QSAR TOOLBOX

The OECD QSAR Toolbox for Grouping Chemicals into Categories

OECD QSAR Toolbox v.3.4

Predicting developmental and reproductive toxicity of Diuron (CAS 330-54-1) based on DART categorization tool and DART SAR model

Outlook

- **Background**
- Objectives
- The exercise
- Workflow

Background

- This is a step-by-step presentation designed to take the user through the workflow for filling data gap for reproductive and developmental toxicity by read-across based on an analogue approach.

Outlook

- Background
- **Objectives**
- The exercise
- Workflow

Objectives

This presentation demonstrates a number of functionalities of the Toolbox:

- Identify analogues of target chemical by applying DART scheme via two different ways:
  - For primary categorization
  - For subcategorization
- Retrieve experimental results available for those analogues.
- Fill data gaps by read across.
- Support read-across prediction by DART model.

Outlook

- Background
- Objectives
- **The exercise**
- Workflow

The exercise

- In this exercise we will predict the developmental and reproductive (DART) toxicity of **3-(3,4-dichlorophenyl)-1,1-dimethylurea CAS 330-54-1 (Diuron).**

- Two scenarios for defining the initial category of similar analogues will be applied:
  - Initial category identified by endpoint specific DART scheme
  - Initial category identified by empiric Organic functional group (OFG) with followed by subcategorization by DART scheme

- Gather available experimental data for the target chemical and identified analogues.

- Apply read across prediction based on analogue approach.

- Apply external DART model.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**

Workflow

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

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - **Input**

Chemical 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.

Chemical Input
Ways of Entering a Chemical

User Alternatives for Chemical ID:

A. Single target chemical
- Chemical Name
- Chemical Abstract Services (CAS) number (#)
- SMILES (simplified molecular information line entry system) notation/InChi
- Drawing chemical structure
- Select from User List/Inventory/Databases
- Chemical IDs such as EC number, Einecs number
- Query Tool

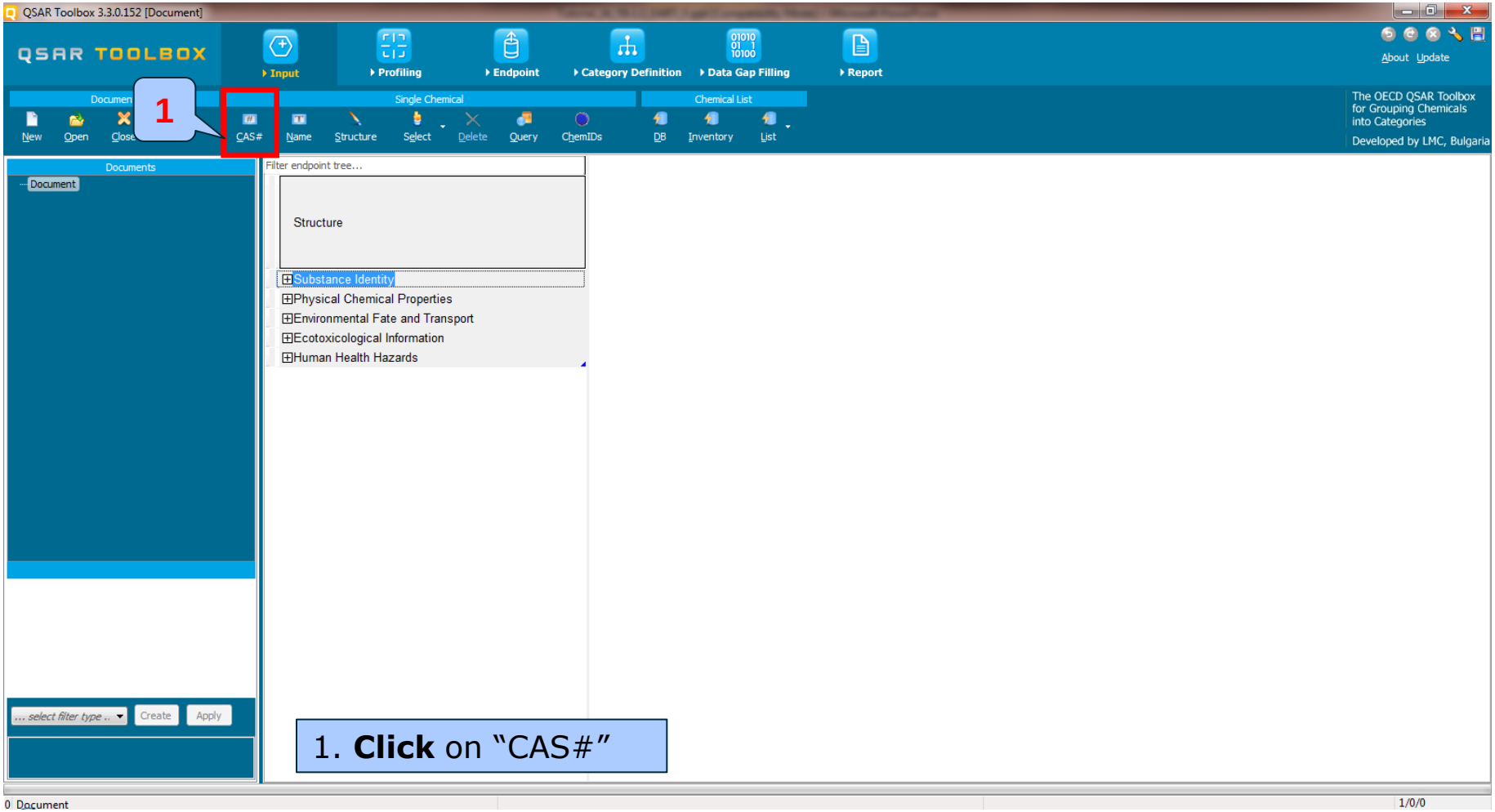
B. Group of chemicals
- User List/Inventory
- Specialized Databases

Getting Started

- Open the Toolbox.
- The six modules in the workflow are seen listed next to "QSAR TOOLBOX".
- Click on "Input" (see next screen shot).

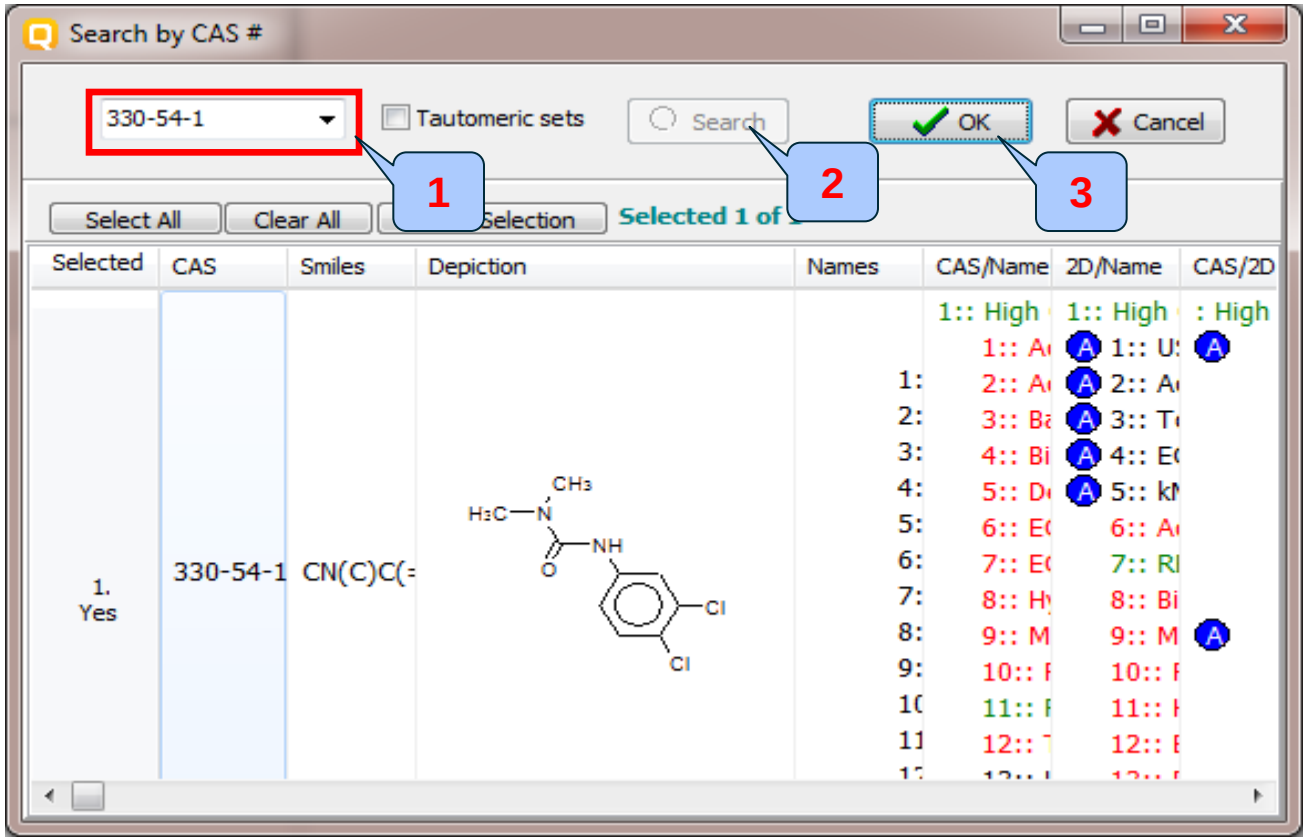
Chemical Input Screen

Input target chemical by CAS#



Chemical Input Screen

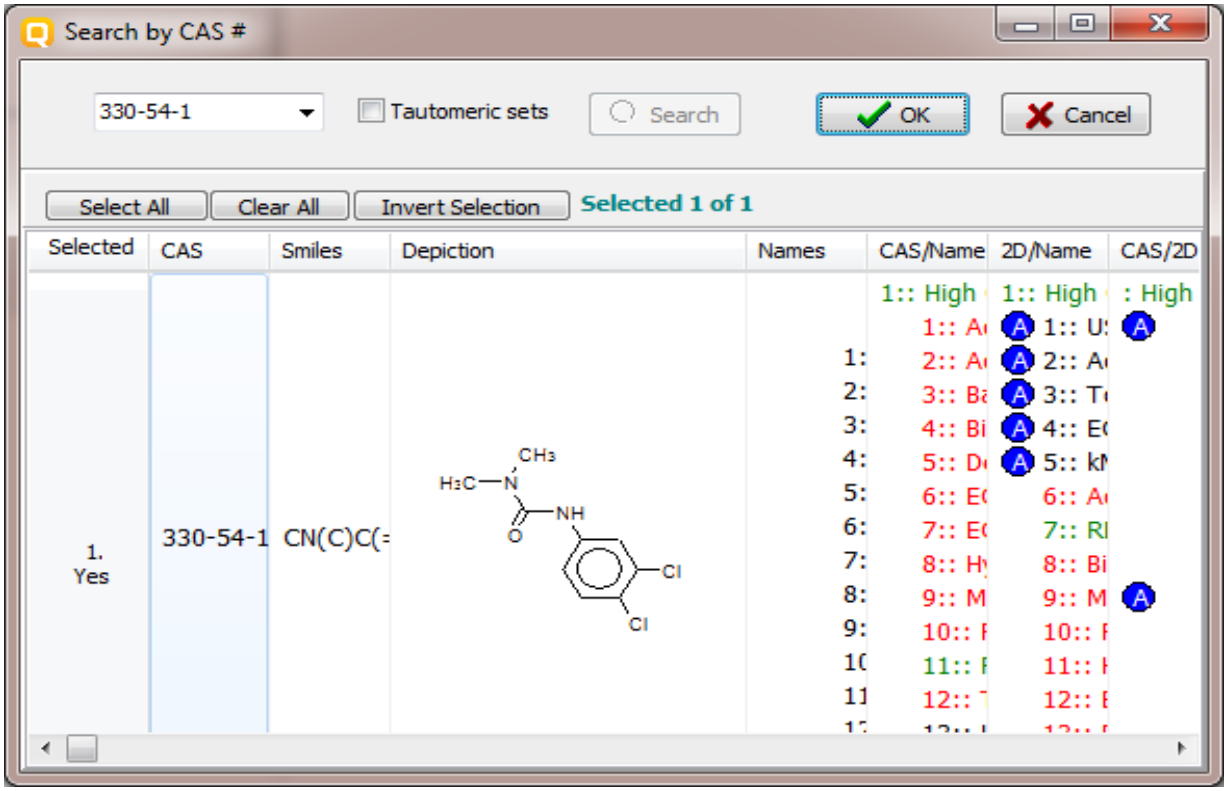
Enter CAS# 330-54-1



1. **Enter** the CAS# in the blank field; 2. **Click** "Search" button; 3. **Press** "OK"

Chemical Input
Target chemical identity

The Toolbox now searches the databases to find out if the CAS# you entered is linked to a molecular structure stored in the Toolbox. It is displayed as a 2-demensional depiction.

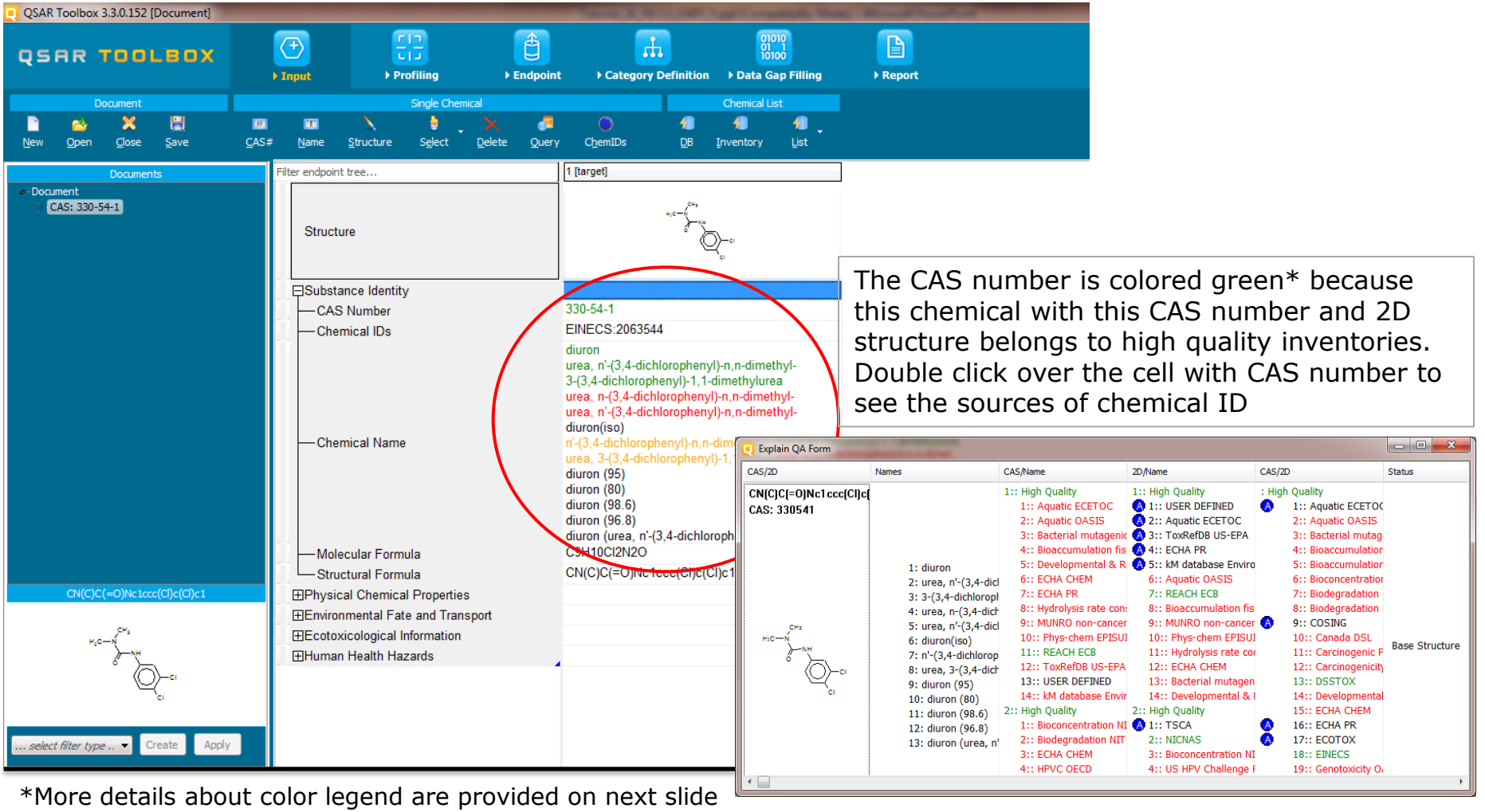


Chemical Input
Target chemical identity

- Double click "Substance Identity" displays the chemical identification information.
- The user should note that existing names of the target chemical are presented in different colours. This indicates the reliability of relation CAS-Name-SMILES for the target chemical (see next screen shots).
- The workflow on the first module is now complete, and the user can proceed to the next module.

Chemical Input

Target chemical identity



The CAS number is colored green* because this chemical with this CAS number and 2D structure belongs to high quality inventories. Double click over the cell with CAS number to see the sources of chemical ID

*More details about color legend are provided on next slide

Chemical Input
Chemical identity

The colour code indicates the reliability of the chemical identifier:

- **Green**: There is a high reliability between the identifier and the structure. This colour is applied if the identifier is the same in several quality assured databases.
- **Yellow:** There is only a moderate reliability between the identifier and the structure. The colour is applied if the identifier is the same in several databases for which the quality assurance could not be established.
- **Red**: There is a poor reliability between the identifier and the structure. The colour is applied if the identifier is allocated to different structures in different databases.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - Input
  - **Profiling**

Profiling
Overview

- "Profiling" refers to the electronic process of retrieving relevant information on the target compound, other than environmental fate, ecotoxicity and toxicity data, which are stored in the Toolbox database.
- Available information includes likely mechanism(s) of action, as well as observed or simulated metabolites.

Profiling
Side-Bar to Profiling

- For most of the profilers, background information can be retrieved by highlighting one of the profilers and clicking on "View".

- Detailed information for DART v. 1.0 scheme (Developmental and Reproductive toxicity) is provided on next slide.

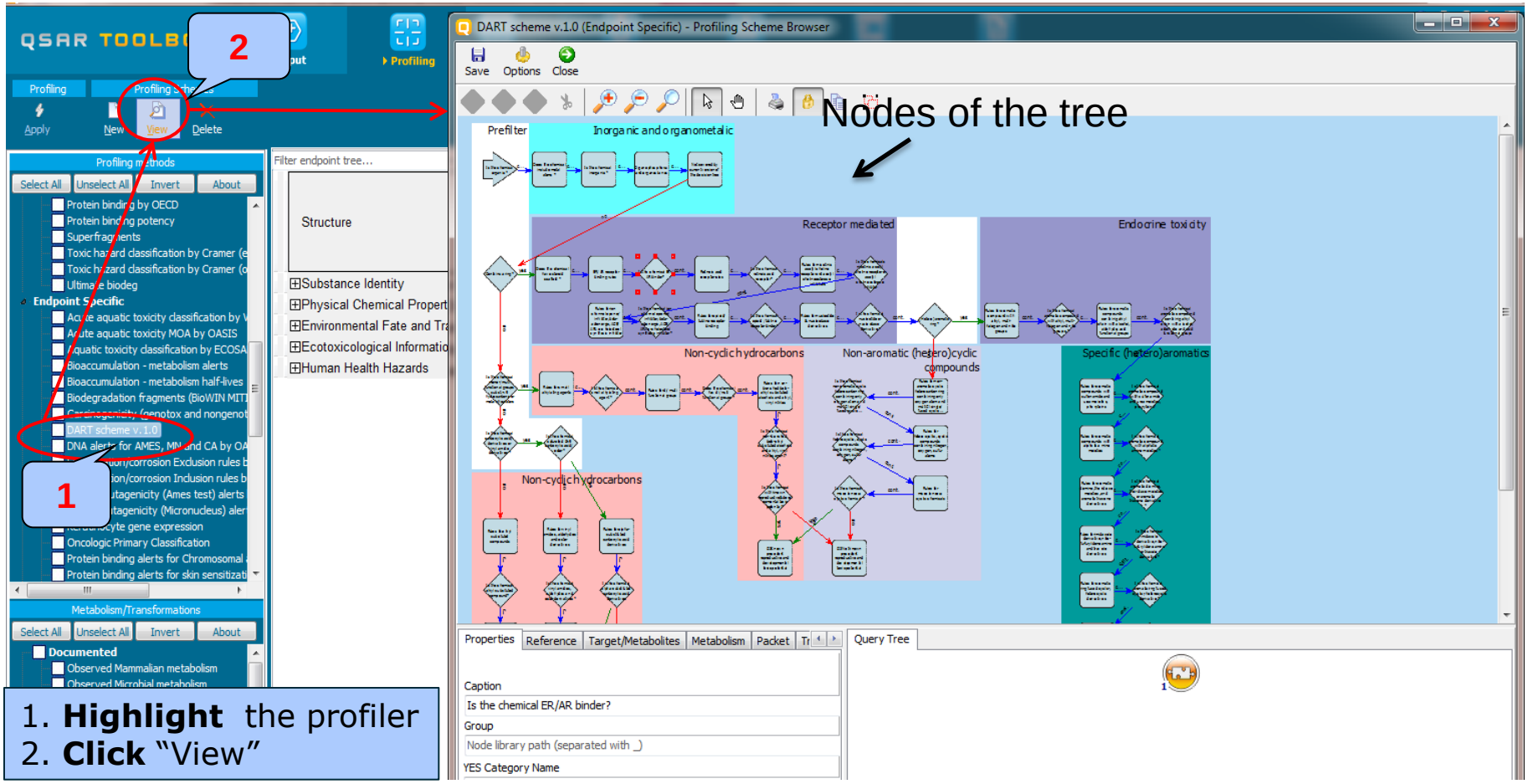
Developmental and Reproductive Toxicity (DART)

Background

- DART scheme is an adaptation of a framework for identifying chemicals with structural features associated with the potential to act as reproductive or developmental toxicants outlined in the journal of Wu S et all [1].
- It is implemented as a pilot endpoint specific scheme, developed on the basis of the combination of known modes of action (MOA) and associated structural features.
- DART scheme include 25 categories and 125 sub-categories organized as a decision scheme. Definition of the categories are based on a detailed review of 716 chemicals that have been evaluated for their DART potential effect. Mechanistic interpretation and reliability is provided for each category.
- DART scheme is implemented as a profiling/categorization tool and as a SAR model.
- It can be used both as a component of a screening system to identify chemicals of potential concern, and as part of weight of evidence decisions based on structure-activity relationships (SAR), to fill data gaps without generating additional test data.
- Output of the scheme indicate that the chemical of interest is associated with chemical structures:
  - known to have DART - *Known precedent reproductive and developmental toxic potential*
  - not known to have DART - *Not known precedent reproductive and developmental toxic potential*
  - which have structural features outside the domain of the DART decision tree - *Not cover by the decision tree.*

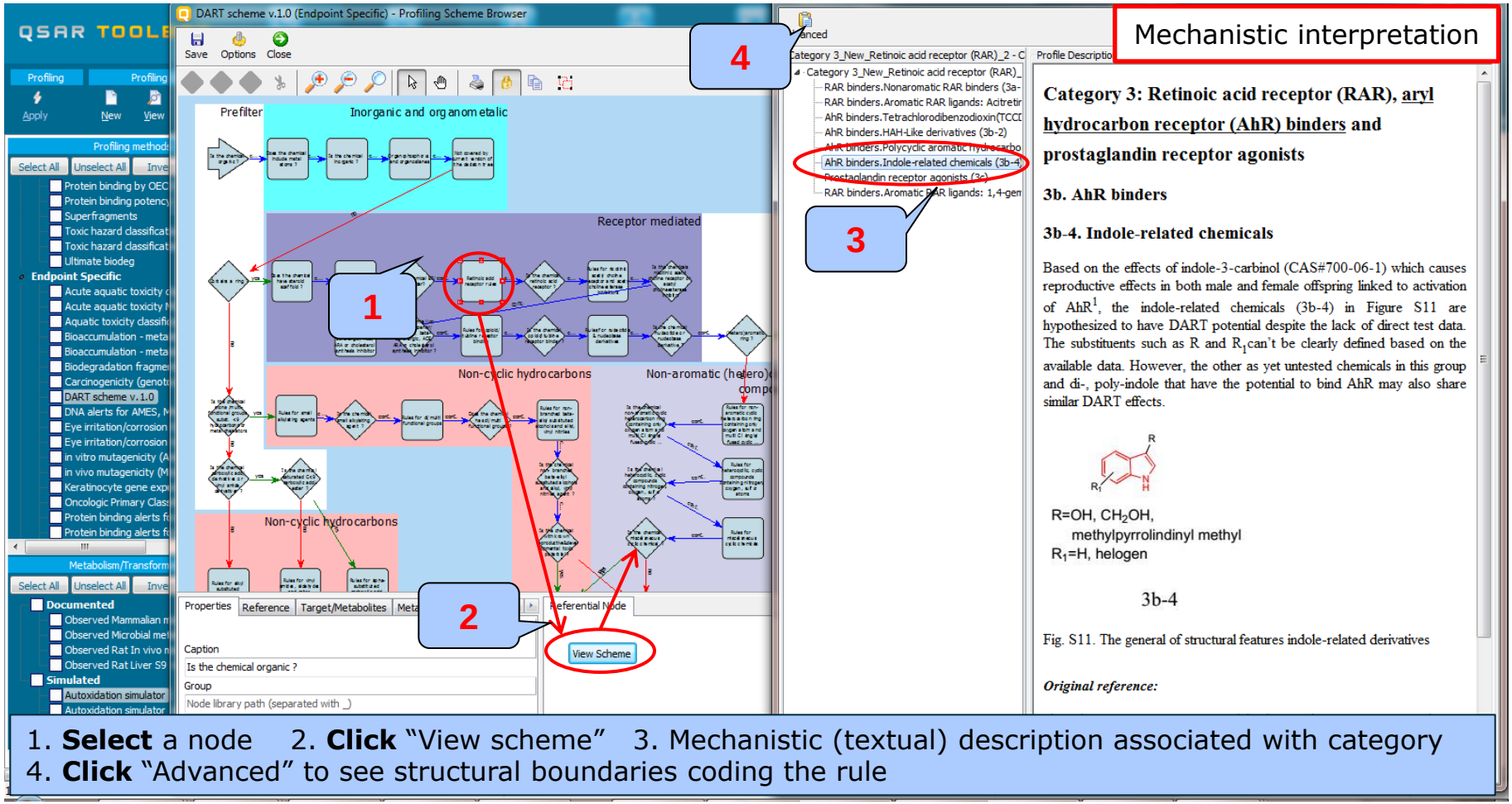
1. Wu S, Fisher J, Naciff J, Laufersweiler M, Lester C, Daston G, Blackburn K. Framework for identifying chemicals with structural features associated with the potential to act as developmental or reproductive toxicants. Chem Res Toxicol. 2013 Dec 16;26(12):1840-61.

Profiling
Overview of DART scheme



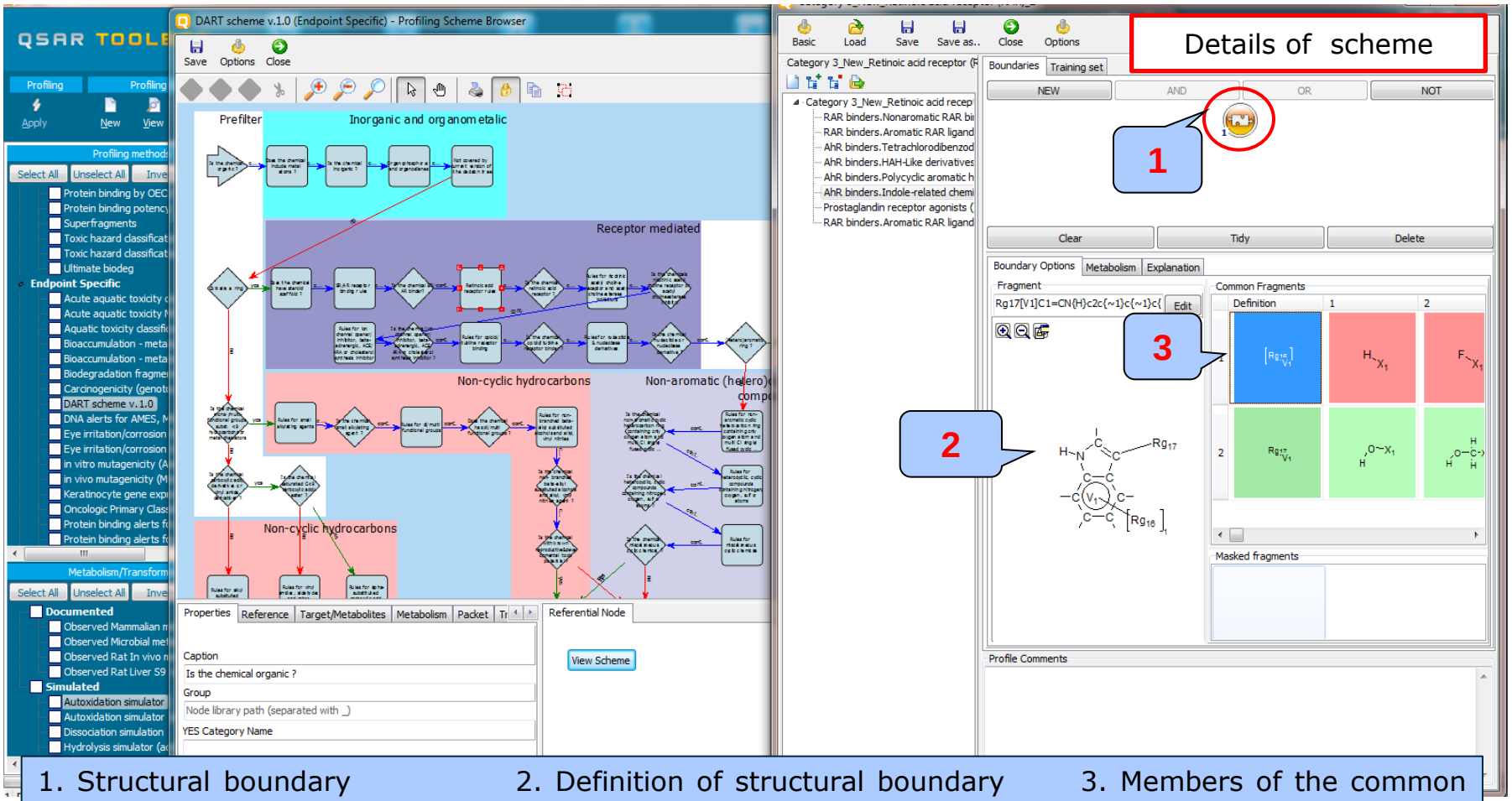
1. **Highlight** the profiler
2. **Click** "View"

Profiling
Overview of DART scheme



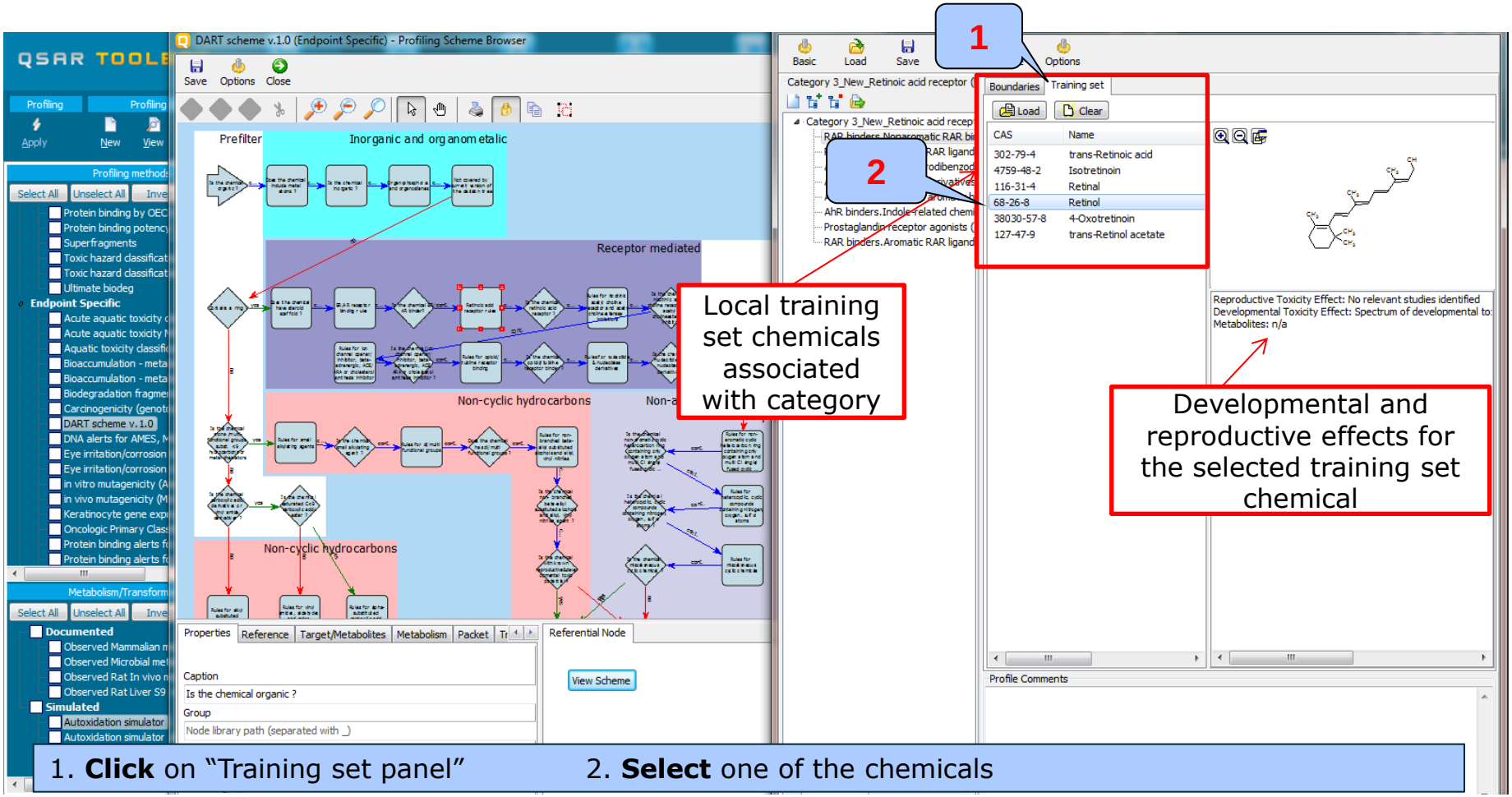
1. **Select** a node 2. **Click** "View scheme" 3. Mechanistic (textual) description associated with category
4. **Click** "Advanced" to see structural boundaries coding the rule

Profiling
Overview of DART scheme



1. Structural boundary fragment
2. Definition of structural boundary
3. Members of the common

Profiling
Overview of DART scheme



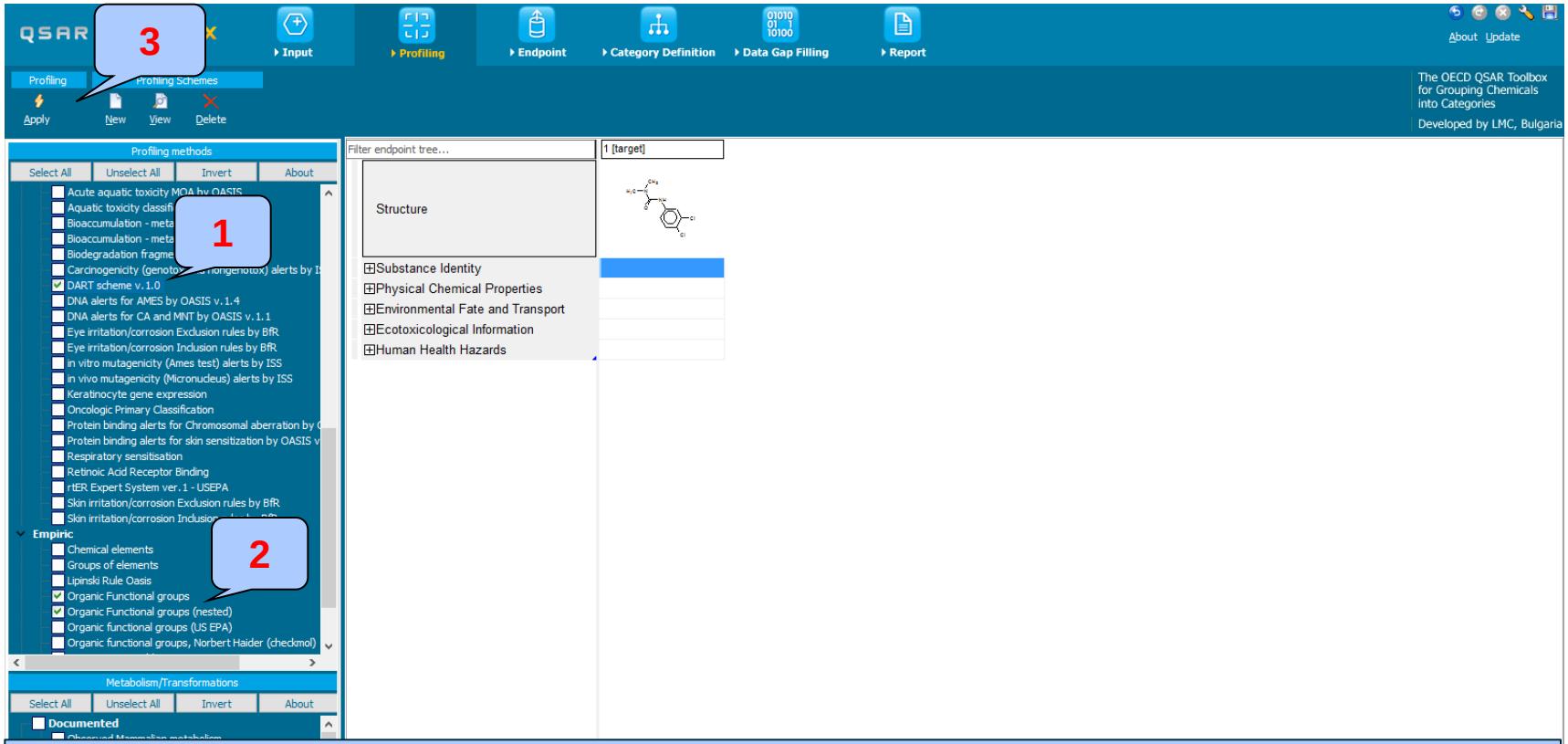
Profiling

Application of endpoint specific profiling schemes

- In this example profiling by DART and OFG is applied in order to analyse the potential to cause DART toxicity and to identify the general structural fragments available in the molecule, which further could be used for categorization.

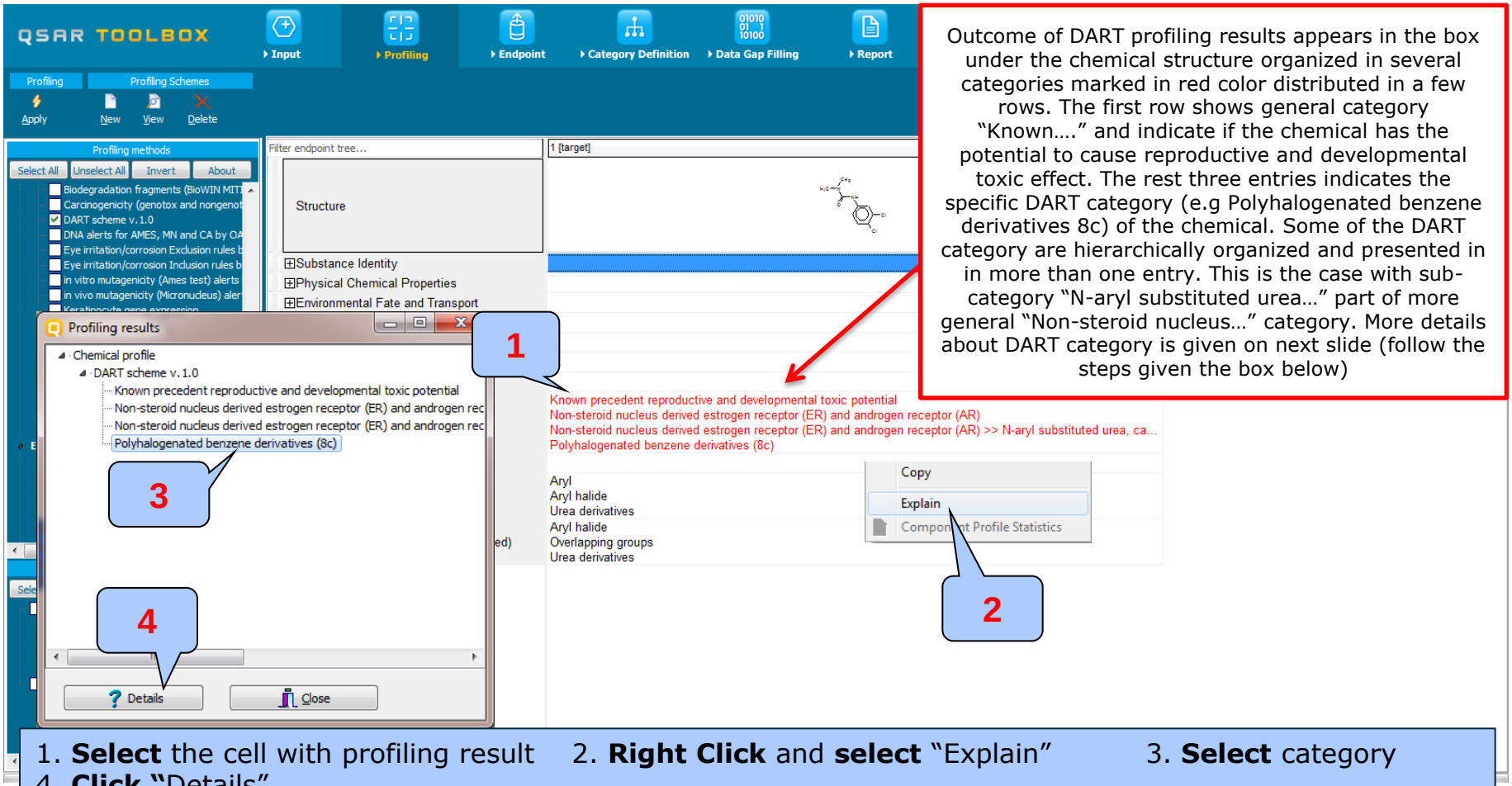
- Follow the steps:
  - Select DART v1.0
  - Select two Organic functional group profilers
  - Click Apply

Profiling
Apply related profiling schemes



1. **Select** the "DART v1.0 profiling scheme"
2. **Select** two "Organic functional groups" profilers
3. **Click** "Apply"

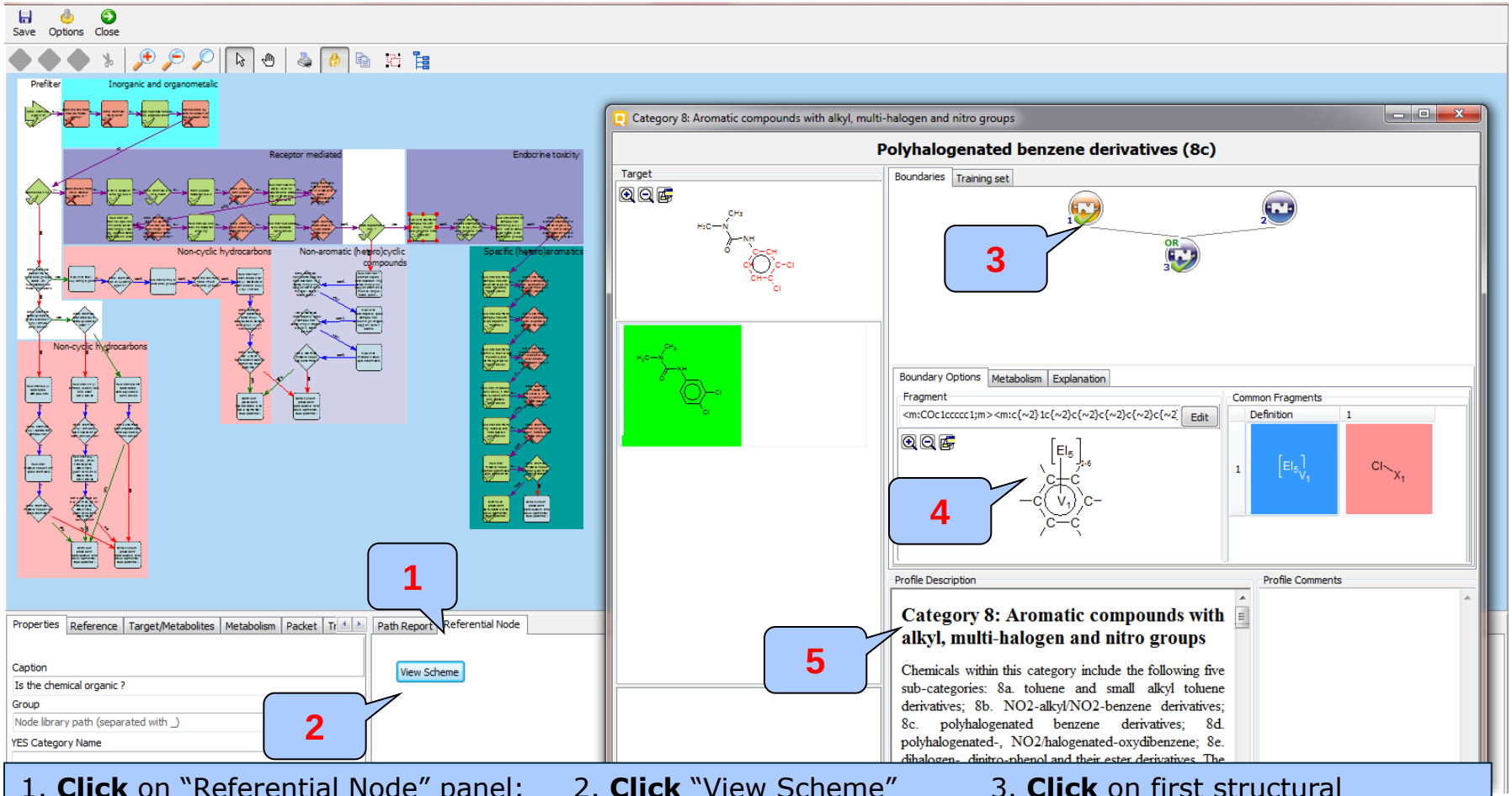
Profiling
Outcome of profiling results



Outcome of DART profiling results appears in the box under the chemical structure organized in several categories marked in red color distributed in a few rows. The first row shows general category "Known…." and indicate if the chemical has the potential to cause reproductive and developmental toxic effect. The rest three entries indicates the specific DART category (e.g Polyhalogenated benzene derivatives 8c) of the chemical. Some of the DART category are hierarchically organized and presented in in more than one entry. This is the case with sub-category "N-aryl substituted urea…" part of more general "Non-steroid nucleus…" category. More details about DART category is given on next slide (follow the steps given the box below)

1. **Select** the cell with profiling result
2. **Right Click** and **select** "Explain"
3. **Select** category
4. **Click** "Details"

Profiling
Explain of profiling results



1. **Click** on "Referential Node" panel; 2. **Click** "View Scheme" 3. **Click** on first structural boundary 4. Structural fragment coding the rule 5. Textual description of the rule

Profiling
Interpretation of profiling results

- Profiling result shows that the target chemical is classified as:
  - "Known precedent for DART effect" based on classification into two DART sub-categories: "Polyhalogenated benzenes" and "N-aryl substituted ureas". Will be further investigated.
  - "Aryl, Urea derivatives and Aryl halide" by OFG, which will be used further for identifying analogues.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - Input
  - Profiling
  - **Endpoint**

Endpoint
Overview

- "Endpoint" refers to the electronic process of retrieving the environmental fate, ecotoxicity 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).

Endpoint
Case study

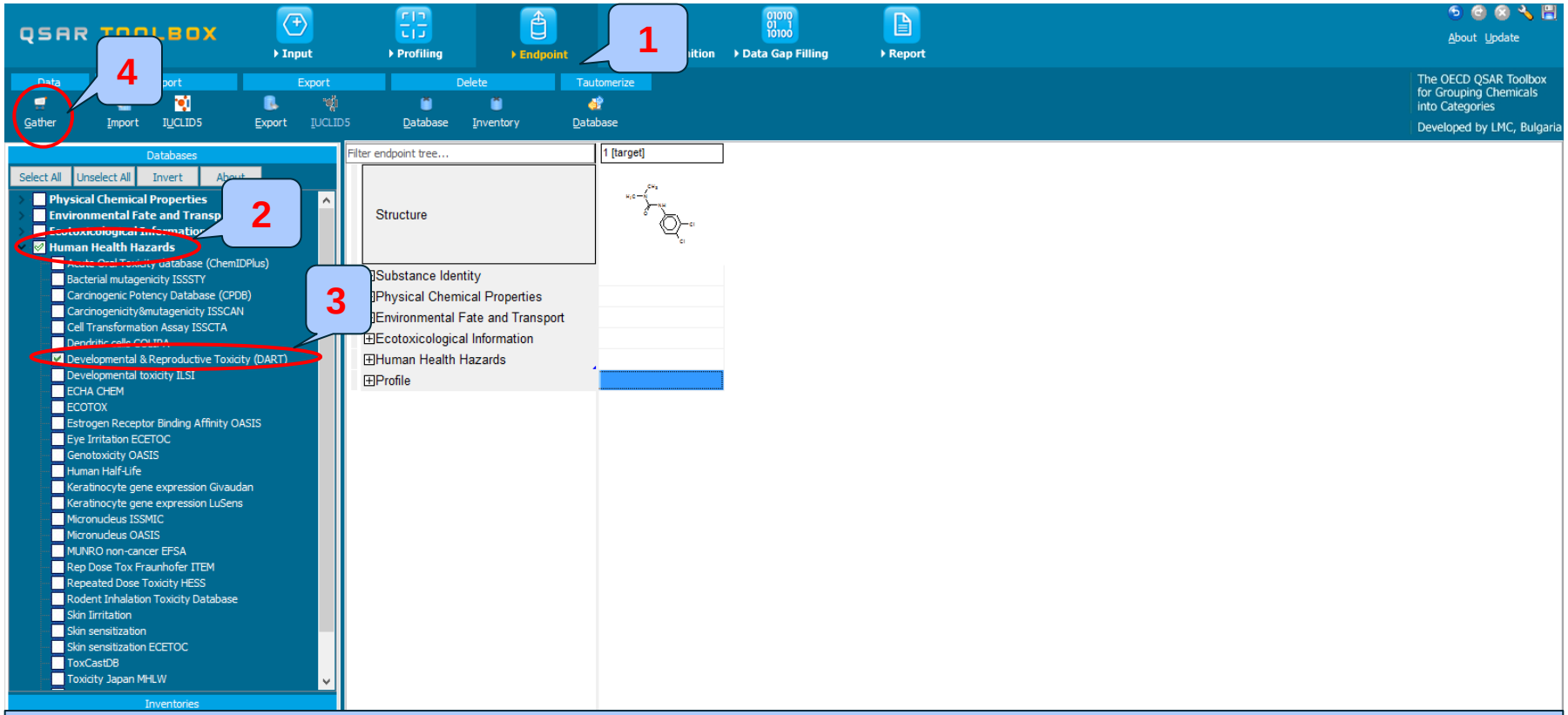
- In this example, we limit our data gathering to two toxicity endpoints: developmental and reproductive toxicity.

- DART database has been implemented into the Toolbox 3.3

- Developmental and Reproductive Toxicity (DART) database – 716 chemicals with 1430 data points separated as follows:
    - Developmental toxicity (716 data points)
    - Reproductive toxicity (714 data points)

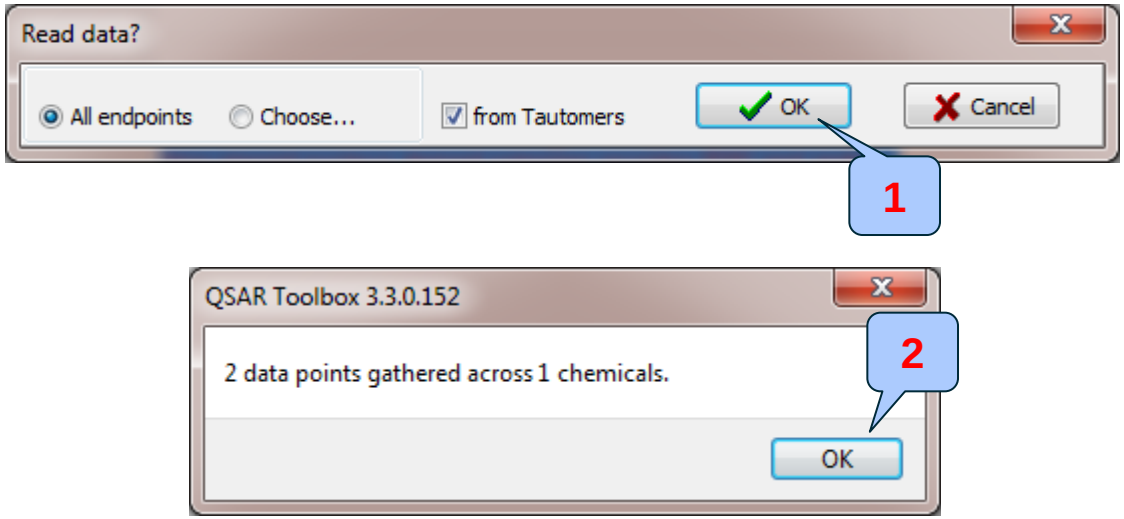
- In this example, we collect data from the DART database containing experimental results for developmental and reproductive toxicity(DART).
    - Click on "Endpoint" in the Toolbox workflow.
    - Expand the "Human Health Hazards" section
    - Click on the box to select that database.
    - Click on "Gather data" (see next screen shot).

Endpoint
Gather data



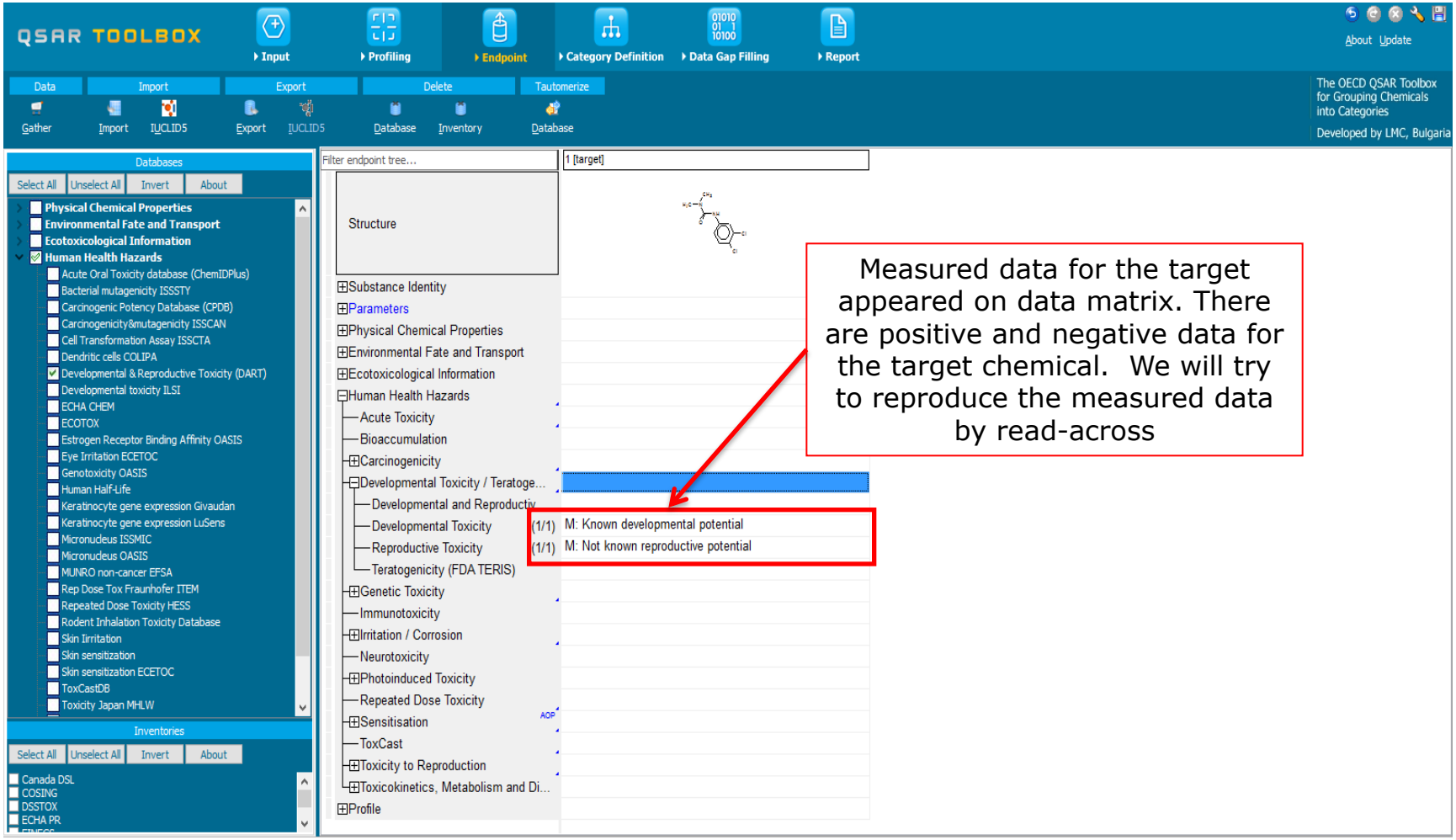
1. **Go** to "Endpoint"
2. **Expand** the "Human Health Hazards" section
3. **Select** database related to the target endpoint: "Developmental & Reproductive Toxicity (DART)"
4. **Click** "Gather"

Endpoint
Gather data



1. **Click** "OK" to extract data from database
2. The message informs you that 2 data points are gathered for the target chemical. **Click** "OK"

Endpoint
Gather data



Recap

- In the first module, you have entered the target chemical being sure of the correctness of the structure.

- In the second module, you have profiled the target chemical and found that the target could cause DART effect. It is categorized as known precedent for developmental and reproductive toxicity. This is due to the chemical pertaining to two chemical classes associated with DART toxicity.

- In the third module, you have found that there are two experimental data for the target structure: positive developmental and negative reproductive. We will try to reproduce them using read across analysis.

- Before proceeding with the "Data Gap Filling" module, the user should define a category with similar analogues. Two scenarios are played for identifying analogues:
  - DART scheme used as a categorization tool (used for phase I, see next slides)
  - DART scheme used in subcategorization procedure (used for phase II)

- Click on "Category Definition" to move to the next module.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - Input
  - Profiling
  - Endpoint
  - **Category definition**
    - **Overview**
    - Scenario 1
    - Scenario 2

Category Definition
Grouping methods

- The forthcoming 4 slides provide basic information about definition and procedure of "Category definition".
- 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.
- Detailed information about grouping chemical (Chapter 4) could be found in document "Manual for Getting started" published on OECD website:

  http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm

Basic guidance for category formation and assessment

Suitable categorization phases:

1. Structure-related profilers.
2. Endpoint specific profilers (for sub-cat).
3. Additional structure-related profilers, if needed to eliminate dissimilar chemicals (to increase the consistency of category) (e.g. chemical elements).

Performing categorization:

1. Categorization phases should be applied successively.
2. The application order of the phases follows three general stages but variations within them are case specific.
3. More than one category can be used within one phase for forming one final category.
4. Some of the main phases could be skipped if consistency of category members is reached.

Graphical illustration of suitable categorization phases is shown on next slide.

Suitable Categorization/Assessment Phases

Phase I. Structure based

- US EPA Categorization
- OECD Categorization
- Organic functional group
- Structural similarity
- ECOSAR

Broad grouping
Endpoint Non-specific

Repeating Phase I due to Multifunctionality of chemicals

Phase II. Mechanism based

- DNA binding mechanism
- Protein binding mechanism
- Genotoxicity/carcinogenicity
- DART v1.0
- Cramer rules
- Verhaar rule
- Skin/eye irritation corrosion rules
- Repeated dose profiler (NITE)

Subcategorization
Endpoint Specific

Metabolism accounted for

Phase III. Eliminating dissimilar chemicals

Apply Phase I – for structural dissimilarity
Filter by test conditions – for Biological dissimilarity

Subcategorization
Endpoint Specific

Category Definition

Grouping methods – phase I

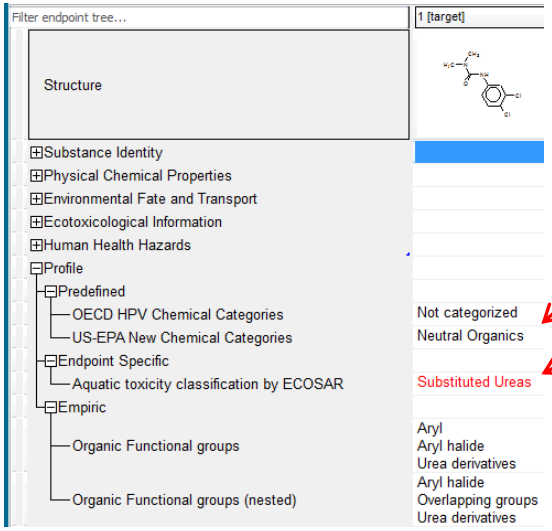
Suitable Categorization/Assessment Phases

Phase I. Structure based

- US EPA Categorization
- OECD Categorization
- Organic functional group
- Structural similarity
- ECOSAR

Broad grouping
Endpoint Non-specific

Phase I categorization in Toolbox



Structural similarity, Dice ACF, 50%

127 analogues are identified. It is not recommended to use "Neutral organic" as phase I*

9 analogues are identified**

8 analogues are identified

8 analogues are identified

3 analogue are identified

*Neutral organic category include chemicals having different functionalities as alcohols, ketones, ethers etc. In this respect the basic principle illustrated on slide 41 that structurally similar chemicals may elicit similar effects would not be preserved, because Neutral organic mixed many different functionalities.

**OFG is used for primary categorization, because the two basic functionalities available within the molecule: "Ureas" and "Aryl halides" will be preserved in the group of identified analogues, while the ECOSAR omits the "Aryl halide" functionality and identifies "Substituted ureas" only.

Category Definition

Grouping methods – phase II

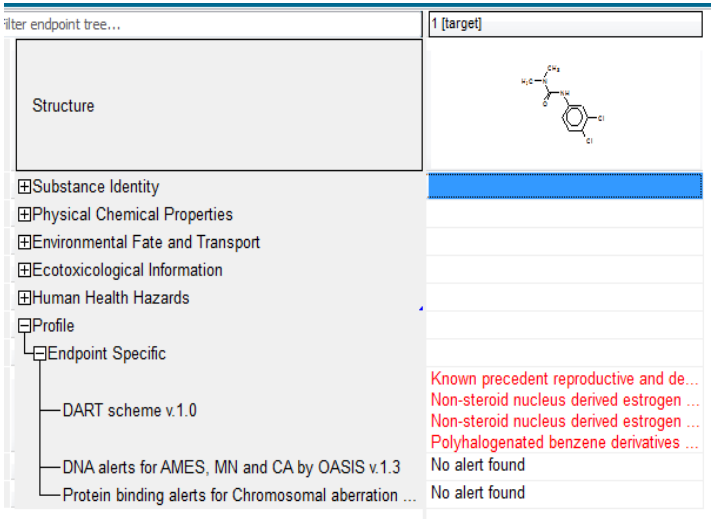
Suitable Categorization/Assessment Phases

Phase II. Mechanism based

- DNA binding mechanism
- Protein binding mechanism
- DART v1.0
- Repeated dose profiler (NITE)

Subcategorization
Endpoint specific

Phase II categorization in Toolbox



5 analogues are identified

In this case it is not reasonable to use DNA or Protein binding profiler for categorization.

Category Definition
Grouping methods

- Based on these classifications and basic guidance for grouping chemicals explained on the previous slides, two scenarios for identifying similar analogues have been applied in further read-across analysis:
  - Scenario 1: DART v1.0 scheme used as a categorization tool (applied for primary categorization - phase I).
  - Scenario 2: OFG is used for primary categorization with forthcoming subcategorization by DART scheme (applied for subcategorization - phase II).
- Identifying analogues based on two scenarios mentioned above will be applied in further read-across analysis.

Outlook

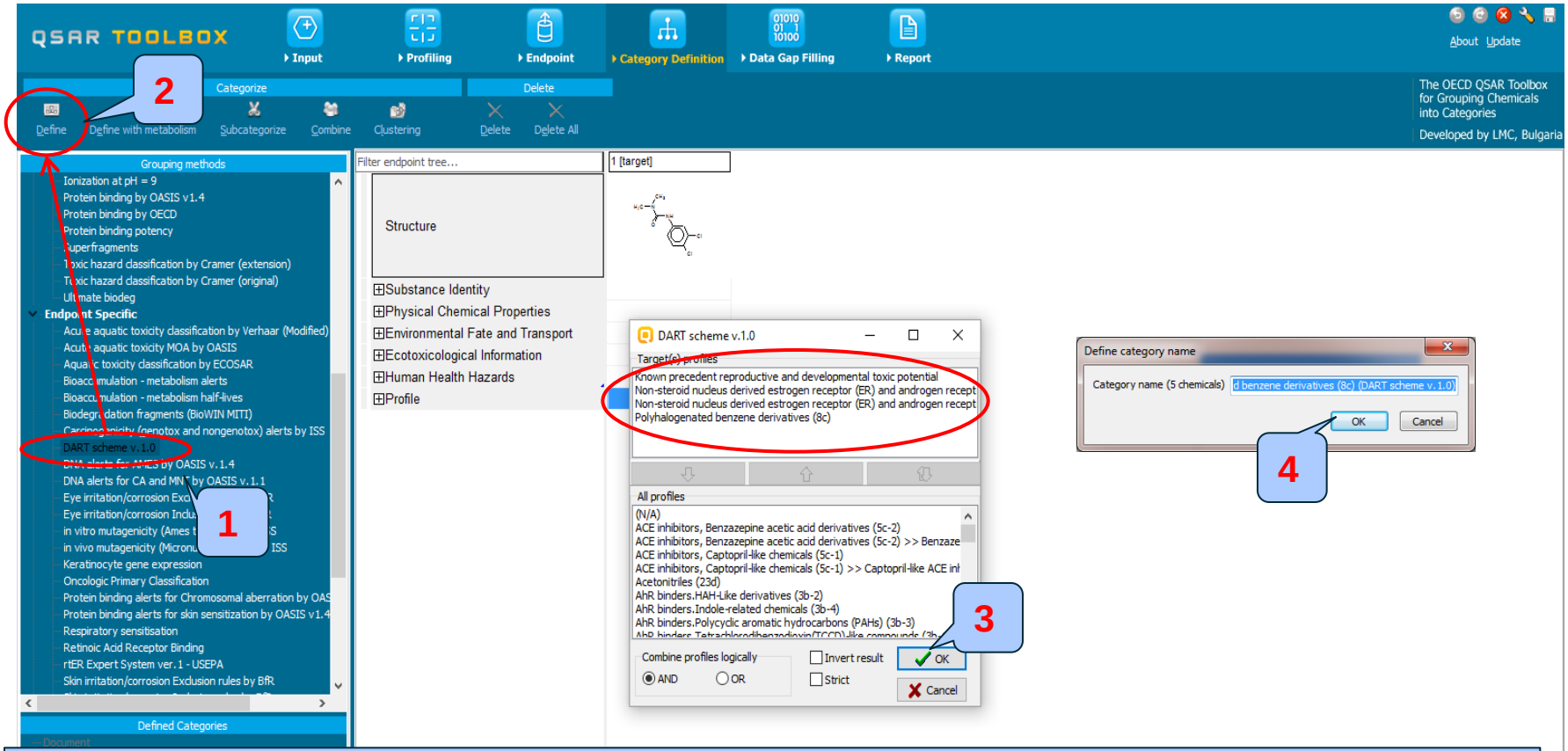
- Background
- Objectives
- The exercise
- **Workflow**
  - Input
  - Profiling
  - Endpoint
  - **Category definition**
    - Overview
    - **Scenario 1**
    - Scenario 2

Category Definition
Grouping methods

- Based on these classifications and basic guidance for grouping chemicals explained on the previous slides, two scenarios for identifying similar analogues have been applied in further read-across analysis:
  - Scenario 1: DART v1.0 scheme used as a categorization tool (applied for primary categorization - phase I).
  - Scenario 2: OFG is used for primary categorization with forthcoming subcategorization by DART scheme (applied for subcategorization - phase II).

Category Definition

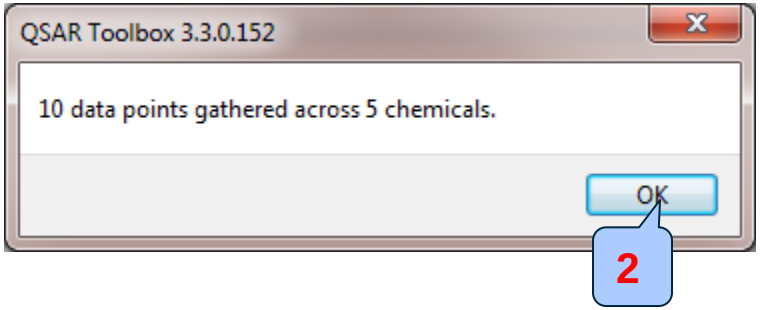
Scenario 1: DART scheme used as a categorization tool



1. **Highlight** the "DART scheme v1.0";
2. **Click** "Define";
3. **Click** "OK" to confirm the defined categories for the target chemical;
4. Five analogues are identified. **Click** "OK"

Category Definition

Scenario 1: DART scheme used as a categorization tool

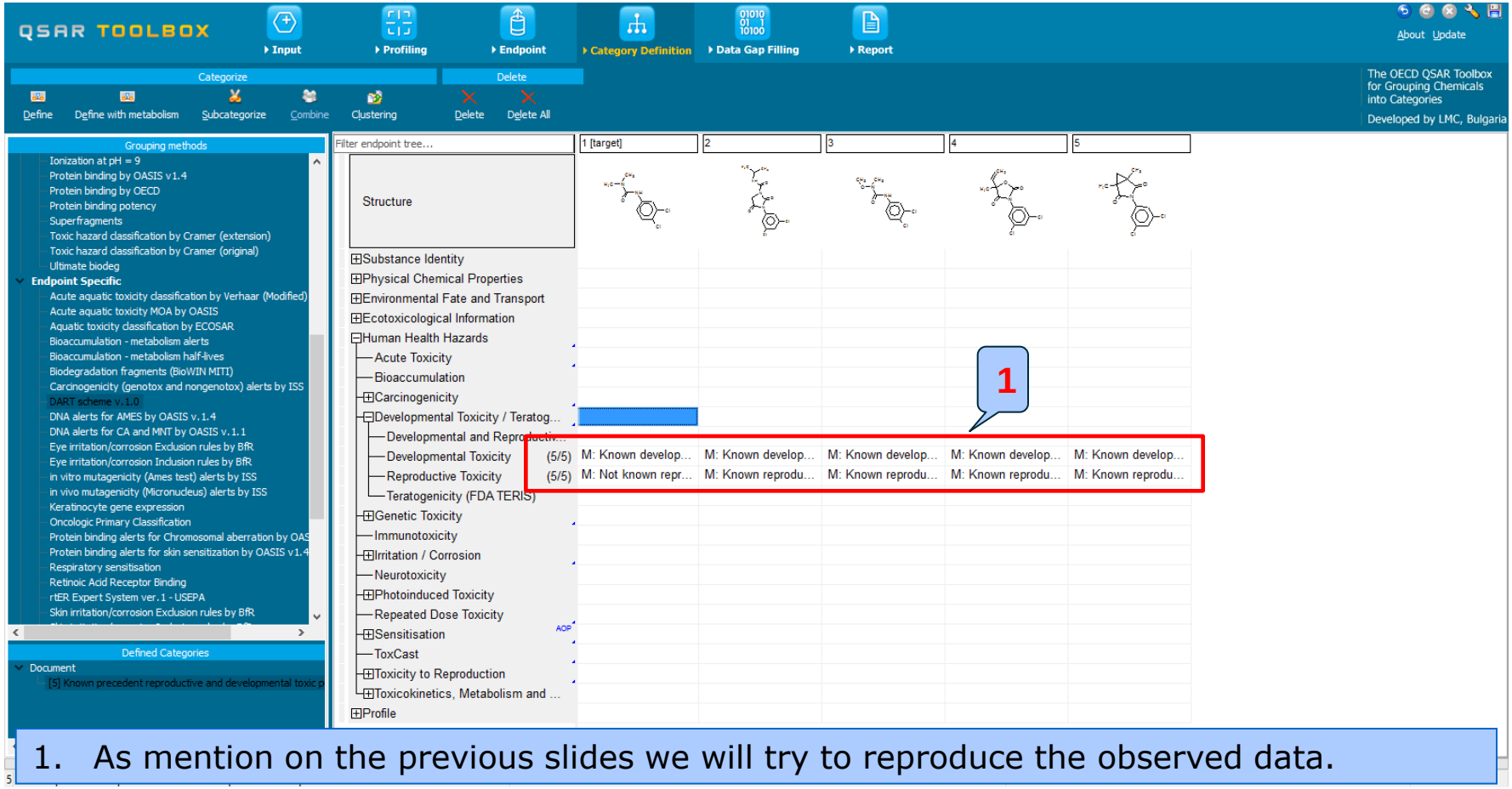


1. **Click** "OK" to extract data for the analogues from DART database
2. 10 data points are gathered for the identified 5 analogues. **Click** "OK"

Category Definition

Scenario 1: DART scheme used as a categorization tool

The experimental results for the analogues appeared on datamatrix



1. As mention on the previous slides we will try to reproduce the observed data.

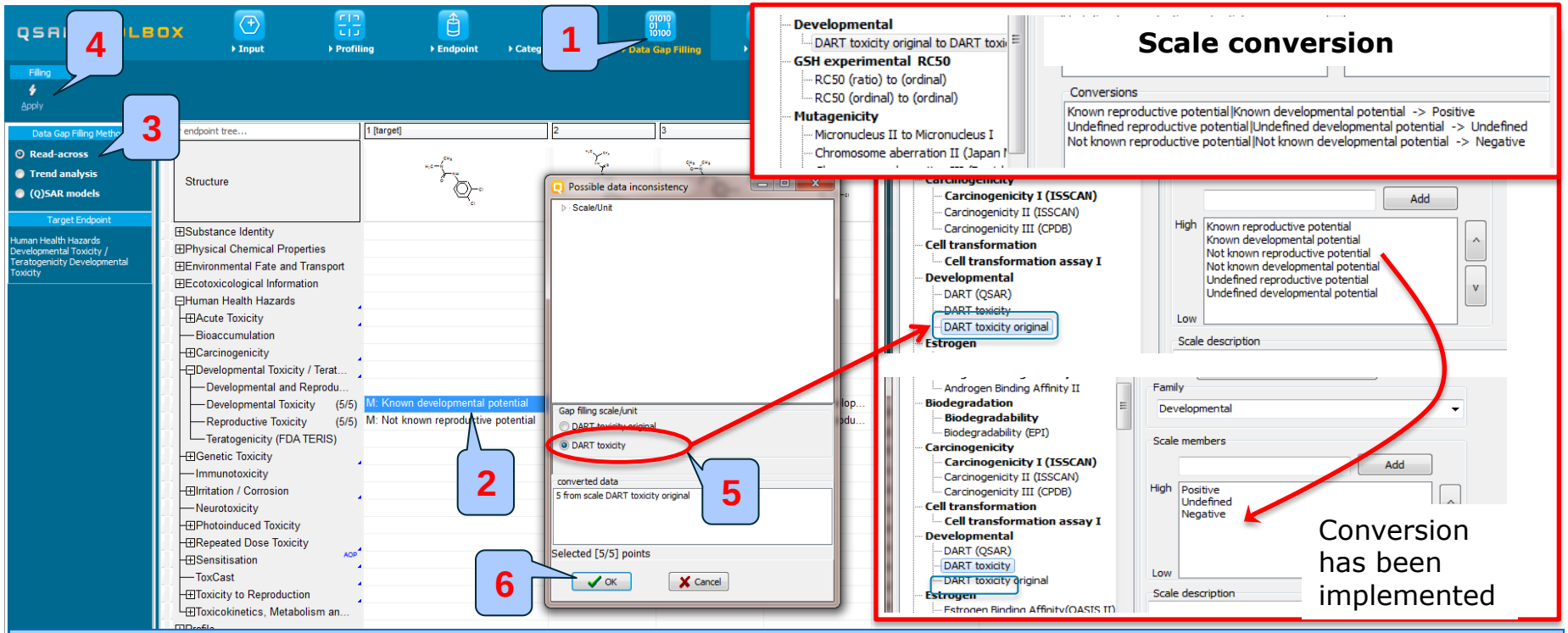
Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - Input
  - Profiling
  - Endpoint
  - **Category definition**
    - Overview
    - **Scenario 1**
    - **Apply read-across**
    - Scenario 2

Read-across applied for developmental tox

Scenario 1: DART scheme used as a categorization tool

Investigated endpoint: Developmental toxicity

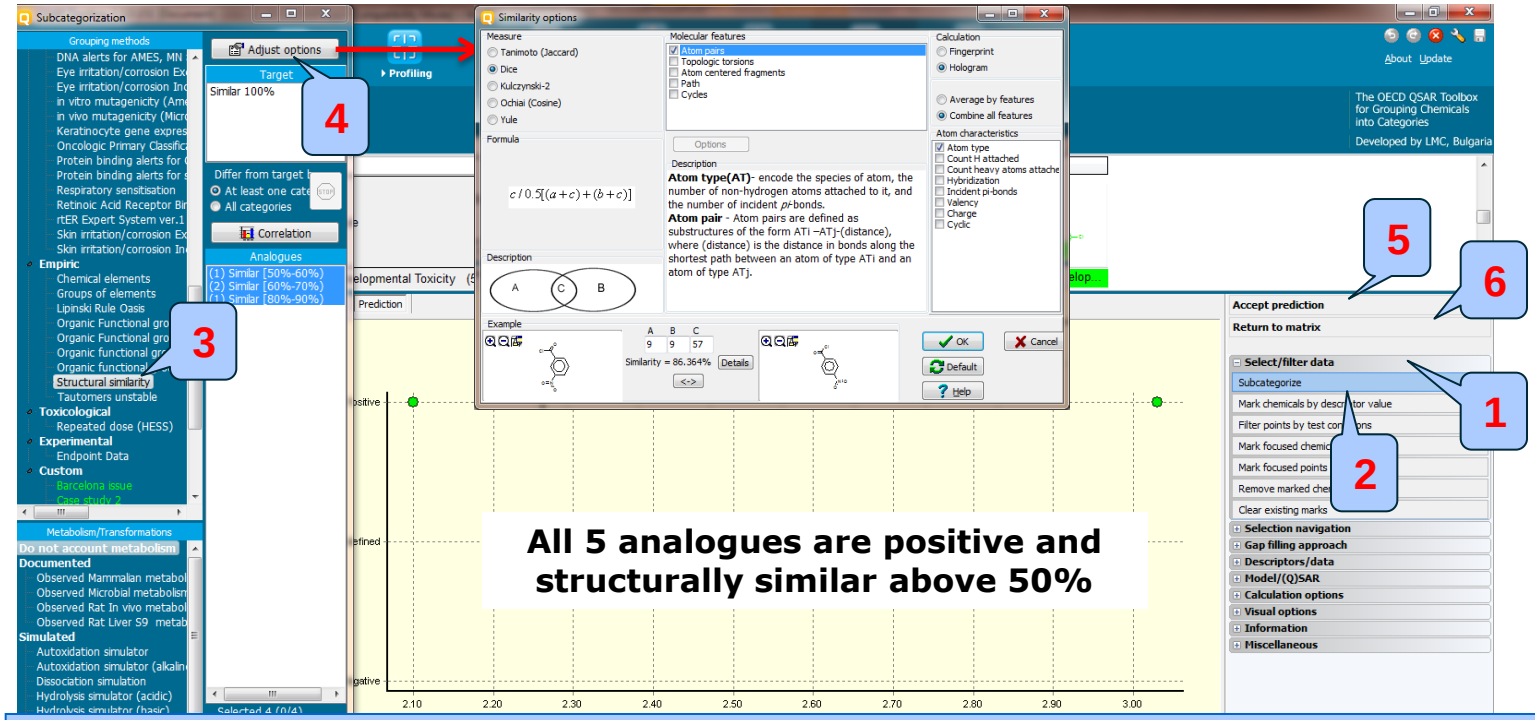


1. **Click** on "Data Gap filling" 2. **Click** on the cell corresponding to "Developmental Toxicity" endpoint
3. **Select** "Read-across" 4. **Click** "Apply" 5. A window indicating data inconsistency appears. DART data has been implemented into the system with original scale called "DART toxicity original". A less informative scale "DART toxicity" has been introduced. Also a scale conversion from original (more informative) to the less informative scale has been implemented. In our case we will use the less informative scale. Select "DART toxicity" scale. 6. **Click** "OK"

Read-across applied for developmental tox

Scenario 1: DART scheme used as a categorization tool

Investigated endpoint: Developmental toxicity



All analogues are positive and similar with respect to DART endpoint. Let's check how similar are they with respect to "Structural similarity". Follow the steps:

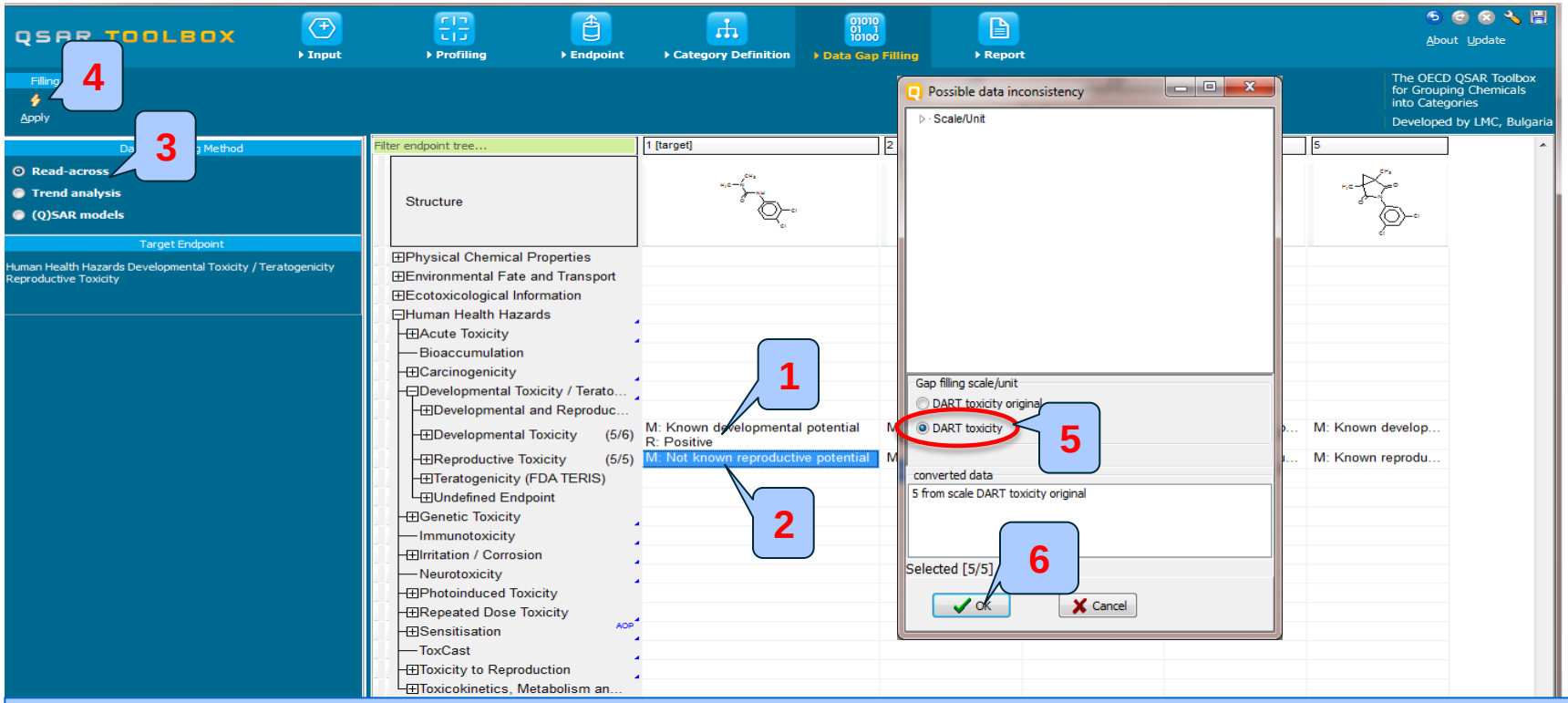
1. **Open** "Select/filter data"; 2. **Click on** "Subcategorize" ; 3. **Select** "Structural similarity"; The following similarity options are used in the subcategorization. **Click** on "Adjust options" button (4) to see the options: Dice, Atom pairs and atom type as atom characteristics are selected only.

The analysis shows that all analogues are similar above 50 % with respect to the target. Continue the workflow with accept the prediction. Follow the steps: 5. **Click** "Accept prediction"; 6. **Click** "Return to matrix"

Read-across applied for reproductive tox

Scenario 1: DART scheme used as a categorization tool

Investigated endpoint: Reproductive toxicity

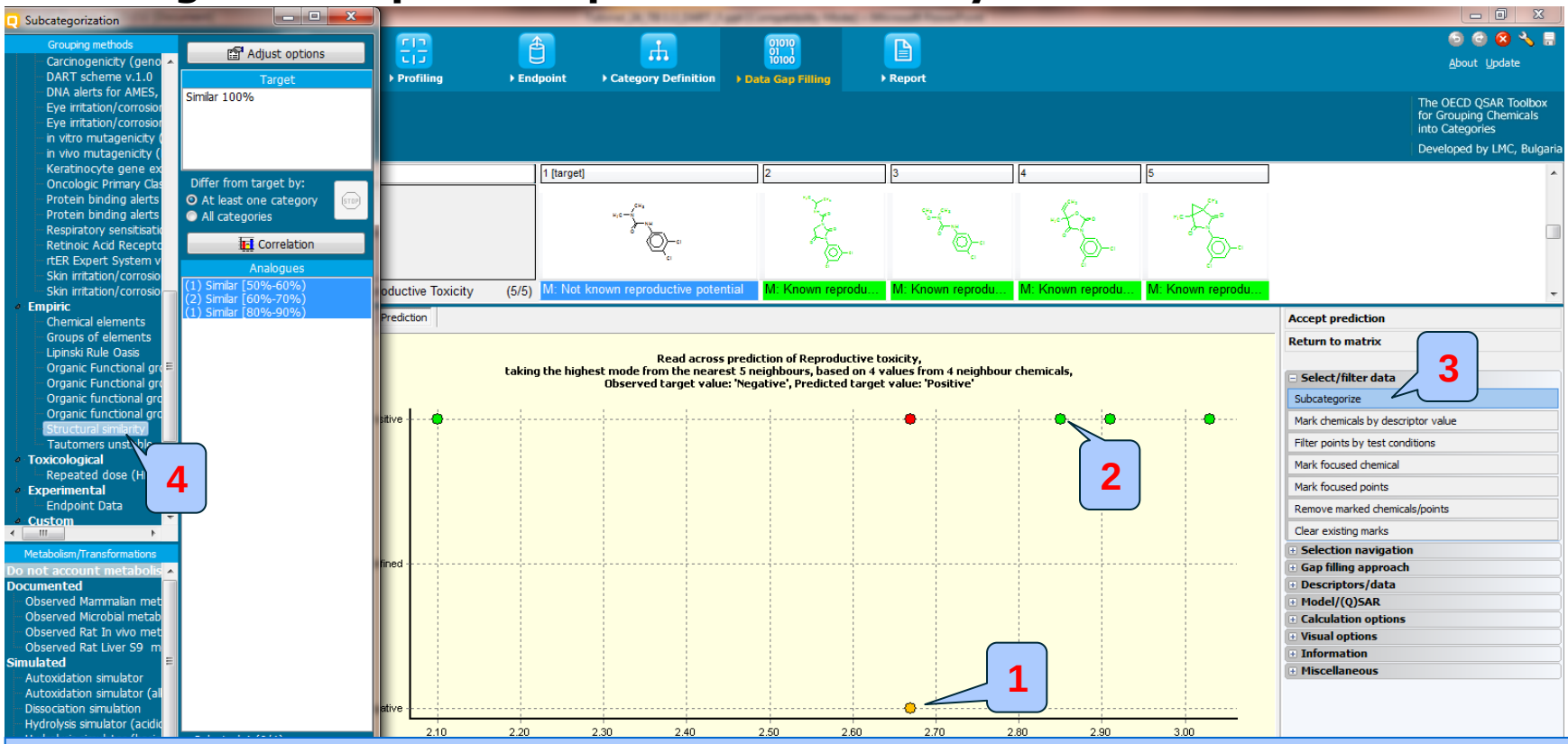


1. Positive prediction obtained for endpoint "Developmental toxicity" reproduces the positive observed data. Further read-across analysis continues with next endpoint "Reproductive toxicity". Follow the steps:
2. **Select** the cell corresponding to "Reproductive toxicity" endpoint; 3. **Select** "Read-across"; 4. **Click** "Apply";
5. In our case less informative scale is used: "DART toxicity"; 6. **Click** "OK"

Read-across applied for reproductive tox

Scenario 1: DART scheme used as a categorization tool

Investigated endpoint: Reproductive toxicity

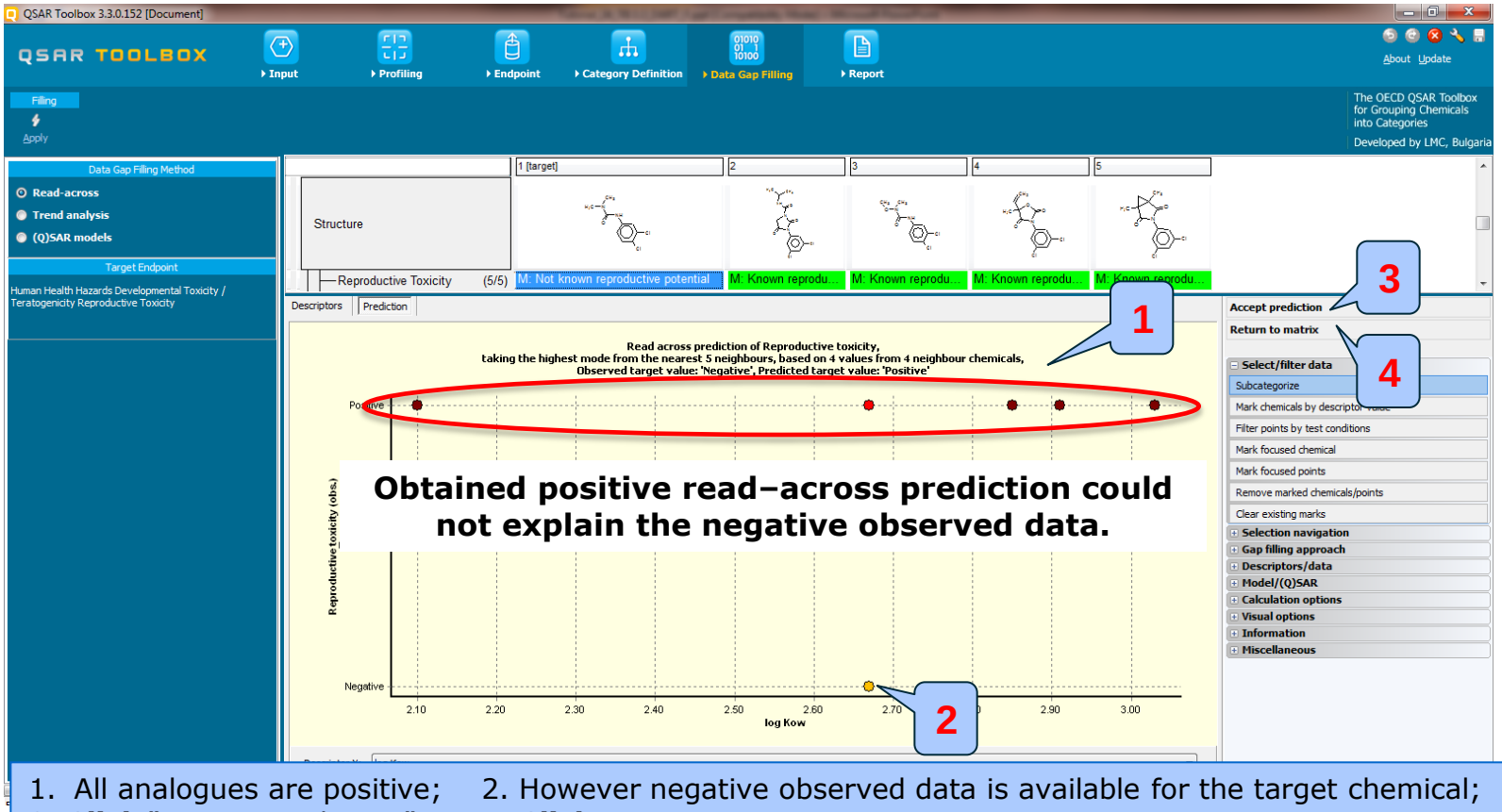


1. Negative observed data is available for the target chemical are the analogues with respect to structural similarity (options used in the read-across prediction are displayed on slide 54, same as in the previous example). All 4 analogues are similar above 50% with respect to the target chemical.
2. All analogues are positive. Lets check how similar
3. **Open** "Subcategorize"
4. **Select** "Structural similarity"

Read-across applied for reproductive tox

Scenario 1: DART scheme used as a categorization tool

Investigated endpoint: Reproductive toxicity



1. All analogues are positive; 2. However negative observed data is available for the target chemical;
3. **Click** "Accept prediction"; 4. **Click** Return to matrix

Recap

- In this step of the workflow scenario 1 has been applied:
  - Scenario 1: DART scheme applied as a primary categorization

- Read-across results shows:
  - Positive prediction for developmental toxicity reproduces positive observed data.
  - Positive prediction for reproductive toxicity could not explain the negative observed data.

- The workflow continues with second categorization scenario.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - Input
  - Profiling
  - Endpoint
  - **Category definition**
    - Overview
    - Scenario 1
    - Apply read-across
    - **Scenario 2**

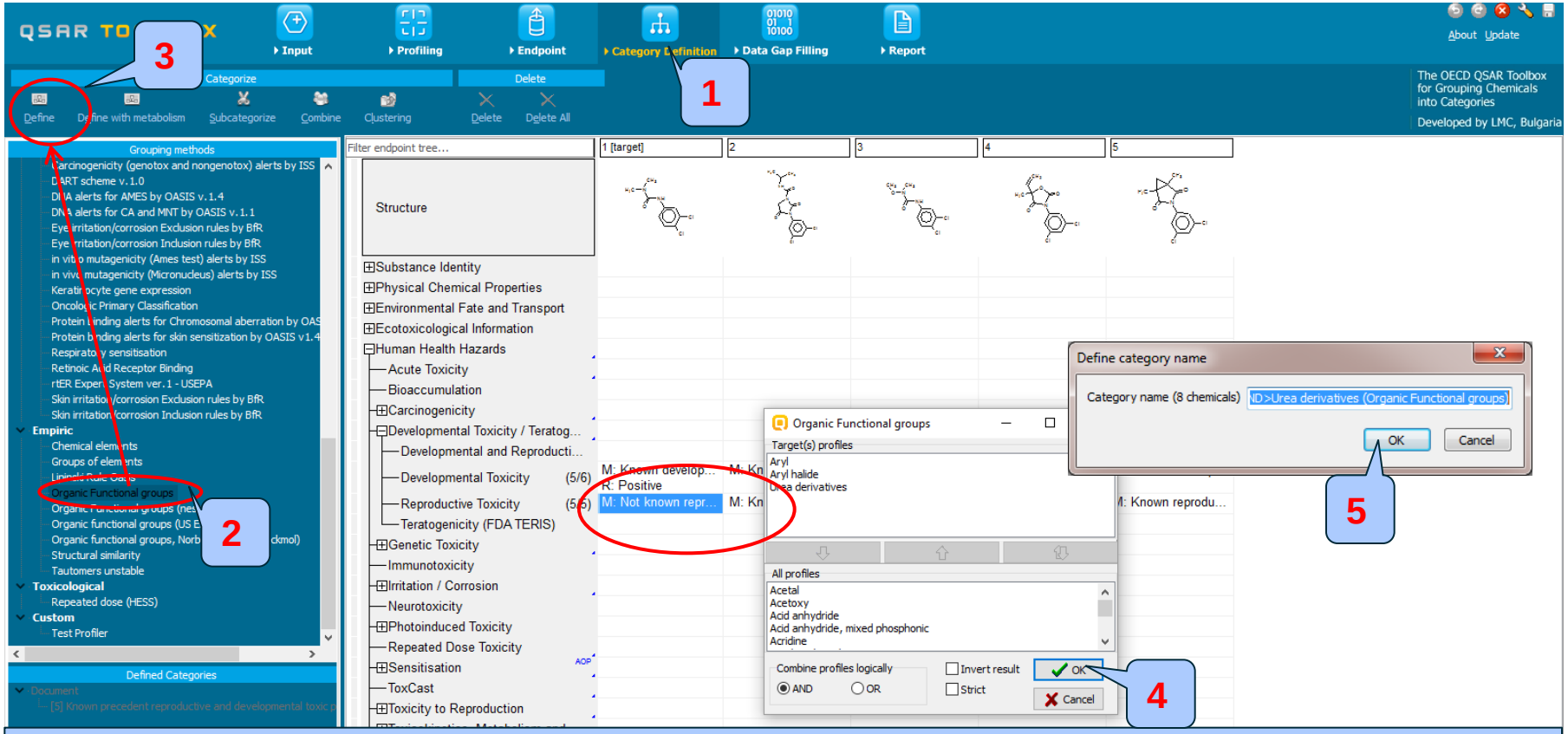
Category Definition
Grouping methods

- Based on these classifications and basic guidance for grouping chemicals explained on the previous slides, two scenarios for identifying similar analogues have been applied in further read-across analysis:
  - Scenario 1: DART v1.0 scheme used as a categorization tool (applied for primary categorization - phase I).
  - Scenario 2: OFG* is used for primary categorization with forthcoming subcategorization by DART scheme (applied for subcategorization - phase II).

*OFG is used for primary categorization in this case, because the two basic functionalities available within the molecule: "Ureas" and "Aryl halides" will be preserved in the group of identified analogues, while the ECOSAR categorization omits "Aryl halide" functionality and identifies "Substituted ureas" only (see slide 44).

Category Definition

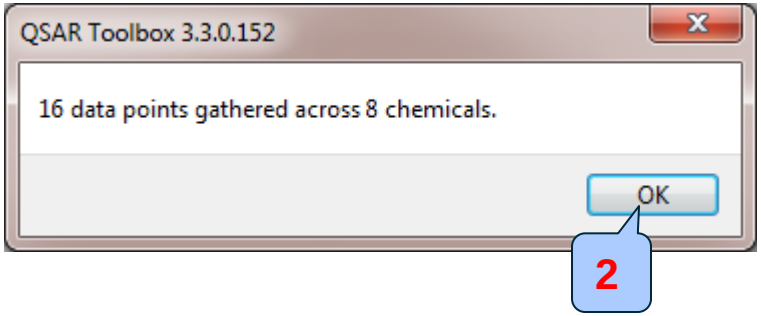
Scenario 2: OFG is used for primary categorization with forthcoming subcategorization by DART scheme



1. **Back** to "Category definition"; 2. **Select** "OFG"; 3. **Click** "Define"
4. **Click** "OK" to confirm the identified categories for the target chemical
5. Eight analogues are identified. **Click** "OK"

Category Definition

Scenario 2: OFG is used for primary categorization

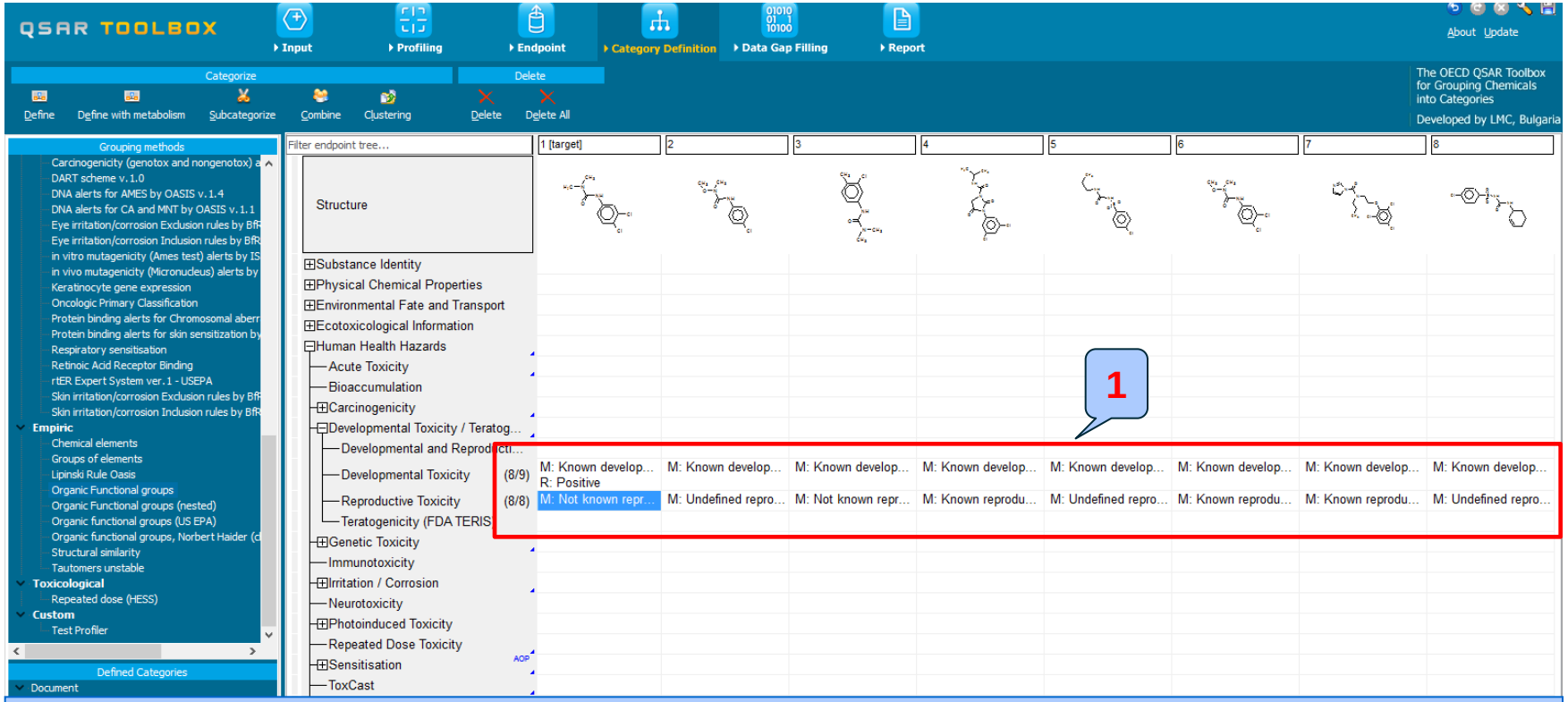


1. **Click** "OK" to extract data for the analogues from DART database
2. 16 data points are gathered for the identified 8 analogues. **Click** "OK"

Category Definition

Scenario 2: OFG is used for primary categorization

The experimental results for the analogues appeared on datamatrix

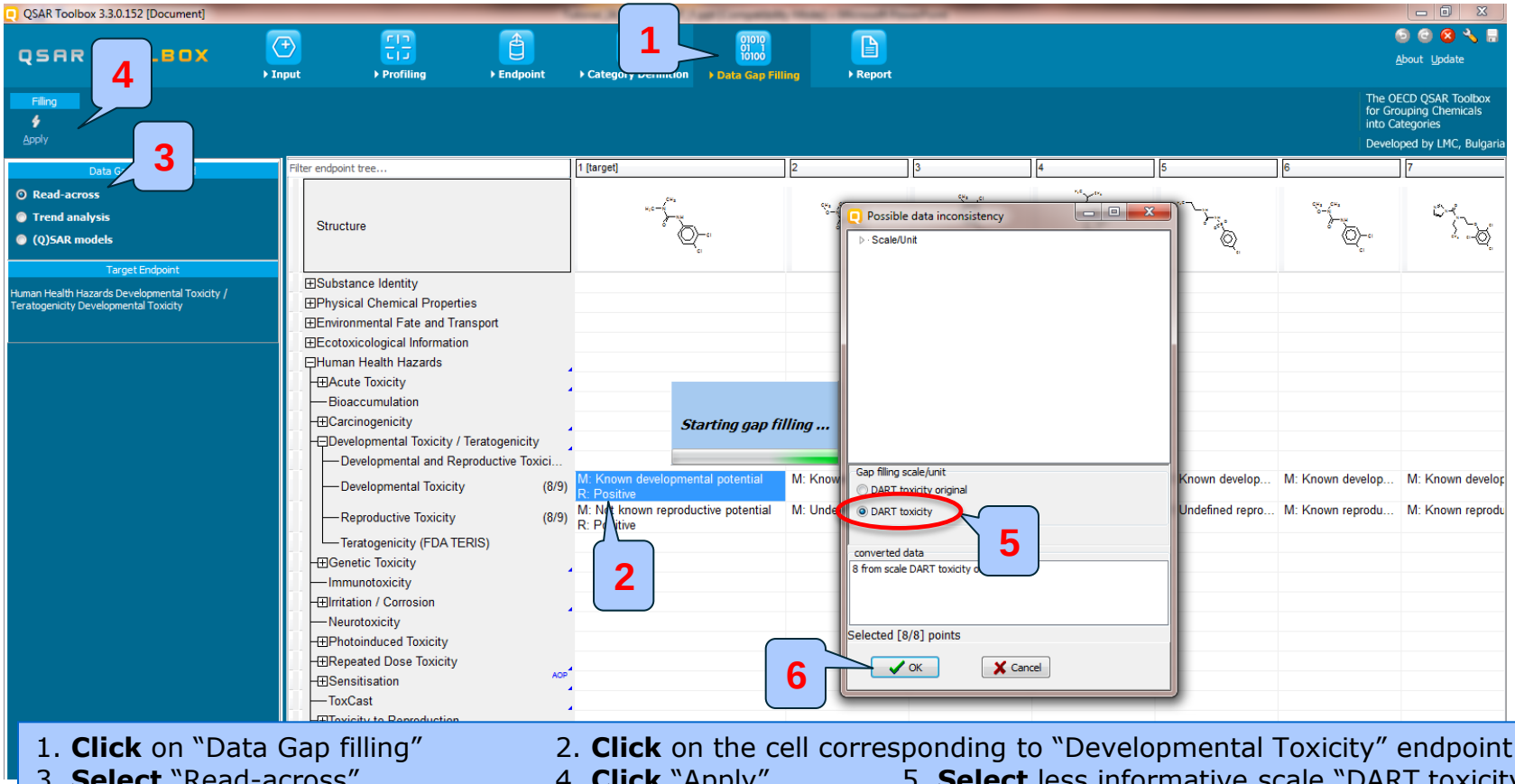


1. As mention on the previous slides there are observed data for the target, according to the two endpoints. We will try to reproduce it.

Read-across applied for developmental tox

Scenario 2: OFG is used for primary categorization

Investigated endpoint: Developmental toxicity



1. **Click** on "Data Gap filling"
2. **Click** on the cell corresponding to "Developmental Toxicity" endpoint
3. **Select** "Read-across"
4. **Click** "Apply"
5. **Select** less informative scale "DART toxicity"
6. **Click** "OK"

Read-across applied for developmental tox

Scenario 2: OFG is used for primary categorization

Investigated endpoint: Developmental toxicity

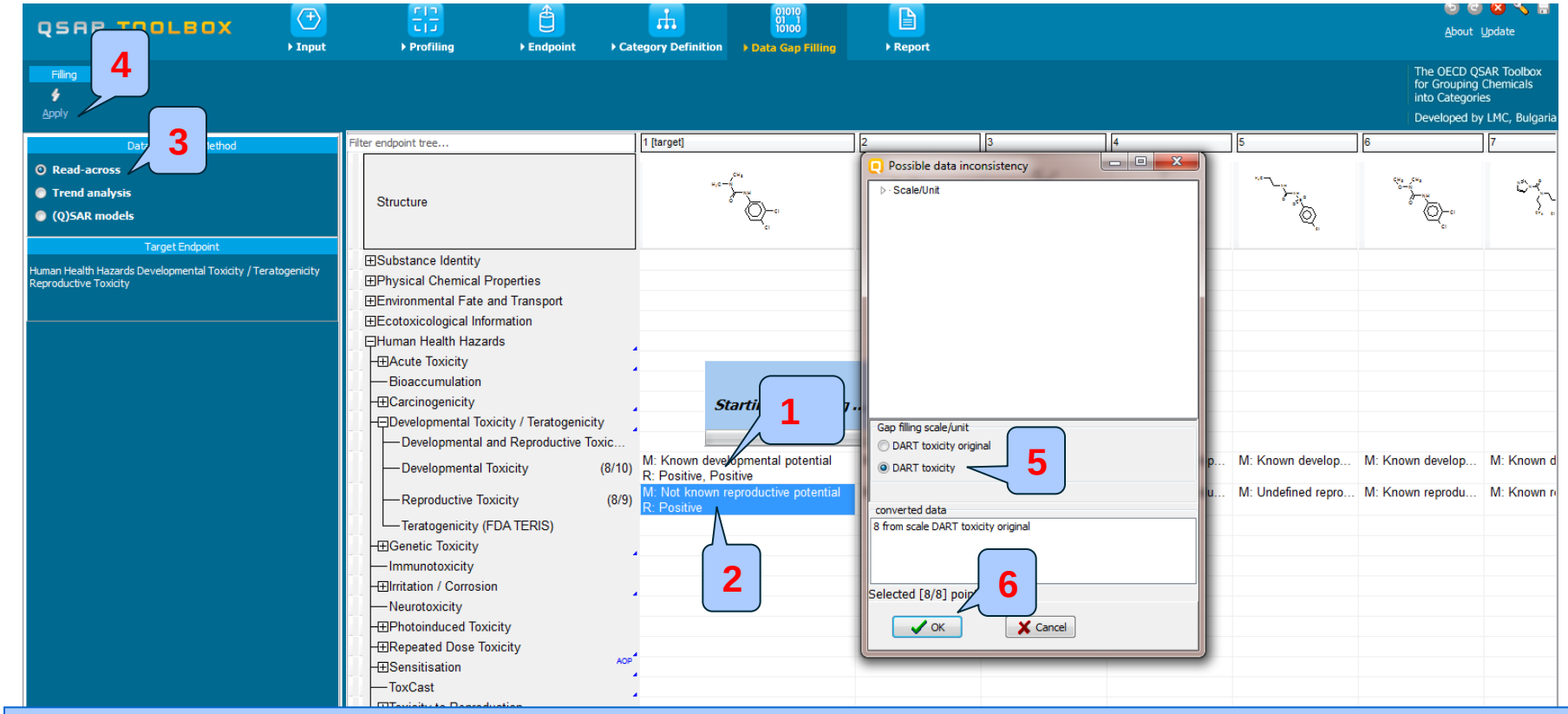
All analogues are positive. Let's check how similar are they with respect to endpoint specificity through applying DART profiling scheme. Follow the steps:

1. **Open** "Select/filter data"; 2. **Select** "Subcategorize" 3. **Select** "DART scheme v1.0". There are 2 analogues having different DART toxicity that the target. They will be eliminated; 4. **Click** "Remove". Now all analogues are positive and consistent with respect to structure (OFG) and endpoint specificity (DART); 5. **Click** "Accept prediction"; 6. **Click** "Return to matrix"

Read-across applied for developmental tox

Scenario 2: OFG is used for primary categorization

Investigated endpoint: Reproductive toxicity

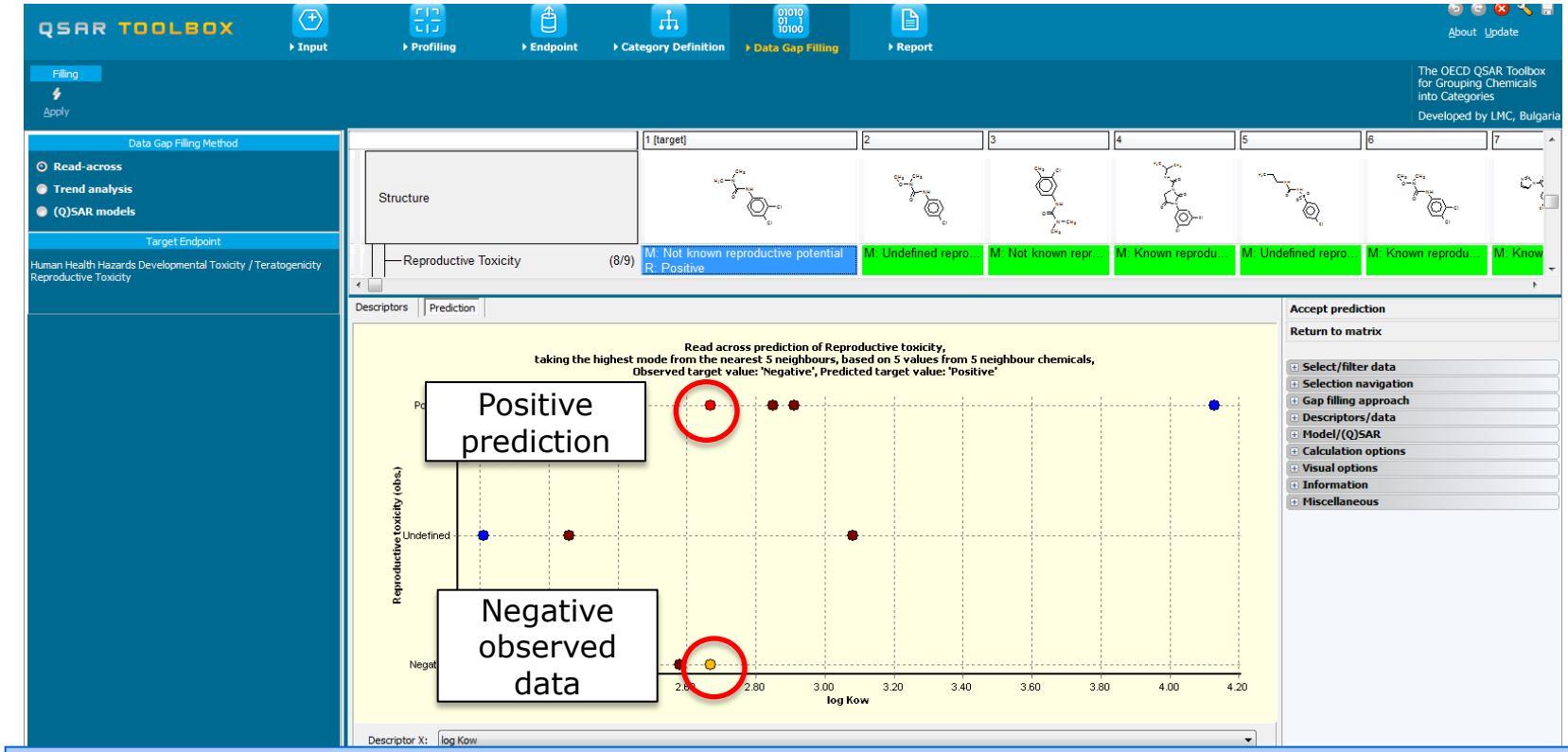


1. Positive prediction obtained for endpoint "Developmental toxicity" reproduces the positive observed data. Further read-across analysis continues with next endpoint "Reproductive toxicity". Follow the steps:
2. **Select** the cell corresponding to "Reproductive toxicity" endpoint; 3. **Select** "Read-across"; 4. **Click** "Apply"
5. In our case less informative scale is used: "DART toxicity" 6. **Click** "OK"

Read-across applied for reproductive tox

Scenario 2: OFG is used for primary categorization

Investigated endpoint: Reproductive toxicity

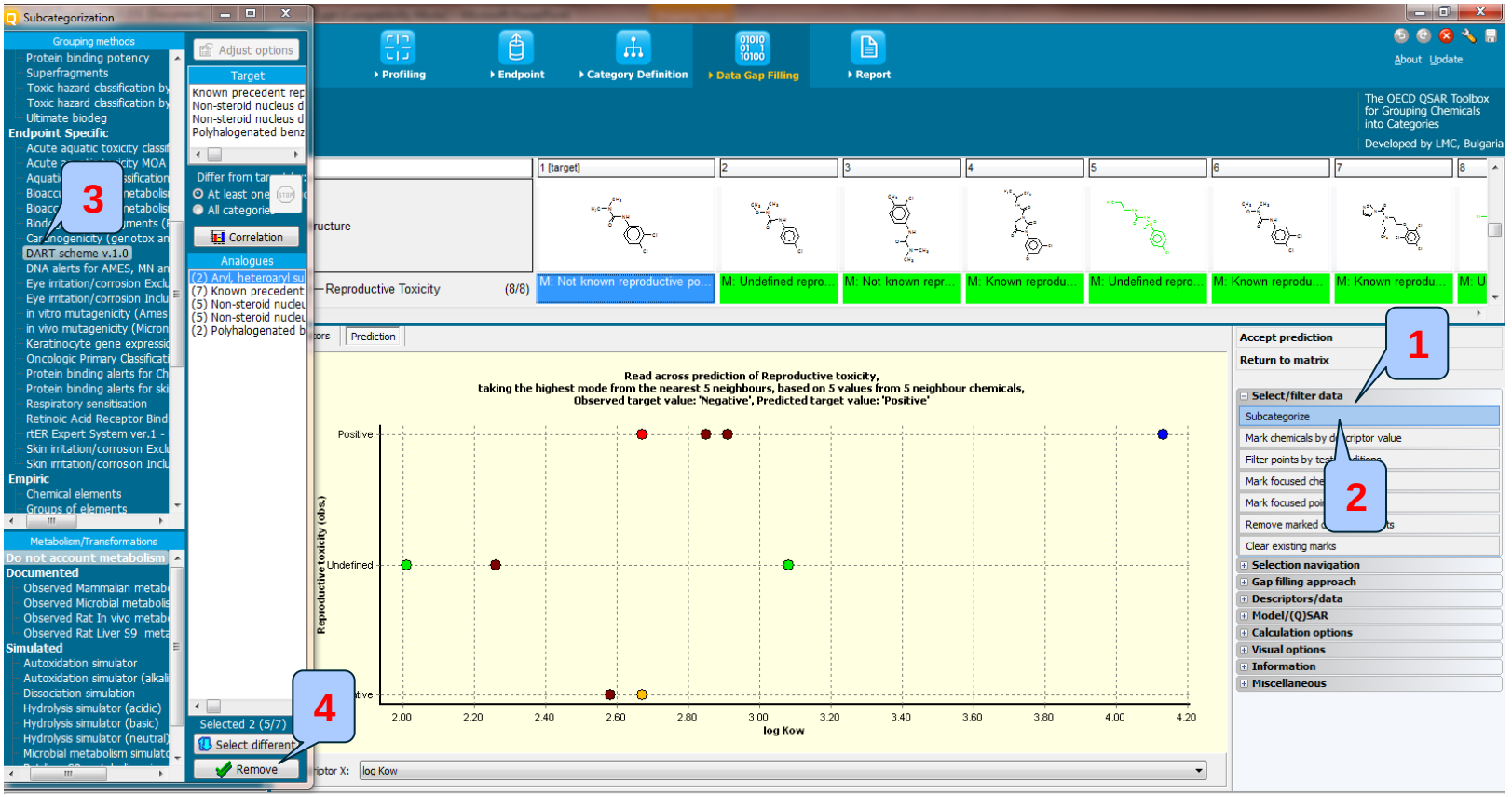


1. The obtained read-across prediction could not reproduce the negative observed data of the target. It is not straightforward and not reliable due to variable endpoint data of the analogues. The purpose of the further workflow is to subcategorize and refine the initial category of analogues. Follow the steps illustrated on next slide.

Read-across applied for reproductive tox

Scenario 2: OFG is used for primary categorization

Investigated endpoint: Reproductive toxicity

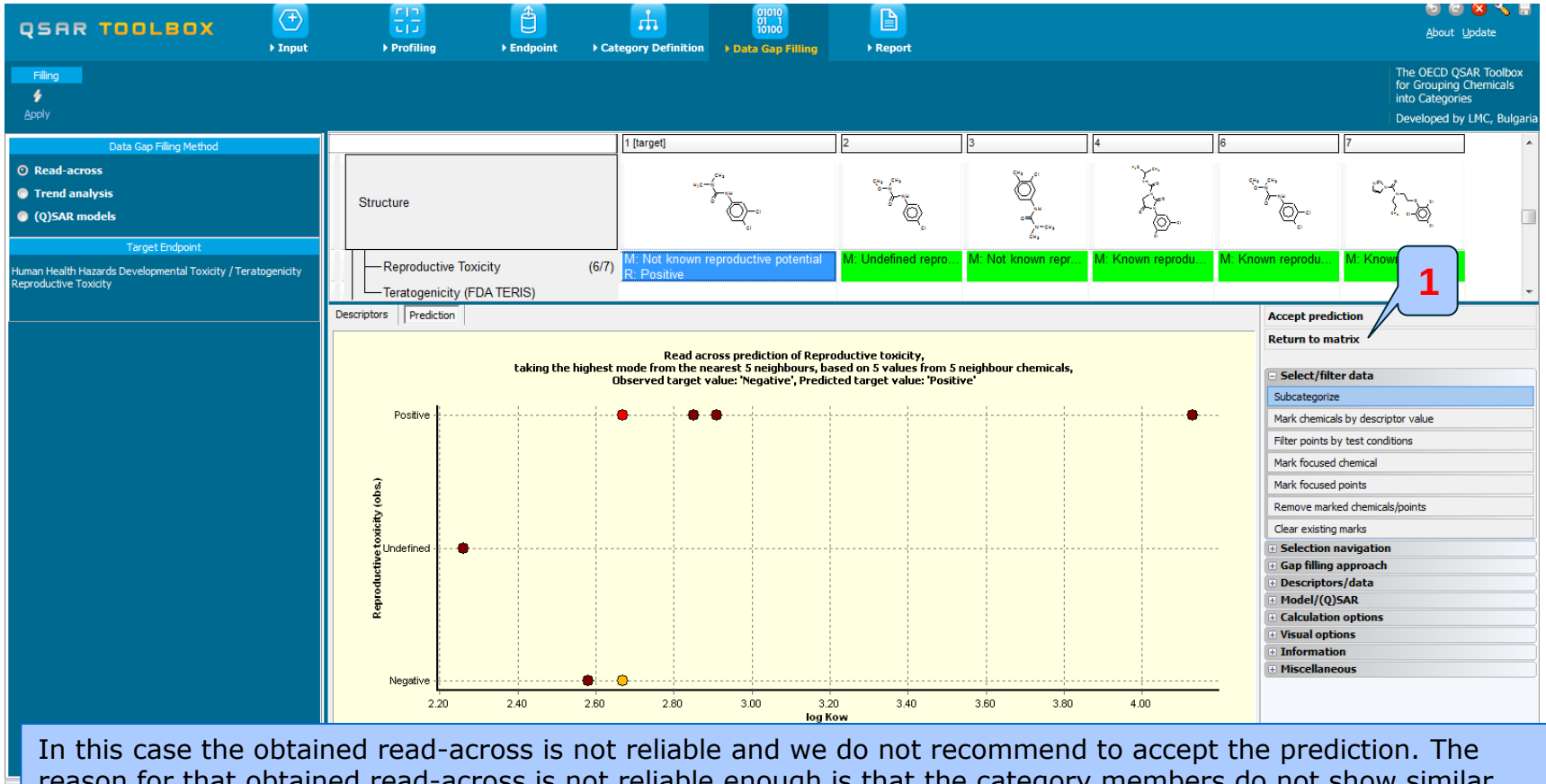


1. **Open** "Select filter data" 2. **Select** "Subcategorize" 3. **Select** "DART scheme"; 4. **Click** "Remove" to eliminate the analogues with different DART toxicity

Read-across applied for reproductive tox

Scenario 2: OFG is used for primary categorization

Investigated endpoint: Reproductive toxicity



In this case the obtained read-across is not reliable and we do not recommend to accept the prediction. The reason for that obtained read-across is not reliable enough is that the category members do not show similar test results, but note that it is not replicating the experimental data. 1. **Click** "Return to matrix"

Recap

- In this step of the workflow two scenarios for identifying analogues are played:
  - Scenario 1: DART scheme applied as a primary categorization.
  - Scenario 2: DART scheme applied in subcategorization procedure.

- Read-across results of scenario 2 shows:
  - The obtained positive prediction for developmental toxicity reproduces the positive observed data.
  - The obtained read-across prediction for reproductive toxicity is not reliable and could not be accepted. The reason for that obtained read-across is not reliable enough is that the category members do not show similar test results, but not that it is not replicating the experimental data.

- The further workflow continues with applying of external DART SAR model.

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - Input
  - Profiling
  - Endpoint
  - **Category definition**
    - Overview
    - Scenario 1
    - Scenario 2
  - **Support the prediction by DART model**

DART SAR model

- DART SAR model is based on profiling results obtained by DART profiling scheme v 1.0. The profiling scheme is implemented following Wu S. paper [1]. The SAR model follows the same organization as DART profiling scheme with an exception of prediction result. It is used for identifying chemicals with structural features associated with the potential to act as reproductive or developmental toxicants.

- The prediction outcome from the SAR model provides more general information for potential of chemical to cause DART toxicity, while the DART profiling scheme provides specific information about the DART category associated with the specific chemical class.

- The outcome from the model is "Known precedent reproductive and developmental toxic potential", when the chemical meet the structural criteria of the model; "Not known precedent reproductive and developmental toxic potential", if the toxic potential of the input chemical is not known and "Not covered by current version of the decision tree" if the identified structural features are not object of the DART tree.

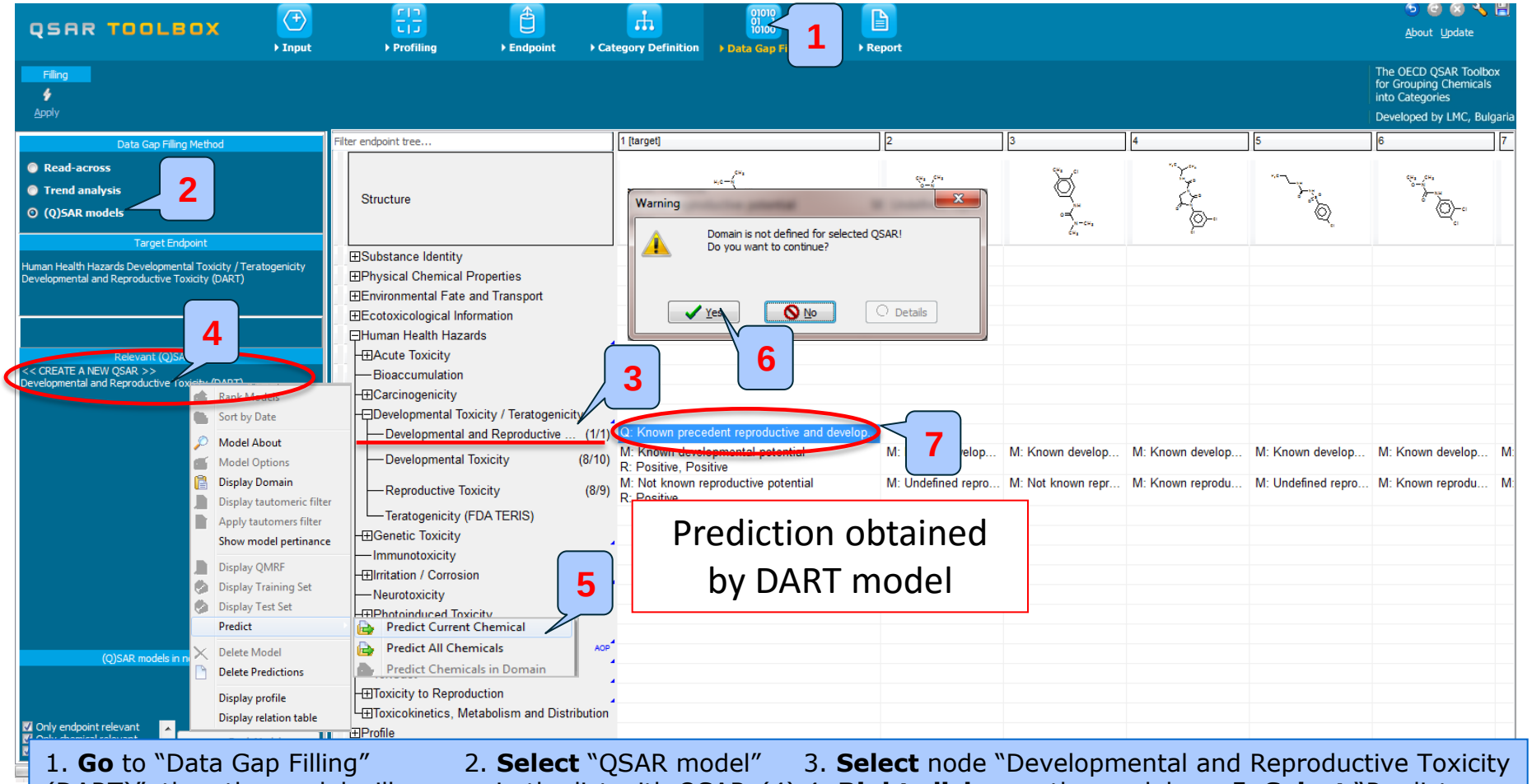
- Prediction outcome from DART model is provided for repro/developmental toxicity as joint effect, while read-across analysis is performed for each of the DART toxicity independently (previous exersice).

- The corresponding category is displayed in the report generated for the obtained prediction.

---

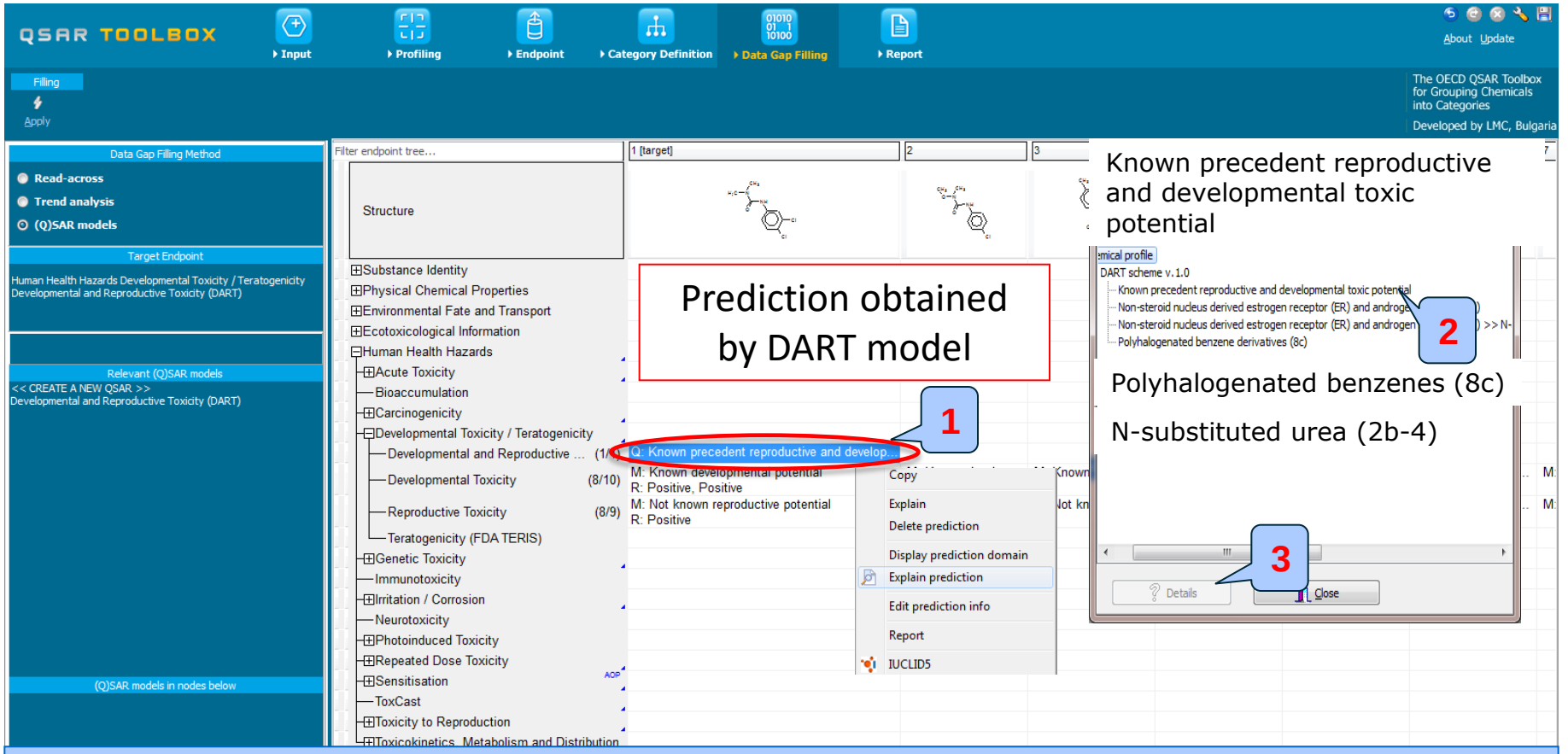
1. Wu S, Fisher J, Naciff J, Laufersweiler M, Lester C, Daston G, Blackburn K. Framework for identifying chemicals with structural features associated with the potential to act as developmental or reproductive toxicants. Chem Res Toxicol. 2013 Dec 16;26(12):1840-61.

DART SAR model - overview



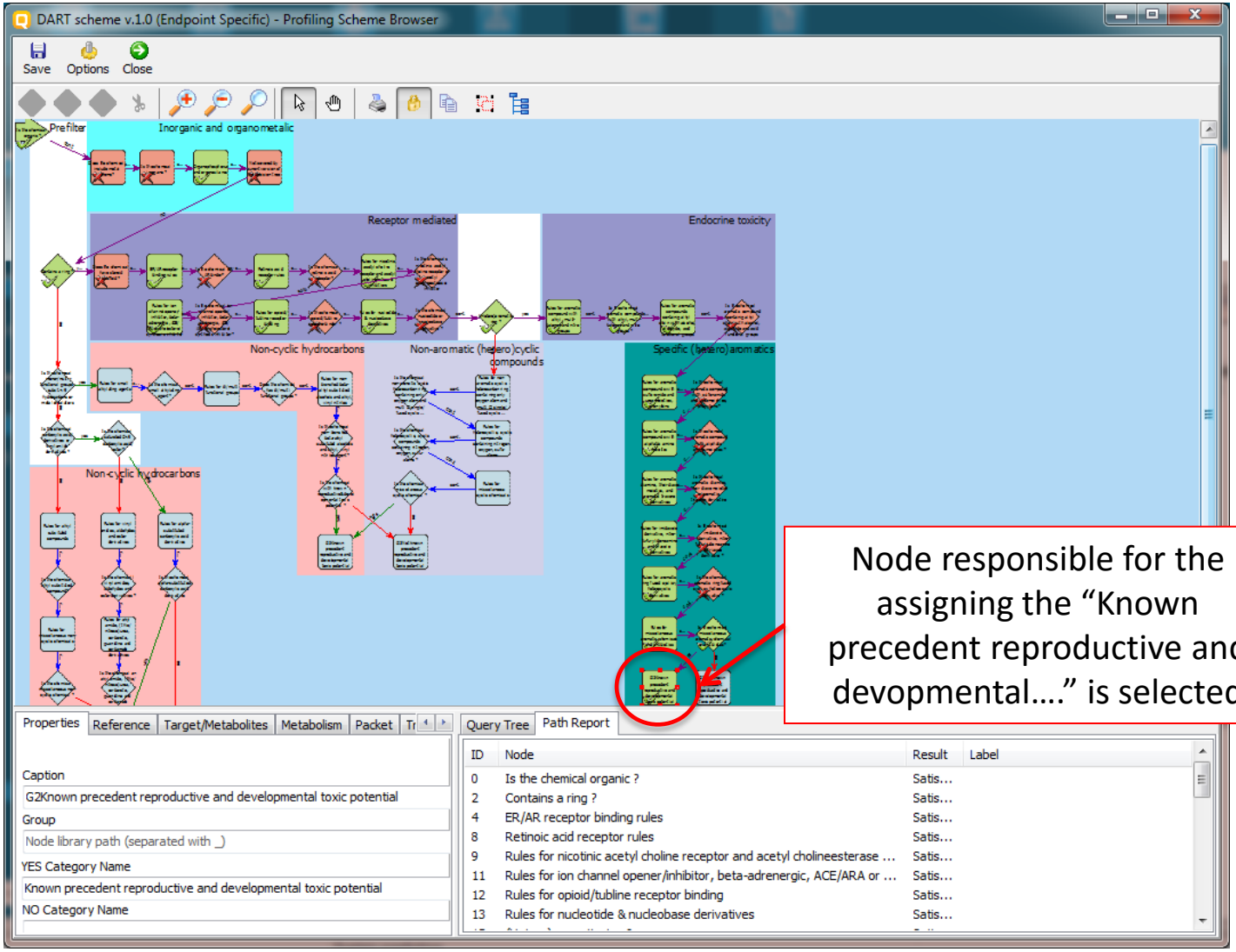
1. **Go** to "Data Gap Filling" 2. **Select** "QSAR model" 3. **Select** node "Developmental and Reproductive Toxicity (DART)", then the model will appear in the list with QSARs(4) 4. **Right click** over the model 5. **Select** "Predict Current Chemical" 6. The software informs the user that the domain is not defined. **Click** "Yes"
7. Prediction obtained by DART model appears on datamatrix

DART SAR model - overview

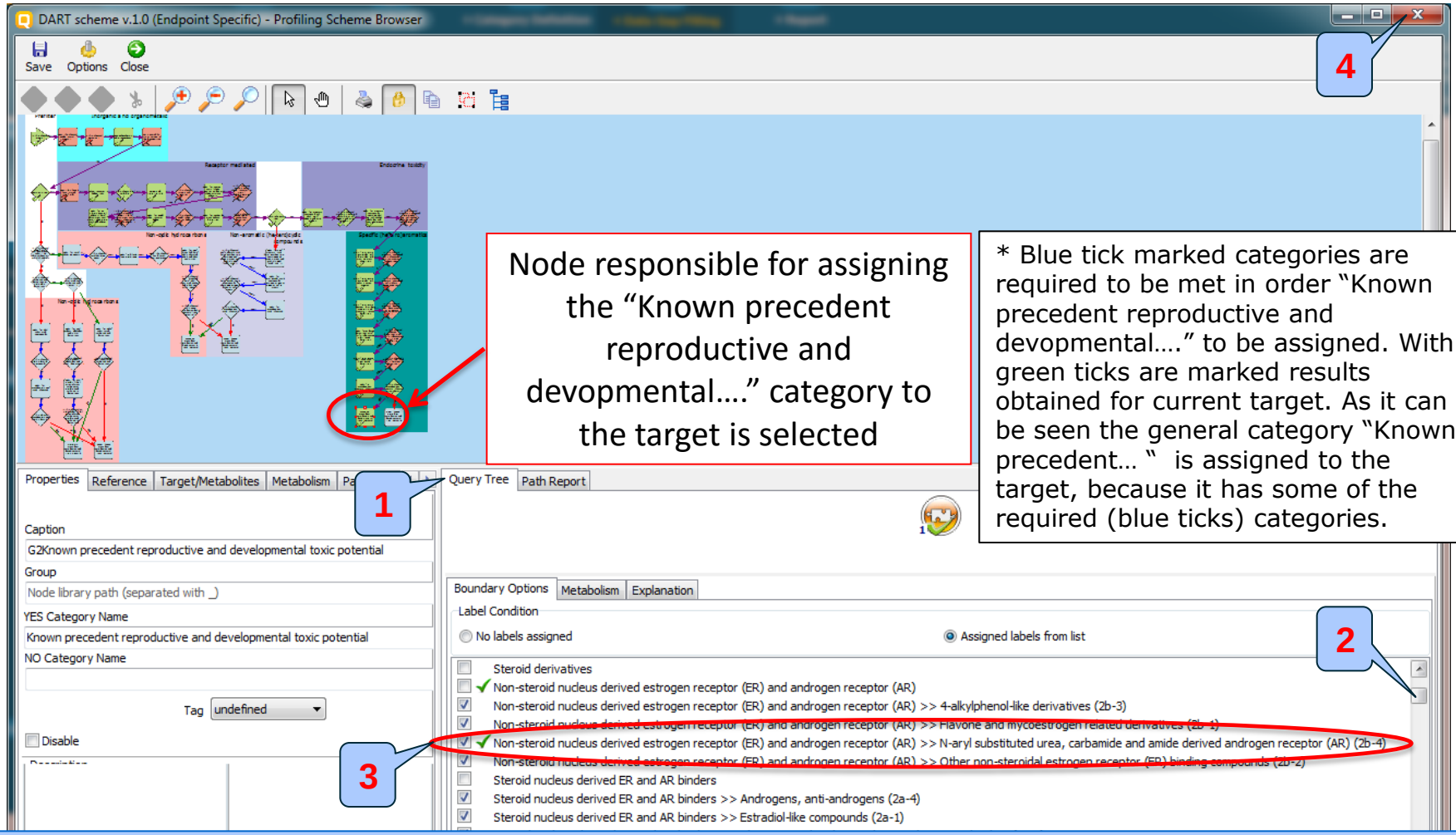


1. **Right click** over the prediction result and select "Explain prediction".
2. **Select** first category "Known…"
3. **Click** "Details"

DART SAR model – Explain results



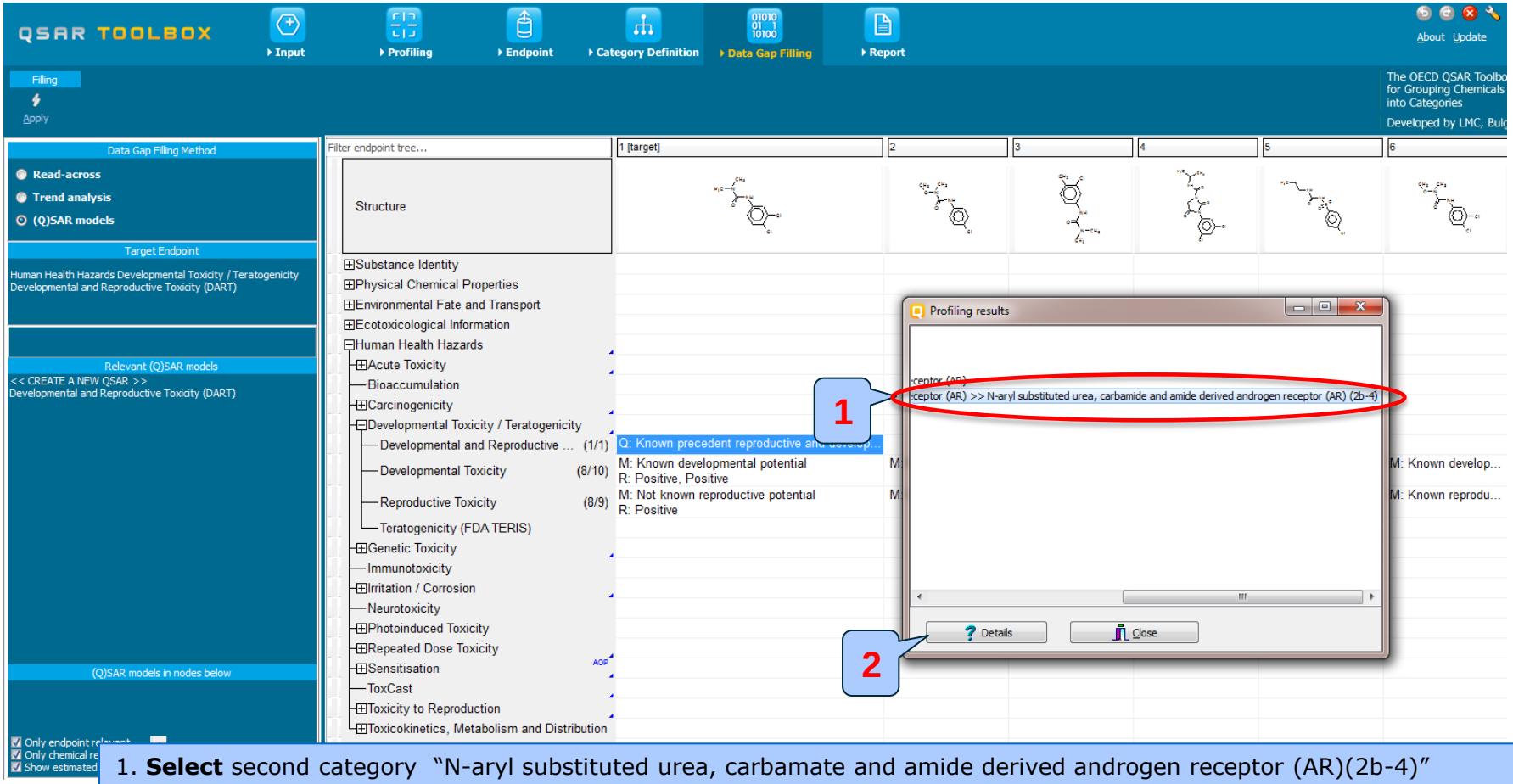
DART SAR model – Explain results



1. **Click** on "Query tree" panel 2. **Scroll down** and **find** category marked with green tick
3. Marked category "Non-stereoid nucleus derived estrogen receptor (ER) ……(2b-4)" is responsible for DART effect*. The category "Known precedent repro….. potential" obtained as a DART SAR results is due to this category and the other "Polyhalogenated benzene derivatives (8c)" category. More details for both schemes is provided on next slides 4. **Close** the window

DART SAR model – Explain results

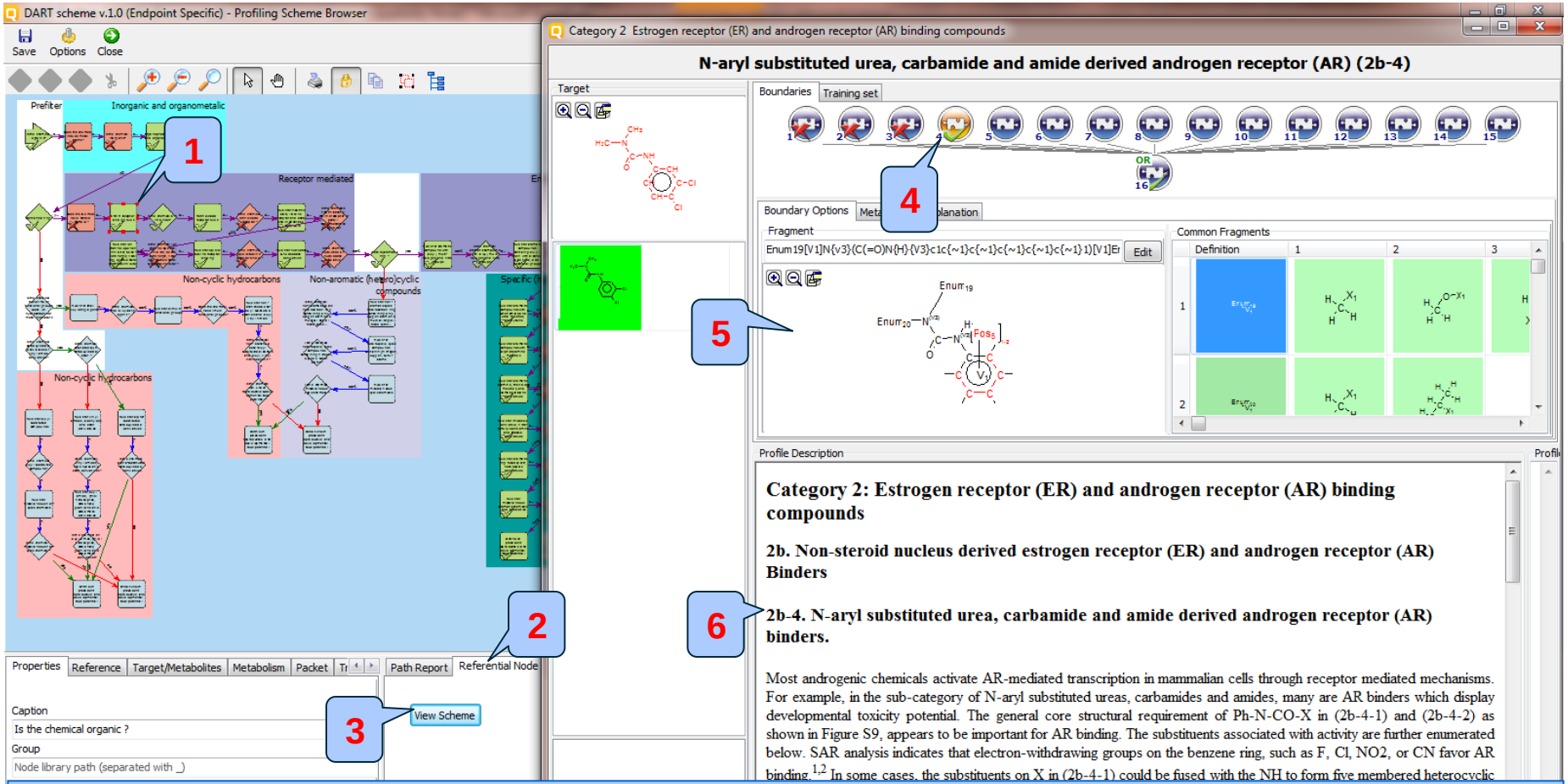
N-aryl subsituted urea(2b-4)



1. **Select** second category "N-aryl substituted urea, carbamate and amide derived androgen receptor (AR)(2b-4)"
2. **Click** "Details"

DART SAR model – Explain results

N-aryl subsituted urea(2b-4)



1. Node responsible for DART effect, which assigns the respective category is selected 2. **Click** "Referential node" tab
3. **Click** "View Scheme" 4. **Select** green marked boundary #4. It is marked green because the target met the criteria of the boundary. 5. Definition of structural boundary 6. Textual description associated with DART category
Details about the category is presented on next slide

DART SAR model – Explain results

N-aryl subsituted urea(2b-4)

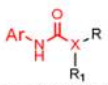
Profile Description

Category 2: Estrogen receptor (ER) and androgen receptor (AR) binding compounds

2b. Non-steroid nucleus derived estrogen receptor (ER) and androgen receptor (AR) Binders

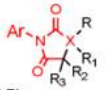
2b-4. N-aryl substituted urea, carbamide and amide derived androgen receptor (AR) binders.

Most androgenic chemicals activate AR-mediated transcription in mammalian cells through receptor mediated mechanisms. For example, in the sub-category of N-aryl substituted ureas, carbamides and amides, many are AR binders which display developmental toxicity potential. The general core structural requirement of Ph-N-CO-X in (2b-4-1) and (2b-4-2) as shown in Figure S9, appears to be important for AR binding. The substituents associated with activity are further enumerated below. SAR analysis indicates that electron-withdrawing groups on the benzene ring, such as F, Cl, NO2, or CN favor AR binding.[1,2] In some cases, the substituents on X in (2b-4-1) could be fused with the NH to form five membered heterocyclic ring moieties (2b-4-2), exemplified by imidazolidine-2,4-dione, oxazolidine-2,4-dione, and pyrrolidine-2,5-dione derivatives. For these cyclic compounds, X can be nitrogen, oxygen or carbon, and substituents (R, R_1, R_2 and R_3) are listed in Figure S9 in structure (2b-4-2). These N-aryl substituted heterocyclic ring derivatives, such as vinclozolin (CAS# 50471-44-8), iprodione (CAS# 36734-19-7) and proyamidone (CAS# 32809-16-8), have a range of developmental and reproductive effects linked to activity as an anti-androgen.[3, 4] Another androgen antagonist is prochloraz (CAS# 67747-09-5), which appears to have a distinct pattern of toxicity. Prochloraz had been reported to have multiple effects on the development of male rodents, and *in vitro* data show not only anti-androgen but also anti-estrogen effects as well as interaction with Ah receptors and inhibition of aromatase[5, 6].



Ar=Ph, alkyl (C1-C3)-Ph, alkyl-,Cl-Ph, mono-, di-Cl-Ph, CF_3,NO_2-Ph
X=C; R=alkyl; R_1=Me
X=N; R=Me, OMe, ClCH2-; R_1=Me, iPr
X=O; R=Alkyl (C1-C3); R_1=None

2b-4-1



Ar=di-Cl-Ph
X=O; R=R_1=none; R_2, R_3=alkyl, vinyl
X=N; R=alkylNHCO; R_1=none; R_2, R_3=alkyl, H
X=C; R,R_1,R_2,R_3 can be alkyl or fused cyclo-alkyl

2b-4-2

Fig. S9. The general structural features of N-aryl substituted ureas, carbamides and amides like chemicals.

Original reference:

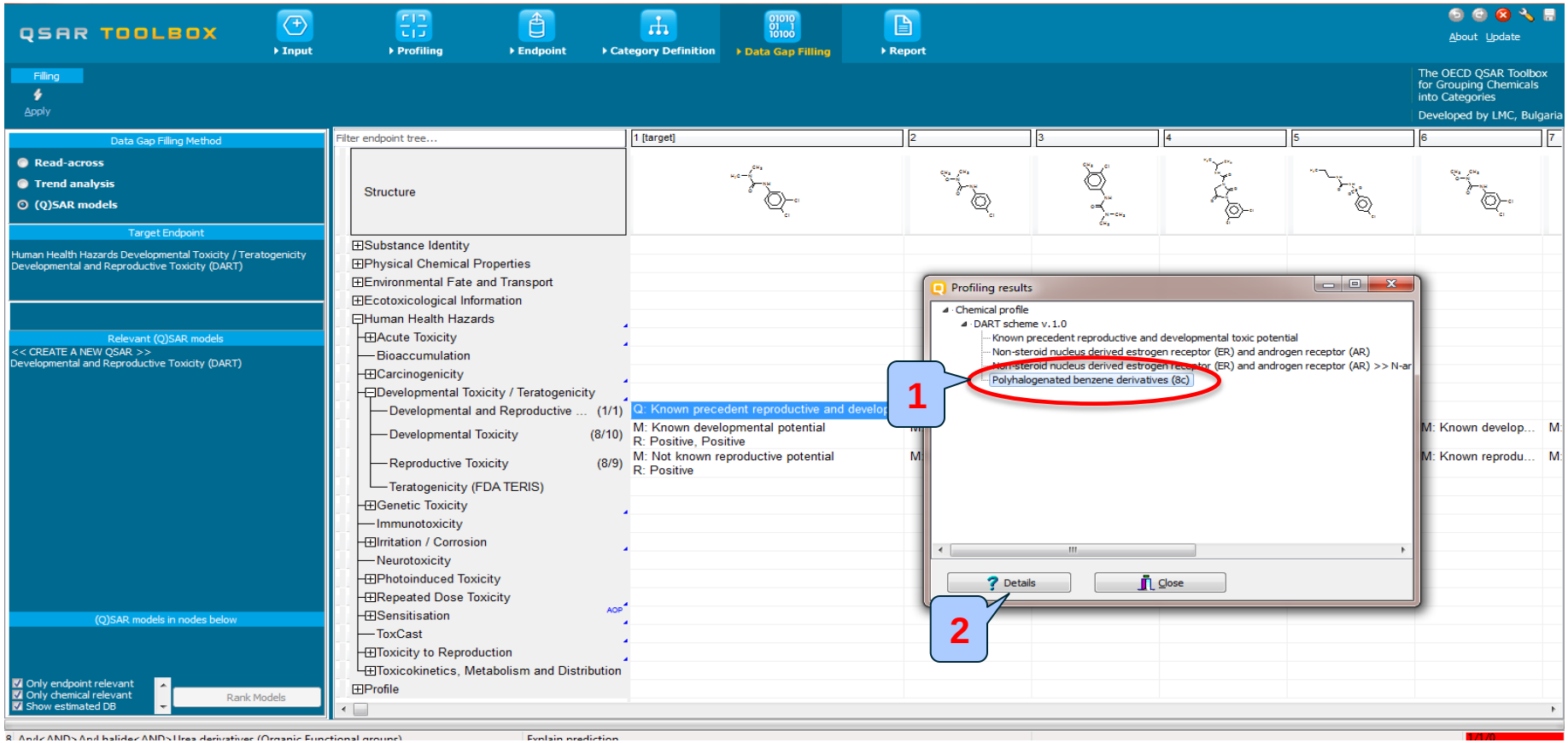
Shengde, W., Joan, F., Jorge N., Michael L., Cathy L., George D., and Karen B.,(2013) Framework for identifying chemicals with structural features associated with the potential to act as developmental or reproductive toxicants. Chem.Res.Toxicol. 26(12), 1840-1861.

References cited in the original article:

1. Fang, H., Tong, W., Branham, W., Moland, C., Dial, S., Hong, H., Xie, Q., Perkins, R., Owens, W.,Sheehan, D. (2003) Study of 202 natural, synthetic, and environmental chemicals for binding to the androgen receptor. *Chem. Res. Toxicol. 16*, 1338-1358

DART SAR model – Explain results

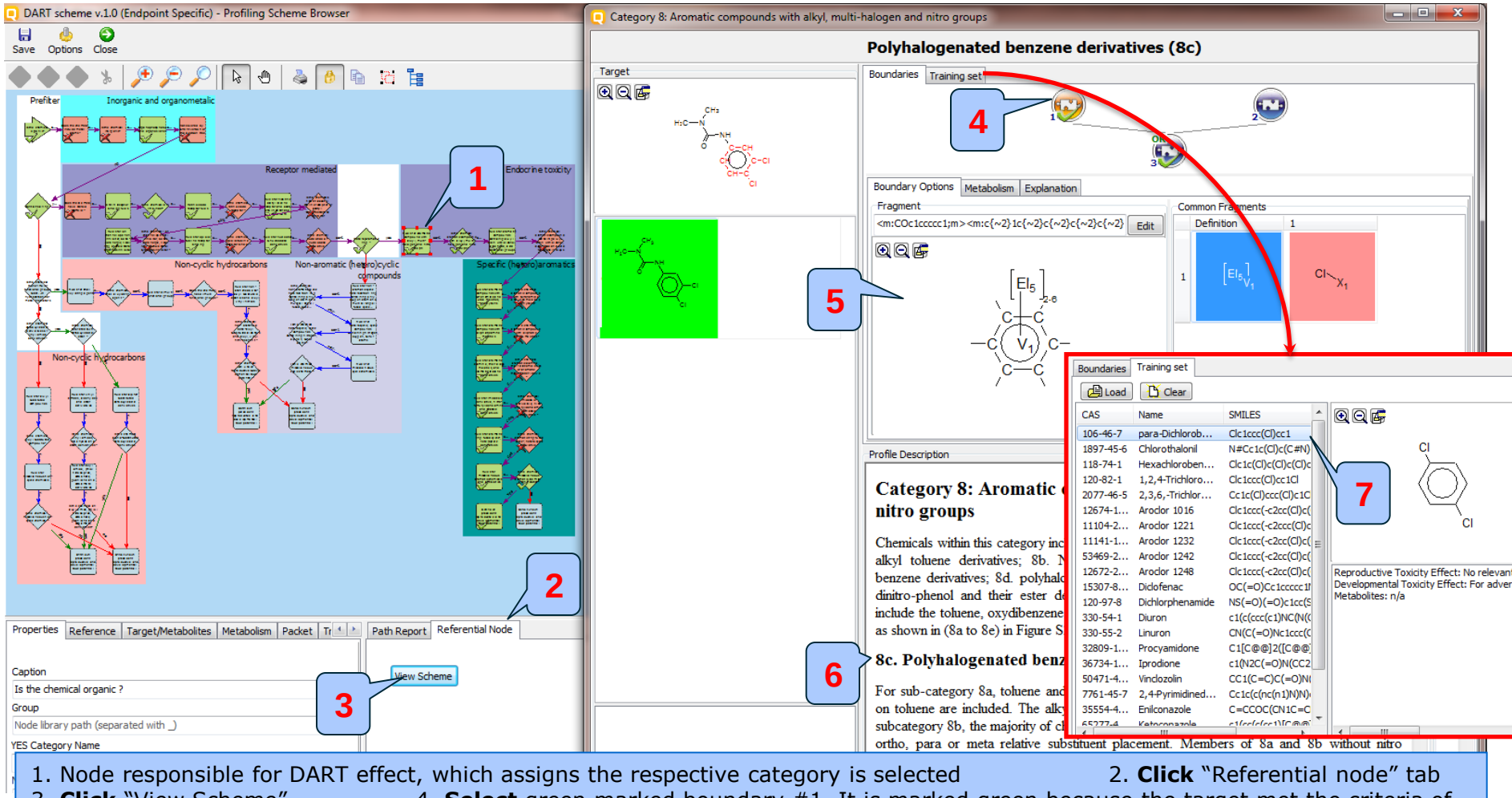
Polyhalogenated benzene derivatives (8c)



1. **Select** second category "Polyhalogenated benzene derivatives (8c)"
2. **Click** "Details"

DART SAR model – Explain results

Polyhalogenated benzene derivatives (8c)



1. Node responsible for DART effect, which assigns the respective category is selected
2. **Click** "Referential node" tab
3. **Click** "View Scheme"
4. **Select** green marked boundary #1. It is marked green because the target met the criteria of the boundary.
5. Definition of structural boundary
6. Textual description associated with DART category Details about the category is presented on next slide
7. Training set associated with category

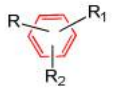
DART SAR model – Explain results
Polyhalogenated benzene derivatives (8c)

Category 8: Aromatic compounds with alkyl, multi-halogen and nitro groups

Chemicals within this category include the following five sub-categories: 8a. toluene and small alkyl toluene derivatives; 8b. NO2-alkyl/NO2-benzene derivatives; 8c. polyhalogenated benzene derivatives; 8d. polyhalogenated-, NO2/halogenated-oxydibenzene; 8e. dihalogen-, dinitro-phenol and their ester derivatives. The general core structures of these chemicals include the toluene, oxydibenzene and phenol ring with alkyl, halogen and/or nitro substituents as shown in (8a to 8e) in Figure S24.

8c. Polyhalogenated benzene derivatives

For sub-category 8a, toluene and a single alkyl chain substituent (< 5 carbon atoms) present on toluene are included. The alkyl substituents can be at ortho, para or meta-positions. For subcategory 8b, the majority of chemicals are mono-, di-, tri-nitrobenzene or nitrotoluene with ortho, para or meta relative substituent placement. Members of 8a and 8b without nitro substituents appear to be primarily developmental toxicants, while addition of a nitro group may be associated with a distinctive pattern of male reproductive toxicity. Sub-category 8c includes multi-chlorinated benzene derivatives containing from 2 to 6 chlorine atoms. Other possible substituents include methyl or nitrile groups. The members of this class included here are primarily developmental toxicants (Appendix 1). Sub-category 8d includes multi-substituted oxydibenzene with halogen or halogen/nitro substituents. Because these chemicals normally do not readily form co-planar structures, they are not anticipated to bind to the AhR, (see section 3b-2), nor do they have a mode of action dependent on the AhR interaction.[29]



$R=R_1=R_2=Cl$
(# of Cls from 1 to 6)

8c

Fig. S24. The structural scope of alkyl substituted benzene, alkyl/NO2-substituted benzene, polyhalogenated benzene, oxydibenzene, poly-halogenated, poly-NO2phenol and their esters

Original reference:

Shengde, W., Joan, F., Jorge N., Michael L., Cathy L., George D., and Karen B.,(2013) Framework for identifying chemicals with structural features associated with the potential to act as developmental or reproductive toxicants. Chem.Res.Toxicol. 26(12), 1840-1861.

References cited in the original article:

1. ATSDR (Agency for Toxic Substances and Disease Registry) (2004) Toxicological profile for polybrominated biphenyls and polybrominated diphenyl ethers.

Interpretation of SAR results

- Obtained SAR results show, that target chemical may elicit developmental and reproductive toxic potential based on belonging to the two DART toxic categories:
  - N-aryl subsituted urea(2b-4)
  - Polyhalogenated benzene derivatives (8c)
- Both DART toxic categories are characterized with mechanistic interpretation and training set chemicals with observed DART data.

Summary

- Toolbox 3.3 includes two application of DART scheme as discussed by Wu S paper:
  - DART scheme that can be used as a profiler for category formation or
  - DART SAR model for obtaining results based on DART profiling scheme and DART training set database.

Outlook

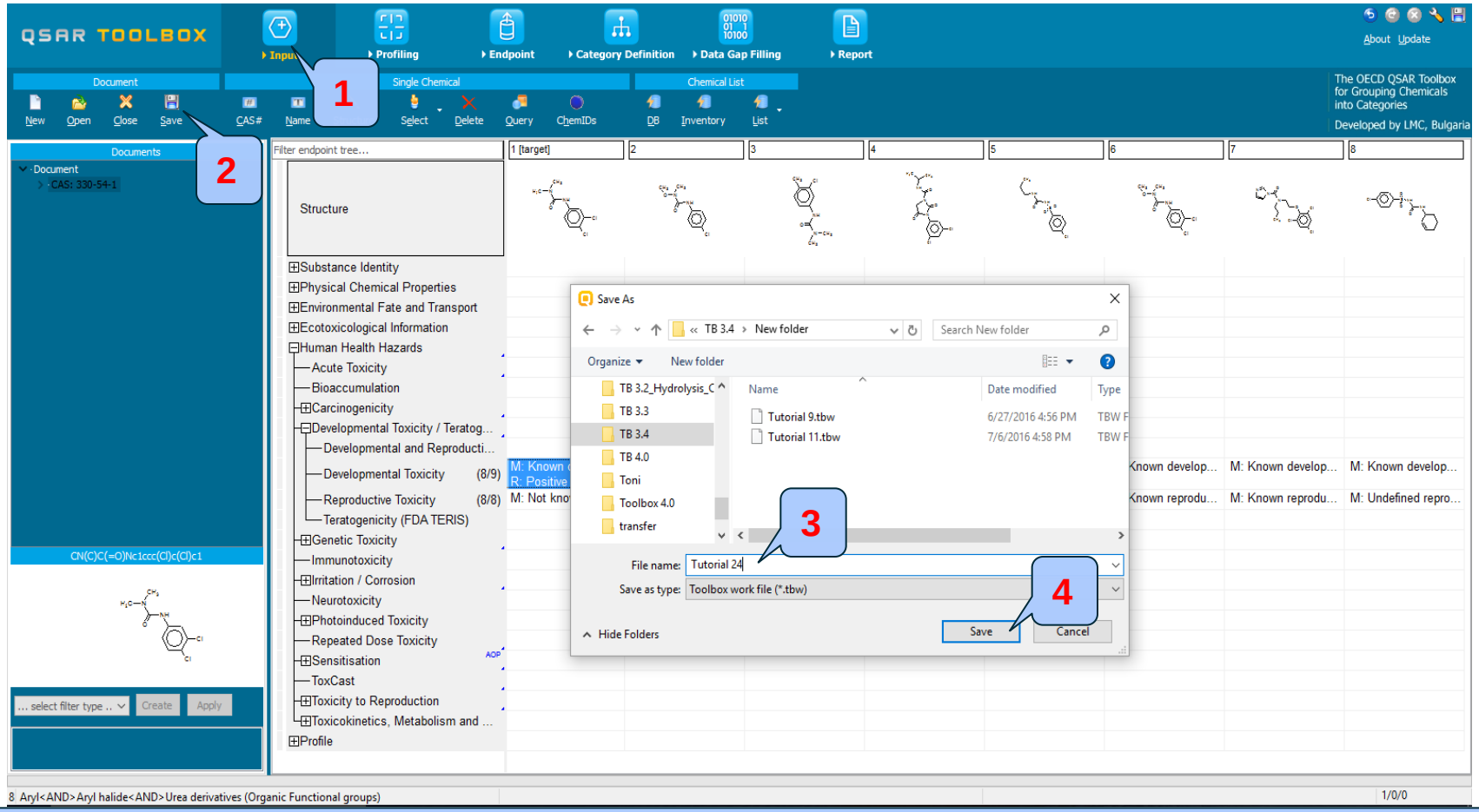
- Background
- Objectives
- The exercise
- **Workflow**
  - Input
  - Profiling
  - Endpoint
  - Category definition
  - Support the prediction by DART model
- **Save predictions**

Saving the prediction result

- This functionality allows storing/restoring the current state of Toolbox documents including loaded chemicals, experimental data, profiles, predictions etc., on the same computer. The functionality is implemented based on saving the sequence of actions that led to the current state of the Toolbox document and later executing these actions in the same sequence in order to get the same result(s).

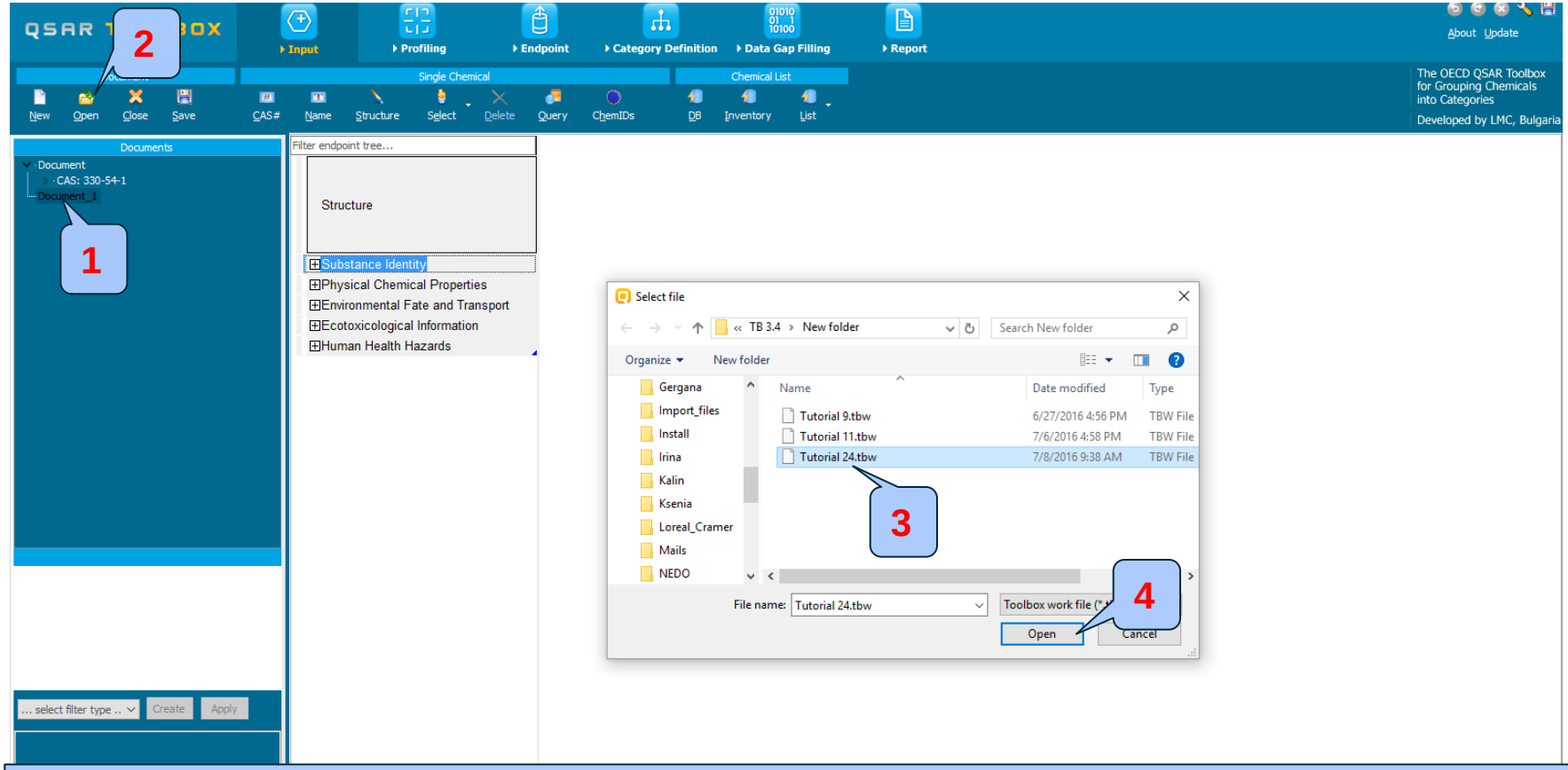
- Saving/Loading the file with TB prediction is shown on next screenshots.

Saving the prediction result



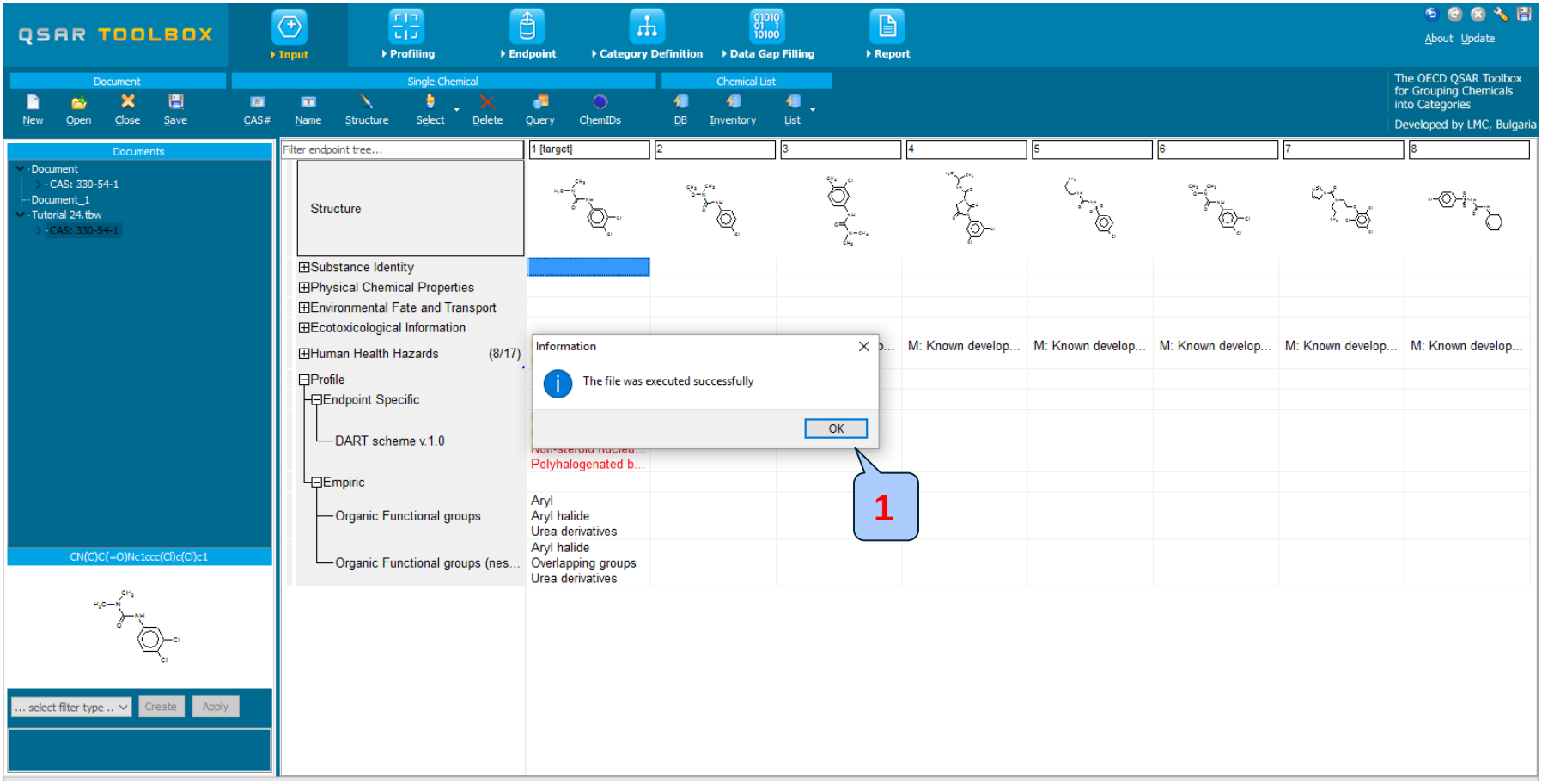
1. **Go** to "Input" section; 2.**Click** on "Save" button; 3. **Browse** and **put** name of the file;
4. **Click** "Save" button

Open saved file



1. **Create** new document; 2. **Click** "Open"; 3. **Find** and **select file**; 4. **Click** "Open"

Open saved file



The file is opened successfully 1. **Click** "OK"