QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

OECD (Q)SAR Toolbox v.4.4.1

Tutorial illustrating quantitative metabolic information and related functionalities

Outlook

- **Aim**
- Background
- Keywords
- Examples

Aim

The implementation of quantitative metabolic information and related functionalities in the Toolbox aims to expand and facilitate the use of metabolic information.

Outlook

- Aim
- **Background**
- Keywords
- Examples

Background

The documented/simulated metabolic information available in Toolbox is expanded by adding quantitative data and developing tools for using this type of information for grouping or pruning existing categories.

Outlook

- Aim
- Background
- **Keywords**
- Examples

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

- Aim
- Background
- Keywords
- **Examples for:**
  - **Visualizing quantitative data within Toolbox user interface**
  - Application of quantitative metabolic data in data gap filling

Visualizing quantitative data within Toolbox user interface:
Steps

- Chemical input
- Profiling

Chemical Input

- 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.

Chemical Input
Ways of Entering a Chemical

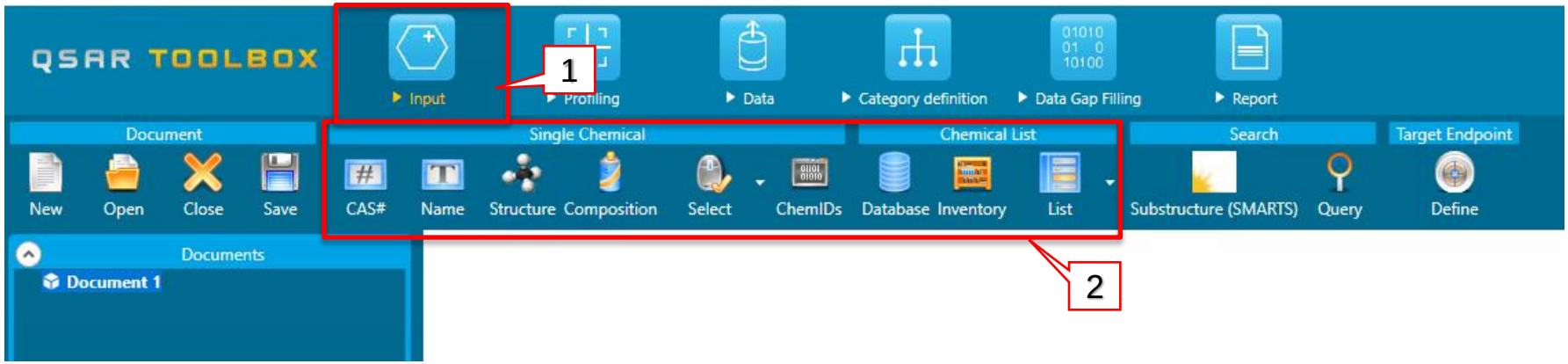
Single target chemical

- Chemical Name
- Chemical Abstract Services (CAS) number (#)
- SMILES (simplified molecular information line entry system) notation
- Chemical with defined composition
- Drawing chemical structure
- Select from User List/Inventory/Databases

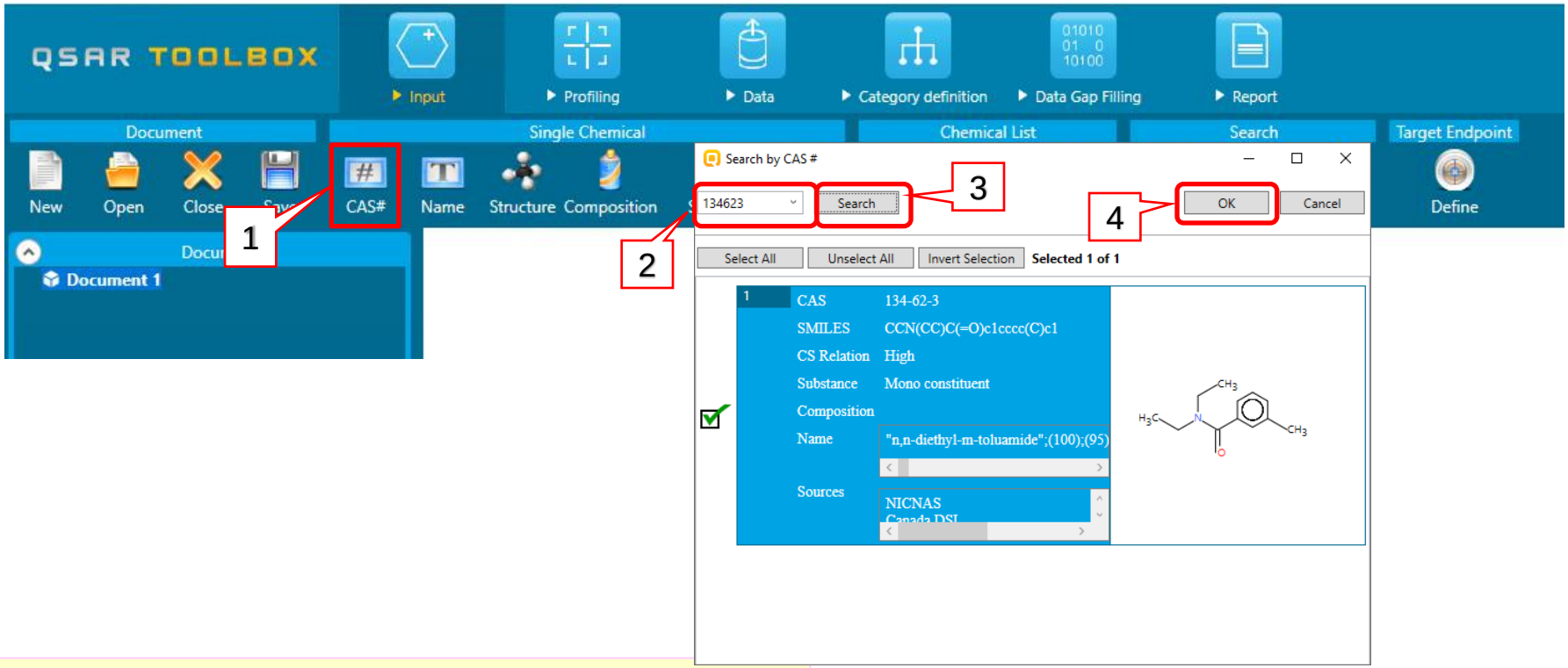
Chemical Input:
Single target chemical

- Open Toolbox.

- Click *Input (1)* to display the main Input section (2).



Single target chemical
CAS RN 134-62-3



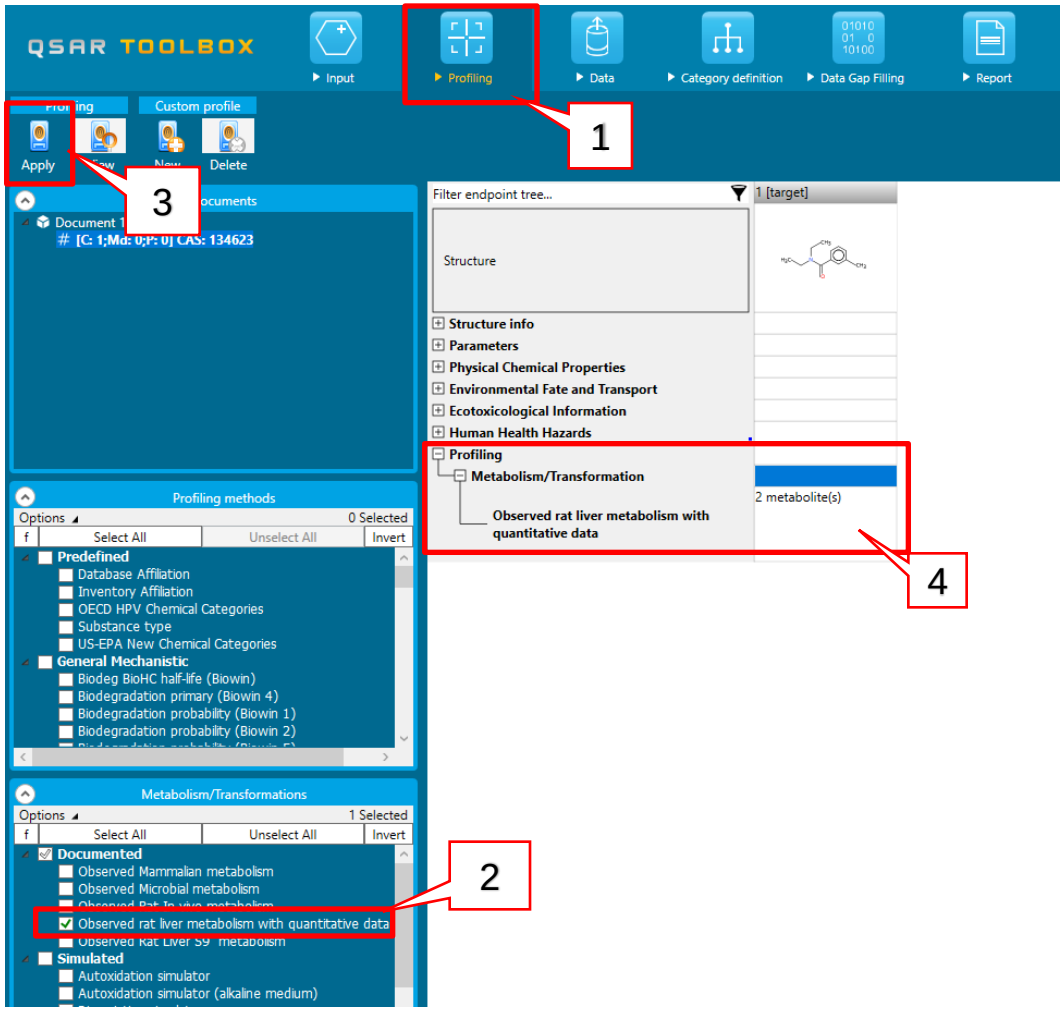
1. Click **CAS#** (1);
2. Type in the **CAS# 134-62-3** (2) ;
3. Click **Search** (3);
4. Click **OK** (4).

Profiling
Overview

- "Profiling" refers to the electronic process of retrieving relevant information on a compound which is stored in the Toolbox, other than its fate and (eco)toxicity data.
- Toolbox has many predefined profilers but it also allows the user to develop new profilers.

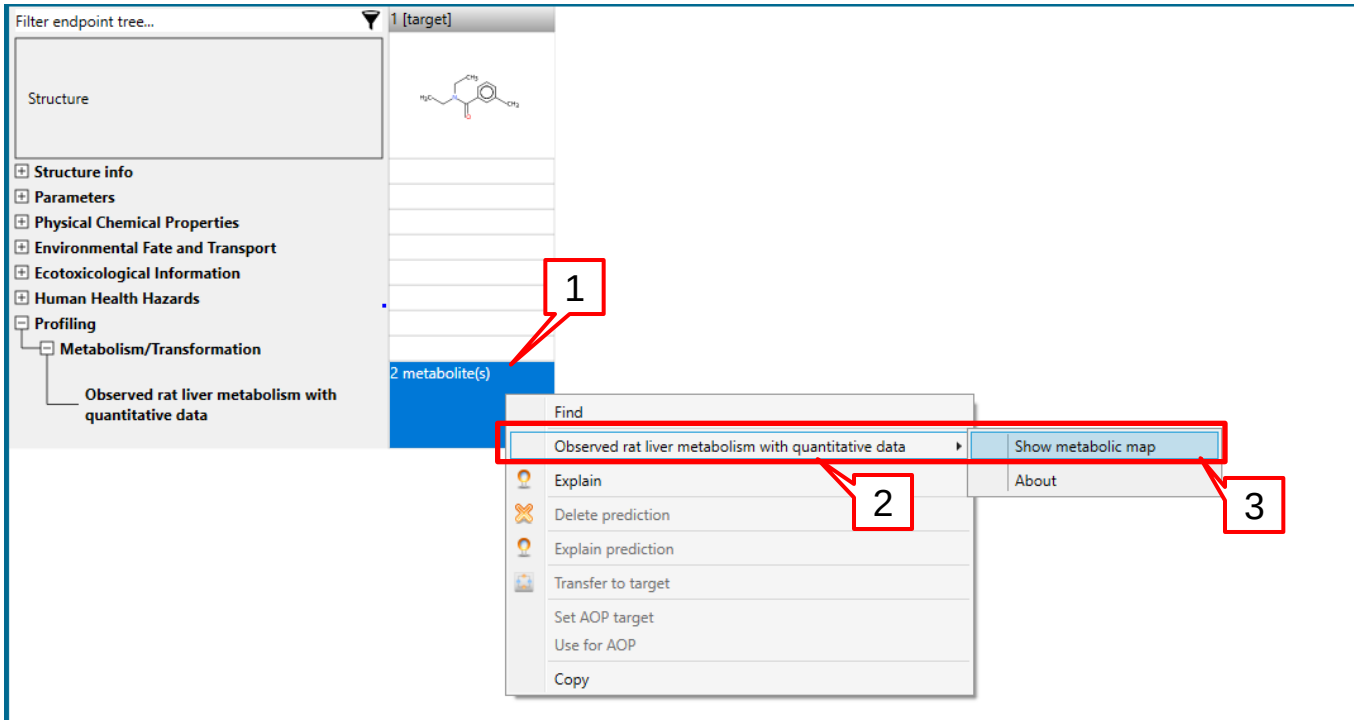
Profiling

1. Select ***Profiling*** (1);
2. Tick **Observed Rat liver metabolism with quantitative data** (2);
3. Click **Apply** (3);
4. Two metabolites are generated (4).

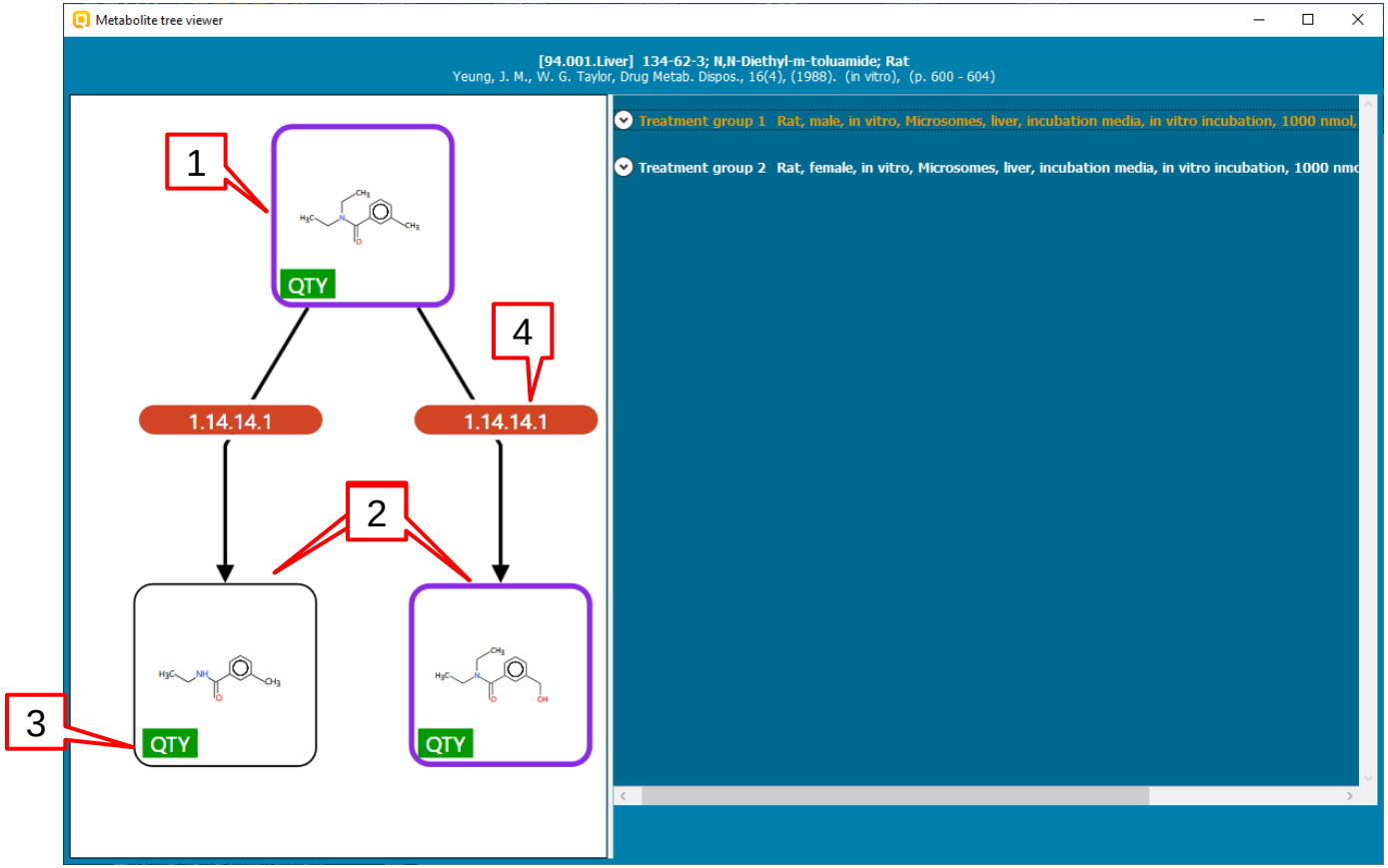


Profiling

1. Right click on the Profiler outcome cell (1);
2. Select **Observed rat liver metabolism with quantitative data** (2) from drop-down menu;
3. Click **Show metabolic map** (3).

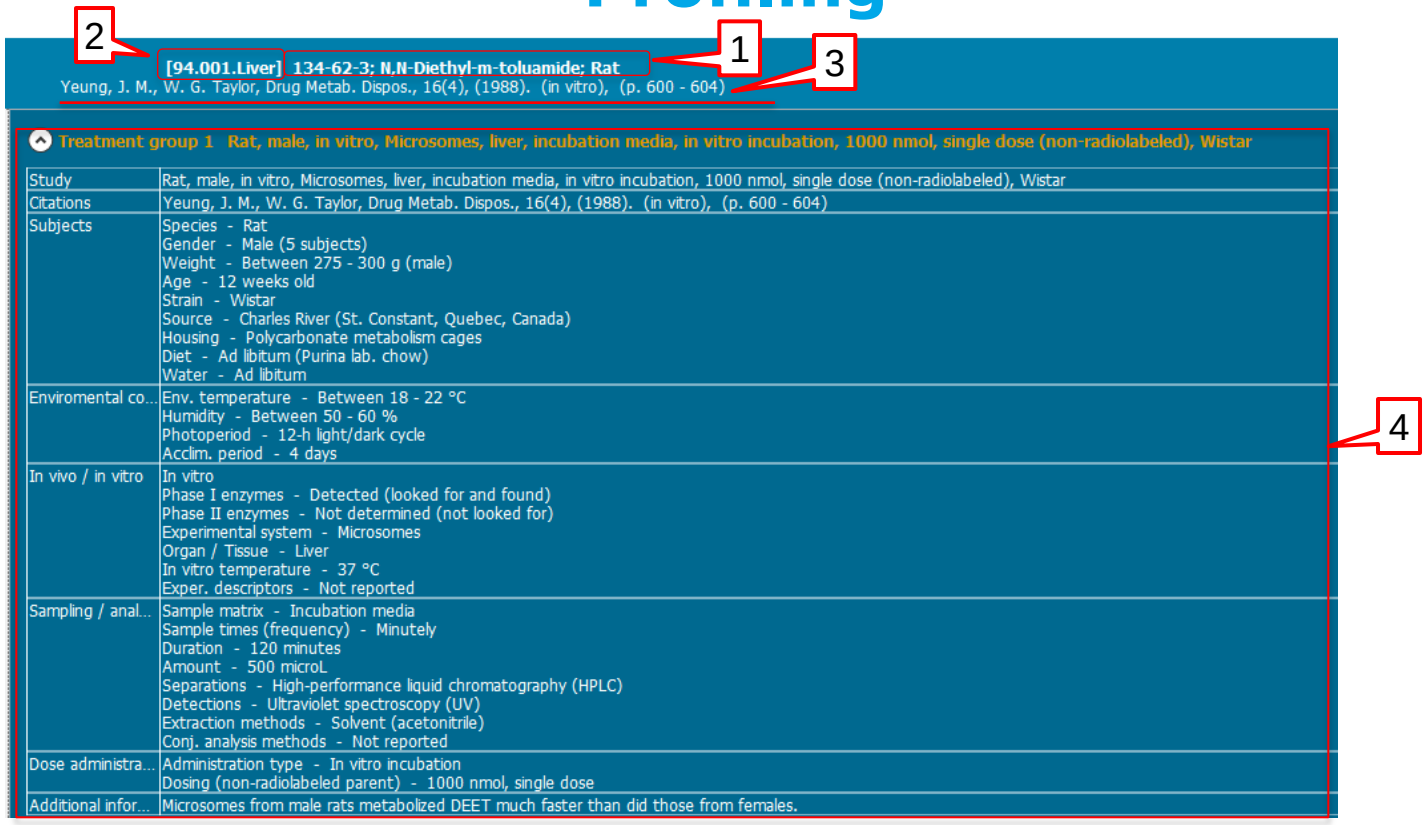


Profiling



- The target (1) and the generated metabolites(2) are shown.
- Quantity label "**QTY"** indicates that there are some quantitative data for the target/metabolite (3)
- Label "1.14.14.1" indicated enzymatic information, which could be seen in METAPATH software(4)

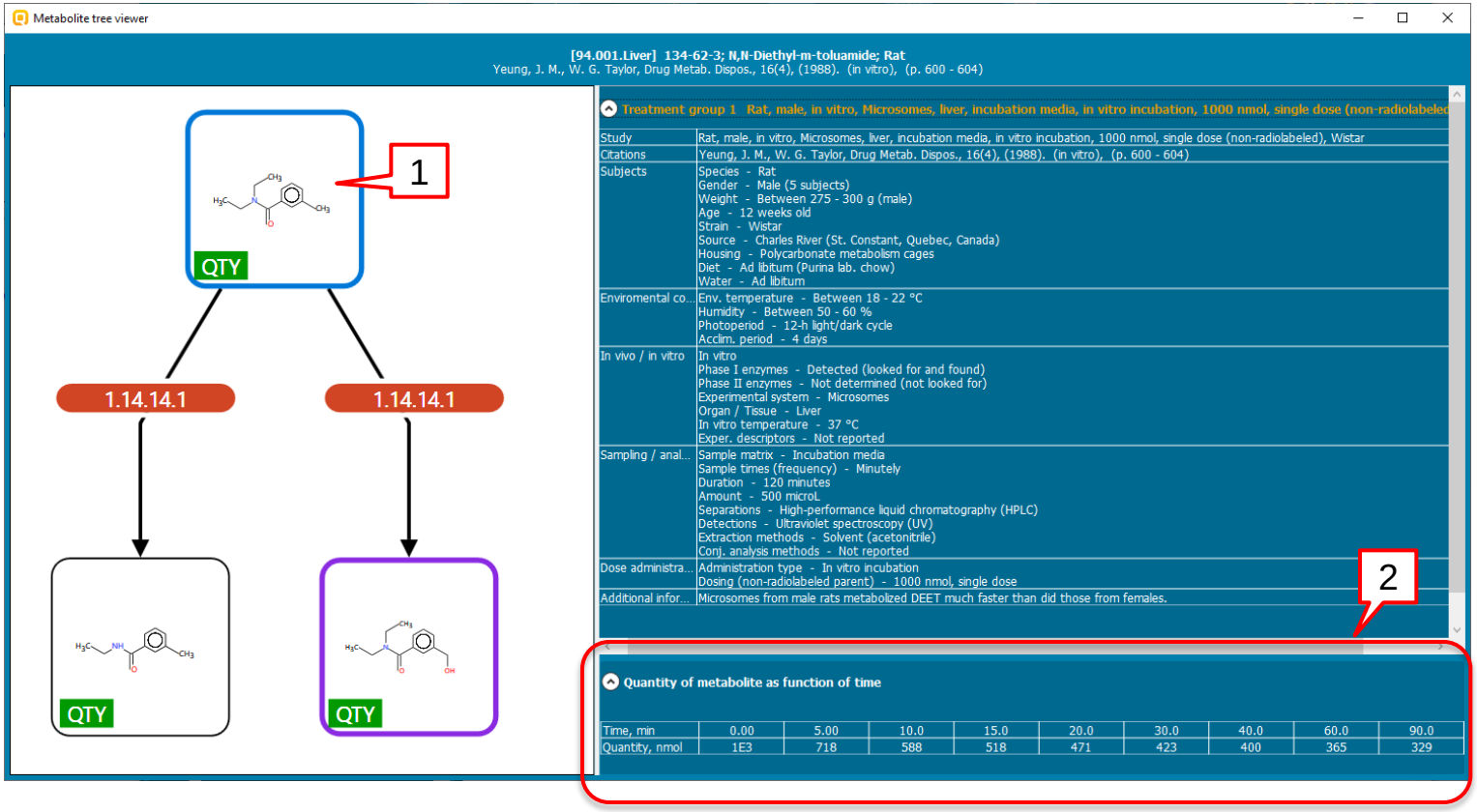
Profiling



The feature of top and right panel are:

- Information about the target chemical (1);
- Map number generated in the METAPATH software(2);
- The reference from which the data is taken is also included (3);
- Detailed information about the treatment groups is displayed upon expansion (4).

Profiling



- Once the treatment group is expanded, make left mouse click on the target/metabolite (1) to see its quantity as a function of time (2).

Outlook

- Aim
- **Background**
- Keywords
- Example for:
  - Visualizing quantitative data within Toolbox user interface
  - **Application of quantitative metabolic data in data gap filling**

Application of quantitative metabolic data in data gap filling:

Steps:

- Input list of chemicals
- Collect experimental data for skin sensitization endpoint
- Data gap filling

Application of quantitative metabolic data in data gap filling

- In this tutorial only a working example illustrating this functionality is shown.
- 13 chemicals with quantitative data are used.
- We are fully aware that this example is not scientifically well defined, however it is used solely to introduce you to this functionality.

Data gap filling
An overview

- Data Gap Filling (DGF) module gives access to three different data gap filling tools:
  - Read-across
  - Trend analysis
  - (Q)SAR models

- 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.

Application of quantative metabolic data in data gap filling

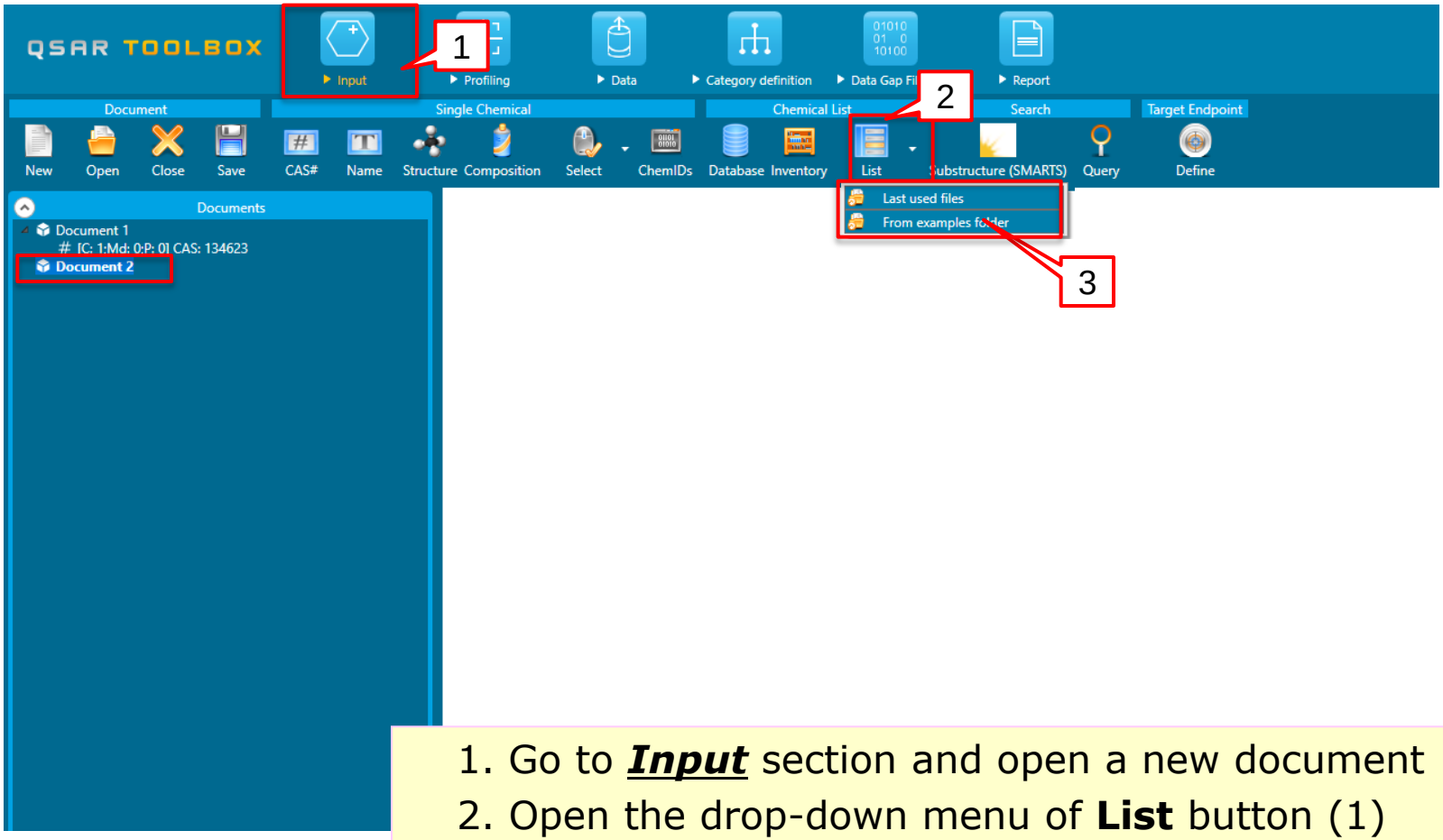
• Quantitative metabolic data can be used to filter analogues in data gap filling.

• Quantities cannot be used directly to filter out chemicals (quantities are not single values, but time series; often data comes in units, which are not convertible - i.e. mol/L vs mol/g protein).

• In this respect a reliable measure that can be used for filtering is the half-life of parent chemicals calculated from quantitative data.

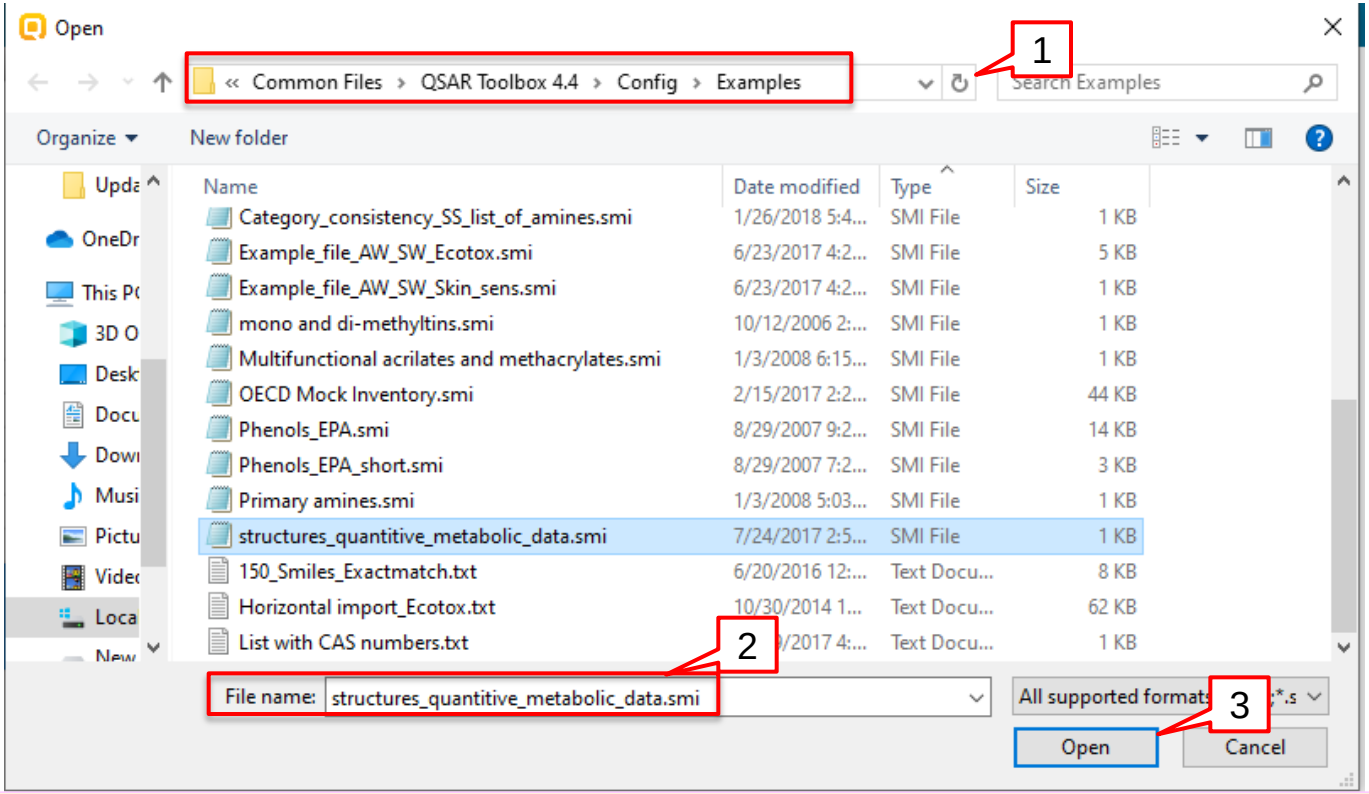
• As a result a new calculator "Half-Life (observed metabolism)" was implemented.

Input list of chemicals



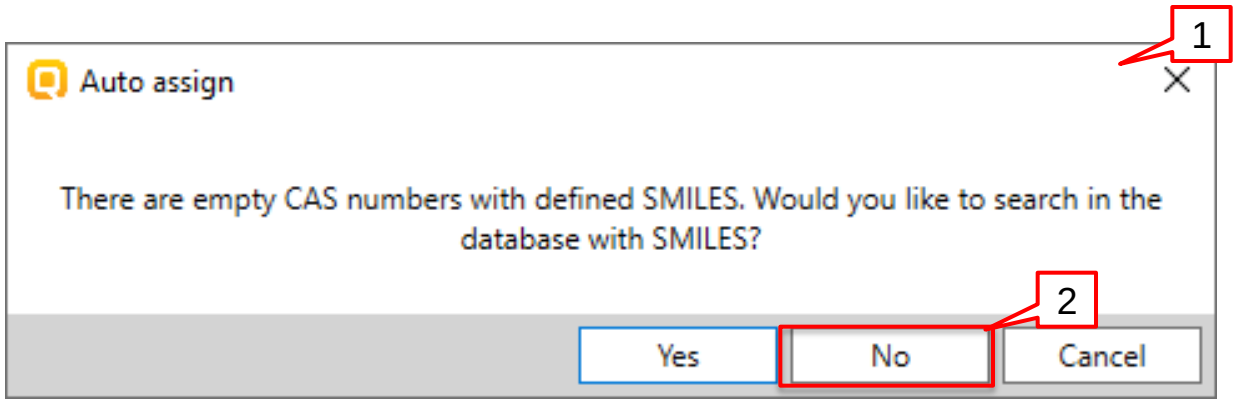
1. Go to ***Input*** section and open a new document
2. Open the drop-down menu of **List** button (1)
3. Select **From Example folder** (2)

Input list of chemicals



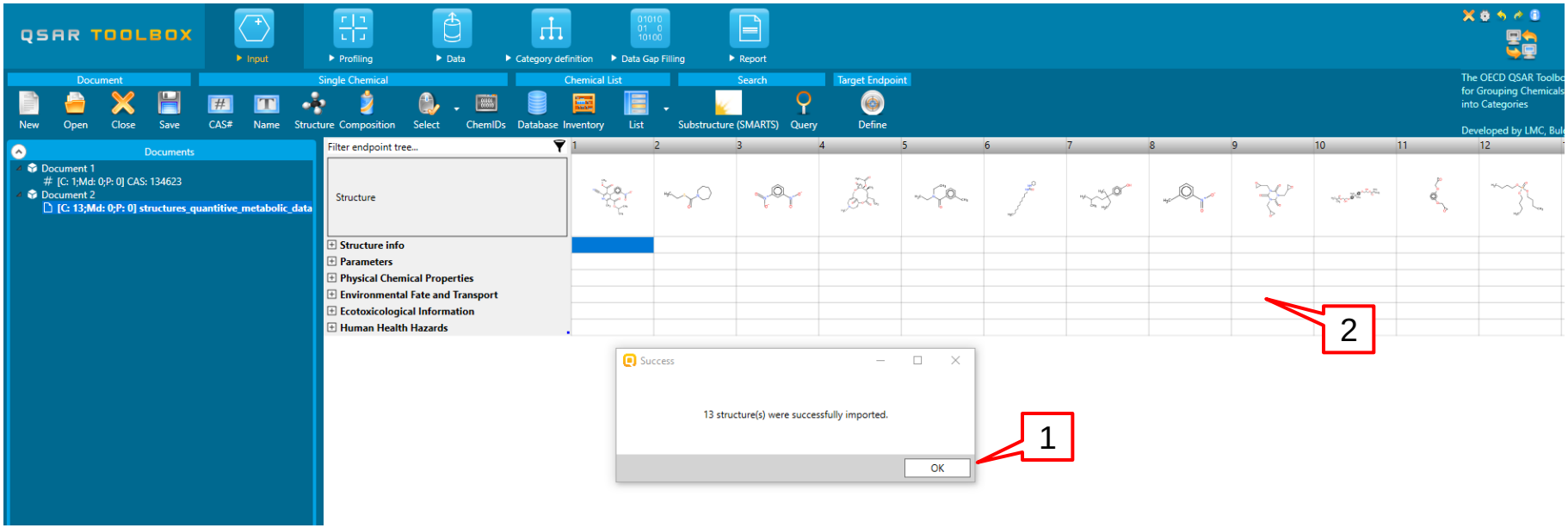
1. Examples folder directory in Toolbox is open (1);
2. Select ***structure_quantative_metabolic_data.smi***(2);
3. Click **Open** (3)

Input list of chemicals



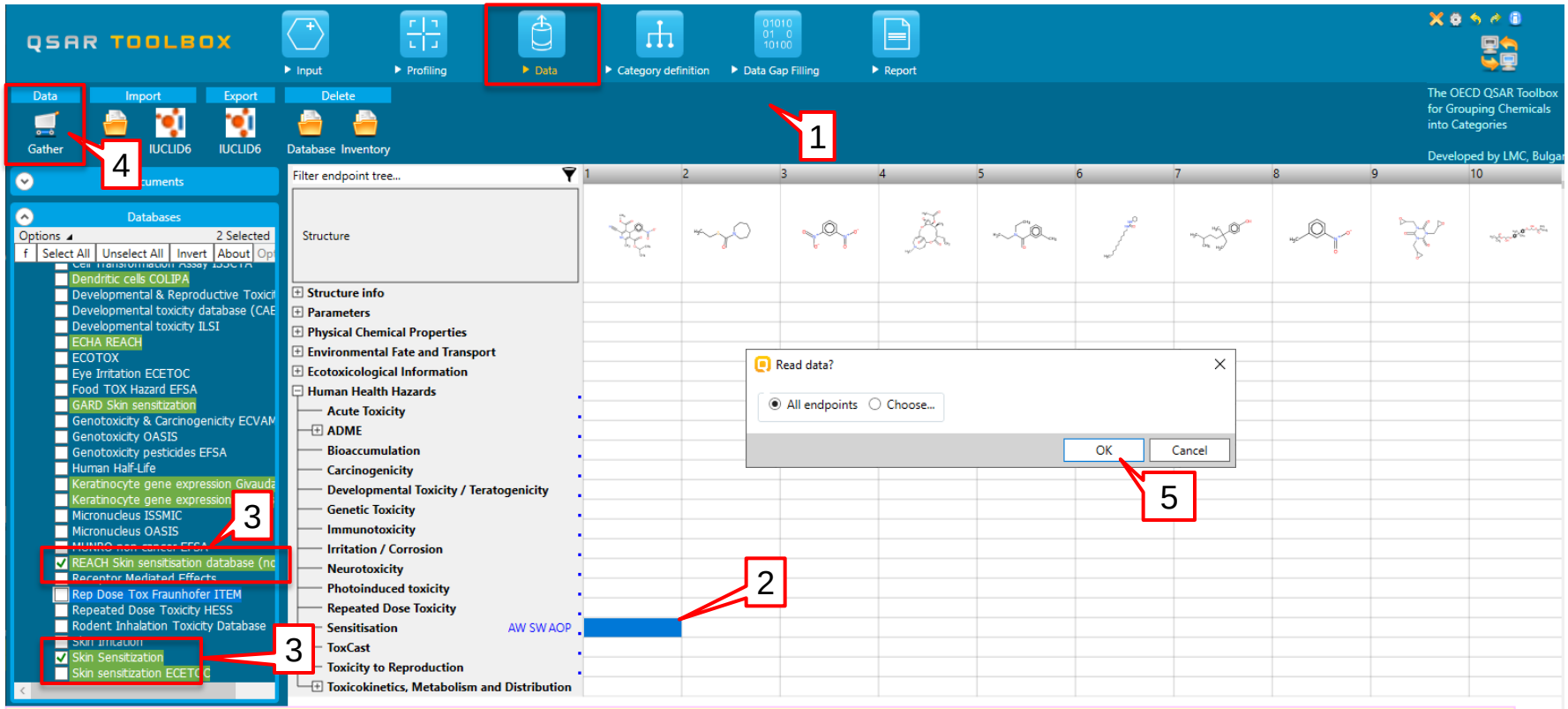
A pop-message (1) informs that some chemicals with defined SMILES have no CAS numbers. If you want to retrieve the CAS numbers from the Toolbox databases click **Yes**, otherwise click **No** (2). Click **No.**

Input list of chemicals



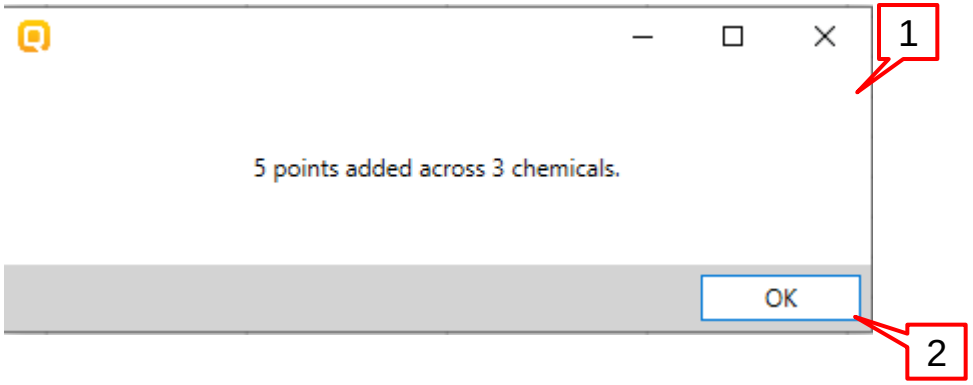
1. A message informing about the successful importing is shown, where you have to click on **OK** (1);
2. 13 chemicals are loaded in the data matrix (2).

Gathering of experimental data for skin sensitization



1. Go to ***Data module*** (1);
2. Click on level "Sensitisation" from the endpoint tree (2)
3. Select **Skin sensitization database** and **REACH Skin sensitisation database (normalised)** (3);
4. Click **Gather** (4), and then click **OK** to collect the data for all endpoints (5).

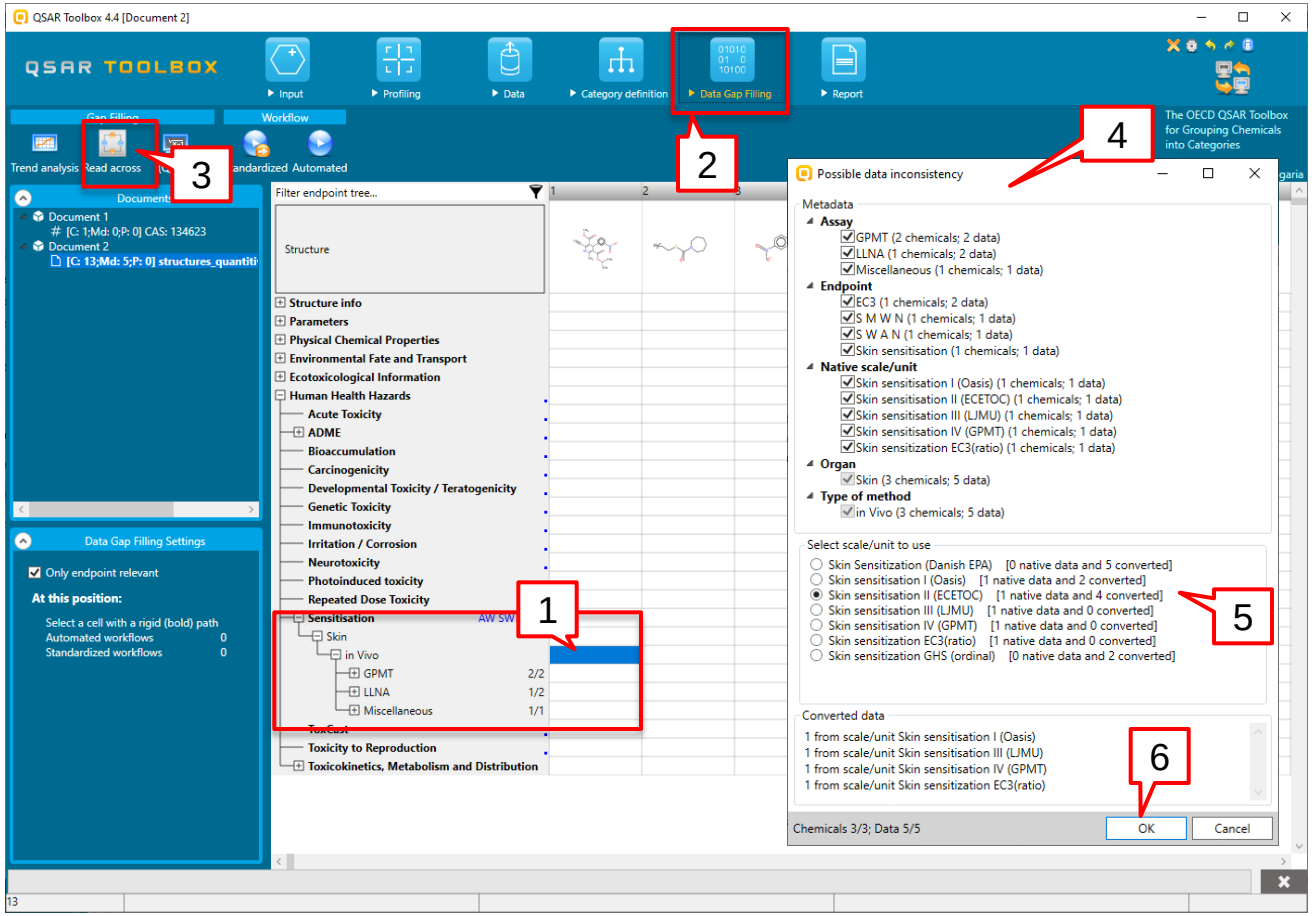
Gathering of experimental data for skin sensitization



1. A pop-up message appears (1)
2. Click **OK** (2);

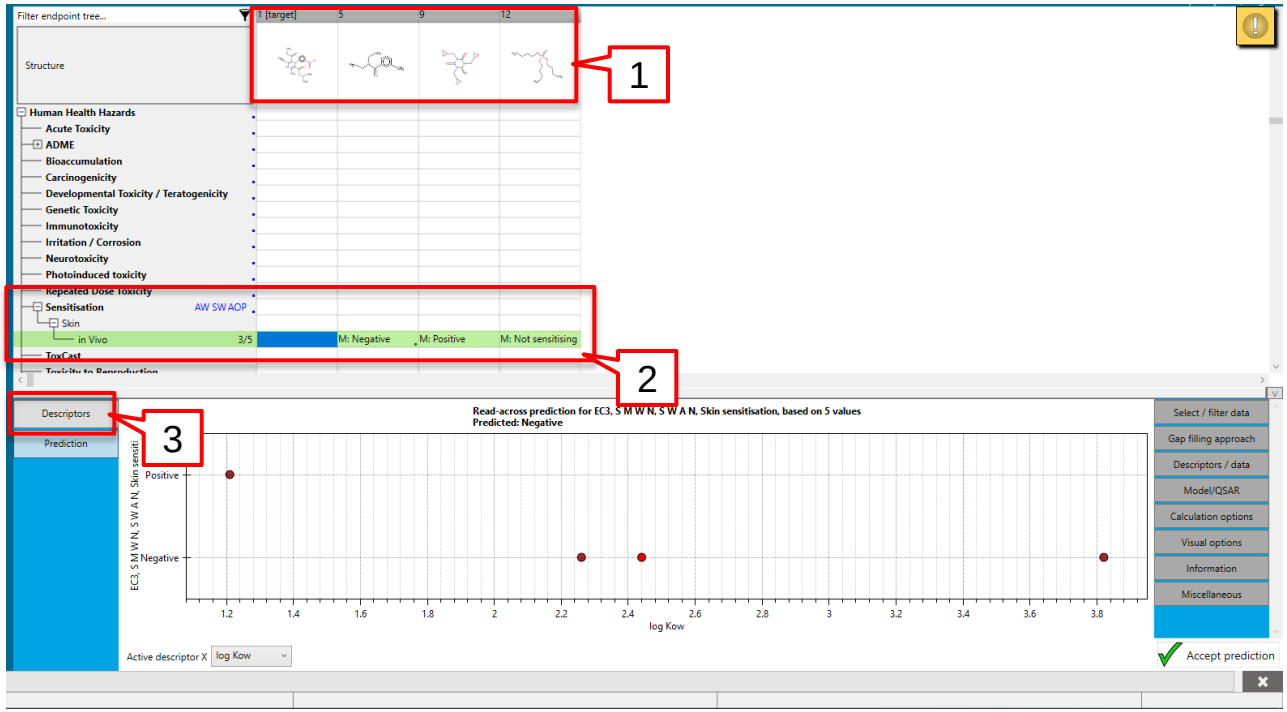
Data gap filling

1. Expand the endpoint tree and go to **Sensitization/Skin/in Vivo** (1);
2. Go to ***Data Gap Filling module*** (2);
3. Click **Read across** (3);
4. In Possible data inconsistency window (4) select **Skin sensitization II(ECETOC)**(5) and click **OK** (6).

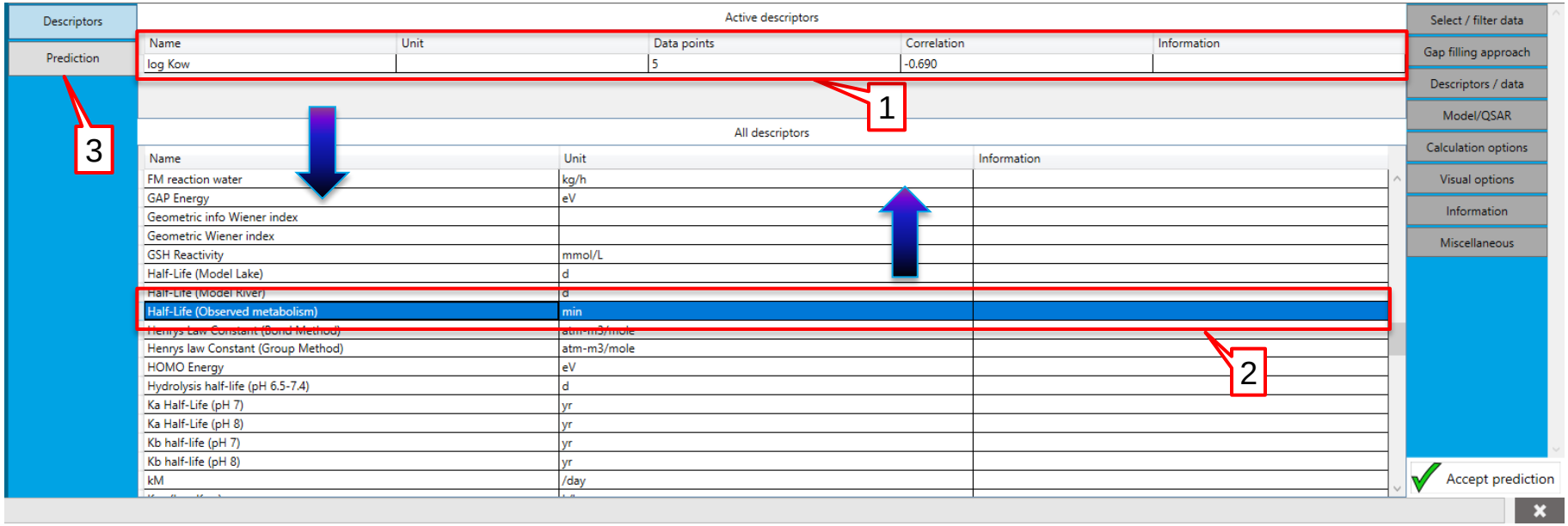


Data gap filling

1. Four chemicals are entered into the read-across; one target and three analogues (1)
2. The experimental data is displayed on the matrix. (2)
3. Select *Descriptors* (3) to change the descriptor on the x axis of the graph



Data gap filling



1. Double left click on the **Active descriptor** Log Kow (1) to shift it to the **All descriptors** list.
2. Then double left click on **Half-life (observed metabolism)** (2) to shift the descriptor to the **Active descriptors** panel, which makes it x-axis descriptor.
3. Click **Prediction** button (3).

Data gap filling

You can compare the observed half-life of the target chemical (1) with the observed half-lives of the analogues (2).