

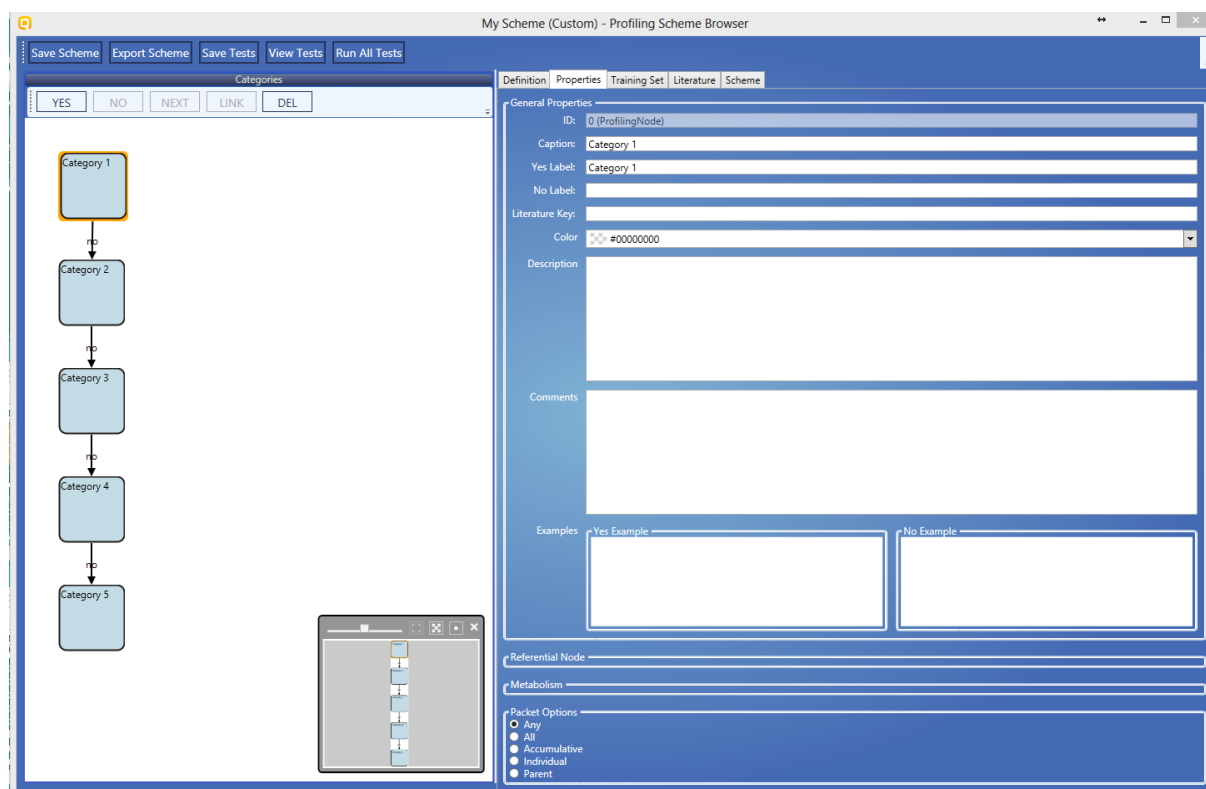
Manual for creating prioritization schemes (using the Profiling Scheme Editor)

Contents

1. Features of the editor.....	3
1.1 Nodes and node manipulations	3
1.2 Node properties	4
1.3 Queries, logic, operations.....	5
1.4 Types of queries	5
1.5 Query content editing	6
1.6 Toolbar buttons	7
2. Creating an example scheme	8
3. Applying an existing scheme and viewing statistics.....	17

1. Features of the editor

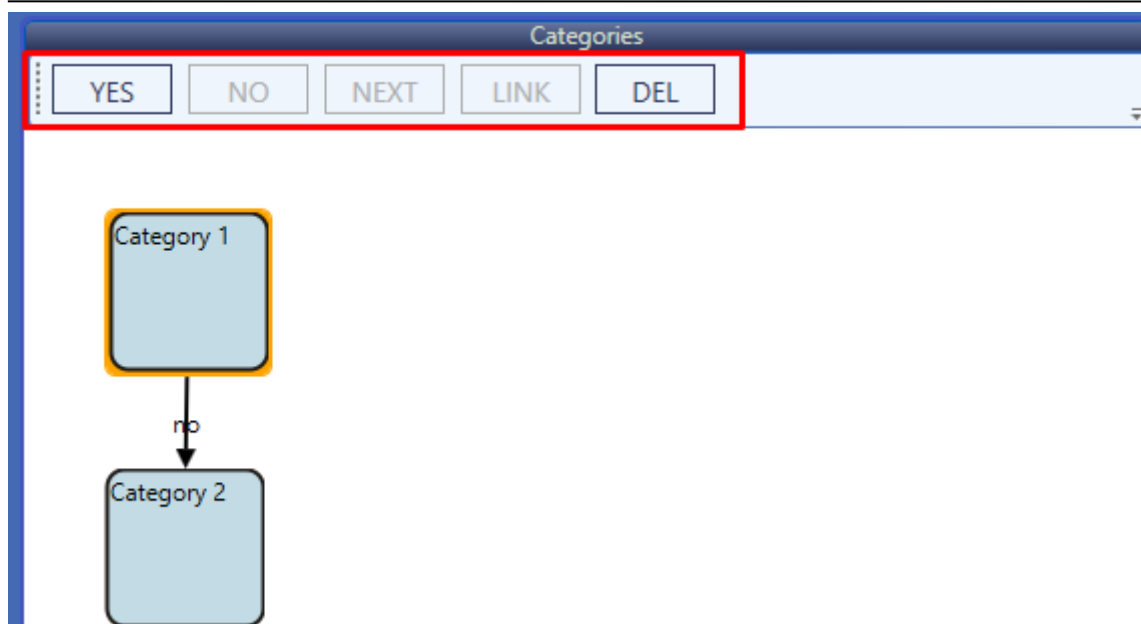
The Profiling Scheme Editor (Browser) gives users the ability to view and edit prioritization schemes. Its features are described in the sections below.



1.1 Nodes and node manipulations

Prioritization schemes in Toolbox 4 are built of logically connected nodes. The nodes themselves are decision blocks navigating each assessed structure through the scheme based on the result of the contained expression. If the structure answers the criteria of the expression, it is passed along the YES route (optionally assigning it a YES label). If it does not answer them, it is passed along the NO route (again, optionally assigning it a NO label). If a NEXT route exists (in place of a YES and/or NO) the structure is passed along that route, regardless of the outcome. It may still be assigned the corresponding label.

The left section of the Editor displays the structure of the prioritization scheme and lets the user manipulate the nodes and their relations. The functions of the buttons in the toolbar are as follows:



YES/NO/NEXT - when a node is selected, a left click on any of these 3 buttons will create a new node, connected to the selected one with the chosen relation. Each of the buttons is enabled only if the relation is possible for the selected node. Right clicking on the buttons provides additional functions, such as adding metabolism or referential nodes or connecting an existing node to the selected one.

LINK - when a link (yes/no/next connection) is selected, right clicking on these buttons shows the available relations that this link can be converted to.

DEL - when a node is selected, clicking on this button will delete it along with any nodes, originating from it

1.2 Node properties

The properties of each node can be viewed and/or modified in the Properties tab. The editable fields are:

The screenshot shows the 'Properties' tab of the Profiling Scheme Editor. The 'General Properties' section contains the following fields:

- ID: 0 (ProfilingNode)
- Caption: Category 1
- Yes Label: Category 1
- No Label: (empty)
- Literature Key: (empty)
- Color: #00000000

Below the 'General Properties' section are three main text areas:

- Description: (empty)
- Comments: (empty)
- Examples: (empty)

The 'Examples' section is divided into two sub-sections: 'Yes Example' and 'No Example', both of which are currently empty.

Caption - descriptive identifier of the node

Yes Label - the label that is assigned to the structure if it fulfills the expression criteria

No Label - the label that is assigned to the structure if it does not fulfill the expression criteria

Description - free text describing the coded logic and the meaning behind the labels

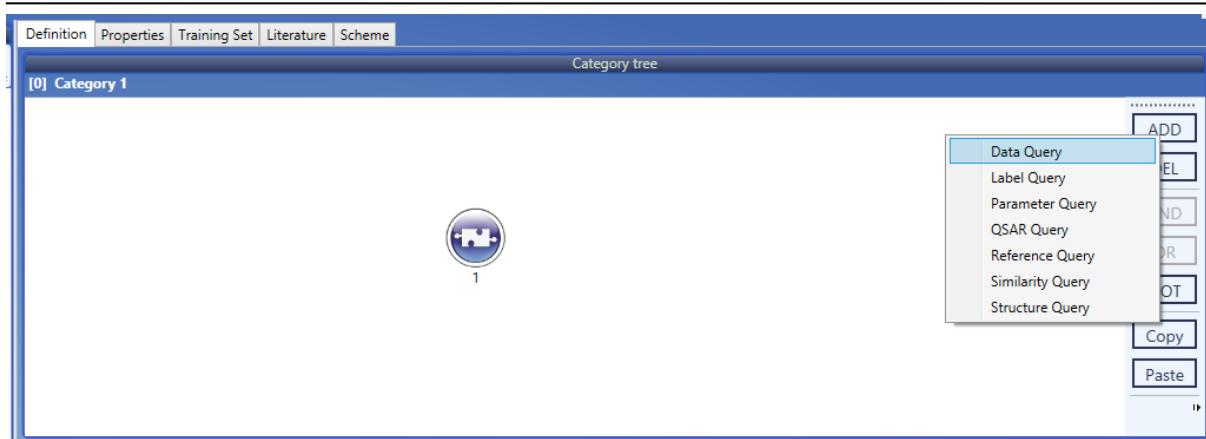
Comments - free text

1.3 Queries, logic, operations

The logical expression coded in each node is displayed on the Definition tab. Queries can be added, deleted and logically combined there. It is also possible to copy the entire query tree and paste it in another node.

1.4 Types of queries

The available queries are listed when the user click on the ADD button:



Data Query - assesses experimental data for the target chemical

Label Query - assesses labels already assigned to the target chemical

Parameter Query - assesses a calculable property of the target chemical

QSAR Query - applies an existing QSAR and assesses the result

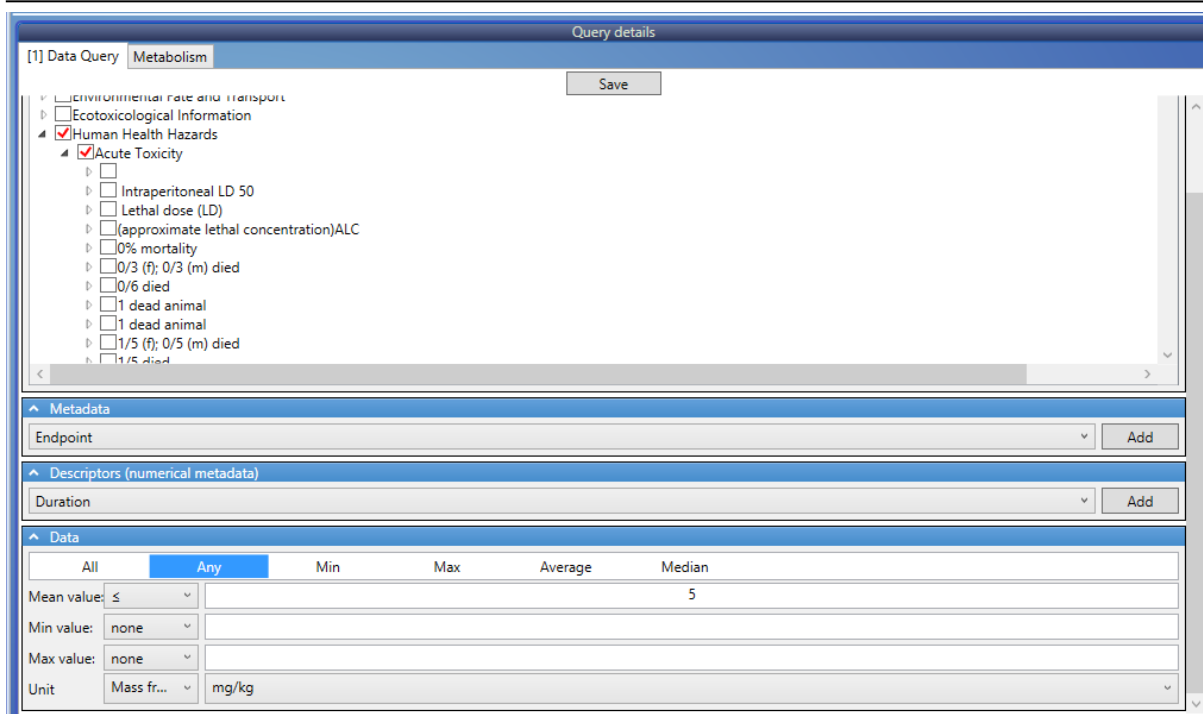
Reference Query - checks the outcome of a referenced prioritization scheme

Similarity Query - assesses the similarity between the examined chemical and a predefined target

Structure Query - assesses the chemical structure of the target chemical

1.5 Query content editing

The contents of each query are displayed in the bottom right part of the editor. Each query has its specific properties.



Query details

[1] Data Query | Metabolism

Save

- Environmental fate and transport
- Ecotoxicological Information
- Human Health Hazards
 - Acute Toxicity
 -
 - Intrapерitoneal LD 50
 - Lethal dose (LD)
 - (approximate lethal concentration)ALC
 - 0% mortality
 - 0/3 (f); 0/3 (m) died
 - 0/6 died
 - 1 dead animal
 - 1 dead animal
 - 1/5 (f); 0/5 (m) died
 - 1/5 died

Metadata

Endpoint Add

Descriptors (numerical metadata)

Duration Add

Data

	All	Any	Min	Max	Average	Median
Mean value:	≤					5
Min value:	none					
Max value:	none					
Unit	Mass fr...		mg/kg			

1.6 Toolbar buttons

The main toolbar buttons serve the following purposes:

Save Scheme - saves the updated scheme to the server

Export Scheme - saves the scheme to a destination specified by the user. Any changes made do not affect the original Toolbox scheme

Save Tests - saves the query tests to the server

View Tests - shows the query tests browser

Run All Tests - executes all query tests in the scheme

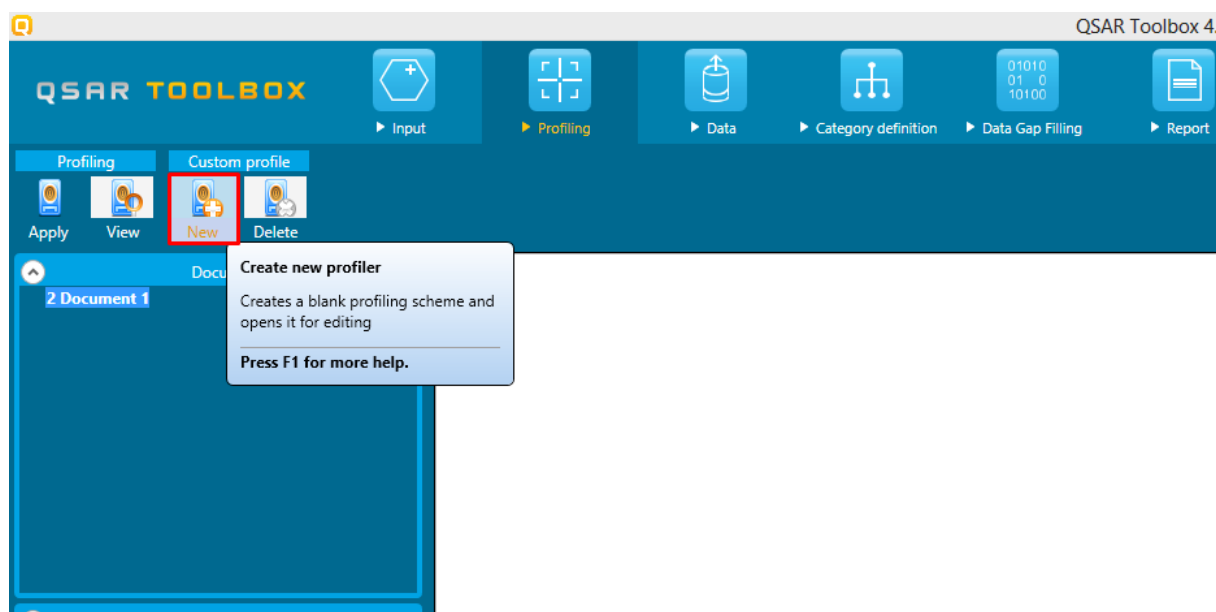
2. Creating an example scheme

To illustrate the process of creating a prioritization scheme, we will use the [GHS classification criteria for acute toxicity](#) and will develop a new scheme implementing its logic. To simplify the example we will focus on oral toxicity only. We will create a scheme with 5 nodes, each one assessing LD50 and assigning the corresponding label according to the following table:

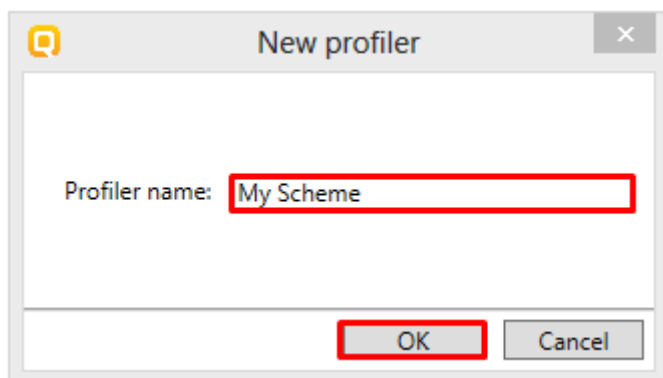
LD50, mg/kg	Label
LD50 ≤ 5	Category 1
5 < LD50 ≤ 50	Category 2
50 < LD50 ≤ 300	Category 3
300 < LD50 ≤ 2000	Category 4
2000 < LD50 ≤ 5000	Category 5
LD50 > 5000	Not acutely toxic

The last category (Not acutely toxic) will be assigned as a NO label by the node assessing for Category 5.

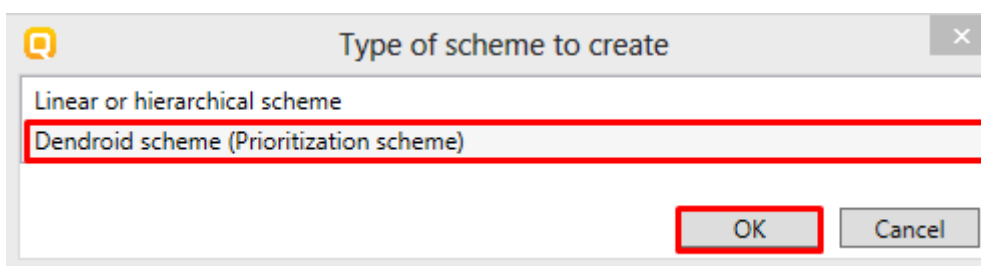
We start by creating a new profiling scheme using the New button on the Profiling tab.



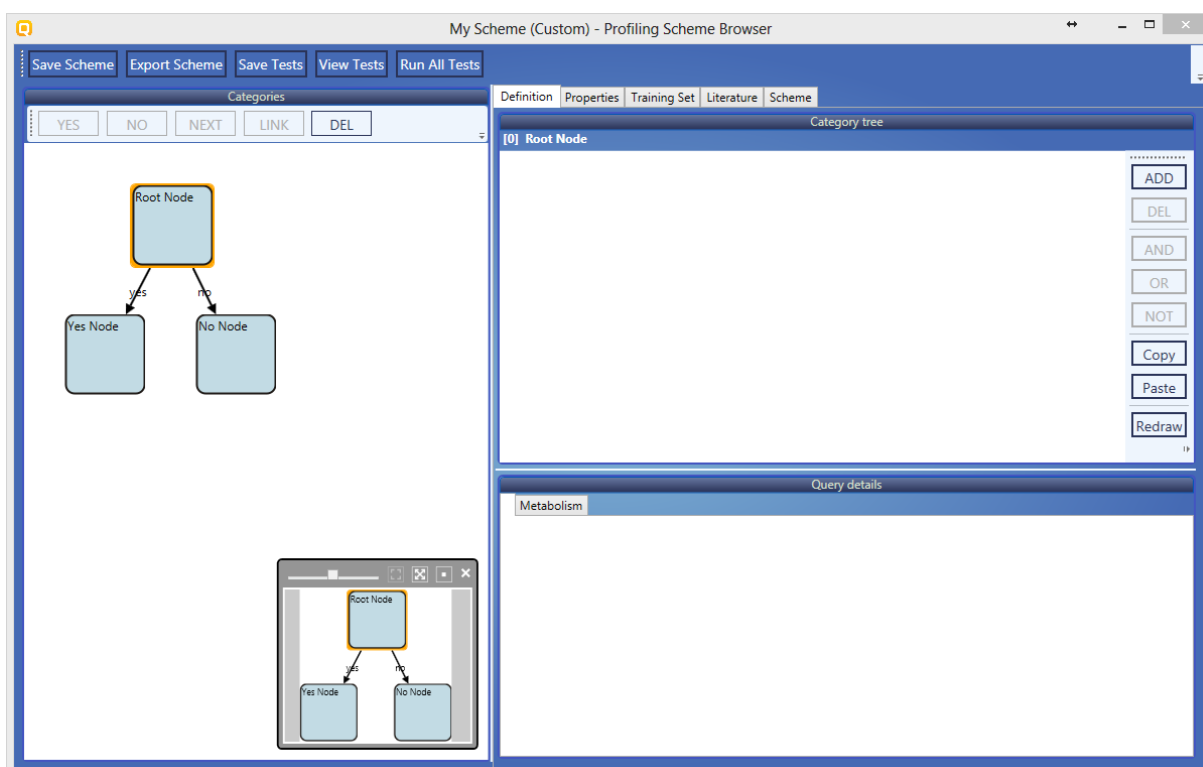
We then enter a name for the new scheme, in this example "My Scheme", and then click OK.



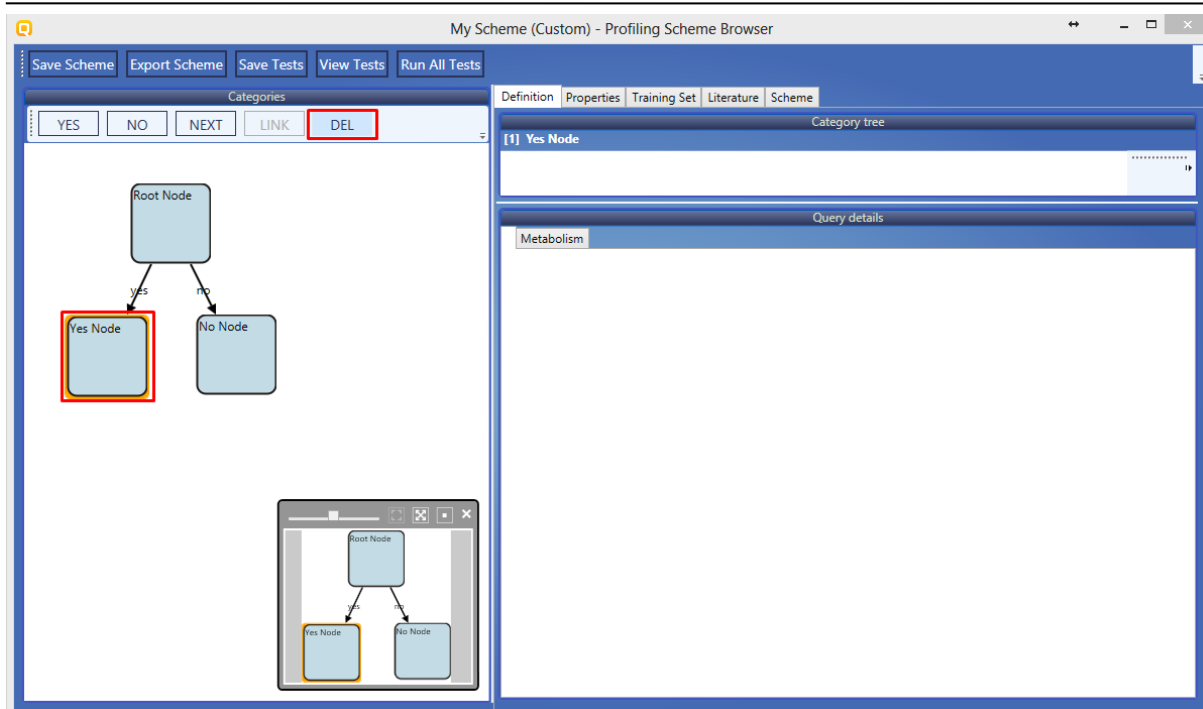
The following screen asks what type of scheme to create. For our example we select "Dendroid scheme (Prioritization scheme)" and click OK.



The new scheme is created and visualized. It is prepopulated with a Root Node as well as a Yes Node and a No Node.

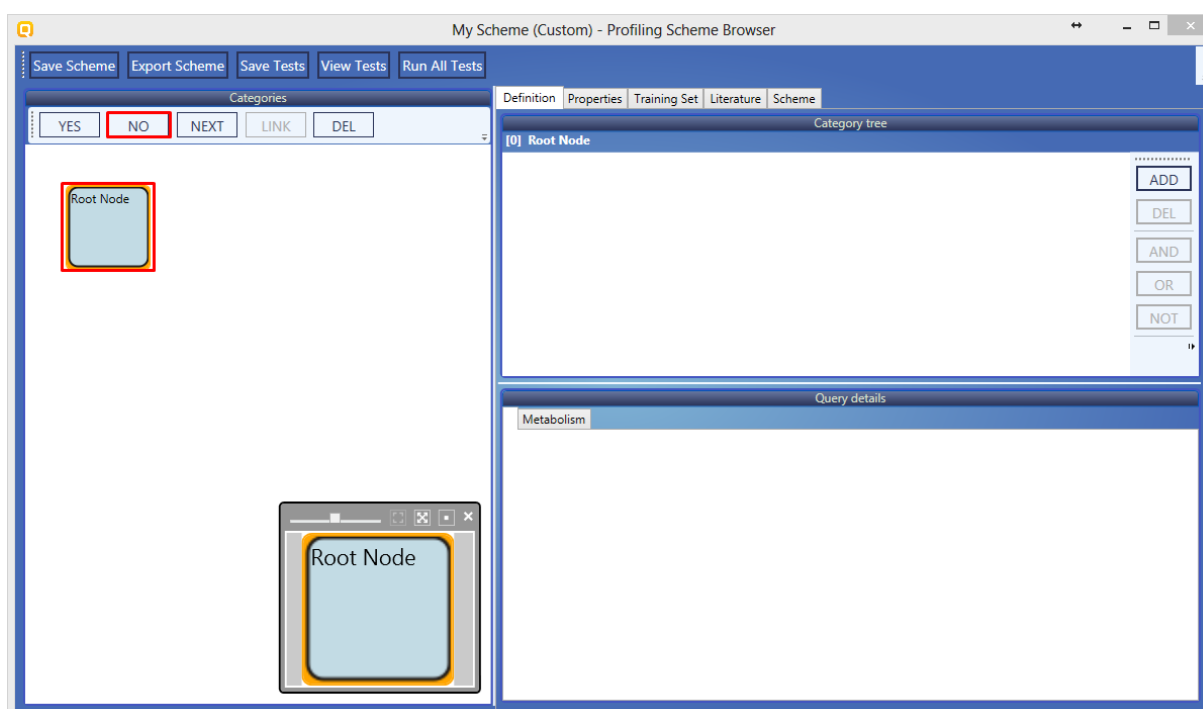


We delete the existing Yes Node by first selecting it and then clicking the DEL button.

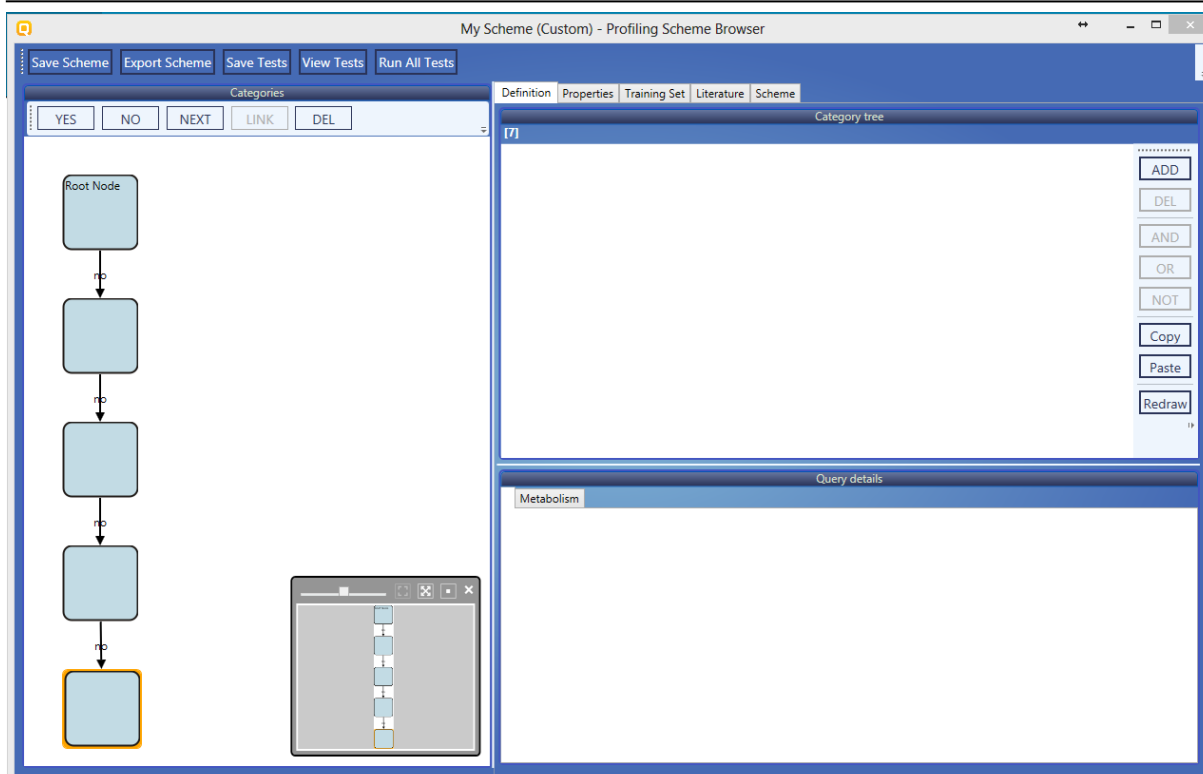


We repeat this for the No Node and are left with the Root Node only.

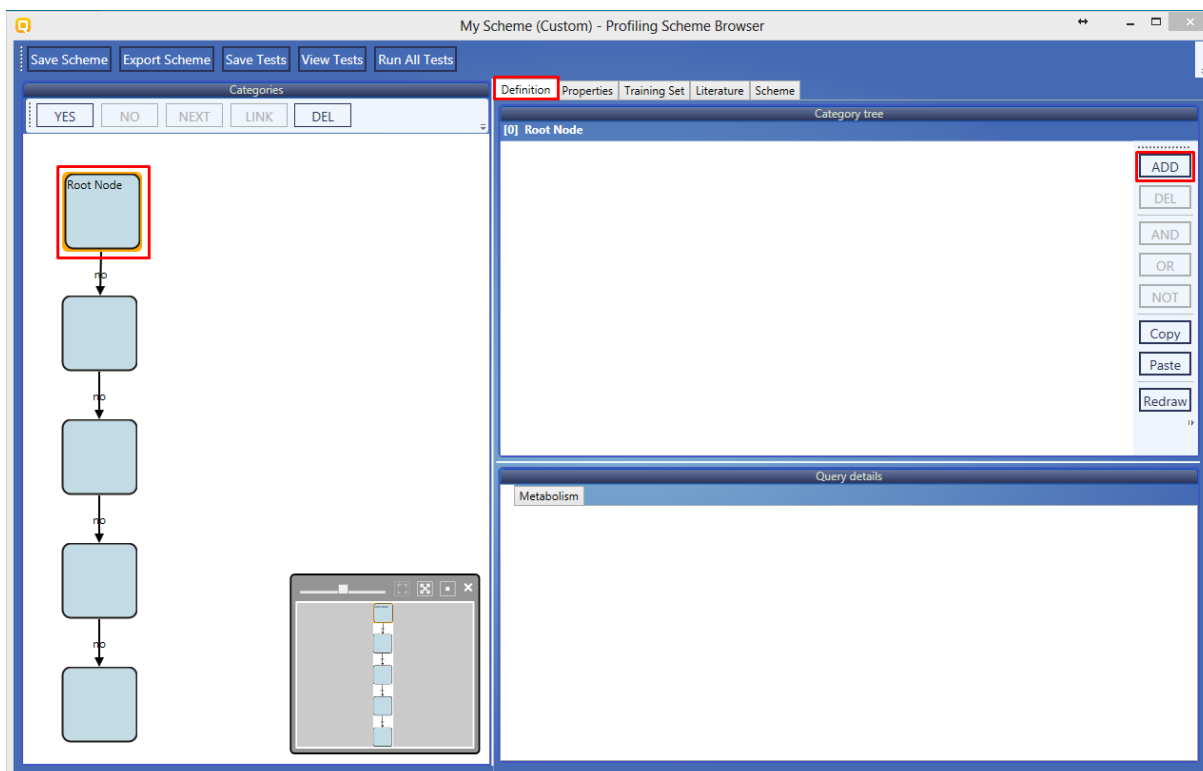
We need to create 4 consecutive NO nodes that will handle each group of chemicals. After selecting the Root Node, we click on the NO button 4 times.



At this stage we have the skeleton of the prioritization scheme. We can start filling in the content of the nodes.



With the Root Node selected, we click the ADD button and then select Data Query.



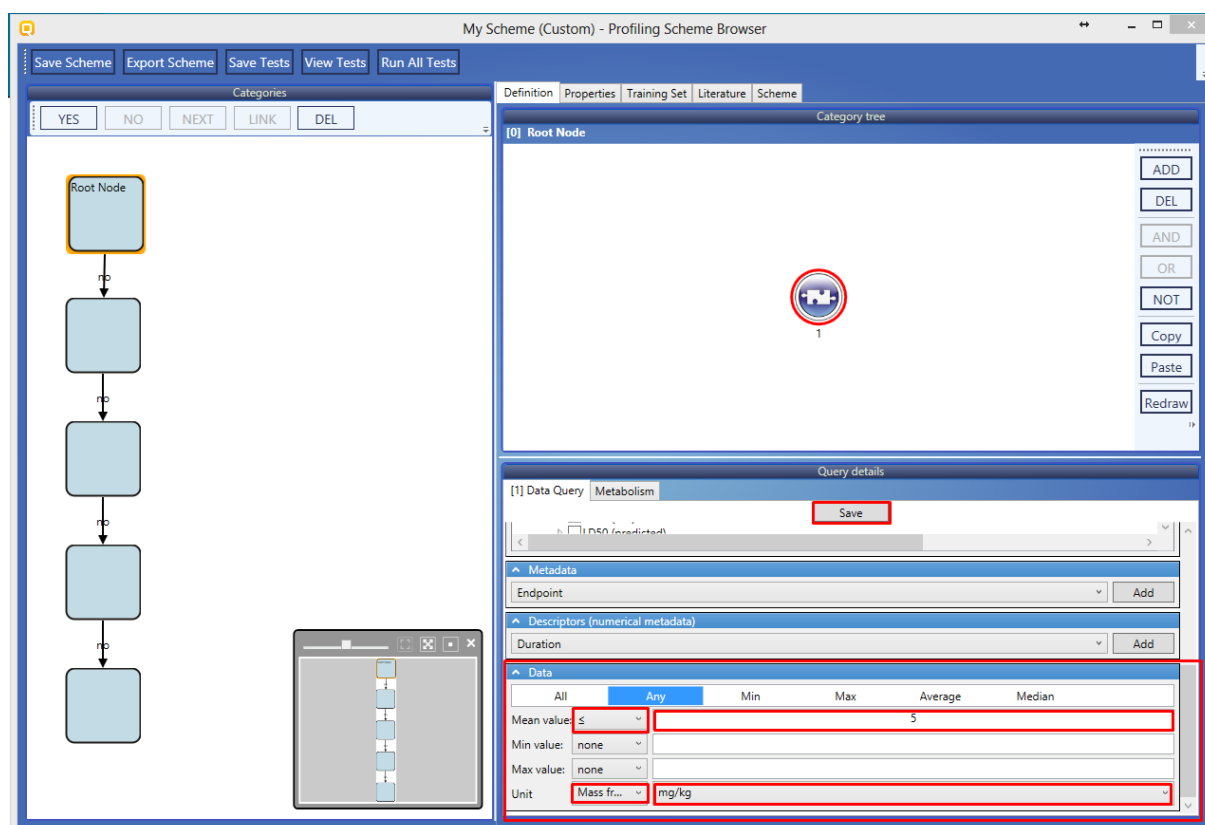
The Data Query editor appears in the bottom right part of the editor. We are going to input the following condition – substances with LD50 less than or equal to 5. To do so we expand the path starting with Human Health Hazards node, followed by the Acute Toxicity node.

The screenshot shows the 'My Scheme (Custom) - Profiling Scheme Browser' interface. On the left, a vertical flowchart shows a 'Root Node' connected to four subsequent nodes. A 'Categories' menu at the top includes 'YES', 'NO', 'NEXT', 'LINK', and 'DEL'. The main workspace is divided into two panes: 'Definition' and 'Properties'. The 'Definition' pane shows a 'Category tree' with a single node '1' circled in red. The 'Properties' pane shows 'Query details' for '[1] Data Query Metabolism'. A tree view in the 'Query details' pane is expanded to 'Human Health Hazards' and 'Acute Toxicity', with 'Acute Toxicity' highlighted by a red box. Below the tree view, a list of sub-categories is shown, including 'Intraperitoneal LD 50', 'Lethal dose (LD)', '(approximate lethal concentration)ALC', '0% mortality', '0/3 (f); 0/3 (m) died', '0/6 died', '1 dead animal', '1 dead animal', and '1/5 (f); 0/5 (m) died'.

We navigate to LD50, expand it and select the Acute Oral toxicity leaf.

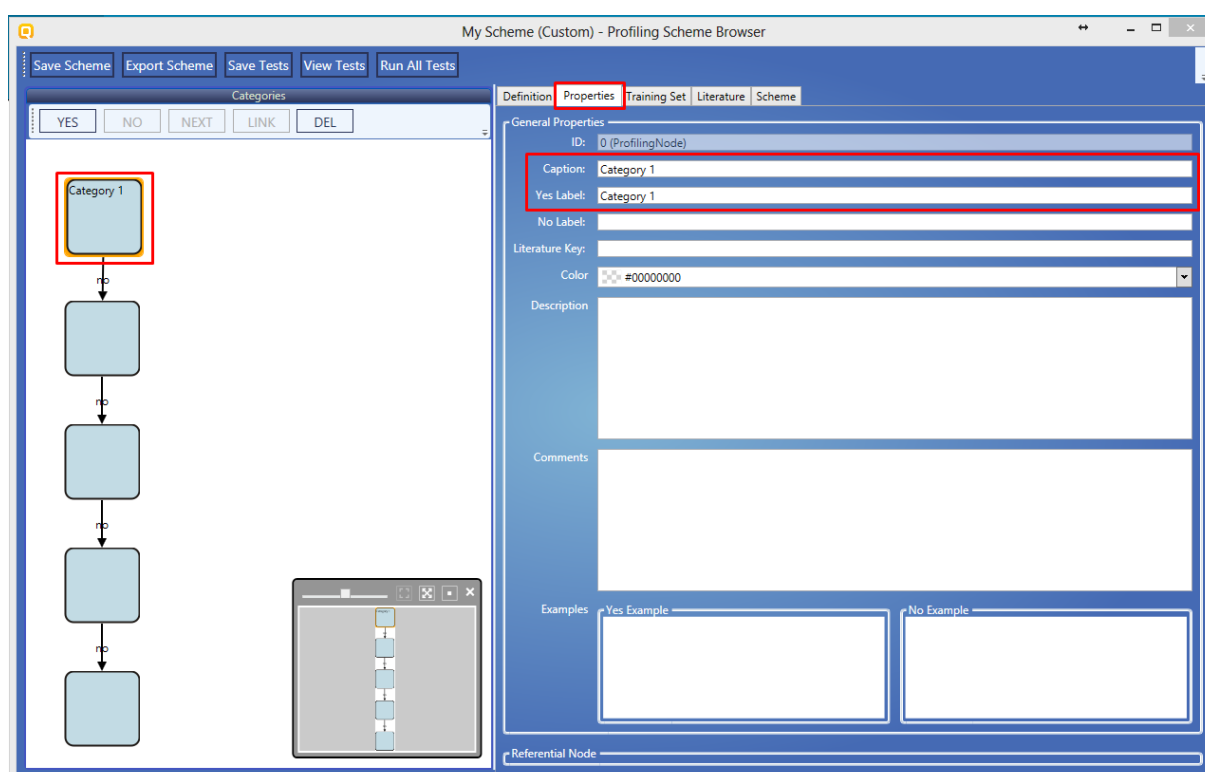
This screenshot shows the same interface as the previous one, but with the 'Query details' pane expanded further. The tree view is now expanded to 'LD50', and the 'Acute Oral toxicity' leaf is selected, highlighted by a red box. The list of sub-categories below the tree view includes 'LD40', 'LD5', 'Acute Oral toxicity', 'ECHA CHEM', 'ECOTOX', 'Rodent Inhalation Toxicity Database', 'Toxicity Japan MHLW', 'ZEBET database', 'LD50 (24 h)', 'LD50 (7 d)', 'LD50 (cut off value)', 'LD50 (cut off)', 'LD50 (cut-off value)', 'LD50 (IHT)', and 'LD50 (IHT) (cut-off value)'.

To enter the exact criteria, we scroll down to the Data section of the query editor, switch to Any, select \leq as the Mean Value operand and then enter 5 in the value field. We also define the unit – Mass fraction, mg/kg. We click on the Save button and the Data Query is now configured.



The screenshot shows the 'My Scheme (Custom) - Profiling Scheme Browser' interface. The 'Query details' section is active, showing a 'Data Query' for 'Metabolism'. The 'Data' section is expanded, and the 'Any' tab is selected. The 'Mean value' is set to \leq with a value of 5. The 'Unit' is set to 'Mass fr...' and 'mg/kg'. A red box highlights the 'Save' button and the 'Data' section.

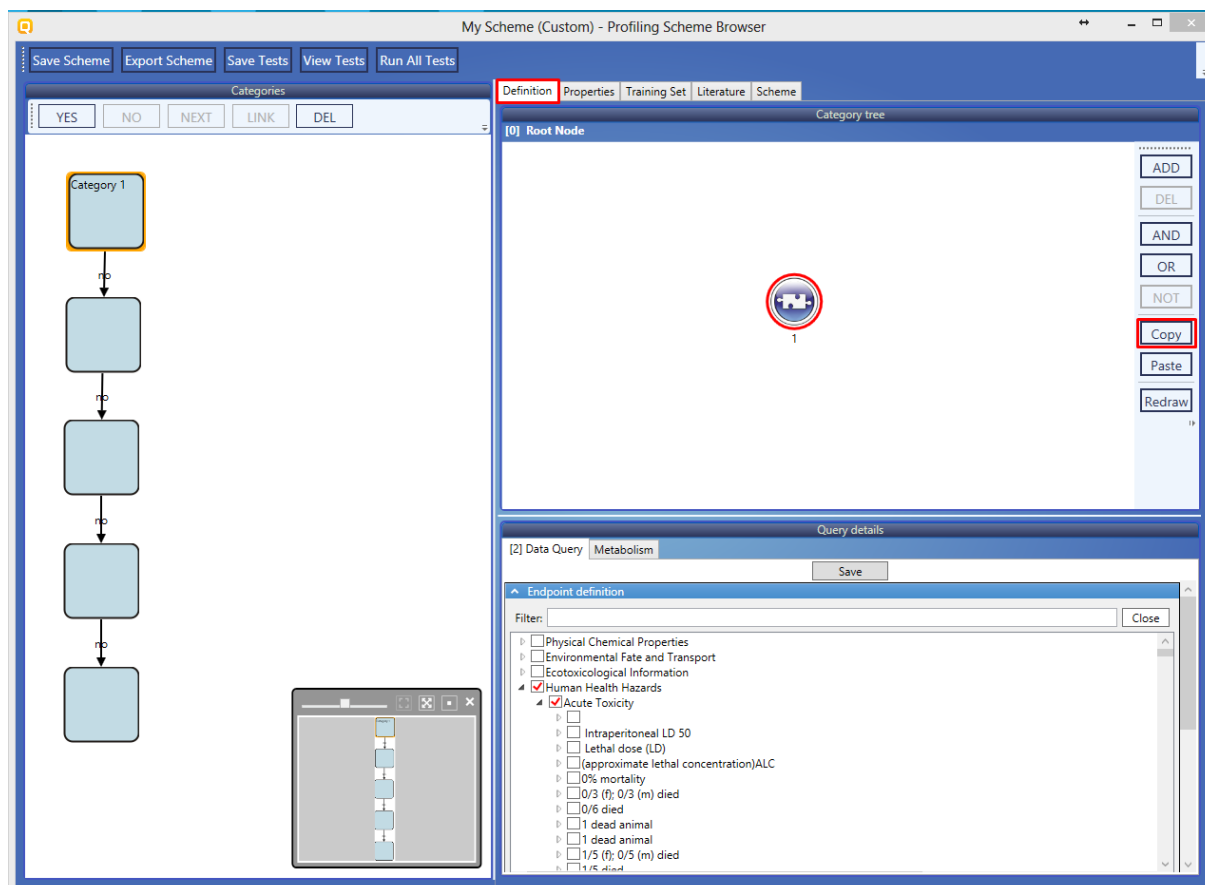
We can now go to the Properties tab to set the node's caption and yes label both to Category 1.



The screenshot shows the 'My Scheme (Custom) - Profiling Scheme Browser' interface with the 'Properties' tab selected. The 'General Properties' section is visible, and the 'Caption' and 'Yes Label' fields are both set to 'Category 1'. A red box highlights these two fields.

All substances that answer the criteria set in the Data Query will be assigned the Category 1 label.

We can now move on to the next node. As the only difference in the query will be the threshold, we can copy it from the Category 1 node. To do so, while the first node is selected, we navigate to the Definition tab and click the Copy button.



We then select the second node and click on Paste.

We then need to update the threshold from 5 to 50 and click on the Save button.

	All	Any	Min	Max	Average	Median
Mean value:	≤				50	
Min value:	none					
Max value:	none					
Unit	Mass fr...	mg/kg				

When done with the query, we can again set the Caption and the Yes Label, this time to Category 2.

The remaining 3 nodes are populated in the same way, setting the thresholds to 300, 2000 and 5000 for Categories 3, 4 and 5 respectively.

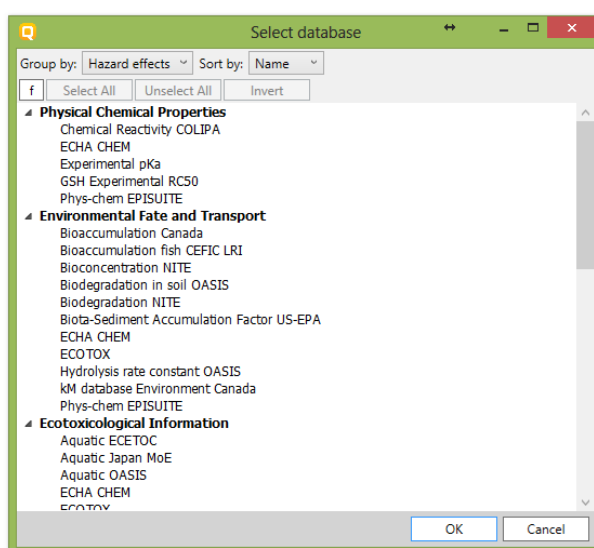
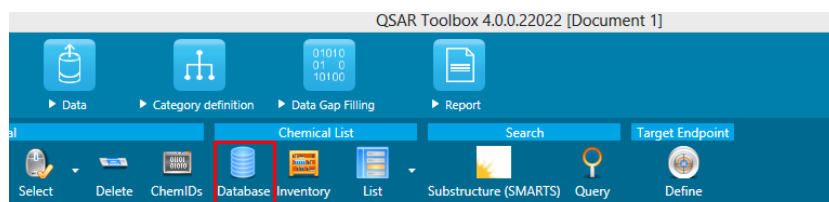
All substances that have LD50 > 5000 have to be assigned the "Not acutely toxic" label. To do so, we enter this string as the No Label in the last node.

We save the scheme by clicking on the Save Scheme button. The "Scheme successfully saved on the server" message should appear and we can close the editor.

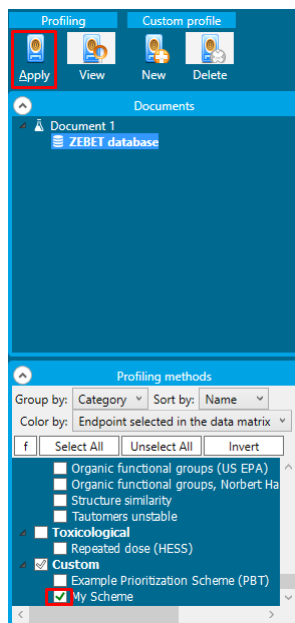
Our new scheme should now appear in the list of schemes.

3. Applying an existing scheme and viewing statistics

To test the newly created scheme, we will apply it to all chemicals within a database with LD50 data. To do so, we go to the Input tab of the QSAR Toolbox and click on the Database button under Chemical List.



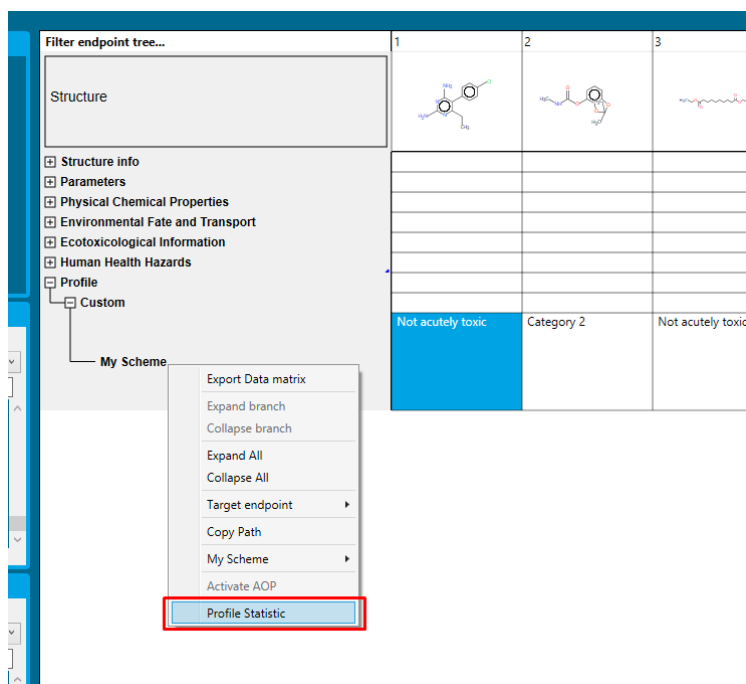
We scroll down and select the ZEBET database from the Human Health Hazards list. After a short while, the database is loaded onto the data matrix and we can go ahead and apply our scheme. We go to the Profiling tab, make sure My Scheme is checked and hit the Apply button to start profiling.



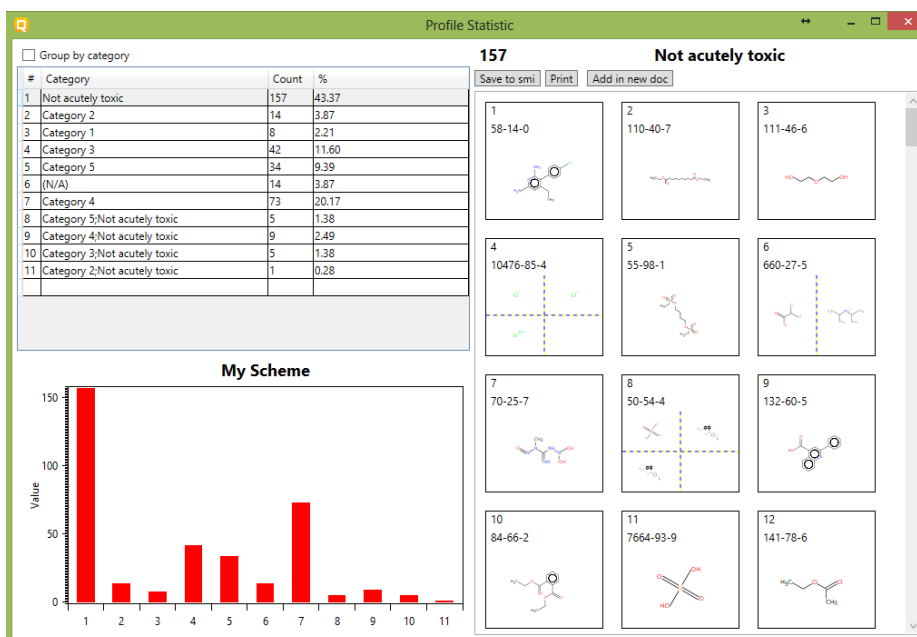
As this is a new profiling scheme, the profiling may take several minutes. After a while the process should finish and the results will be displayed on a new line on the data matrix.

Filter endpoint tree...	1	2	3	4	5	6	7	8	9	10
Structure										
Structure info										
Parameters										
Physical Chemical Properties										
Environmental Fate and Transport										
Ecotoxicological Information										
Human Health Hazards										
Profile										
Custom										
My Scheme	Not acutely toxic	Category 2	Not acutely toxic	Category 1	Category 3	Not acutely toxic	Not acutely toxic	Category 3	Not acutely toxic	Not acutely toxic

At this stage, we can request to see the profiling statistics by right clicking on My Scheme and then clicking on Profile Statistic.

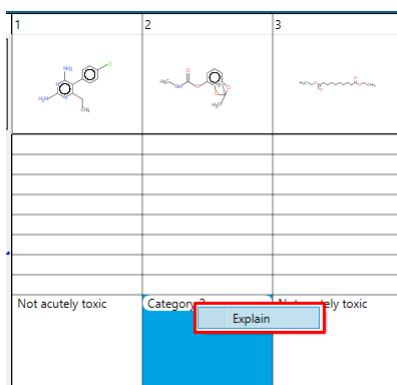





This will bring up the Profile Statistic form showing the distribution of the substances across the defined categories.



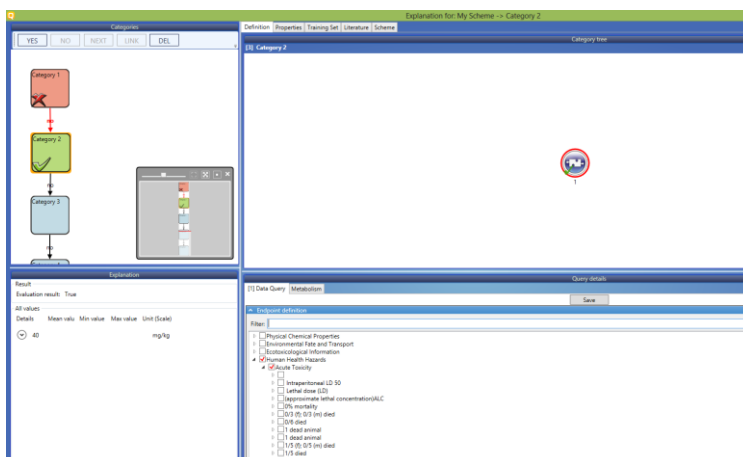
#	Category	Count	%
1	Not acutely toxic	157	43.37
2	Category 2	14	3.87
3	Category 1	8	2.21
4	Category 3	42	11.60
5	Category 5	34	9.39
6	(N/A)	14	3.87
7	Category 4	73	20.17
8	Category 5;Not acutely toxic	5	1.38
9	Category 4;Not acutely toxic	9	2.49
10	Category 3;Not acutely toxic	5	1.38
11	Category 2;Not acutely toxic	1	0.28

We can also request an explanation of an assigned category by right clicking on it and selecting Explain.



 Not acutely toxic	 Category 2	 Not acutely toxic
--	---	--

This will bring up the Profiling Editor but this time it will be in explain mode.



In the example shown, Category 2 is assigned by the green node because LD50 is 40 mg/kg, and the threshold coded in the Data Query is $LD50 \leq 50$ mg/kg.