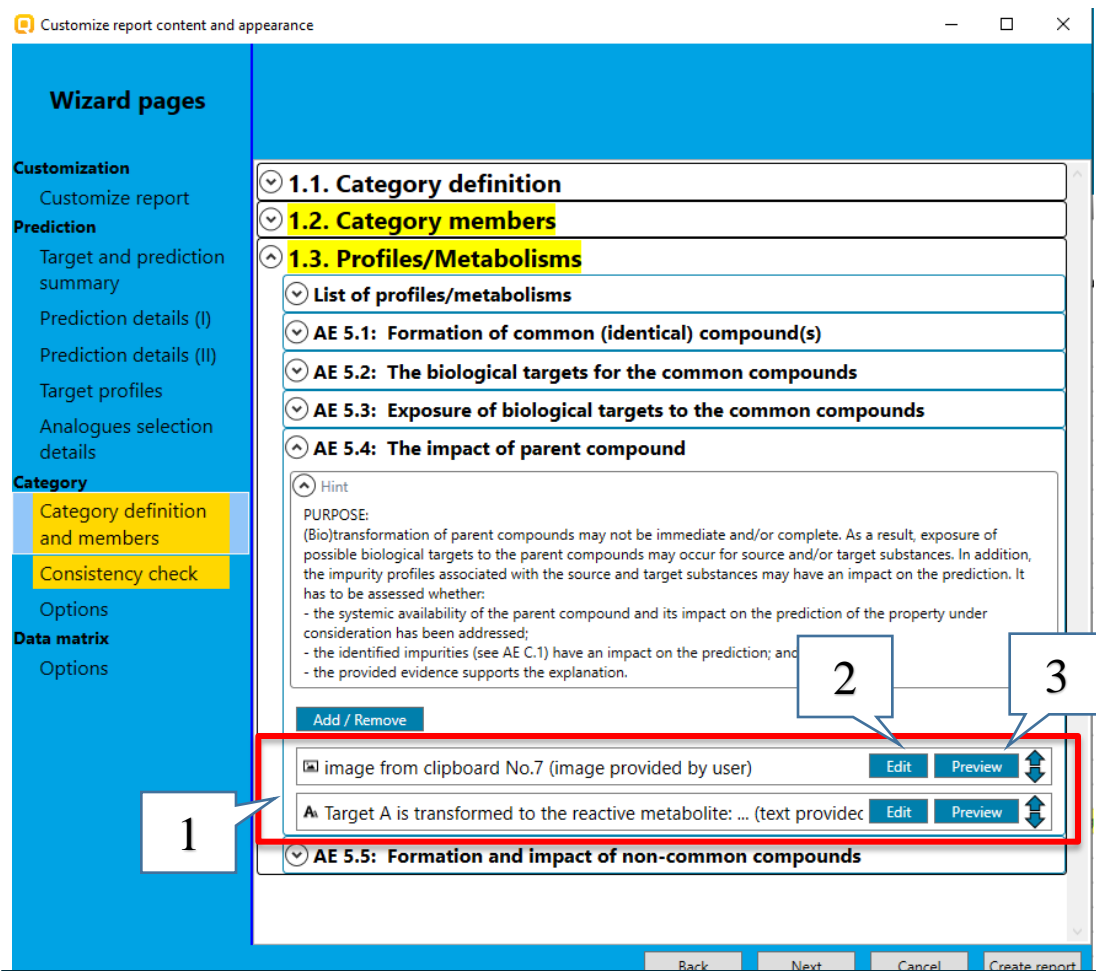
Now text will be also included. Click **Create new** (2), select **Text provided by user** (3) and click **OK** (4). Copy the following example text:

- Target A is transformed to the reactive metabolite: Ethylenediamine based on the result of in vivo rat metabolism
- The reactive metabolite is claimed 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



# Report

## Assessment elements of Scenario 5

Customize report content and appearance

**Wizard pages**

**Customization**  
Customize report

**Prediction**  
Target and prediction summary  
Prediction details (I)  
Prediction details (II)  
Target profiles  
Analogues selection details

**Category**  
Category definition and members  
**Consistency check**  
Options

**Data matrix**  
Options

**2.1. Physicochemical similarity**

**2.2. Structural similarity**

Structural similarity

Comments on structural similarity

**AE C.2: Structural similarity and structural differences within the category**

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

Hint

PURPOSE:  
It has to be assessed whether:  
- the documentation provides an explanation why the category members should behave in a predictable manner (e.g. based on no absorption due to molecular-weight considerations, or lacking reactivity towards biological material, regular pattern in increasing strength of effect due to kinetic differences);  
- it is likely that all category members follow the proposed explanation and where the boundaries of the category are in this respect; and  
- the provided evidence supports the explanation.

Add / Remove

**2.3. Mechanistic similarity**

**2.4. Additional endpoints**

**2.5. Other AEs**

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

**AE C.6: Bias that influences the prediction**

Hint

PURPOSE:  
It has to be assessed whether:  
- it is clear from the documentation how the source substance(s) have been chosen, for example, what methods/tools have been used to map the field of potential source substance(s), which other substances have been considered and why they have been discarded;  
- there are additional, structurally-similar substances which are currently not used in the analogue approach and which arguably could be used;  
- there is readily-available information from these additional substances;  
- this information is biologically significantly different for relevant properties in comparison with the existing analogue(s); and  
- these differences decrease the confidence in the prediction (possibility of underestimation of hazard).

Add / Remove

Back Next Cancel Create report

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

Customize report content and appearance

**Wizard pages**

**Customization**  
Customize report

**Prediction**  
Target and prediction summary  
Prediction details (I)  
Prediction details (II)  
Target profiles  
Analogues selection details

**Category**  
Category definition and members  
**Consistency check**

**Options**  
**Data matrix**  
Options

**2.1. Physicochemical similarity**

**2.2. Structural similarity**

**2.3. Mechanistic similarity**

**2.4. Additional endpoints**

**2.5. Other AEs**

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

**Hint**  
PURPOSE:  
The category justification should include comparison of experimental data for the category members and a clear data matrix. It has to be assessed whether:  
- a data matrix has been provided which lists the category members in a suitable order versus their experimental data (e.g. for REACH information requirements) and which identifies data gaps;  
- the properties of category members across the data matrix are consistent in effects; this has to be assessed in the following dimensions:  
- within the specific property which is under consideration for the prediction;  
- between the property under consideration and related properties (e.g. between 28-day and 90-day repeated-dose toxicity studies; reproductive toxicity screening tests; and pre-natal developmental toxicity studies);  
- characteristics across all relevant properties (e.g. different reactivity towards genetic material may indicate different reactivity towards biological macromolecules which may influence the prediction for a 90-day repeated-dose toxicity study);  
- the effects reported for the property under consideration differ in strength for the source substance and whether a basis for this difference is provided; and  
- the underlying data support the provided conclusions and explanations.

**Add / Remove**

**AE C.6: Bias that influences the prediction**

Back Next Cancel Create report

### **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

Customize report content and appearance

**Wizard pages**

**Customization**  
Customize report

**Prediction**  
Target and prediction summary  
Prediction details (I)  
Prediction details (II)  
Target profiles  
Analogues selection details

**Category**  
Category and mem **1**  
Consistency check

**Data matrix**  
Options

**2.1. Physicochemical similarity**

**2.2. Structural similarity**

**Structural similarity**

Structure similarity profilers

Options 5 Selected

Select All Unselect All Invert

☒ Plausible

- ☒ Chemical elements
- ☐ Groups of elements
- ☐ Lipinski Rule Oasis
- ☒ Organic functional groups
- ☐ Organic functional groups (nested)
- ☒ Organic functional groups (US EPA)
- ☒ Organic functional groups, Norbert Halder (checkmol)
- ☒ Structure similarity

☐ Unclassified

Add / Remove

Table with calculated structural similarity

Table with profiling results for "Organic functional groups,"

Table with profiling results for "Organic functional groups"

Table with profiling results for "Chemical elements"

Table with profiling results for "Structure similarity"

Table with profiling results for "Organic functional groups (US EPA)"

**Comments on structural similarity**

**AE C.2: Structural similarity and structural differences within the category**

**Hint**

**PURPOSE:**  
The aim of this AE is to verify that all category members indeed meet the criteria for structural similarities and allowed structural differences used for the category description. It has to be assessed whether:

- the structural similarities identified apply to all category members; and
- there are structural differences which are allowed within the category.

**3**

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

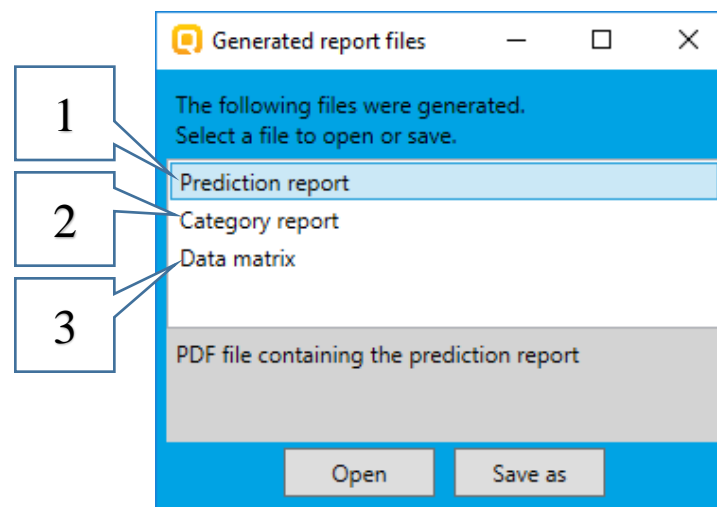
## Generation report

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

### Category report

#### QSAR Toolbox prediction for single chemical

(In accordance with RAAF scenario 5)

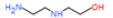

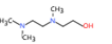
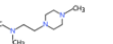
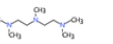

The selected RAAF scenario is specified in the first page

#### QSAR Toolbox report for category

(In accordance with RAAF scenario 5)

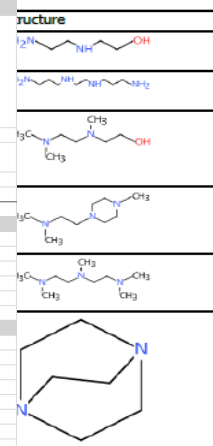
Date: 15 Apr 2020  
Author(s):  
Contact details:

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

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T						
		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'-ethylenbis-				(dimethylamino)ethyl(methylamino)et				N,N,4-trimethylpiperazine-1-ethylamine				bis[2-(dimethylaminoethyl)](methyl)amine				TEDA					
Other identifier																											
SMILES		NCCNCCO				NCCCNCCNCCN				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			
Profiles																											
Profiles used for grouping/subcategorization																											
Using of "in vivo Rat metabolism simulator"																											
US-EPA New Chemical Categories		metabolite #1; Aliphatic Amines				metabolite #1; Aliphatic Amines				metabolite #1; Aliphatic Amines				metabolite #1; Aliphatic Amines				metabolite #1; Aliphatic Amines				metabolite #1; 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)					

### Data matrix report

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





# 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!