

QSAR TOOLBOX

The OECD QSAR Toolbox
for Grouping Chemicals
into Categories

User manual

Database Import Wizard

Document history

Version	Comment
Version 1.0	Database Import Wizard for version 2.1 of the QSAR Toolbox

Issue date: April 2011

Language: English

If you have questions or comments that relate to this document, please send them to

ehscont@oecd.org or visit the QSAR Toolbox discussion forum at

https://community.oecd.org/community/toolbox_forum

Table of Contents

Document history	1
Document history	2
1 Executive summary	4
2 QSAR Toolbox data model.....	4
3 Import layouts.....	5
3.1 Vertical layout.....	5
3.2 Horizontal layout	6
4 Endpoint tree path	7
5 Building the dynamic tree.....	8
6 Running the import wizard	12
6.1 Vertical	13
6.2 Horizontal	16
Appendix I: Preparing a file for horizontal import	19
Appendix II: Import example for a database with ecotoxicological information	31
Appendix III: Import example for database with Human health hazards	
Information	47
Appendix IV: Import example for database with Human health hazards	
Information	59

1 Executive summary

The QSAR Toolbox Database Import Wizard, together with the IUCLID 5 import wizard (see guidance document "IUCLID 5 Import/Export via Webservices"), is the entry point for importing custom user data to the QSAR Toolbox database. It can import XLS files (Excel 97-2003 version) as well as TXT (UNICODE) plain text files. Both file types pertain to how the data is read by QSAR Toolbox, but not how the data is parsed afterwards.

2 QSAR Toolbox data model

The QSAR Toolbox operates with the following data model:

Data point record

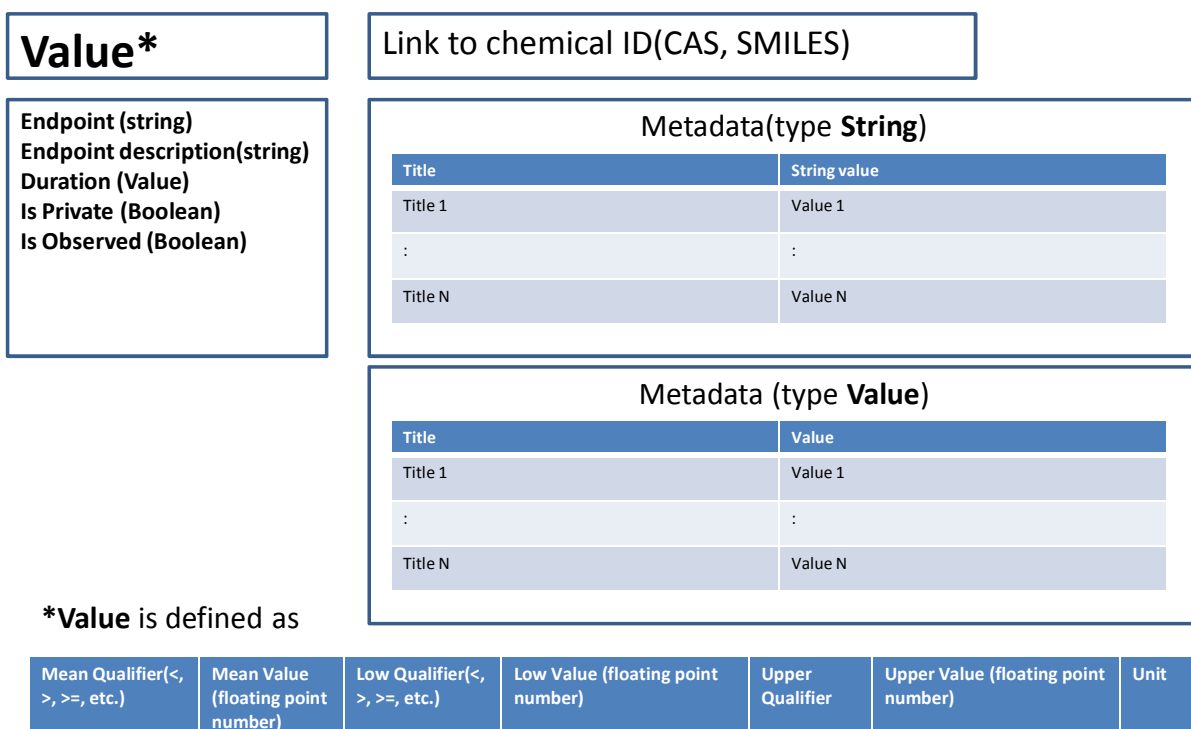


Figure 1: Database structure of the QSAR Toolbox

The Import's function is to translate the information in a file (be it XLS or TXT), separate it in different chunks (see the figure above) and write them into the database. The information consists of the chemical connected with numerical and meta-data. In other words the point of the import is to define a list of data points (the number that the user sees in the data-matrix and uses for gap-filling) with its corresponding metadata, namely the additional information on duration, test organisms, endpoint etc. In order to properly parse the information the import expects one of two file layouts as outlined below.

3 Import layouts

The two layouts the QSAR Toolbox can parse are the so called Vertical layout and the Horizontal layout. The Horizontal layout has each data point, with its corresponding **chemical** and **metadata**, defined in a single row. In a way each **row is a single record** (hence "horizontal"). The Vertical layout on the other hand can have **multiple records on each row** with the metadata for each record defined on a **column by column basis** (hence "vertical").

3.1 Vertical layout

This layout is used where there is a list of chemicals and a result for each chemical, but all results have the same metadata. So the chemical is defined in the first columns, and the next columns are used for the data points. For each data column there is one set of metadata. This means the vertical layout can import multiple values for a chemical.

	A	B	C	D	E
1	CAS	NAME	SMILES	Experiment 1	Experiment
2	CAS# 1	NAME 1	SMILES 1	Value 1	Value 1
3	CAS# 2	NAME 2	SMILES 2	Value 2	Value 2
4					
5					
6	CAS# K	NAME K	SMILES K	Value K	Value K

Figure 2: Vertical layout

Figure 2 illustrates the format of an XLS file for import. The first three columns represent the chemical identity information and column D and E represent results from two different "experiments" (a package of metadata such as Organ, Duration, Temperature, Dose, Species, Endpoint etc.).

3.2 Horizontal layout

This layout is used when each data point is defined in a row. Here the user specifies in which column is the data, the metadata and the type of metadata.

	A	B	C	D	E	F	G	H
1	CAS	NAME	SMILES	Param descriptor 1	Param descriptor 2		Param descriptor N	VALUE
2	CAS# 1	NAME 1	SMILES 1	Param descriptor value [1,1]	Param descriptor value [2,1]		Param descriptor value [1,1]	Value 1
3	CAS# 2	NAME 2	SMILES 2	Param descriptor value [1,2]	Param descriptor value [2,2]		Param descriptor value [1,2]	Value 2
4								
5								
6	CAS# K	NAME K	SMILES K	Param descriptor value [1,K]	Param descriptor value [2,K]		Param descriptor value [2,K]	Value K

Figure 3: Horizontal layout

Figure 3 shows how an XLS file could look like for horizontal import. Each row defines a record in its entirety. At import time the user specifies which columns contain chemical identity data (CAS, Name, SMILES), which columns contain the Value (the result that is seen in the Data-matrix and used for Data-gap filling) and which columns contain the metadata (e.g. Organ, Duration, Temperature, Dose, Species, Endpoint etc.).

4 Endpoint tree path

When a data point is imported into the QSAR Toolbox, the database engine needs to assign it to a leaf node in the endpoint tree. However, the way the endpoint tree is constructed differs significantly from version 1.1 of the QSAR Toolbox. For the user the tree looks similar in both versions, but the underlying logic has changed.

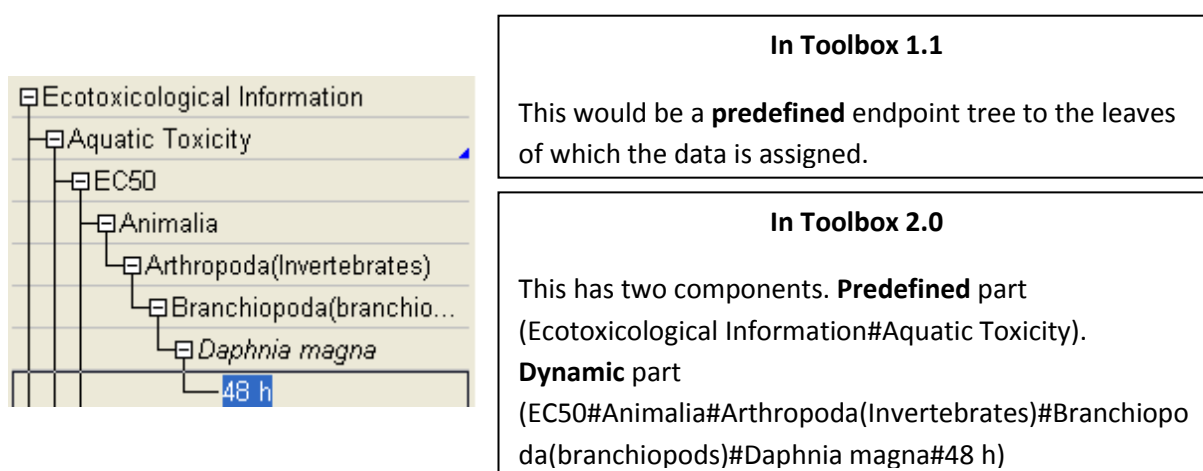


Figure 4: Endpoint tree in Toolbox version 1.1 vs. Endpoint tree in QSAR Toolbox version 2.0

In Toolbox 1.1 the tree displayed in the data-matrix was predefined and data could be imported to any of its leaves. In the QSAR Toolbox 2.0 however the data-matrix displays not only the **predefined** part of the endpoint tree but also builds a dynamic part based on the metadata of the currently displayed data points and/or QSARs.

- ❗ To check which part is predefined and which part is dynamic press the Ctrl key and the predefined part of the tree will be underscored.

5 Building the dynamic tree

The dynamic tree is a feature of the QSAR Toolbox where the endpoint tree is expanded with nodes that organize the data point's metadata. In essence it is a way to visualize the data the user has gathered from the database. The data is assigned to a path (what we call **predefined** path). The **dynamic** part of the tree is a function of the data point's metadata. It is an instruction that the user has given to the QSAR Toolbox software requiring that the data point's metadata are connected with metadata fields in a specific order. The metadata fields and their hierarchy are called the **Set tree hierarchy** feature.

It is important to make the distinction between what is the data point's **endpoint tree path** and on what node the data point is displayed on the data-matrix. The first one is an immutable attribute of the data point, and the latter is an undefined path that is build at runtime based on the **endpoint tree path**, the loaded data point's metadata and the current settings of the **Set tree hierarchy** feature.

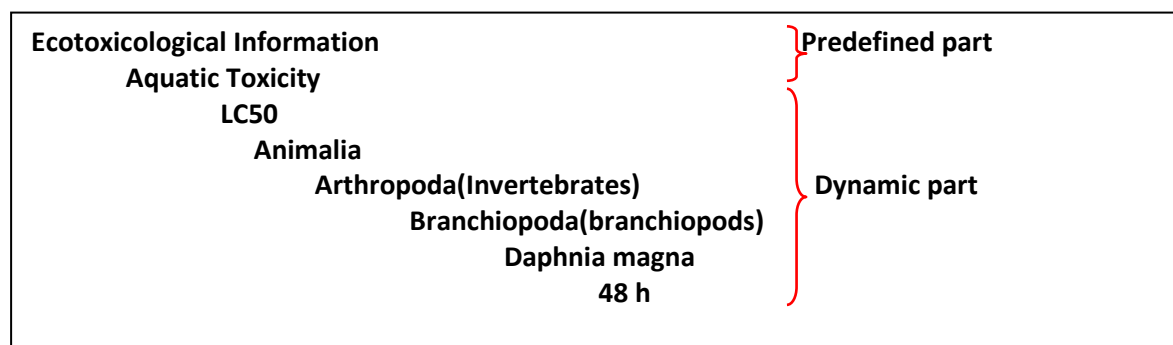
How does the above pertain to the import?

All the **Aquatic Toxicity** data in the QSAR Toolbox is assigned to the **Ecotoxicological Information#Aquatic Toxicity** path. However, when the user installs the QSAR Toolbox and loads data for aquatic toxicity, the entire tree path is shown, for example: **Ecotoxicological Information#Aquatic Toxicity#LC50#Animalia#Arthropoda(Invertebrates)#Branchiopoda(branchiopods)#Daphnia magna#48 h**

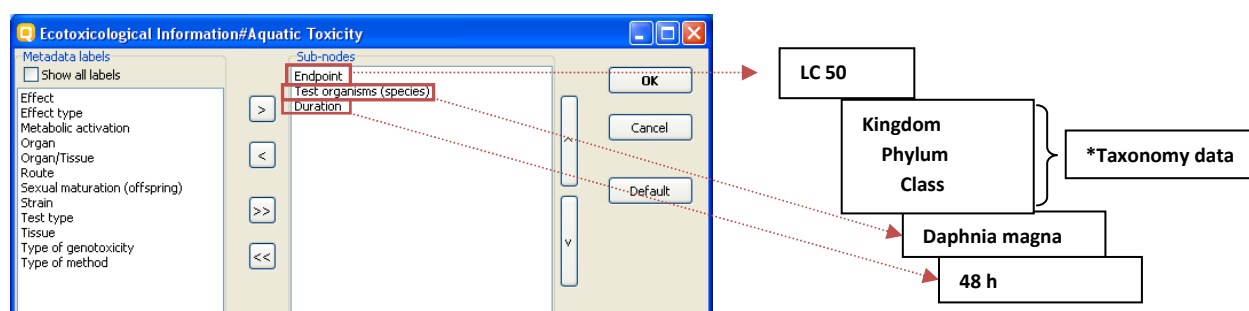
Where do the other fields come from?

The other fields come from the data point's metadata. The metadata fields build the dynamic part of the endpoint tree -

LC50#Animalia#Arthropoda(Invertebrates)#Branchiopoda(branchiopods)#Daphnia magna#48 h. The data point itself is associated to the shallow (predefined part) of the tree - **Ecotoxicological Information#Aquatic Toxicity**.



When a data point is read from the database and needs to be displayed to the data-matrix, the tree is expanded to display the metadata of the data points (Set tree hierarchy feature):



***Taxonomy data** – A large diversity of species has been stored and organized in Toolbox Taxonomy library including more than 12,295 biological species. Species have been distributed in five kingdoms: Animalia, Plantae, Fungi, Protozoa and Monera. Biological information is organized in the following taxa: Kingdom/Phylum/Class. Scientific information is associated automatically to each of the biological species.

Figure 5: Endpoint tree hierarchy

The QSAR Toolbox has default settings regarding which metadata is displayed. The default fields used are Endpoint, Duration, Test organisms (species), Effect, Effect type, Metabolic activation, Sexual maturation (offspring), Strain, Test type, Type of genotoxicity, Type of method, Tissue, Organ, Route. The list with default fields pertains to the whole endpoint tree. For **Ecotoxicological Information#Aquatic Toxicity** the default hierarchy is **Effect#Endpoint#Duration#Test organisms (species)**.

Table 1: Examples of metadata field values

Metadata field	Examples of metadata field values
Endpoint	LC 50, EC10, EC 50, LOEL, NOEL, Skin sensitisation, Carcinogenicity, Ames, Chromosomal aberration, Estrogen receptor binding....
Duration	years, months, days, hours, minutes, seconds...
Test organisms (species)	Daphnia magna, Lepomis symnetricus, Oncorhynchus mykiss, Poecilia reticulata, Tetrahymena pyriformis....
Effect	Immobilization, Mortality, Reproduction....
Effect type	Maternal toxicity, Developmental toxicity, Fetotoxicity, Embryotoxicity
Metabolic activation	with S9, without S9, no S9 info, with and without
Sexual maturation (offspring)	Male, Female, Male/Female...
Strain	TA 98, TA 100, TA 104, New Zealand White, Swiss, Fischer 344/DuCrj
Test type	bacterial reverse mutation assay (e.g. Ames test), in vitro mammalian cell micronucleus test, bacterial gene mutation assay, acute, subacute, chronic, developmental, static, semi-static, flow-through
Type of genotoxicity	Gene mutation, Chromosomal aberration, DNA damage and/or repair, genome mutation
Type of method	in vivo, in vitro, other
Organ	Lung, Liver
Route	oral, inhalation, dermal, implantation, intramuscular, intraperitoneal

This means that if a user wants to have a consistent display of the endpoint tree, he/she must import the metadata to the aforementioned fields before defining additional fields. For instance the above long path has two components – predefined and dynamic:

- Predefined: **Ecotoxicological Information#Aquatic Toxicity**
- Dynamic: **LC50#Animalia#Arthropoda(Invertebrates)#Branchiopoda(branchiopods)#Daphnia magna#48 h**

In the Import wizard these two translate to **Endpoint tree path** (Figure 5) and metadata (Figure 6). The

Animalia#Arthropoda(Invertebrates)#Branchiopoda(branchiopods) part is a separate feature in which Kingdom#Phylum#Class information is inserted before the field **Test organisms (species)**. In this respect the user should take extra caution to import species information to the **Test organisms (species)** metadata field.

Endpoint path

Endpoint tree path	ation#Aquatic Tox
Ecotoxicological Information#Aquatic Tox	
Ecotoxicological Information#Aquatic Tox	
Ecotoxicological Information#Aquatic Tox	
Ecotoxicological Information#Aquatic Tox	
Ecotoxicological Information#Aquatic Tox	
Ecotoxicological Information#Aquatic Tox	

Figure 5: Endpoint tree path

Endpoint	Duration	Mean value/Scale value	Duration	Unit	Test organisms (species)
LC50	48			h	Daphnia magna
LC50	48			h	Daphnia magna
LC50	48			h	Daphnia magna
LC50	48			h	Daphnia magna
LC50	48			h	Daphnia magna
LC50	48			h	Daphnia magna
LC50	48			h	Daphnia magna
LC50	48			h	Daphnia magna
LC50	48			h	Daphnia magna

Figure 6: Metadata

 Metadata information should always be provided for the fields **Test organisms (species)**, **Duration** and **Endpoint**.

6 Running the import wizard

The import wizard is organised in a three step process:

First step (Figure 7): The first window of the import wizard outlines the open file control [1], the file review pane [2], database name edit box[3], the used decimal and thousands separators and the import as inventory check box. It is very important that the thousands and decimal separators are properly set while importing. Especially with TXT file this could lead to erroneous parsing of results.

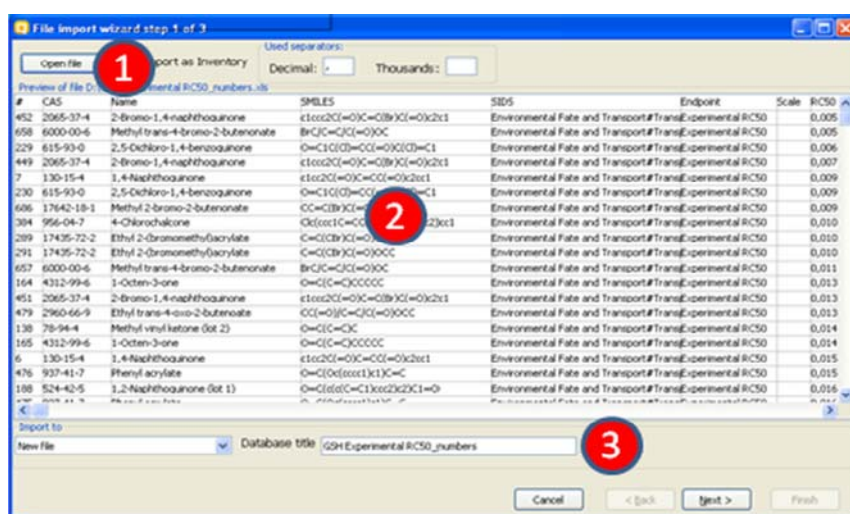


Figure 7: Import wizard step one

Second step: During the second step, the file's layout is selected which leads to two separate code paths, vertical or horizontal.

6.1 Vertical

File import wizard step 2 of 3

File layout:
☒ Vertical ☐ Horizontal ☒ I have a header row [What do different layouts mean?](#)

Define new region

CAS
Chemical name
SMILES
Undefined

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

CAS	Name	SMILES	RC50 value
CAS-37-4	Chemical name	thoquinone	SMILES
5000-00-6	Methyl trans-4-bromo-2-butenonate	<chem>BrC/C=C/C(=O)OC</chem>	0,0050
515-93-0	2,5-Dichloro-1,4-benzoquinone	<chem>O=C1C(Cl)=CC(=O)C(Cl)=C1</chem>	0,0053
2065-37-4	2-Bromo-1,4-naphthoquinone	<chem>O=C1C(Cl)=CC(=O)C(Cl)=C1</chem>	0,0063
130-15-4	1,4-Naphthoquinone	<chem>c1ccc2C(=O)C=C(C(=O)c2cc1</chem>	0,0070
515-93-0	2,5-Dichloro-1,4-benzoquinone	<chem>O=C1C(Cl)=CC(=O)C(Cl)=C1</chem>	0,0090
17642-18-1	Methyl 2-bromo-2-butenonate	<chem>CC=C(Br)C(=O)OC</chem>	0,0091
956-04-7	4-Chlorochalcone	<chem>Clc(ccc1C=CC(=O)c(ccc2)cc2)cc1</chem>	0,0097
17435-72-2	Ethyl 2-(bromomethyl)acrylate	<chem>C=C(CBr)C(=O)OCC</chem>	0,010
17435-72-2	Ethyl 2-(bromomethyl)acrylate	<chem>C=C(CBr)C(=O)OCC</chem>	0,010
5000-00-6	Methyl trans-4-bromo-2-butenonate	<chem>BrC/C=C/C(=O)OC</chem>	0,011

Cancel < Back Next > Finish

Figure 8: Import wizard step 2 (vertical import)

The second step window (Figure 8) contains the **CAS**, **Chemical name** and **SMILES** columns. Here is also a button that invokes the Scales definitions editor in case the user wants to import categorical data that has no available scale.



The type of the column is specified by clicking on the column and then selecting its type (CAS/Chemical name/SMILES) from the list box above.

To remove designations click a column and then click on **Undefined** from the list box.

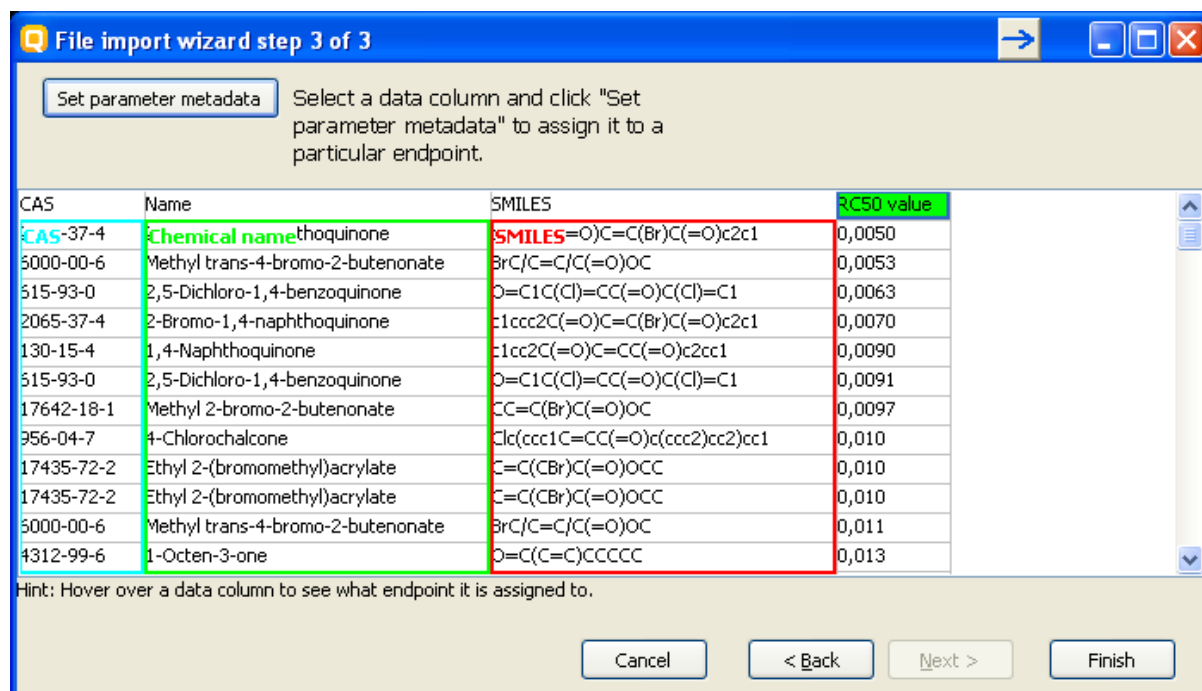


Figure 9: Import wizard step 3 (vertical import)

The third step of the vertical import (Figure 9) is where the user specifies the meaning of the different results columns. In the example above it is only one (RC50 value), but the import can handle multiple columns at once. To set the column metadata double-click on the column or click on the column and press **"Set parameter metadata"**. This will bring up the metadata editor (Figure 10).

☐ Scale data
Unit:

☐ Private data

Data tree position

- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
 - Aquatic Toxicity
 - Sediment toxicity
 - Terrestrial Toxicity
- Human health hazards

Metadata fields:

#1: Effect
Value type:
☒ Text
☐ Value

#2: Duration
Value type:
☐ Text
☒ Value
Mean value:
Lower value:
Upper value:
Unit:

#3: Test organisms (species)
Value type:
☒ Text
☐ Value
Abarenicola pacifica

Add field...

OK Cancel

Figure 10: Metadata editor

The most important part of the metadata setup is setting the “**Data tree position**” it should be a leaf from the displayed endpoint tree. If the column contains categorical data the user should check the “**Scale data**” checkbox and specify the Scale for the values in the column.

In the “**Metadata fields**” panel the user can enter a list of the metadata for the data point. There are two types of metadata “**Text**” and “**Value**”. The first is a

simple string while the latter is a **Mean/Lower value/Upper value** combination.



The Vertical import imports the data column without qualifiers and only to the **Mean** part of the data point record. If the user wants to enter qualified numbers or Mean/Low/Max combinations it is recommended to use the **Horizontal layout**.

After all metadata is set for all columns the user should press **Finish**. After the process is finished, an “**Import successful**” message will be displayed and the wizard will close.

6.2 Horizontal

The **Horizontal** layout is selected from a radio-button group in the second step of the import. There are basically two things the user can define, marked (1) and (2) in figure 11 below.

First, in the “**Define new region**” panel are the **CAS, NAME, SMILES, Endpoint tree path** and **Data** items. When they are defined the minimum is met.

File import wizard step 2 of 3

File layout:
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low qualifier
☐ Low value
☐ Upper qualifier
☒ Upper value
☒ Unit
☐ Undefined

Define metadata
☐ is value

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

1

2

CAS	Name	SMILES	Endpoint tree path	Data value	DIMENSIONS	Endpoint	Duration	Mean value/
50328	Chem	<chem>c1cc2c(c3c(c1)ccc3)ccc2</chem>	Environmental Fate and Transport#Bioac	2,69019608	og(L/kg wet)	BCF	48	
50328		<chem>c1cc2c(c3c(c1)ccc3)ccc2</chem>	Environmental Fate and Transport#Bioac	2,9253330558	og(L/kg wet)	BCF	48	
50328		<chem>c1cc2c(c3c(c1)ccc3)ccc2</chem>	Environmental Fate and Transport#Bioac	3,3636119799	og(L/kg wet)	BCF	48	
50328		<chem>c1cc2c(c3c(c1)ccc3)ccc2</chem>	Environmental Fate and Transport#Bioac	3,4243915544	og(L/kg wet)	BCF	48	

Cancel < Back Next > Finish

Figure 11: Import wizard step 2 (horizontal import)

The type of the column is specified by clicking on the column and then clicking its type (CAS/Chemical name/SMILES) from the list box in the **Define new region** panel or selecting a metadata field label from the list box in the **Metadata** panel. To remove designations click a column and then click on **Undefined** from the list box.

The definition of the **Data** region has its particularities. The data-record can contain categorical data - which would require two columns, one for the Scale, and one for the Value of the record. On the other hand the data-record can contain a value – which in the QSAR Toolbox is a packet of Mean/Min/Max value plus corresponding qualifiers and a Unit. These two are combined in the **Define value** panel, part of the **Define new region** panel.



When defining the data record the user should define:

- I. For categorical data – Scale** column for the scale name. This should contain a name of a scale **exactly** as defined in the scales list in the QSAR Toolbox database. The **Mean value/Scale value** column should contain a value that is **exactly** one of the scale's members
- II. For value data –** At least one column for the **Mean/Min/Max** values. Qualifiers and Unit are optional

Second is the "**Metadata**" panel. The data there is not mandatory for the import but can be used to import additional data to the data value. Defining numerical metadata is just like the data-record value definition process.

The third stage of the horizontal import is for review purposes. It is recommended that the user should look again at each column that is to be imported. The data record regions will be marked with color and text. The metadata fields are marked with a bold text over the first row for each column. The metadata of type value will be marked with <name of metadata>.<data_subtype> (for example "**Duration.Units**").

Where all columns are set and double-checked, the user should press **Finish**. After the process is finished, an "**Import successful**" message will be displayed and the wizard will close.

Appendix I: Preparing a file for horizontal import

File layout

A file prepared for import should follow the layout as shown in chapter 3.

Horizontal layout - this layout is used when the file is in the form where a row defines a single data point. Here the user specifies which column is the data, which column is metadata and the type of metadata.

	A	B	C	D	E	F	G	H
1	CAS	NAME	SMILES	Param descriptor 1	Param descriptor 2		Param descriptor N	VALUE
2	CAS# 1	NAME 1	SMILES 1	Param descriptor value [1,1]	Param descriptor value [2,1]		Param descriptor value [1,1]	Value 1
3	CAS# 2	NAME 2	SMILES 2	Param descriptor value [1,2]	Param descriptor value [2,2]		Param descriptor value [1,2]	Value 2
4								
5								
6	CAS# K	NAME K	SMILES K	Param descriptor value [1,K]	Param descriptor value [2,K]		Param descriptor value [2,K]	Value K

Figure 1. File layout for horizontal import

Figure 1 illustrates the format of an XLS file for horizontal import. Each row defines a record in its entirety. At import time the user specifies which columns has chemical identity information (CAS, Name, SMILES), which columns contain the values (what is seen in the Data-matrix and used in Data-gap filling) and which columns contain the metadata (Organ, Duration, Temperature, Dose, Species, Endpoint etc.).

Horizontal example – compact view

Ecotoxicological example

	Endpoint	Test organisms (species)	Duration	Duration Unit	Solubility	Unit	Effect	Lifestage	Water type	Salinity	Temperature	pH	Water alkalinity	Water Hardness	DATA QUALITY	Author	Year	Title	Reference source	Comm
1	EC 50	Ceriodaphnia	48	h			Immobils	Juvenile	Freshwater		23				2.2	WARNE	1999			
2	EC 50	Daphnia mag	1	h			Physiol	Juvenile	Freshwater		20				3	JANSSEN	1993			End point
3	EC 50	Daphnia mag	24	h			Mortality	Juvenile	Freshwater		20				3	JANSSEN	1993			
4	EC 50	Daphnia mag	48	h			Mortality	Juvenile	Freshwater		20				3	JANSSEN	1993			
5	EC 50	Vibrio fischeri	0.25	h			Behavior	Adult	Saltwater						2.2	Calleja	1994			
6	EC 50	Ampelisca al	24	h			Mortality		Saltwater						2.2	Calleja	1994			
7	EC 50	Scrobicularia	24	h			Mortality		Freshwater						2.2	Calleja	1994			
8	EC 50	Brachionus c	24	h			Mortality		Freshwater						2.2	Calleja	1994			
9	EC 50	Pimephales j	456	h			Reprodu	Adult	Freshwater		26	8.1	317		1	Kramer V	1998	Reprodu	Aquat Toxicol	4 End point
10	EC 50	Pimephales j	456	h			Reprodu	Adult	Freshwater		26	8.1	317		1	Kramer V	1998	Reprodu	Aquat Toxicol	4 End point
11	EC 50	Pimephales j	456	h			Mortality	Adult	Freshwater		26	8.1	317		1	Kramer V	1998	Reprodu	Aquat Toxicol	4 Male fish
12	EC 50	Hemicentrotus	2.25	h			Reprodu	Egg	Saltwater		15				3	WYNBEF	1989			Egg refe
13	EC 50	Penaeus van	48	h			Mortality	Juvenile	Saltwater	20	28.5				3	REYES	1996			
14	EC 50	Hyalella aztec	240	h			Mortality		Freshwater		22				1	Phipps G	1995	Relative	Arch Env Conta	
15	EC 50	Chironomus	240	h			Mortality		Freshwater		22				1	Phipps G	1995	Relative	Arch Env Conta	
16	EC 50	Vibrio fischeri	0.25	h			Behavior	Adult	Saltwater						2.2	Calleja	1994			
17	EC 50	Ampelisca al	24	h			Mortality		Saltwater						2.2	Calleja	1994			
18	EC 50	Scrobicularia	24	h			Mortality		Freshwater						2.2	Calleja	1994			
19	EC 50	Brachionus c	24	h			Mortality		Freshwater						2.2	Calleja	1994			
20	EC 50	Vibrio fischeri	0.25	h			Behavior	Adult	Saltwater						2.2	Calleja	1994			
21	EC 50	Ampelisca al	24	h			Mortality		Saltwater						2.2	Calleja	1994			
22	EC 50	Scrobicularia	24	h			Mortality		Freshwater						2.2	Calleja	1994			
23	EC 50	Brachionus c	24	h			Mortality		Freshwater						2.2	Calleja	1994			
24	EC 50	Vibrio fischeri	0.25	h			Behavior	Adult	Saltwater						2.2	Calleja	1994			
25	EC 50	Ampelisca al	24	h			Mortality		Saltwater						2.2	Calleja	1994			
26	EC 50	Scrobicularia	24	h			Mortality		Freshwater						2.2	Calleja	1994			
27	EC 50	Brachionus c	24	h			Mortality		Freshwater						2.2	Calleja	1994			
28	EC 50	Vibrio fischeri	0.25	h			Behavior	Adult	Saltwater						2.2	Calleja	1994			

Content of file for import:

Any file to be imported should contain:

- fields with structural information (chemical identity)
- a field with the Endpoint tree path
- fields with endpoint data information (experimental results and metadata)

 The endpoint tree path should point to a leaf of the **predefined** endpoint tree. For more information check chapter 4: **Endpoint tree path** of this document.

The screenshot shows a Microsoft Excel spreadsheet titled 'Ecotoxicological examples.xls [Compatibility Mode] - Microsoft Excel'. The spreadsheet is divided into several columns, with red boxes highlighting specific areas:

- Structural information:** Columns A, B, and C, containing CAS, NAME, and SMILES respectively.
- Endpoint tree path:** Column D, containing the endpoint tree path.
- Experimental Data:** Columns E, F, and G, containing Data Mean value, Scale, and Data Unit respectively.
- Metadata information:** Columns H through R, containing endpoint, test organisms, duration, duration unit, solubility, unit, effect, life stage, water type, salinity, and temperature.
- Endpoint data region:** A red box at the bottom highlights the region containing the endpoint data region.

CAS	NAME	SMILES	Endpoint tree path	Data Mean value	Scale	Data Unit	Endpoint	Test organisms (species)	Duration	Duration Unit	Solubility	Unit	Effect	Life stage	Water type	Salinity	Temperature
50-00-0	FORMAL C=O		Ecotoxicological	12.98	>	mg/L	EC 50	Ceriodaphnia	48	h			Immobilized Juvenile	Freshwater			23
50-00-0	FORMAL C=O		Ecotoxicological	39	>	mg/L	EC 50	Daphnia maj	1	h			Physiology Juvenile	Freshwater			20
50-00-0	FORMAL C=O		Ecotoxicological	57	>	mg/L	EC 50	Daphnia maj	24	h			Mortality Juvenile	Freshwater			20
50-00-0	FORMAL C=O		Ecotoxicological	29	>	mg/L	EC 50	Daphnia maj	48	h			Mortality Juvenile	Freshwater			20
50-06-6	PHEXOB C11=O/C		Ecotoxicological	2095.767	>	mg/L	EC 50	Vibrio fischer	0.25	h			Behavior Adult	Saltwater			
50-06-6	PHEXOB C11=O/C		Ecotoxicological	10009.113	<	mg/L	EC 50	Ampelica al	24	h			Mortality	Saltwater			
50-06-6	PHEXOB C11=O/C		Ecotoxicological	1212.5406	<	mg/L	EC 50	Scrobiculata	24	h			Mortality	Freshwater			
50-06-6	PHEXOB C11=O/C		Ecotoxicological	5178.729	<	mg/L	EC 50	Brachionus c	24	h			Mortality	Freshwater			
50-28-2	BETA-ES c12c(C)P		Ecotoxicological	0.00012	<	mg/L	EC 50	Pimephales	456	h			Reproduci Adult	Freshwater			26
50-28-2	BETA-ES c12c(C)P		Ecotoxicological	0.000251	<	mg/L	EC 50	Pimephales	456	h			Reproduci Adult	Freshwater			26
50-28-2	BETA-ES c12c(C)P		Ecotoxicological	0.00115	>	mg/L	EC 50	Pimephales	456	h			Mortality Adult	Freshwater			26
50-29-3	DOT C(C)C(C)C		Ecotoxicological	70	>	mg/L	EC 50	Hemicentrot	2.25	h			Reproduci Egg	Saltwater			15
50-29-3	DOT C(C)C(C)C		Ecotoxicological	0.0087	>	mg/L	EC 50	Penseus van	48	h			Mortality Juvenile	Saltwater	20		28.5
50-29-3	DOT, P.P C(C)C(C)C		Ecotoxicological	0.00007	>	mg/L	EC 50	Hyalella aspei	240	h			Mortality	Freshwater			22
50-29-3	DOT, P.P C(C)C(C)C		Ecotoxicological	0.00123	>	mg/L	EC 50	Chironomus	240	h			Mortality	Freshwater			22
50-48-6	AMITRYP c12c(C)C		Ecotoxicological	21.580942	>	mg/L	EC 50	Vibrio fischer	0.25	h			Behavior Adult	Saltwater			
50-48-6	AMITRYP c12c(C)C		Ecotoxicological	36.89287	>	mg/L	EC 50	Ampelica al	24	h			Mortality	Saltwater			
50-48-6	AMITRYP c12c(C)C		Ecotoxicological	0.776692	>	mg/L	EC 50	Scrobiculata	24	h			Mortality	Freshwater			
50-48-6	AMITRYP c12c(C)C		Ecotoxicological	0.804431	>	mg/L	EC 50	Brachionus c	24	h			Mortality	Freshwater			
50-54-4	QUINIDH c12c(C)C		Ecotoxicological	836.684	>	mg/L	EC 50	Vibrio fischer	0.25	h			Behavior Adult	Saltwater			
50-54-4	QUINIDH c12c(C)C		Ecotoxicological	273.3837	>	mg/L	EC 50	Ampelica al	24	h			Mortality	Saltwater			
50-54-4	QUINIDH c12c(C)C		Ecotoxicological	8.291145	>	mg/L	EC 50	Scrobiculata	24	h			Mortality	Freshwater			
50-54-4	QUINIDH c12c(C)C		Ecotoxicological	8.65462	>	mg/L	EC 50	Brachionus c	24	h			Mortality	Freshwater			
50-63-5	CHLORO c11NC(C)C		Ecotoxicological	866.712	>	mg/L	EC 50	Vibrio fischer	0.25	h			Behavior Adult	Saltwater			
50-63-5	CHLORO c11NC(C)C		Ecotoxicological	2042.964	>	mg/L	EC 50	Ampelica al	24	h			Mortality	Saltwater			
50-63-5	CHLORO c11NC(C)C		Ecotoxicological	11.65934	>	mg/L	EC 50	Scrobiculata	24	h			Mortality	Freshwater			
50-63-5	CHLORO c11NC(C)C		Ecotoxicological	4.38515	>	mg/L	EC 50	Brachionus c	24	h			Mortality	Freshwater			
50-78-2	ACETYLE C(=O)O		Ecotoxicological	26.12175	>	mg/L	EC 50	Vibrio fischer	0.25	h			Behavior Adult	Saltwater			
50-78-2	ACETYLE C(=O)O		Ecotoxicological	381.918	>	mg/L	EC 50	Ampelica al	24	h			Mortality	Saltwater			
50-78-2	ACETYLE C(=O)O		Ecotoxicological	177.9882	>	mg/L	EC 50	Scrobiculata	24	h			Mortality	Freshwater			
50-78-2	ACETYLE C(=O)O		Ecotoxicological	141.41775	>	mg/L	EC 50	Brachionus c	24	h			Mortality	Freshwater			
51-28-5	DINITROI c1(O)C(N)C		Ecotoxicological	0.5	>	mg/L	NOEC	Oncorhynchus	720	h	8710 mg/L		Growth	Early life stage	Freshwater		12
51-28-5	DINITROI c1(O)C(N)C		Ecotoxicological	0.8	>	mg/L	NOEC	Oncorhynchus	720	h	8710 mg/L		Growth	Early life stage	Freshwater		17
51-28-5	DINITROI c1(O)C(N)C		Ecotoxicological	0.9	>	mg/L	NOEC	Cyprinodon v	672	h	8710 mg/L		Growth	Early life stage	Saltwater	15	22
51-28-5	DINITROI c1(O)C(N)C		Ecotoxicological	1.1	>	mg/L	NOEC	Oncorhynchus	1440	h	8710 mg/L		Growth	Early life stage	Freshwater		7
51-28-5	DINITROI c1(O)C(N)C		Ecotoxicological	1.3	>	mg/L	NOEC	Oncorhynchus	1440	h	8710 mg/L		Mortality	Early life stage	Freshwater		7
51-28-5	DINITROI c1(O)C(N)C		Ecotoxicological	1.6	>	mg/L	NOEC	Oncorhynchus	720	h	8710 mg/L		Mortality	Early life stage	Freshwater		17
51-28-5	DINITROI c1(O)C(N)C		Ecotoxicological	1.8	>	mg/L	NOEC	Oncorhynchus	720	h	8710 mg/L		Growth	Embryo	Freshwater		17
51-28-5	DINITROI c1(O)C(N)C		Ecotoxicological	1.9	>	mg/L	NOEC	Oncorhynchus	720	h	8710 mg/L		Mortality	Early life stage	Freshwater		12
51-28-5	DINITROI c1(O)C(N)C		Ecotoxicological	1.9	>	mg/L	NOEC	Cyprinodon v	672	h	8710 mg/L		Growth	Early life stage	Saltwater	15	22
51-28-5	DINITROI c1(O)C(N)C		Ecotoxicological	1.9	>	mg/L	NOEC	Cyprinodon v	672	h	8710 mg/L		Growth	Early life stage	Saltwater	25	22

Structural information (chemical identity)

1. CAS field – Column with CAS Registry numbers
2. Name field – Column with Chemical Name(s)
3. SMILES field – Column with SMILES string

1 CAS	2 NAME	3 Smiles
50-00-0	FORMALDEHYDE	C=O
50-00-0	FORMALDEHYDE	C=O
50-00-0	FORMALDEHYDE	C=O
50-00-0	FORMALDEHYDE	C=O
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C
50-28-2	BETA-ESTRADIOL, 1	c12c(C{P+}3C{P+})(C{P-}4C
50-28-2	BETA-ESTRADIOL, 1	c12c(C{P+}3C{P+})(C{P-}4C
50-28-2	BETA-ESTRADIOL, 1	c12c(C{P+}3C{P+})(C{P-}4C
50-29-3	DDT	C(Cl)(Cl)(Cl)C(c1ccc(Cl)cc
50-29-3	DDT	C(Cl)(Cl)(Cl)C(c1ccc(Cl)cc
50-29-3	DDT, P, P'-	C(Cl)(Cl)(Cl)C(c1ccc(Cl)cc

- Field with Endpoint tree path (*Predefined path*)

Endpoint tree path
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity
Ecotoxicological Information#Aquatic Toxicity

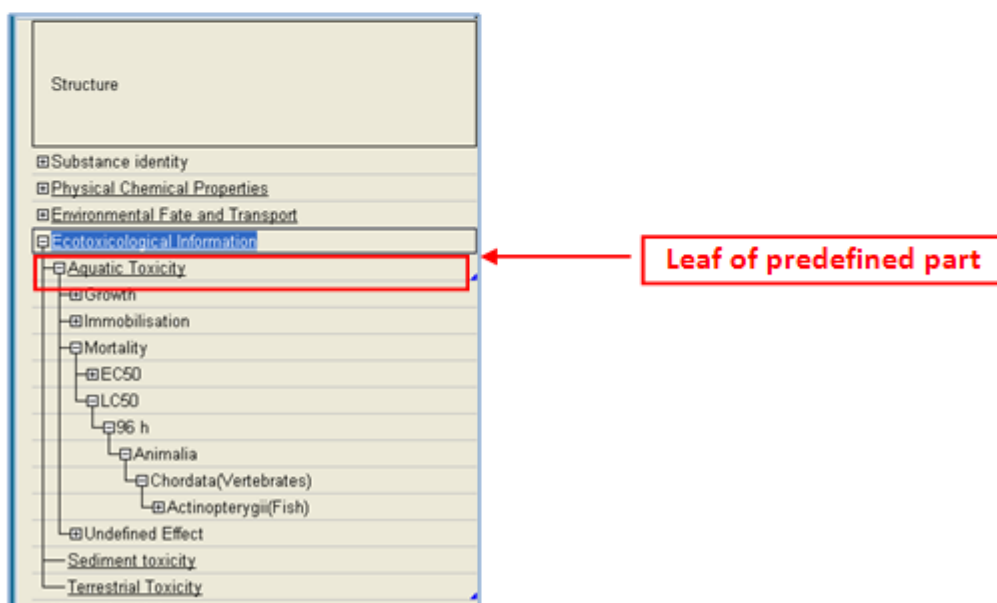
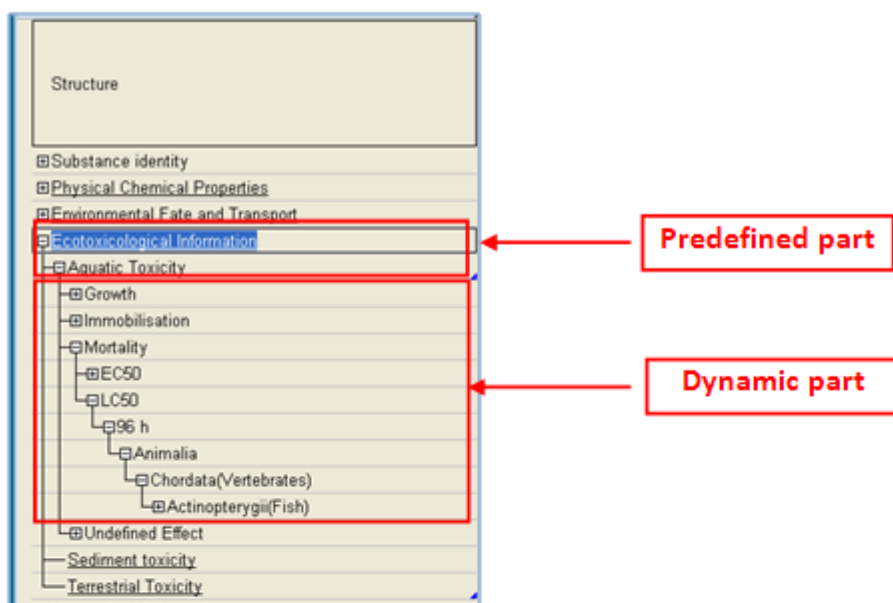
Important: The endpoint tree path should point to a leaf of the **predefined** endpoint tree. For more information check chapter 4 **Endpoint tree path** of this document.

Endpoint tree path

This is a field containing the endpoint tree path related to the endpoint for which experimental data is to be imported. The endpoint tree is separated in a predefined and a dynamic part:

- The predefined part determines the basic structure of the endpoint tree, e.g. substance identity, physical chemical properties, environmental fate and transport, ecotoxicological information, human health hazard. (*ordinary endpoint data is related to this predefined part*)
- The dynamic part of the endpoint tree builds additional layers of the tree depending on the data implemented in the databases. (*ordinary metadata is related to dynamic part*)

Endpoint data is assigned to a predefined part of the endpoint tree, i.e. it is assigned to a **leaf** of the endpoint tree:



- ! The predefined part can be visualized by holding the Ctrl button and clicking on the Endpoint tree. The predefined part of the tree is underlined.

Endpoint data information

The QSAR Toolbox operates with the following data structure:

Data point record

Value*

Endpoint (string)
 Endpoint description(string)
 Duration (Value)
 Is Private (Boolean)
 Is Observed (Boolean)

Link to chemical ID(CAS, SMILES)

Metadata(type String)

Title	String value
Title 1	Value 1
:	:
Title N	Value N

Metadata (type Value)

Title	Value
Title 1	Value 1
:	:
Title N	Value N

***Value** is defined as

Mean Qualifier(<, >, >=, etc.)	Mean Value (floating point number)	Low Qualifier(<, >, >=, etc.)	Low Value (floating point number)	Upper Qualifier	Upper Value (floating point number)	Unit
--------------------------------	------------------------------------	-------------------------------	-----------------------------------	-----------------	-------------------------------------	------

The user should therefore organize his/her data as follows:

Mean Qualifier(<, >, >=, etc.)	Mean Value (floating point number)	Low Qualifier(<, >, >=, etc.)	Low Value (floating point number)	Upper Qualifier	Upper Value (floating point number)	Unit
--------------------------------	------------------------------------	-------------------------------	-----------------------------------	-----------------	-------------------------------------	------

Data Mean value/Scale value	Data Mean qualifier	Data Unit
12.98	>	mg/L
39	>	mg/L
57	>	mg/L
29	>	mg/L
2995.767	>	mg/L
10009.113	<	mg/L
1212.2406	<	mg/L
5178.729	<	mg/L
0.00012	<	mg/L
0.000251	>	mg/L
0.00115	>	mg/L
70	>	mg/L
0.0087	>	mg/L
0.00007	>	mg/L

When the user defines data of a **Categorical type** (e.g. positive/negative) s/he should follow a scale definition:

1. The file for import should have a **Scale** column for the scale name. This should contain a name of a scale **exactly** as defined in the scales list in the QSAR Toolbox database for example "Gene mutation I" is a scale from the QSAR Toolbox database scale list (*red colored fields*):

Scales definition...

Scales list | Scale conversions

Scale list

New

Default scale

Name: Gene mutation I

Type: Ordinal

Families:

Mutagenicity

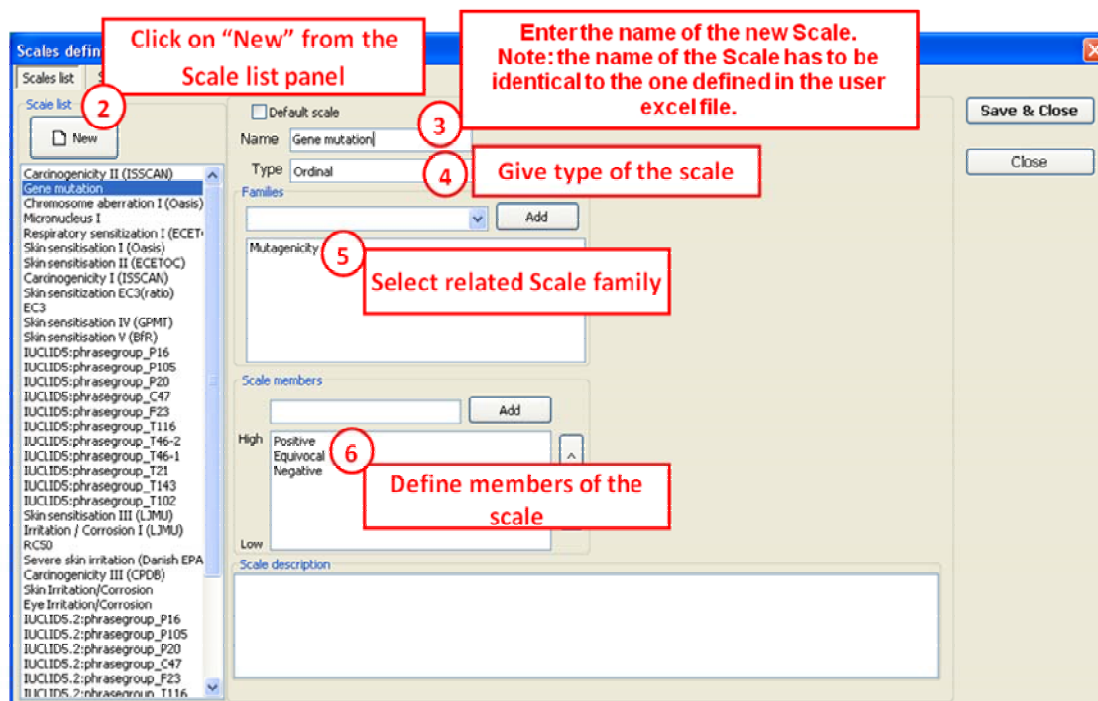
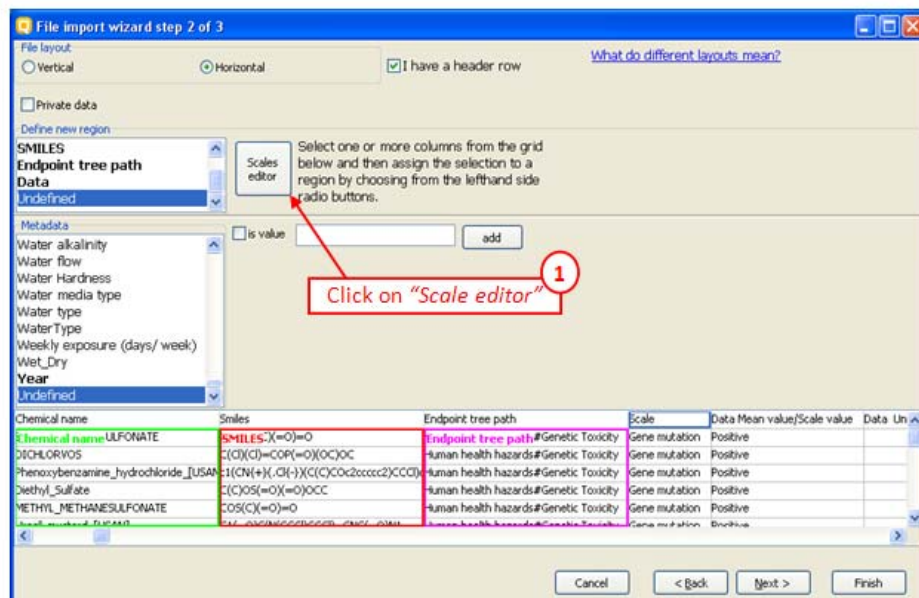
Scale members:

High: Positive, Equivocal, Negative

	A	B	C	D	E	F
	CAS	Chemical name	Smiles	Endpoint tree path	Scale	Data Mean value/Scale value
1						
2	62500	ETHYL_METHA	C(C)OS(C)(=O)	Human health hazard	Gene mutation I	Positive
3	62737	DICHLORVOS	C(Cl)(Cl)=COP(=O)	Human health hazard	Gene mutation I	Positive
4	63923	Phenoxybenzai	c1(CN(=O)C(=O)c2ccccc2	Human health hazard	Gene mutation I	Positive
5	64875	Diethyl_Sulfate	C(C)OS(=O)(=O)C	Human health hazard	Gene mutation I	Positive
6	66273	METHYL_METH	COS(C)(=O)=O	Human health hazard	Gene mutation I	Positive
7	66751	Uracil_mustard	C1(=O)C(N)CC(=O)N	Human health hazard	Gene mutation I	Positive
8	67209	1-[(S-NITROFUR	C1(C=NN2C(=O)N2C(=O)N	Human health hazard	Gene mutation I	Equivocal
9	70257	Methylnitronitro	C(=N)N(C)N(=O)=O	Human health hazard	Gene mutation I	Positive
10	70348	1-FLUORO-2,4-	c1(F)c(N(=O)=O)c2ccccc2	Human health hazard	Gene mutation I	Positive
11	74964	ETHYL_BROME	C(C)Br	Human health hazard	Gene mutation I	Positive
12	75252	TRIBROMOMETI	C(Br)(Br)Br	Human health hazard	Gene mutation I	Positive
13	75263	2-Bromopropan	C(C)C(Br)	Human health hazard	Gene mutation I	Positive
14	75478	TRIDODIMETHA	C(O)(O)O	Human health hazard	Gene mutation I	Positive
15	75558	PROPYLENIMINI	C1(C)CN1	Human health hazard	Gene mutation I	Negative
16	75912	TERT-BUTYL_H	C(C)(C)C(C)OO	Human health hazard	Gene mutation I	Positive
17	78342	DIOXATHION	C1(S(=S)(OCC)CCN1	Human health hazard	Gene mutation I	Positive
18	78875	1,2-DICHLOROF	C(C)Cl(Cl)	Human health hazard	Gene mutation I	Positive

2. The **Mean value/Scale value** column should contain only values that scale members (*blue colored fields*)

3. A new scale could be defined during the process of import:



⚠ When defining the data record the user should define:

I. For categorical data – Scale column for the scale name. This should contain a name of a scale **exactly** as defined in the scales list in the QSAR Toolbox database or as defined by the user. The **Mean value/Scale value** column should contain only values

that scale members

II. For value data – At least one column for the **Mean/Min/Max** values. Qualifiers and Unit are optional

Metadata information

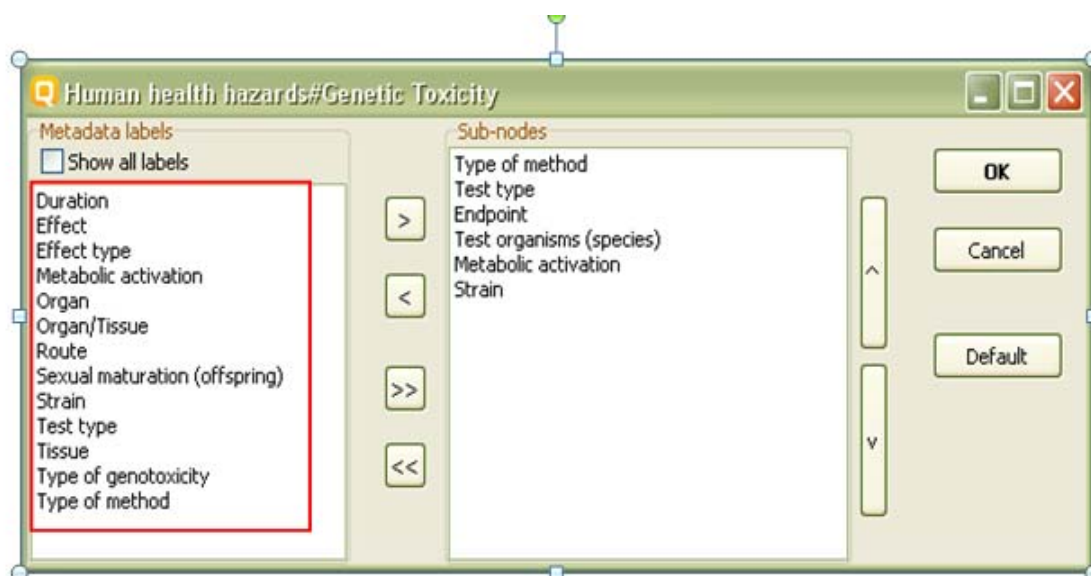
These are fields for metadata such as Duration, Organ, Tissue, Route of administration etc., which are not mandatory for the import:

*Metadata
information*

Endpoint	Test organisms (species)	Duration	Duration Unit	Solubility	Unit	Effect	Lifestage	Water type	Salinity	Temperature	pH	Water alkalinity	Water Hardness	DATA QUALITY
EC 50	Ceriodaphnia	48	h			Immobilis	Juvenile	Freshwater		23				2.2
EC 50	Daphnia mag	1	h			Physiolog	Juvenile	Freshwater		20				3
EC 50	Daphnia mag	24	h			Mortality	Juvenile	Freshwater		20				3
EC 50	Daphnia mag	48	h			Mortality	Juvenile	Freshwater		20				3
EC 50	Vibrio fischer	0.25	h			Behavior	Adult	Saltwater						2.2
EC 50	Ampelisca al	24	h			Mortality		Saltwater						2.2
EC 50	Scrobicularia	24	h			Mortality		Freshwater						2.2
EC 50	Brachionus c	24	h			Mortality		Freshwater						2.2
EC 50	Pimephales f	456	h			Reproduc	Adult	Freshwater		26	8.1	317		1
EC 50	Pimephales f	456	h			Reproduc	Adult	Freshwater		26	8.1	317		1
EC 50	Pimephales f	456	h			Mortality	Adult	Freshwater		26	8.1	317		1
EC 50	Hemicentrotu	2.25	h			Reproduc	Egg	Saltwater		15				3
EC 50	Penaeus van	48	h			Mortality	Juvenile	Saltwater	20	28.5				3
EC 50	Hyalella aztei	240	h			Mortality		Freshwater		22				1

Metadata fields can be used for building the dynamic part of the endpoint tree.

The feature is accessible through the “Set the hierarchy” option. The user sets what categories metadata he/she wants displayed on the endpoint tree and their hierarchy. Then, when the data are loaded the tree is build with the contents of the corresponding fields.



Metadata field	Examples of metadata field values
Endpoint	LC 50, EC10, EC 50, LOEL, NOEL, Skin sensitisation, Carcinogenicity, Ames, Chromosomal aberration, Estrogen receptor binding....
Duration	years, months, days, hours, minutes, seconds...
Test organisms (species)	Daphnia magna, Lepomis symnetricus, Oncorhynchus mykiss, Poecilia reticulata, Tetrahymena pyriformis....
Effect	Immobilization, Mortality, Reproduction....
Effect type	Maternal toxicity, Developmental toxicity, Fetotoxicity, Embryotoxicity
Metabolic activation	with S9, without S9, no S9 info, with and without
Sexual maturation (offspring)	Male, Female, Male/Female...
Strain	TA 98, TA 100, TA 104, New Zealand White, Swiss, Fischer 344/DuCrj
Test type	bacterial reverse mutation assay (e.g. Ames test), in vitro mammalian cell micronucleus test, bacterial gene mutation assay, acute, subacute, chronic, developmental, static, semi-static, flow-through
Type of genotoxicity	Gene mutaion, Chromosomal aberration, DNA damage and/or repair, genome mutation
Type of method	in vivo, in vitro, other
Organ	Lung, Liver
Route	oral, inhalation, dermal, implantation, intramuscular,

Metadata field	Examples of metadata field values
	intraperitoneal

 When defining the metadata record the user should organize his/her own metadata by using existing field labels. This ensures that the same fields do not multiply under different names. For example effects from the experiment should be placed in a column named "Effect", the type of the test method such as "Ames test" should be placed in a column named "Test type" etc.

The metadata fields can be organized in "Set tree hierarchy" panel. Right click on endpoint tree to set the hierarchy.

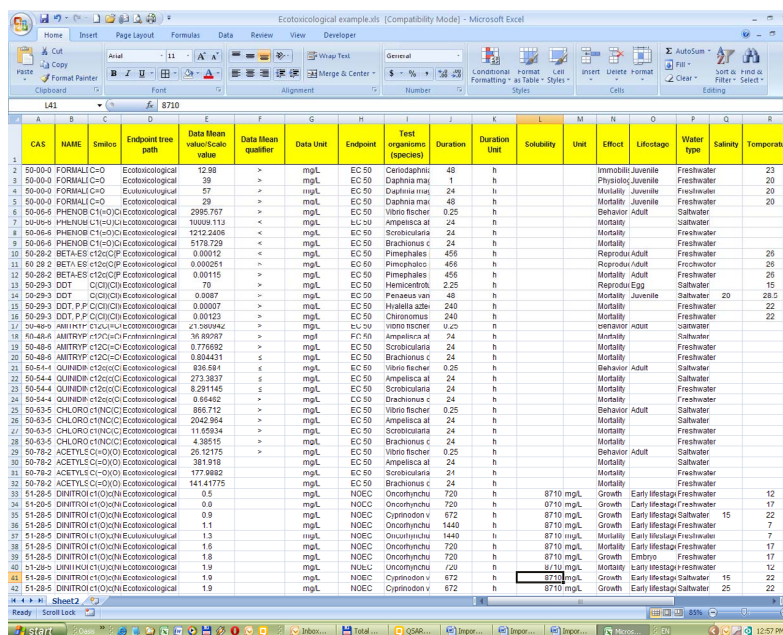
Appendix II: Import example for a database with ecotoxicological information

 The example below uses a file that is already prepared. Guidance on how to prepare file for horizontal import can be found in Appendix I.

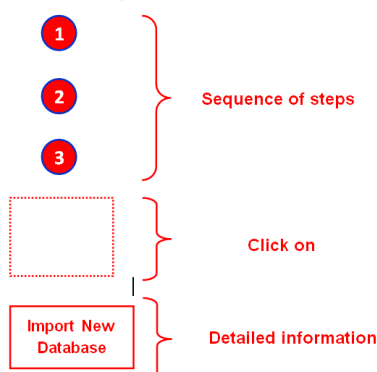
The destination for example files is [Install folder]\Examples

The default path is: C:\Program Files\QSAR Toolbox\QSAR Toolbox

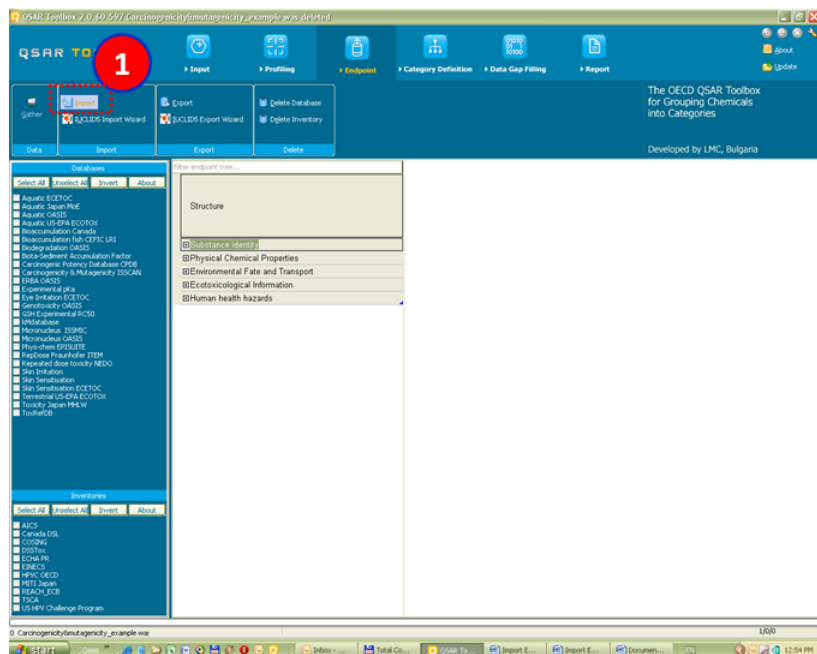
2.1\Examples\Ecotoxicological example.xls



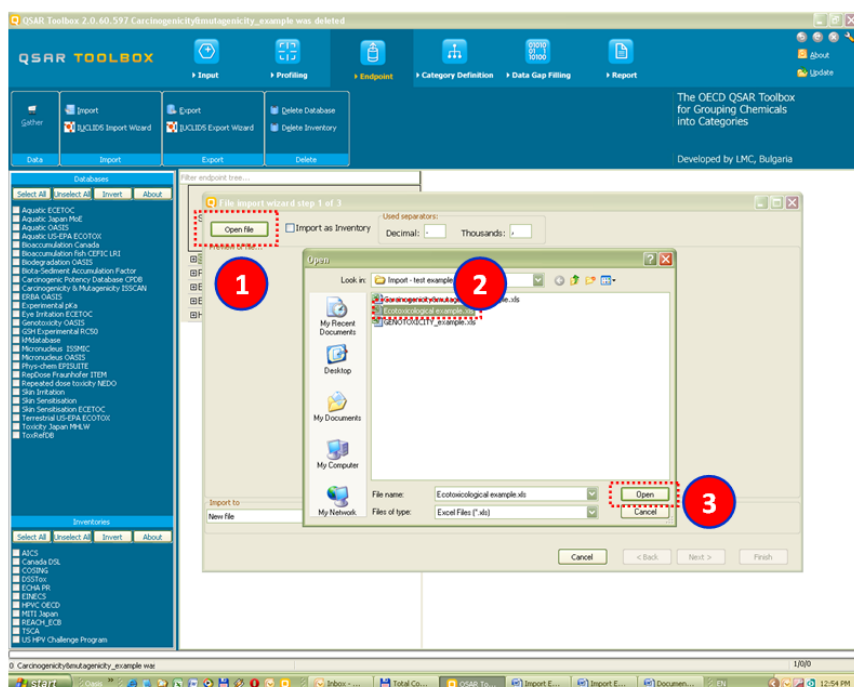
Legend:



Import New Database



Open example file



File is opened

File import wizard step 1 of 3

Open file ☐ Import as Inventory

Used separators:
Decimal: Thousands:

Preview of file D:\Transfer\Import - test example\Ecotoxicological example.xls

CAS	NAME	Smiles	Endpoint tree path	Data Mean value/Scale value	Data Mean qu
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Tox12.98	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Tox139	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Tox157	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Tox129	>	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox12995.767	>	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox10009.113	<	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox1212.2406	<	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox15178.729	<	
50-28-2	BETA-ESTRADIOL, 17	c12c(C{P+}3C{P+})(C{P-}4C{P-})(C{P-}Ecotoxicological Information#Aquatic Tox10.00012	<		
50-28-2	BETA-ESTRADIOL, 17	c12c(C{P+}3C{P+})(C{P-}4C{P-})(C{P-}Ecotoxicological Information#Aquatic Tox10.000251	>		
50-28-2	BETA-ESTRADIOL, 17	c12c(C{P+}3C{P+})(C{P-}4C{P-})(C{P-}Ecotoxicological Information#Aquatic Tox10.00115	>		
50-29-3	DDT	C(Cl)(Cl)C(c1ccc(Cl)cc1)c1ccc(Cl)cc1	Ecotoxicological Information#Aquatic Tox170	>	
50-29-3	DDT	C(Cl)(Cl)(Cl)C(c1ccc(Cl)cc1)c1ccc(Cl)cc1	Ecotoxicological Information#Aquatic Tox10.0087	>	
50-29-3	DDT, P,P'	C(Cl)(Cl)(Cl)C(c1ccc(Cl)cc1)c1ccc(Cl)cc1	Ecotoxicological Information#Aquatic Tox17e-005	>	
50-29-3	DDT, P,P'	C(Cl)(Cl)(Cl)C(c1ccc(Cl)cc1)c1ccc(Cl)cc1	Ecotoxicological Information#Aquatic Tox10.00123	>	
50-48-6	AMITRYPTYLIN	c12C(=CCCN(C)C)c3c(cccc3)CCc1ccc2	Ecotoxicological Information#Aquatic Tox121.580942	>	
50-48-6	AMITRYPTYLIN	c12C(=CCCN(C)C)c3c(cccc3)CCc1ccc2	Ecotoxicological Information#Aquatic Tox136.89287	>	
50-48-6	AMITRYPTYLIN	c12C(=CCCN(C)C)c3c(cccc3)CCc1ccc2	Ecotoxicological Information#Aquatic Tox10.776692	>	
50-48-6	AMITRYPTYLIN	c12C(=CCCN(C)C)c3c(cccc3)CCc1ccc2	Ecotoxicological Information#Aquatic Tox10.804431	<	

Import to
New File ☐ Database title Ecotoxicological example

Cancel < Back **Next >** Finish

Import to: The user has to choose whether to import the file as a new database or whether to add the data to a database already existing in the Toolbox.

Database title : Ecotoxicological example.
Note: the user could change the name

⚠ It is very important that the thousands and decimal separators are properly set while importing. Especially with TXT file this could lead to erroneous parsing of data values.

Step 2: Horizontal import

File import wizard step 2 of 3

File layout:
☒ Vertical
☐ Horizontal **1**

☒ I have a header row [What do different layouts mean?](#)

Define new region

CAS
 Chemical name
 SMILES
 Undefined

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

CAS	NAME	Smiles	Endpoint tree path	Data Mean value/Scale value	Data Mean qua
CAS-0	FORMALDEHYDE	SMILES	Ecotoxicological Information#Aquatic Toxi12.98	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi39	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi57	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi29	>	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Toxi2995.767	>	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Toxi10009.113	<	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Toxi1212.2406	<	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Toxi5178.729	<	
50-28-2	BETA-ESTRADIOL, 17	-12c(C{P+}3C{P+}(C{P-}4C{P-}(C{C{P-	Ecotoxicological Information#Aquatic Toxi0.00012	<	
50-28-2	BETA-ESTRADIOL, 17	-12c(C{P+}3C{P+}(C{P-}4C{P-}(C{C{P-	Ecotoxicological Information#Aquatic Toxi0.000251	>	
50-28-2	BETA-ESTRADIOL, 17	-12c(C{P+}3C{P+}(C{P-}4C{P-}(C{C{P-	Ecotoxicological Information#Aquatic Toxi0.00115	>	
50-29-3	DDT	C(Cl)(Cl)(Cl)C(c1ccc(Cl)cc1)c1ccc(Cl)cc1	Ecotoxicological Information#Aquatic Toxi70	>	
50-29-3	DDT	C(Cl)(Cl)(Cl)C(c1ccc(Cl)cc1)c1ccc(Cl)cc1	Ecotoxicological Information#Aquatic Toxi0.0087	>	
50-29-3	DDT, P,P'-	C(Cl)(Cl)(Cl)C(c1ccc(Cl)cc1)c1ccc(Cl)cc1	Ecotoxicological Information#Aquatic Toxi7e-005	>	
50-29-3	DDT, P,P'-	C(Cl)(Cl)(Cl)C(c1ccc(Cl)cc1)c1ccc(Cl)cc1	Ecotoxicological Information#Aquatic Toxi0.00123	>	
50-48-6	AMITRYPTYLIN	-12C(=CCCN(C)C)C3c(cccc3)CCc1cccc2	Ecotoxicological Information#Aquatic Toxi21.580942	>	
50-48-6	AMITRYPTYLIN	-12C(=CCCN(C)C)C3c(cccc3)CCc1cccc2	Ecotoxicological Information#Aquatic Toxi36.89287	>	
50-48-6	AMITRYPTYLIN	-12C(=CCCN(C)C)C3c(cccc3)CCc1cccc2	Ecotoxicological Information#Aquatic Toxi0.776692	>	

2

Cancel < Back **Next >** Finish

Select: CAS

File import wizard step 2 of 3

File layout:
☐ Vertical
☒ Horizontal

☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

CAS **2**
 Chemical name
 SMILES
 Endpoint tree path

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata

☐ is value add

Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Undefined

CAS	NAME	Smiles	Endpoint tree path	Data Mean value/Scale value	Data Mean qua
CAS-0	FORMALDEHYDE	SMILES	Ecotoxicological Information#Aquatic Toxi12.98	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi39	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi57	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi29	>	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Toxi2995.767	>	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Toxi10009.113	<	

1

Cancel < Back **Next >** Finish

Select: *Chemical name*

File import wizard step 2 of 3

File layout:
☐ Vertical
☒ Horizontal
☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region:
 CAS
 Chemical name **2**
 SMILES
 Endpoint tree path

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
☐ is value

CAS	NAME	Smiles	Endpoint tree path	Data Mean value/Scale value	Data Mean qua
CAS-0	Chemical name	SMILES	Ecotoxicological Information#Aquatic Toxi12.98	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi39	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi57	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi29	>	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Toxi2995.767	>	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Toxi10000.113	>	

Cancel < Back Next > Finish

Select: *SMILES*

File import wizard step 2 of 3

File layout:
☐ Vertical
☒ Horizontal
☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region:
 CAS
 Chemical name
 SMILES **2**
 Endpoint tree path
 Data
 Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
☐ is value

CAS	NAME	Smiles	Endpoint tree path	Data Mean value/Scale value	Data Mean qua
CAS-0	Chemical name	SMILES 1	Ecotoxicological Information#Aquatic Toxi12.98	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi39	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi57	>	
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Toxi29	>	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Toxi2995.767	>	
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Toxi10000.113	>	

Cancel < Back Next > Finish

Select : Endpoint tree path

File import wizard step 2 of 3

File layout:
☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region:
 SMILES
 Endpoint tree path (2)
 Data
 Undefined

Scales editor: Select one or more columns from the grid below and then assign the selection to a region by choosing from the left hand side radio buttons.

Metadata:
☐ is value [] add

CAS	NAME	Smiles	Endpoint tree path	Data Mean value/Scale value	Data Mean qua
50-00-0	FORMALDEHYDE	C=O	Endpoint tree path	12.98	>
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Tox	39	>
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Tox	57	>
50-00-0	FORMALDEHYDE	C=O	Ecotoxicological Information#Aquatic Tox	29	>
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox	2995.767	>
50-06-6	PHENOBARBITAL	C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox	112	>

Cancel < Back Next > Finish

Important: The endpoint tree path should point to a leaf of the predefined endpoint tree. For more information check chapter 4 **Endpoint tree path** in this document.

Select: DATA Mean Value/Scale value

File import wizard step 2 of 3

File layout:
☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region:
 SMILES
 Endpoint tree path (2)
 Data
 Undefined

Define value:
☐ Scale ☐ Low qualifier ☐ Upper value
☐ Mean qualifier ☐ Low value ☐ Unit
☒ Mean value/Scale value ☐ Upper qualifier ☐ Undefined

Scales editor: Select one or more columns from the grid below and then assign the selection to a region by choosing from the left hand side radio buttons.

Metadata:
☐ is value [] add

Smiles	Endpoint tree path	Data Mean value/Scale value	Data Mean qualifier	Data Unit	Endpoint Test organisms
SMILES	Endpoint tree path	Data Mean value/Scale value			
C=O	Ecotoxicological Information#Aquatic Tox	39		mg/L	EC 50 Ceriodaphnia affinis
C=O	Ecotoxicological Information#Aquatic Tox	57		mg/L	EC 50 Daphnia magna
C=O	Ecotoxicological Information#Aquatic Tox	29		mg/L	EC 50 Daphnia magna
C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox	2995.767		mg/L	EC 50 Vibrio fischeri
C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox	112		mm/L	EC 50 Annelina shubla

Cancel < Back Next > Finish

Select: DATA Mean qualifier

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale
☒ Mean qualifier
☐ Mean value/Scale value
☐ Low qualifier
☐ Low value
☐ Upper value
☐ Unit
☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
☐ is value add

Smiles	Endpoint tree path	Data Mean value/Scale value	Data Mean qualifier	Data Unit	Endpoint	Test organisms
SMILES	Endpoint tree path	Data Mean value/Scale value	Data Mean qualifier	Data Unit	EC 50	Ceriodaphnia affinis
C=O	Ecotoxicological Information#Aquatic Tox	39		mg/L	EC 50	Daphnia magna
C=O	Ecotoxicological Information#Aquatic Tox	57		mg/L	EC 50	Daphnia magna
C=O	Ecotoxicological Information#Aquatic Tox	29		mg/L	EC 50	Daphnia magna
C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox	2995.767		mg/L	EC 50	Vibrio fischeri
C1=CC=C(C=C1)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox	10000.112		mg/L	EC 50	Amniontes abdita

Cancel < Back Next > Finish

Select: DATA Unit

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low qualifier
☐ Low value
☒ Upper value
☒ Unit
☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
☐ is value add

Smiles	Endpoint tree path	Data Mean value/Scale value	Data Mean qualifier	Data Unit	Endpoint	Test organisms
SMILES	Endpoint tree path	Data Mean value/Scale value	Data Mean qualifier	Data Unit	EC 50	Ceriodaphnia affinis
C=O	Ecotoxicological Information#Aquatic Tox	39		mg/L	EC 50	Daphnia magna
C=O	Ecotoxicological Information#Aquatic Tox	57		mg/L	EC 50	Daphnia magna
C=O	Ecotoxicological Information#Aquatic Tox	29		mg/L	EC 50	Daphnia magna
C1(=O)C(c2ccccc2)(CC)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox	2995.767		mg/L	EC 50	Vibrio fischeri
C1=CC=C(C=C1)C(=O)NC(=O)N1	Ecotoxicological Information#Aquatic Tox	10000.112		mg/L	EC 50	Amniontes abdita

Cancel < Back Next > Finish

-  I. The type of the column is specified by clicking on the column and then clicking its type (CAS/Chemical name/SMILES) from the list box in the **Define new region** panel or selecting a metadata field label from the list box in the **Metadata** panel. To remove designations click a column and then click on **Undefined** from the list box.
- II. All fields defined with the **Define new region** panel are color-distinguished from the fields of the **Metadata** panel.

Add test Duration data

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale
☐ Mean qualifier
☐ Mean
☐ Low qualifier
☐ Low value
☐ Upper value
☒ Unit
☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
 Duration
 Effect
 Effect additional
 Effect details
 Effect type
 Elimination constant - k2
 Elimination phase duration
 Elimination > 5xT50
 Embryotoxic / teratogenic effects
 Endpoint

is value ☐ add

Data	Mean qualifier	Data Unit	Endpoint	Test organisms (species)	Duration	Duration Unit	Solubility	Unit	Effect	Lifestage	Water type	Salinity	Temp
>	Data.Mean qualifier	Data.Unit	EC 50	Ceriodaphnia affinis	48	h			Immobilisation	Juvenile	Freshwater		23
>		mg/L	EC 50	Daphnia magna	1	h			Physiology	Juvenile	Freshwater		20
>		mg/L	EC 50	Daphnia magna	24	h			Mortality	Juvenile	Freshwater		20
>		mg/L	EC 50	Daphnia magna	48	h			Mortality	Juvenile	Freshwater		20
>		mg/L	EC 50	Vibrio fischeri	0.25	h			Behavior	Adult	Saltwater		
>		mg/L	EC 50	Amalica shulka	74	h			Mortality		Saltwater		

Cancel < Back Next > Finish

File import wizard step 2 of 3

File layout:
☐ Vertical
☒ Horizontal
☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region:

SMILES
 Endpoint tree path
 Data
 Undefined

Define value:
☐ Scale
☐ Low qualifier
☐ Upper value
☐ Mean qualifier
☐ Low value
☒ Unit
☐ Mean value/Scale value
☐ Upper qualifier
☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata:
 Duration
 Effect
 Effect additional
 Effect details
 Effect type
 Elimination constant - k2
 Elimination phase duration
 Elimination>5xT50
 Embryotoxic / teratogenic effects
 Endpoint

Define value:
☐ Scale
☐ Upper qualifier
☐ Mean qualifier
☐ Upper value
☒ Mean value/Scale value
☐ Low qualifier
☐ Undefined
☐ Low value

is value add

Data	Mean qualifier	Data Unit	Endpoint	Test organisms (species)	Duration	Mean value/Scale value	Duration Unit	Solubility	Unit	Effect	Lifestage	V
>	mg/L	EC 50	Ceriodaphnia affinis	48						Immobilisation	Juvenile	F
>	mg/L	EC 50	Daphnia magna	1						Physiology	Juvenile	F
>	mg/L	EC 50	Daphnia magna	24						Mortality	Juvenile	F
>	mg/L	EC 50	Daphnia magna	48						Mortality	Juvenile	F
>	mg/L	EC 50	Vibrio fischeri	0.25						Behavior	Adult	S
>	mg/L	EC 50	Amalicia shubia	24						Mortality		S

Cancel < Back Next > Finish

Add test "Duration unit"

File import wizard step 2 of 3

File layout:
☐ Vertical
☒ Horizontal
☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region:

SMILES
 Endpoint tree path
 Data
 Undefined

Define value:
☐ Scale
☐ Low qualifier
☐ Upper value
☐ Mean qualifier
☐ Low value
☒ Unit
☐ Mean value/Scale value
☐ Upper qualifier
☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata:
 Dry Weight Percent
 Duration
 Effect
 Effect additional
 Effect details
 Effect type
 Elimination constant - k2
 Elimination phase duration
 Elimination>5xT50
 Embryotoxic / teratogenic effects

Define value:
☐ Scale
☐ Upper qualifier
☐ Mean qualifier
☐ Upper value
☒ Mean value/Scale value
☐ Low qualifier
☐ Undefined
☐ Low value

is value add

Data	Mean qualifier	Data Unit	Endpoint	Test organisms (species)	Duration	Mean value/Scale value	Duration Unit	Solubility	Unit	Effect	Lifestage	V
>	mg/L	EC 50	Ceriodaphnia affinis	48						Immobilisation	Juvenile	F
>	mg/L	EC 50	Daphnia magna	1						Physiology	Juvenile	F
>	mg/L	EC 50	Daphnia magna	24						Mortality	Juvenile	F
>	mg/L	EC 50	Daphnia magna	48						Mortality	Juvenile	F
>	mg/L	EC 50	Vibrio fischeri	0.25						Behavior	Adult	S
>	mg/L	EC 50	Amalicia shubia	24						Mortality		S

Cancel < Back Next > Finish

Add undefined parameter - "Solubility"

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale
☐ Low qualifier
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low value
☐ Upper value
☒ Unit
☐ Upper qualifier
☐ Undefined

Metadata
 Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Undefined

Define value
☐ Scale
☐ Upper qualifier
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low qualifier
☒ is value
☐ Upper value
☐ Unit
☐ Undefined
☐ Low value

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Data	Mean qualifier	Data Unit	Endpoint	Test organisms (species)	Duration	Mean value/Scale value	Duration	Unit	Solubility	Unit	Effect	Lifestage	
Data	Mean qualifier	Data Unit	EC 50	Ceriodaphnia affinis	48			h			Immobilisation	Juvenile	F
>		mg/L	EC 50	Daphnia magna	1			h			Physiology	Juvenile	F
>		mg/L	EC 50	Daphnia magna	24			h			Mortality	Juvenile	F
>		mg/L	EC 50	Daphnia magna	48			h			Mortality	Juvenile	F
>		mg/L	EC 50	Vibrio fischeri	0.25			h			Behavior	Adult	F
>		mg/L	EC 50	Amelara shiba	24			h			Mortality		

Cancel < Back Next > Finish

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale
☐ Low qualifier
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low value
☐ Upper value
☒ Unit
☐ Upper qualifier
☐ Undefined

Metadata
 Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Undefined

Define value
☐ Scale
☐ Upper qualifier
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low qualifier
☒ is value
☐ Upper value
☐ Unit
☐ Undefined
☐ Low value

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

☐ is value Solubility add

Define "Solubility", click "add"

Data	Mean qualifier	Data Unit	Endpoint	Test organisms (species)	Duration	Mean value/Scale value	Duration	Unit	Solubility	Unit	Effect	Lifestage	
Data	Mean qualifier	Data Unit	EC 50	Ceriodaphnia affinis	48			h			Immobilisation	Juvenile	F
>		mg/L	EC 50	Daphnia magna	1			h			Physiology	Juvenile	F
>		mg/L	EC 50	Daphnia magna	24			h			Mortality	Juvenile	F
>		mg/L	EC 50	Daphnia magna	48			h			Mortality	Juvenile	F
>		mg/L	EC 50	Vibrio fischeri	0.25			h			Behavior	Adult	F
>		mg/L	EC 50	Amelara shiba	24			h			Mortality		

Cancel < Back Next > Finish

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale
☐ Low qualifier
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low value
☐ Upper value
☒ Unit
☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
 Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Solubility

☐ is value Solubility add

5

Check "is value"

Data	Mean qualifier	Data Unit	Endpoint	Test organisms (species)	Duration	Mean value/Scale value	Duration	Unit	Solubility	Unit	Effect	Lifestage	
Data	Mean qualifier	Data Unit	EC 50	Ceriodaphnia affinis	48			h			Immobilisation	Juvenile	F
>		ng/L	EC 50	Daphnia magna	1			h			Physiology	Juvenile	F
>		ng/L	EC 50	Daphnia magna	24			h			Mortality	Juvenile	F
>		ng/L	EC 50	Daphnia magna	48			h			Mortality	Juvenile	F
>		ng/L	EC 50	Vibrio fischeri	0.25			h			Behavior	Adult	F
>		ng/L	EC 50	Amalica shubia	24			h			Mortality	Adult	F

Cancel < Back Next > Finish

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale
☐ Low qualifier
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low value
☐ Upper value
☒ Unit
☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
 Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Solubility

☒ is value Solubility add

8

Select "Mean value/Scale value"

7

6

Data	Mean qualifier	Data Unit	Endpoint	Test organisms (species)	Duration	Mean value/Scale value	Duration	Unit	Solubility	Mean value/Scale value	Unit	Effect	
Data	Mean qualifier	Data Unit	EC 50	Ceriodaphnia affinis	48			h				Imm	
>		ng/L	EC 50	Daphnia magna	1			h				Phys	
>		ng/L	EC 50	Daphnia magna	24			h				Mort	
>		ng/L	EC 50	Daphnia magna	48			h				Mort	
>		ng/L	EC 50	Vibrio fischeri	0.25			h				Beha	
>		ng/L	EC 50	Amalica shubia	24			h				Mort	

Cancel < Back Next > Finish

Select "Unit" of defined "Solubility" parameter

File import wizard step 2 of 3

File layout
☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale ☐ Low qualifier ☐ Upper value
☐ Mean qualifier ☐ Low value ☒ Unit
☐ Mean value/Scale value ☐ Upper qualifier ☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Solubility
 Undefined

Define value
☐ Scale ☐ Upper qualifier ☒ is value Solubility
☐ Mean qualifier ☐ Upper value
☐ Mean value/Scale value ☒ Unit
☐ Low qualifier ☐ Undefined
☐ Low value

Endpoint	Test organisms (species)	Duration	Mean value/Scale value	Duration	Unit	Solubility	Mean value/Scale value	Solubility	Unit	Effect	Lifestage
EC 50	Ceriodaphnia affinis	48			h					Immobilisation	Juvenile
EC 50	Daphnia magna	1			h					Physiology	Juvenile
EC 50	Daphnia magna	24			h				Mortality	Juvenile	
EC 50	Daphnia magna	48			h				Mortality	Juvenile	
EC 50	Vibrio fischeri	0.25			h				Behavior	Adult	
EC 50	Amelara shubia	24			h				Mortality		

Cancel < Back Next > Finish

Scroll bar

File import wizard step 2 of 3

File layout
☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale ☐ Low qualifier ☐ Upper value
☐ Mean qualifier ☐ Low value ☒ Unit
☐ Mean value/Scale value ☐ Upper qualifier ☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Solubility
 Undefined

Define value
☐ Scale ☐ Upper qualifier ☒ is value Solubility
☐ Mean qualifier ☐ Upper value
☐ Mean value/Scale value ☒ Unit
☐ Low qualifier ☐ Undefined
☐ Low value

Effect	Lifestage	Water type	Salinity	Temperature	pH	Water alkalinity	Water Hardness	DATA QUALITY	Author	Year
Immobilisation	Juvenile	Freshwater		23				2.2	WARNE AND SCHIFKO	1999
Physiology	Juvenile	Freshwater		20				3	JANSSEN AND PERSOONE	1993
Mortality	Juvenile	Freshwater		20				3	JANSSEN AND PERSOONE	1993
Mortality	Juvenile	Freshwater		20				3	JANSSEN AND PERSOONE	1993
Behavior	Adult	Saltwater						2.2	Calleja	1994

Cancel < Back Next > Finish

Scroll the bar to review data column designations. Click on "Next" button

File import wizard step 2 of 3

File layout
☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale ☐ Low qualifier ☐ Upper value
☐ Mean qualifier ☐ Low value ☒ Unit
☐ Mean value/Scale value ☐ Upper qualifier ☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Solubility
 Undefined

Define value
☐ Scale ☐ Upper qualifier ☒ is value Solubility
☐ Mean qualifier ☐ Upper value
☐ Mean value/Scale ☒ Unit
☐ Low qualifier ☐ Undefined
☐ Low value

Author	Year	Title	Reference source	Comment
WARNE AND SCHIFKO	1999			
JANSSEN AND PERSOONE	1993			End point is fluorescence od Daphnia with
JANSSEN AND PERSOONE	1993			
JANSSEN AND PERSOONE	1993			
Calleja	1994			

Cancel < Back Next > **1** Finish

Click on "Finish" button

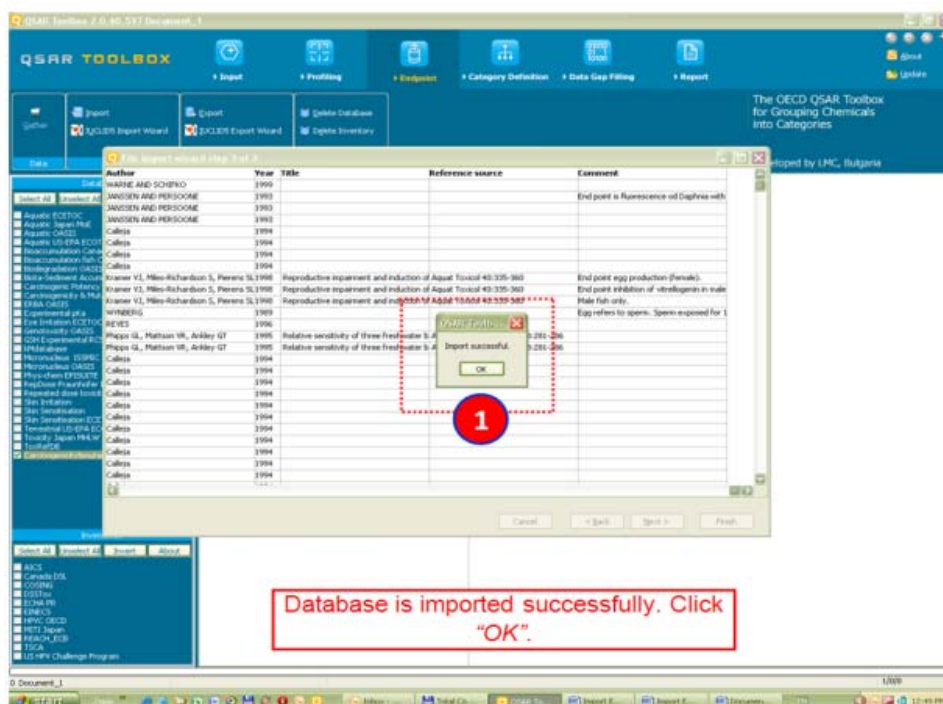
File import wizard step 3 of 3

Author	Year	Title	Reference source	Comment
WARNE AND SCHIFKO	1999			
JANSSEN AND PERSOONE	1993			End point is fluorescence od Daphnia with
JANSSEN AND PERSOONE	1993			
JANSSEN AND PERSOONE	1993			
Calleja	1994			
Calleja	1994			
Calleja	1994			
Calleja	1994			
Kramer VJ, Miles-Richardson S, Pierens SL	1998	Reproductive impairment and induction of Aquat Toxicol 40:335-360		End point egg production (female).
Kramer VJ, Miles-Richardson S, Pierens SL	1998	Reproductive impairment and induction of Aquat Toxicol 40:335-360		End point inhibition of vitellogenin in male
Kramer VJ, Miles-Richardson S, Pierens SL	1998	Reproductive impairment and induction of Aquat Toxicol 40:335-360		Male fish only.
WYNBERG	1989			Egg refers to sperm. Sperm exposed for 1
REYES	1996			
Phipps GL, Mattson VR, Ankley GT	1995	Relative sensitivity of three freshwater b	Arch Env Contam Toxicol 28:281-286	
Phipps GL, Mattson VR, Ankley GT	1995	Relative sensitivity of three freshwater b	Arch Env Contam Toxicol 28:281-286	
Calleja	1994			
Calleja	1994			
Calleja	1994			
Calleja	1994			
Calleja	1994			
Calleja	1994			
Calleja	1994			
Calleja	1994			
Calleja	1994			
Calleja	1994			
Calleja	1994			
Calleja	1994			

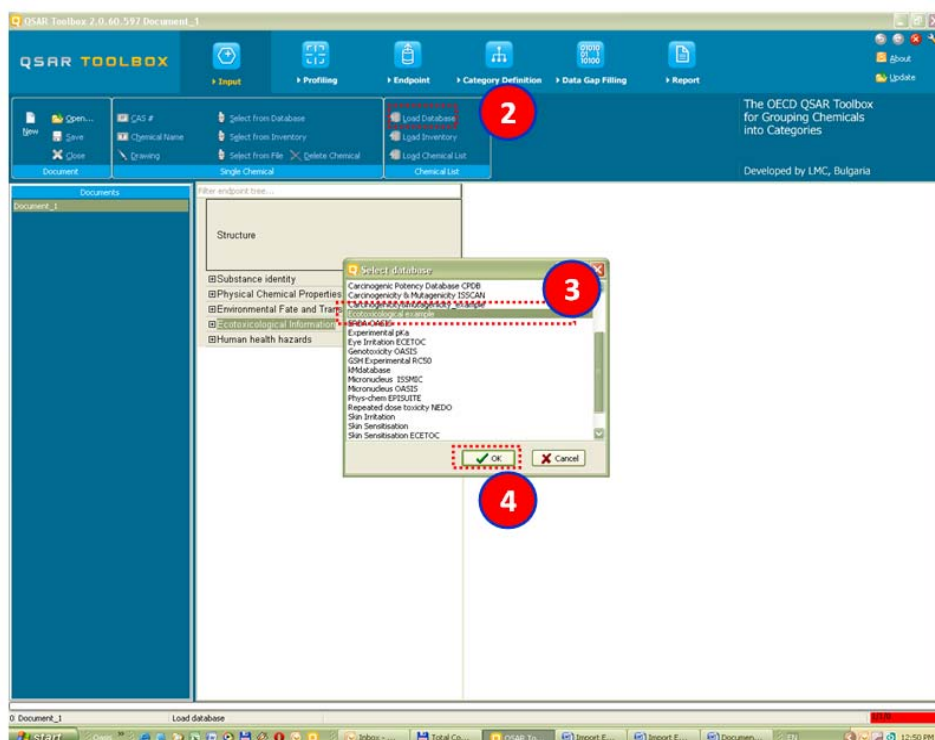
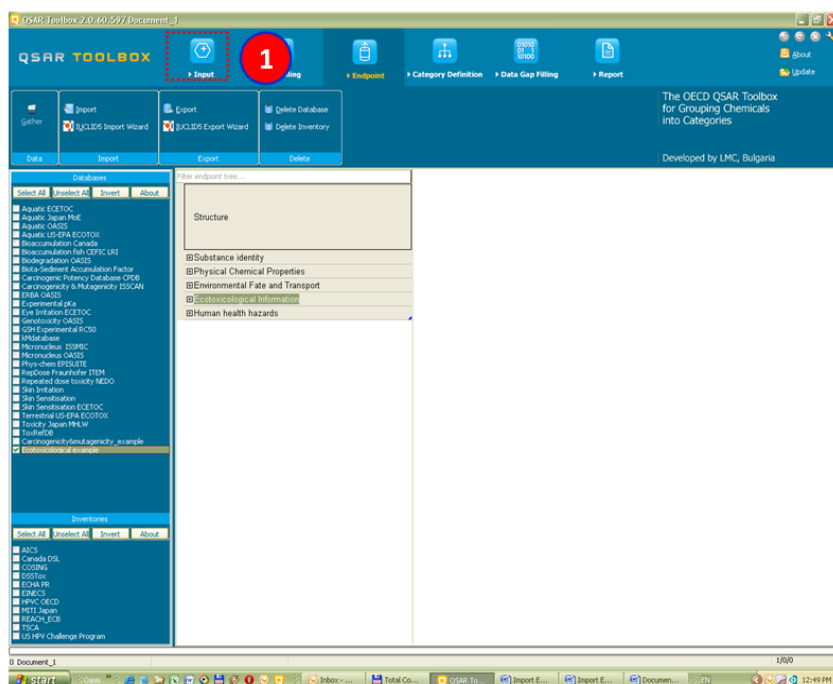
Cancel < Back Next > **1** Finish



The database is in the process of being imported. For large databases this might take considerable time.



Load imported database



The screenshot displays the QSAR Toolbox software interface. The top menu bar includes 'Input', 'Profiling', 'Grouping' (highlighted with a red dashed box and a red circle with the number 5), 'Property Definition', 'Data Gap Filling', and 'Report'. The left sidebar has 'Data' (highlighted with a red dashed box and a red circle with the number 7), 'Input', 'Export', 'Delete Database', and 'Delete Inventory'. The main window is divided into three panes. The left pane shows a 'Databases' list with 'Ecological endpoints' selected (highlighted with a red dashed box and a red circle with the number 6). The middle pane shows a tree view of 'Water endpoint toxic...' with 'Structure' selected. The right pane displays chemical structures and associated data for various endpoints, including 'Aquatic Toxicity' and 'Human health hazards'. The bottom status bar shows '39 Ecotoxicological endpoints'.

Appendix III: Import example for database with Human health hazards Information



The example uses a file that is already prepared. Guidance on how to prepare file for horizontal import can be found in Appendix I.

The destination for example files is [Install folder]\Examples

The default path is: C:\Program Files\QSAR Toolbox\QSAR Toolbox 2.1\Examples\ Carcinogenicity&mutagenicity_example.xls

CAS	ChemName	Smiles	Endpoint tree path	Scale	Data Mean value/Scale value	Data Unit	Endpoint	Type of method
92177-18-8	Nitrozo-2-oxopropylamine	NC(=O)C(=O)N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
15973-39-9	4-Nitro-5-oxo-1,2,3,4-tetrahydronaphthalene	C1=CC(=CC=C1)C(=O)N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
76881-18-4	1,4-Nitro-3,4,5-trimethoxybenzene	COC1=C(C(=C(C=C1)OC)OC)OC	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
105-09-9	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
51179-11-4	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
71785-67-7	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
114382-83-4	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
10595-26-6	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
66587-10-3	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
114279-28-2	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
101016-48-4	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
25081-31-6	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Equival		Summary carcinogenicity	
29921-49-9	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
70416-19-7	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
59884-51-5	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
17-9	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
69658-91-9	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
15219-30-1	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
44647-91-2	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
4215-16-8	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
56247-3	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
7516-26-9	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
81795-97-5	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
26541-51-6	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
23292-30-4	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
59-43-4	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
2444-63-3	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
23062-74-4	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
115-80-1	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Equival		Summary carcinogenicity	
64221-21-1	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
72599-59-5	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
65-89-1	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Equival		Summary carcinogenicity	
22105-20-0	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
3098-62-2	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
14688-28-4	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
116558-85-4	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Equival		Summary carcinogenicity	
89837-35-4	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
6462-71-7	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
102-77-2	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
149-28-1	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
87690-02-0	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
2227-09-2	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Equival		Summary carcinogenicity	
92-84-2	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
103-75-0	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
106-45-2	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
90-49-1	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	
105650-23-4	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Positive		Summary carcinogenicity	
56393-22-7	Nitrosodiazene	N=O	Human health hazardsCarcinogenicity	Carcinogenicity (I) (RSC4)	Negative		Summary carcinogenicity	

1

2

3

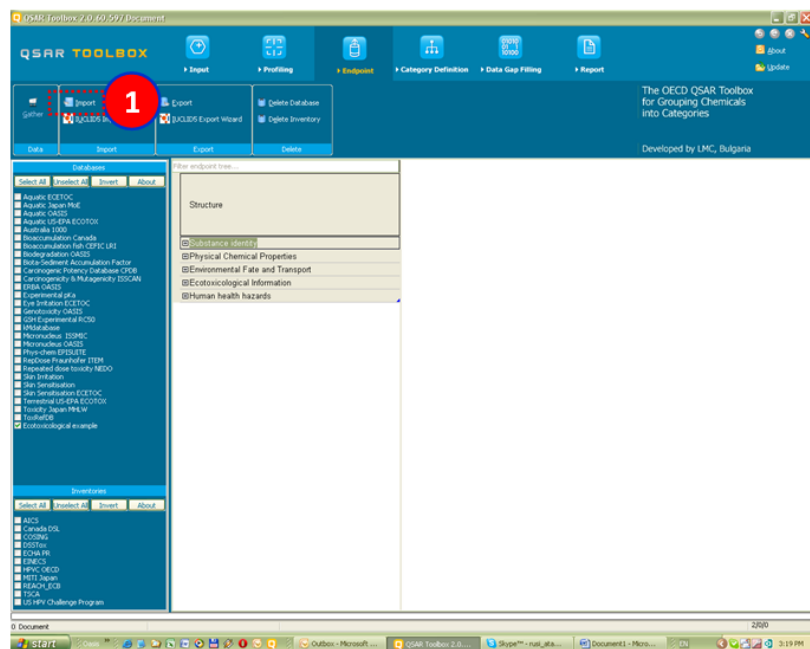
Sequence of steps

Click on

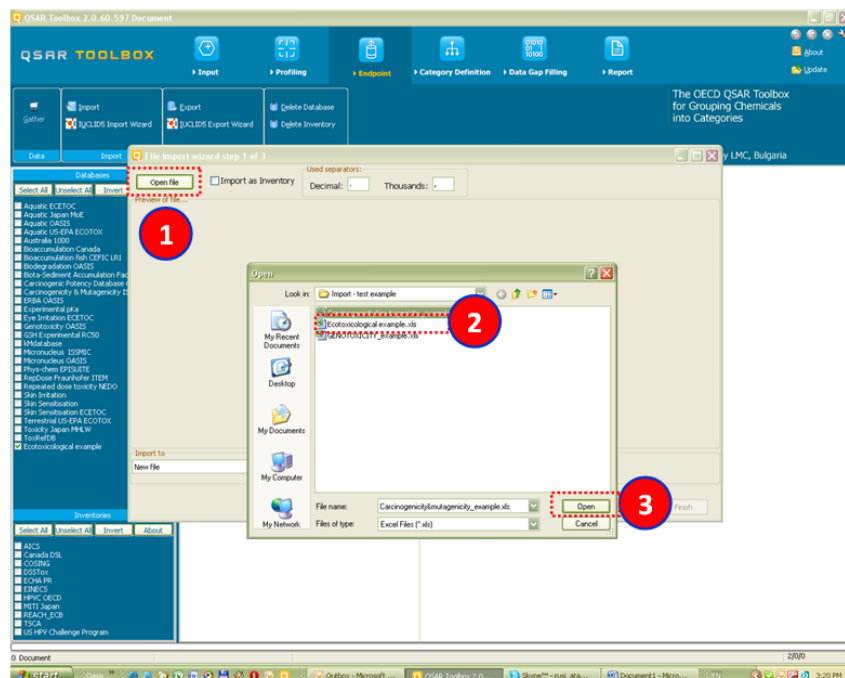
Import New Database

Detailed information

Import New Database



Open example file



File is open

File import wizard step 1 of 3

Used separators:
 ☐ Import as Inventory
 Decimal: Thousands:

Preview of file D:\Transfer\Import - test example\Carcinogenicity\mutagenicity_example.xls

CAS	ChemName	Smiles	Endpoint tree path	Scale	Data Mea
92177-49-6	Nitroso-2-oxopropylethanolamine	<chem>N(CC(C)=O)(CCO)N=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
15973-99-6	di(N-Nitroso)-perhydropyrimidine	<chem>C1CCN(CN1N=O)N=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
75881-18-4	1-Nitroso-3,4,5-trimethylpiperazine	<chem>C1C(N(C(CN1N=O)C)C)C</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
625-89-8	N-Nitrosobis(2,2,2-trifluoroethyl) amine	<chem>N(N(CC(F)(F)F)CC(F)(F)F)=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
51715-17-4	Nitroschloridiazepoxide	<chem>c1(C2c3c(ccc(C)C)N=C(N(C)N=O)CN=2</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
73785-40-7	N-Nitrosocimetidine	<chem>O=NN(C)C(=N(C#N)NCCSCC1=C(N=CN1)</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
114282-83-6	N-Nitrosodithiazine	<chem>N1(CSCSC1)N=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
10595-95-6	Nitrosoethylmethylamine	<chem>CN(CC)N=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
55557-02-3	N-Nitrosoguvacoline	<chem>OC(=O)C1=CCCN(C1)N=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
42579-28-2	1-Nitrosohydantoin	<chem>C1(NC(CN1N=O)=O)=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
30310-80-6	Nitrosohydroxyproline	<chem>C(=O)O)C1CC(CN1N=O)O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
25081-31-6	Nitrosiminodiacetic acid	<chem>N(N(CC(O)=O)CC(O)=O)=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Equivoca
26921-68-6	N-Nitrosomethyl-(2-hydroxyethyl) amine	<chem>CN(CCO)N=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
70415-59-7	N-Nitrosomethyl-(3-hydroxypropyl) amine	<chem>CN(CCCO)N=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
55984-51-5	N-Nitrosomethyl-(2-oxopropyl) amine	<chem>CN(CC(C)=O)N=O</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
9	N-Nitrosomethyl-(2-tosyloxyethyl) amine	<chem>NC(COS(=O)(=O)C1C=CC(C)=CC=1)CN</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
69658-91-9	3-Nitrosomethylaminopyridine	<chem>C1=CC=C(C(=N1)N(N=O)C</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
16219-99-1	4-Nitrosomethylaminopyridine	<chem>C1=CC(=CC=N1)N(N=O)C</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
55557-03-4	Nitrosomethylphenidate	<chem>C1=CC=C(C(C(=O)OC)C2N(N=O)CCCC2</chem>	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative

Import to:
 Database title:

Import to: The user has to choose whether to import the file as a new database or whether to add the data to a database already existing in the Toolbox.

Database title:
Carcinogenicity&mutagenicity_example
 Note: the user could change the name

⚠ It is very important that the thousands and decimal separators are properly set while importing. Especially with TXT file this could lead to erroneous parsing of data values.

Step 2: Horizontal import

File import wizard step 2 of 3

File layout
☒ Vertical
☐ Horizontal **1**

☒ I have a header row [What do different layouts mean?](#)

Define new region

CAS
Chemical name
SMILES
Undefined

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

CAS	ChemName	Smiles	Endpoint tree path	Scale	Data Mean
92177-49-6	Nitroso-2-oxopropylethanolamine	SMILES=O(CCO)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
15973-99-6	di(N-Nitroso)-perhydropyrimidine	C1CCN(CN1N=O)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
75881-18-4	1-Nitroso-3,4,5-trimethylpiperazine	C1C(N(C(CN1N=O)C)C)C	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
625-89-8	N-Nitrosobis(2,2,2-trifluoroethyl) amine	WN(C(C(F)(F)F)CC(F)(F)F)=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
51715-17-4	Nitrosochloridiazepoxide	C1(C2c3c(ccc(C)C)N=C(N(C)N=O)CN=2	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
73785-40-7	N-Nitrosocimetidine	O=NN/C(=N/C#N)NCCSCC1=C(N=CN1	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
114282-83-6	N-Nitrosodithiazine	N1(CSCSC1)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
10595-95-6	Nitrosoethylmethylaniline	CN(CC)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
55557-02-3	N-Nitrosoguvacoline	OC(=O)C1=CCCN(C1)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
42579-28-2	1-Nitrosopyridine	C1(NC(CN1N=O)=O)=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
30310-80-6	Nitrosohydroxyproline	C(=O)O(C)C1CC(CN1N=O)O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
25081-31-6	Nitrosiminodiacetic acid	N(N(C(C(=O)O)CC(C(=O)O)=O)=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Equivocal
26921-68-6	N-Nitrosomethyl-(2-hydroxyethyl) amine	CN(CCO)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
70415-59-7	N-Nitrosomethyl-(3-hydroxypropyl) amine	N(N(CCCO)C)O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
55984-51-5	N-Nitrosomethyl-(2-oxopropyl)amine	CN(C(C)O)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
9	N-Nitrosomethyl-(2-tosyloxyethyl) amine	NC(COS(=O)(=O)C1C=CC(C)=CC=1)CN	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
69658-91-9	3-Nitrosomethylaminopyridine	C1=CC=C(C(=N1)N(N=O)C	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
16219-99-1	4-Nitrosomethylaminopyridine	C1=CC(=CC(=N1)N(N=O)C	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative

Cancel < Back **Next >** **2** Finish

Select : CAS

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal

☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region **2**

CAS
Chemical name
SMILES
Endpoint tree path

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata

☐ is value add

Water alkalinity
Water flow
Water Hardness
Water media type
Water type
WaterType
Weekly exposure (days/ week)
Wet_Dry
Year
Undefined

CAS	ChemName	Smiles	Endpoint tree path	Scale	Data Mean
CAS7-49-6	Nitroso-2-oxopropylethanolamine	SMILES=O(CCO)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
15973-99-6	di(N-Nitroso)-perhydropyrimidine	C1CCN(CN1N=O)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
75881-18-4	1-Nitroso-3,4,5-trimethylpiperazine	C1C(N(C(CN1N=O)C)C)C	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
625-89-8	N-Nitrosobis(2,2,2-trifluoroethyl) amine	WN(C(C(F)(F)F)CC(F)(F)F)=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
51715-17-4	Nitrosochloridiazepoxide	C1(C2c3c(ccc(C)C)N=C(N(C)N=O)CN=2	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
73785-40-7	N-Nitrosocimetidine	O=NN/C(=N/C#N)NCCSCC1=C(N=CN1	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative

1

Cancel < Back **Next >** Finish

Select : *Chemical name*

Select : *SMILES*

Page 51 of 70

Select : Endpoint tree path

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Metadata
 Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Undefined

☐ is value add

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

CAS	ChemName	Smiles	Endpoint tree path	Scale	Data Mean
74-49-6	Chemical nameylethanolamine	SMILES=O(CCO)N=O	Endpoint tree path#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
15973-99-6	di(N-Nitroso)-perhydropyrimidine	C1CCN(CN1N=O)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
75881-18-4	1-Nitroso-3,4,5-trimethylpiperazine	C1C(N(C(CN1N=O)C)C)C	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
525-89-8	N-Nitrosobis(2,2,2-trifluoroethyl) amine	N(N(C(C(F)(F)F)CC(F)(F)F)=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
51715-17-4	Nitroschloridiazepoxide	-1(C2c3c(ccc(C)C3)N=C(N(C)N=O)CN=2	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative

Cancel < Back Next > Finish

Important: The endpoint tree path should point to a leaf of the **predefined** endpoint tree. For more information check chapter 4 **Endpoint tree path** in this document.

Select : DATA Scale

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Metadata
 Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Undefined

☐ is value add

Define value
☒ Scale
☐ Low qualifier
☐ Upper value
☐ Mean qualifier
☐ Low value
☐ Unit
☐ Mean value/Scale val
☐ Upper qualifier
☐ Undefined

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

CAS	ChemName	Smiles	Endpoint tree path	Scale	Data Mean
74-49-6	Chemical nameylethanolamine	SMILES=O(CCO)N=O	Endpoint tree path#Carcinogenicity	Data Scale by I (ISSCAN)	Positive
15973-99-6	di(N-Nitroso)-perhydropyrimidine	C1CCN(CN1N=O)N=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
75881-18-4	1-Nitroso-3,4,5-trimethylpiperazine	C1C(N(C(CN1N=O)C)C)C	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Positive
525-89-8	N-Nitrosobis(2,2,2-trifluoroethyl) amine	N(N(C(C(F)(F)F)CC(F)(F)F)=O	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative
51715-17-4	Nitroschloridiazepoxide	-1(C2c3c(ccc(C)C3)N=C(N(C)N=O)CN=2	Human health hazards#Carcinogenicity	Carcinogenicity I (ISSCAN)	Negative

Finish

The Scale Carcinogenicity I (ISSCAN) has been defined in advance. According to this scale chemicals from the database are classified as "Positive" and "Negative". This is the default scale.

Note : Scale editor

File import wizard step 2 of 3

File layout

☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES

Endpoint tree path

Data

Undefined

Metadata

☐ Is value

Water alkalinity

Water flow

Water Hardness

Water media type

Water type

WaterType

Weekly exposure (days/ week)

Wet_Dry

Year

Undefined

1

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

CAS	Chemical name	Smiles	Endpoint tree path	Scale	Data Mean value/Scale v
5450	Chemical name ULFONATE	SMILES: C(=O)O	Endpoint tree path #Genetic Toxicity	Gene mutation I	Positive
52737	DICHLORVOS	C(Cl)(Cl)=COP(=O)(OC)OC	Human health hazards #Genetic Toxicity	Gene mutation I	Positive
53923	Phenoxybenzamine_hydrochloride [USAN]	C1(CN(+)C1)C(C)COC2CCCC2)CCC1	Human health hazards #Genetic Toxicity	Gene mutation I	Positive
54675	Diethyl_Sulfate	C(C)OS(=O)(=O)OCC	Human health hazards #Genetic Toxicity	Gene mutation I	Positive
56273	METHYL_METHANESULFONATE	COS(C)=O	Human health hazards #Genetic Toxicity	Gene mutation I	Positive
56274	Diethyl_methanesulfonate	CC(C)OS(=O)(=O)OCC	Human health hazards #Genetic Toxicity	Gene mutation I	Positive

Cancel < Back Next > Finish

Note : Scale definition

Scales definition...

Scale list | Scale conversions

Scale list

New

Carcinogenicity I (IS5CAN)

Gene mutation I

Chromosome aberration I (Oasis)

Micronucleus I

Respiratory sensitization I (ECET)

Skin sensitization I (Oasis)

Skin sensitization II (ECETO-C)

Carcinogenicity I (IS5CAN)

Skin sensitization EC3(ratio)

EC3

Skin sensitization IV (GPMT)

Skin sensitization V (BFR)

IUCLID5:phrasegroup_P16

IUCLID5:phrasegroup_P105

IUCLID5:phrasegroup_P20

IUCLID5:phrasegroup_C47

IUCLID5:phrasegroup_F23

IUCLID5:phrasegroup_T116

IUCLID5:phrasegroup_T46-2

IUCLID5:phrasegroup_T46-1

IUCLID5:phrasegroup_T21

IUCLID5:phrasegroup_T143

IUCLID5:phrasegroup_T102

Skin sensitization III (LJMU)

Irritation / Corrosion I (LJMU)

RC50

Severe skin irritation (Danish EPA)

Carcinogenicity III (CPDB)

Skin Irritation/Corrosion

Eye Irritation/Corrosion

IUCLID5.2:phrasegroup_P16

IUCLID5.2:phrasegroup_P105

IUCLID5.2:phrasegroup_P20

IUCLID5.2:phrasegroup_C47

IUCLID5.2:phrasegroup_F23

IUCLID5.2:phrasegroup_T116

Name: Carcinogenicity I (IS5CAN)

Type: Ordinal

Families

Carcinogenicity

Scale members

High: Positive, Equivocal, Negative

Low:

Scale description

Save & Close

Close

1

The user should select one of the scales from the Scale list panel or define a new scale.

Select : DATA Main Value/Scale value

File import wizard step 2 of 3

File layout
☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale ☐ Low qualifier ☐ Upper value
☐ Mean qualifier ☐ Value ☐ Unit ☐ Undefined
☒ Mean value/Scale value ☐ qualifier

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
 Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Undefined

Endpoint tree path
 Endpoint tree path#Carcinogenicity
 Human health hazards#Carcinogenicity
 Human health hazards#Carcinogenicity
 Human health hazards#Carcinogenicity
 Human health hazards#Carcinogenicity
 Human health hazards#Carcinogenicity

Scale	Data Mean value/Scale value	Data Unit	Endpoint	Type of method	Type of
Scale	Scale	Scale	Summary carcinogenicity		
Scale	Scale	Scale	Summary carcinogenicity		
Scale	Scale	Scale	Summary carcinogenicity		
Scale	Scale	Scale	Summary carcinogenicity		
Scale	Scale	Scale	Summary carcinogenicity		
Scale	Scale	Scale	Summary carcinogenicity		

Cancel < Back Next > Finish

Select : DATA Unit

File import wizard step 2 of 3

File layout
☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale ☐ Low qualifier ☐ Upper value
☐ Mean qualifier ☐ Value ☐ Unit ☐ Undefined
☐ Mean value/Scale value ☐ qualifier

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
 Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Undefined

Endpoint tree path
 Endpoint tree path#Carcinogenicity
 Human health hazards#Carcinogenicity
 Human health hazards#Carcinogenicity
 Human health hazards#Carcinogenicity
 Human health hazards#Carcinogenicity
 Human health hazards#Carcinogenicity

Scale	Data Mean value/Scale value	Data Unit	Endpoint	Type of method	Type of
Scale	Scale	Scale	Summary carcinogenicity		
Scale	Scale	Scale	Summary carcinogenicity		
Scale	Scale	Scale	Summary carcinogenicity		
Scale	Scale	Scale	Summary carcinogenicity		
Scale	Scale	Scale	Summary carcinogenicity		
Scale	Scale	Scale	Summary carcinogenicity		

Cancel < Back Next > Finish



I. The type of the column is specified by clicking on the column and

then clicking its type (CAS/Chemical name/SMILES) from the list box in the **Define new region** panel or selecting a metadata field label from the list box in the **Metadata** panel. To remove designations click a column and then click on **Undefined** from the list box.

- II. All fields defined with the **Define new region** panel are color-distinguished from the fields of the **Metadata** panel.

Scroll bar

File import wizard step 2 of 3

File layout:
☐ Vertical
☒ Horizontal
☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value:
☐ Scale
☐ Mean qualifier
☐ Mean value/Scale val
☐ Low qualifier
☐ Low value
☐ Upper qualifier
☐ Upper value
☒ Unit
☐ Undefined

Scales editor: Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata:
☐ is value

Type of method	Type of genotoxicity	Test type	Test organisms (species)	Sexual maturation (offspring)	Route	Comment
					other	Summary carcinogenicit;
					other	Summary carcinogenicit;
					other	Summary carcinogenicit;
					other	Summary carcinogenicit;
					other	Summary carcinogenicit;
					other	Summary carcinogenicit;

Buttons: Cancel, < Back, Next >, Finish

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☐ Scale
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low qualifier
☐ Low value
☐ Upper qualifier
☐ Upper value
☒ Unit
☐ Undefined

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

☐ is value

Metadata
 Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Undefined

Sexual maturation (offspring)Route	Comment	Reference source	Author	Year
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008

Cancel < Back Next > Finish

[illegible]

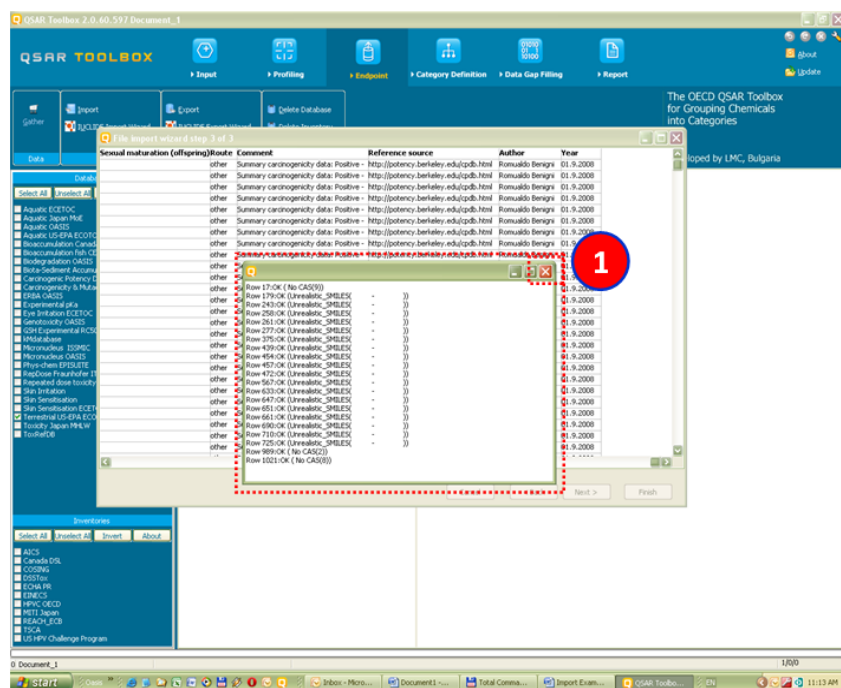
File Import wizard Step 3 of 3

Sexual maturation (offspring)Route	Comment	Reference source	Author	Year
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008
other	Summary carcinogenicity data: Positive -	http://potency.berkeley.edu/cpdb.html	Romualdo Benigni	01.9.2008

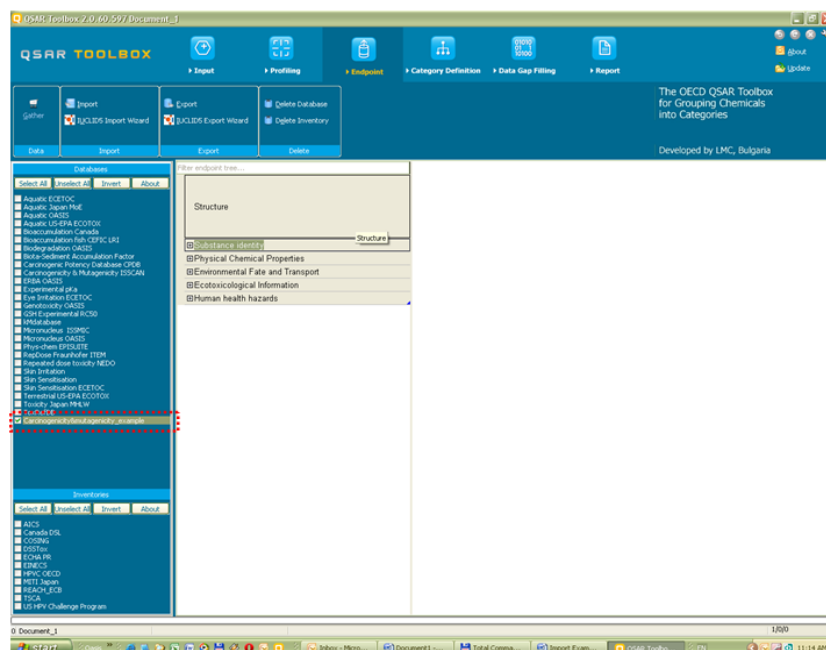
Cancel < Back Next > Finish

[illegible]

Error message log file



Imported database

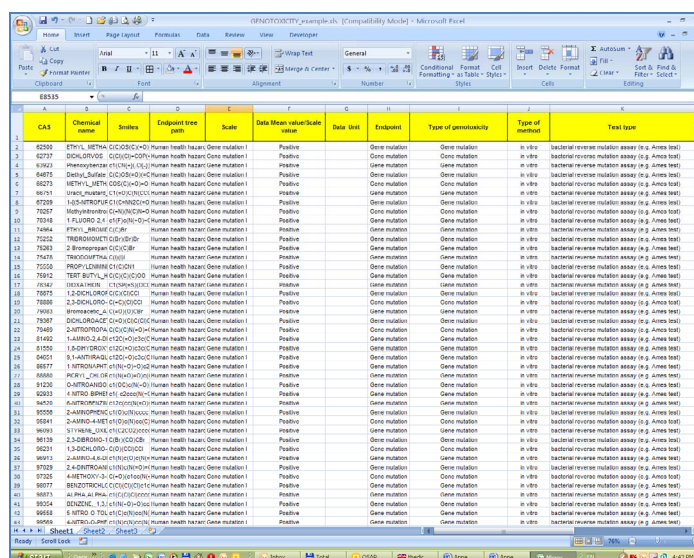


Appendix IV: Import example for database with Human health hazards Information

 The example below uses a file that is already prepared. Guidance on how to prepare file for horizontal import can be found in Appendix I.

The destination for example files is [Install folder]\Examples.

The default path is C:\Program Files\QSAR Toolbox\QSAR Toolbox 2.1\Examples\GENOTOXICITY_example.xls



CAS	Chemical name	Studies	Endpoint time point	Scale	Data Mean value/Scale value	Data unit	Evidence	Type of genotoxicity	Type of method	Test type
62592	ETHYL METHYL CARBAMATE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
62777	DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
43923	Phenylketonamide	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
64875	Diethyl Sulfate	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
68273	METHYL METHYL CARBAMATE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
69751	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
67239	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
76287	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
70348	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
74944	ETHYL METHYL CARBAMATE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
75252	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
76283	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
76425	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
75526	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
75912	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
10447	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
76875	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
76886	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
76903	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
76987	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
76949	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
81482	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
81505	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
84051	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
86577	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
86680	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
81226	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
82823	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
84539	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
85056	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
85841	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
86953	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
86139	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
86231	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
86917	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
87329	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
87325	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
86977	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
86973	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
85054	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
86618	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)
86959	1,2-DICHLORIDE	Human health hazard	Gene mutation	Positive			Gene mutation	Gene mutation	In vitro	biochemical-reverse mutation assay (e.g. Ames test)

1

2

3

Sequence of steps

Click on

Import New Database

Detailed information

[illegible]

The screenshot shows the QSRAR TOOLBOX software interface. The top menu bar includes 'File', 'Edit', 'View', 'Tools', 'Help', 'About', and 'Update'. The main window has a sidebar with 'Databases' and 'Import' sections. The 'Import' section is active, showing a list of databases. The 'EPA ECOTOX' database is selected. The 'Open' dialog box is open, showing the file 'GENOTOXITY_example.xls' selected. Red dashed boxes and blue circles with numbers 1, 2, and 3 highlight the 'Open File' button, the selected file, and the 'Open' button respectively.

File is open

File import wizard step 1 of 3

Used separators:
☐ Open file ☐ Import as Inventory
 Decimal: Thousands:

Preview of file D:\Transfer\Import - test example\GENOTOXICITY_example.xls

CAS	Chemical name	Smiles	Endpoint tree path	Scale	Data Mean value/Scale
62500	ETHYL METHANESULFONATE	C(C)OS(CX=O)=O	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
62737	DICHLORVOS	C(Cl)(Cl)=COP(=O)(OC)OC	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
63923	Phenoxylbenzamine_hydrochloride_[USAN]	C1(CN(+)X(C(-)X(C(C)COC2CCCC2)CCCl)Human health hazards#Genetic Toxicity	Gene mutation I	Positive	
64675	Diethyl_Sulfate	C(C)OS(=O)(=O)OCC	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
66273	METHYL METHANESULFONATE	COS(CX=O)=O	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
66751	Uracil_mustard_[USAN]	C1(=O)C(N(CCC)CCC)=CNC(=O)N1	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
67209	1-[(5-NITROFURFURYLIDENE)AMINO]HYDRAZINE	C1(C=NN2C(=O)NC(=O)C2)=CC=C(N(=O)C(=O)N2)C1	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
70257	Methylnitroinosoguanidine	C(=N)(N(C)N=O)NN(=O)=O	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
70348	1-FLUORO-2,4-DINITROBENZENE	c1(F)c(N(=O)=O)cc(N(=O)=O)cc1	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
74964	ETHYL BROMIDE	C(C)Br	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
75252	TRIBROMOMETHANE	C(Br)(Br)Br	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
75263	2-Bromopropane	C(C)(C)Br	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
75478	TRITODOMETHANE	C(I)(I)I	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
75558	PROPYLENIMINE	C1(C)CN1	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
75912	TERT-BUTYL HYDROPEROXIDE	C(C)(C)(C)OO	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
78342	DIOXATHION	C1(SP(=S)(OCC)OCC)C(SP(=S)(OCC)OCC)C1	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
78875	1,2-DICHLOROPROPANE;_1,2-Dichloro- <i>propan-1,2-diol</i>	C(C)(Cl)CCl	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
78886	2,3-DICHLORO-1-PROPENE	C(=C)(Cl)CCl	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
79083	Bromoacetic_Acid	C(=O)(O)CBr	Human health hazards#Genetic Toxicity	Gene mutation I	Positive

Import to
☐ New file ☒ Database title GENOTOXICITY_example

Cancel < Back **Next >** 1 Finish

Import to: The user has to choose whether to import the file as a new database or whether to add the data to a database already existing in the Toolbox.

Database title : GENOTOXICITY_example
 Note: the user could change the name

⚠ It is very important that the thousands and decimal separators are properly set while importing. Especially with TXT file this could lead to erroneous parsing of data values.

Step 2: Horizontal import

File import wizard step 2 of 3

File layout
☒ Vertical
☐ Horizontal

1

☒ I have a header row

[What do different layouts mean?](#)

Define new region

CAS
 Chemical name
SMILES
 Undefined

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

CAS	Chemical name	Smiles	Endpoint tree path	Scale	Data Mean value/Scale v
52450	ETHYL_METHANESULFONATE	SMILES :C(=O)=O	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
52737	DICHLORVOS	C(Cl)(Cl)=COP(=O)(O)COC	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
53923	Phenoxybenzamine_hydrochloride_[USAN]	C1(CN(+)(X.Cl(-)))(C(C)COC2ccccc2)CCCC1	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
54675	Diethyl_Sulfate	C(C)OS(=O)(=O)OCC	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
56273	METHYL_METHANESULFONATE	COS(C)=O=O	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
56751	Uracil_mustard_[USAN]	C1=(O)(C(N(CCC)CCC)=CNC(=O)N)I	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
57209	1-[(S-NITROFURFURYLIDENE)AMINO]HYDRAZINE	C1C(=NN2C(=O)NC(=O)C2)=CC=C(N(=N)C1=O)N	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
70257	Methylnitrosoguanidine	C(=N)N(C)N=O	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
70348	1-FLUORO-2,4-DINITROBENZENE	C1(F)(C(N(=O)=O)cc(N(=O)=O)cc1	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
74964	ETHYL_BROMIDE	C(C)Br	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
75252	TRIBROMOMETHANE	C(Br)(Br)Br	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
75263	2-Bromopropane	C(C)C)Br	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
75478	TRIOMOMETHANE	C(I)(I)I	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
75558	PROPYLENIMINE	C1C(C)CN1	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
75912	TERT-BUTYL_HYDROPEROXIDE	C(C)(C)OO	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
78342	DIOXATHION	C1(SP(=S)(OCC)OCC)C(SP(=S)(OCC)OCC)S1	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
78875	2,2-DICHLOROPROPANE; "1,2-Dichloro-2-methylpropane"	C(C)(C)C(Cl)C(Cl)	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
78886	2,3-DICHLORO-1-PROPENE	C=C(C)C(Cl)C(Cl)	Human health hazards#Genetic Toxicity	Gene mutation I	Positive

Navigation: < Back Next > Finish

Select : CAS

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row

[What do different layouts mean?](#)

☐ Private data

Define column region
 CAS
 Chemical name
 SMILES
 Endpoint tree path

Metadata
☐ is value
 add

Water alkalinity
 Water flow
 Water Hardness
 Water media type
 Water type
 WaterType
 Weekly exposure (days/ week)
 Wet_Dry
 Year
 Undefined

CAS	name	Smiles	Endpoint tree path	Scale	Data Mean value/Scale v
52737	DICHLORVOS	<chem>C(Cl)(Cl)=COP(=O)(OC)OC</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
53923	Phenoxybenzamine_hydrochloride [USAN]	<chem>ClCN+)(Cl)(C(C)COC2CCCC2)CCC</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
54675	Diethyl_Sulfate	<chem>C(C)OS(=O)(=O)OCC</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
56273	METHYL_METHANESULFONATE	<chem>COS(C)(=O)=O</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
56273	Diethyl_methylsulfonyl_sulfonate [USAN]	<chem>CC(C)OS(=O)(=O)OCC</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive

Cancel < Back Next > Finish

Select : Chemical name

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

CAS
 Chemical name
 SMILES
 Endpoint tree path

Metadata
☐ is value

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

CAS	Chemical name	Smiles	Endpoint tree path	Scale	Data Mean value/Scale v
52737	DICHLORVOS	<chem>C(Cl)(Cl)=COP(=O)(OC)OC</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
53923	Phenoxybenzamine_hydrochloride	<chem>C1(CN(+)C1)C(C)COC2CCCC2CC(C)C</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
54675	Diethyl_Sulfate	<chem>C(C)OS(=O)(=O)OCC</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
56273	METHYL_METHANESULFONATE	<chem>COS(C)(=O)=O</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive

Cancel < Back Next > Finish

Select : SMILES

File import wizard step 2 of 3

File layout
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

CAS
 Chemical name
 SMILES
 Endpoint tree path

Metadata
☐ is value

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

CAS	Chemical name	Smiles	Endpoint tree path	Scale	Data Mean value/Scale v
52737	DICHLORVOS	<chem>C(Cl)(Cl)=COP(=O)(OC)OC</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
53923	Phenoxybenzamine_hydrochloride_USAN	<chem>C1(CN(+)C1)C(C)COC2CCCC2CC(C)C</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
54675	Diethyl_Sulfate	<chem>C(C)OS(=O)(=O)OCC</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive
56273	METHYL_METHANESULFONATE	<chem>COS(C)(=O)=O</chem>	Human health hazards#Genetic Toxicity	Gene mutation I	Positive

Cancel < Back Next > Finish

Select : DATA Scale

File import wizard step 2 of 3

File layout
☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☒ Scale ☐ Low qualifier ☐ Upper value
☐ Mean qualifier ☐ Low value ☐ Unit
☐ Mean value/Scale value ☐ Upper qualifier ☐ Undefined

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
☐ Is value

CAS	Chemical name	Smiles	Endpoint tree path	Scale	Data Mean value/Scale v
52737	DICHLORVOS	<chem>C(Cl)(Cl)=COP(=O)(OC)OC</chem>	human health hazards#Genetic Toxicity	Gene mutation I	Positive
53923	Phenoxybenzamine_hydrochloride_[USAN]	<chem>C1(CN(+)C1C(C)C(C)C2CCCC2)CCCl</chem>	human health hazards#Genetic Toxicity	Gene mutation I	Positive
54675	Diethyl_Sulfate	<chem>CCOS(=O)(=O)OCC</chem>	human health hazards#Genetic Toxicity	Gene mutation I	Positive
56273	METHYL_METHANESULFONATE	<chem>COS(C)(=O)=O</chem>	human health hazards#Genetic Toxicity	Gene mutation I	Positive

Cancel < Back Next > Finish

Select : Endpoint tree path

File import wizard step 2 of 3

File layout
☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value
☒ Scale ☐ Low qualifier ☐ Upper value
☐ Mean qualifier ☐ Low value ☐ Unit
☐ Mean value/Scale value ☐ Upper qualifier ☐ Undefined

Scales editor

Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
☐ Is value

CAS	Chemical name	Smiles	Endpoint tree path	Scale	Data Mean value/Scale v
52737	DICHLORVOS	<chem>C(Cl)(Cl)=COP(=O)(OC)OC</chem>	human health hazards#Genetic Toxicity	Gene mutation I	Positive
53923	Phenoxybenzamine_hydrochloride_[USAN]	<chem>C1(CN(+)C1C(C)C(C)C2CCCC2)CCCl</chem>	human health hazards#Genetic Toxicity	Gene mutation I	Positive
54675	Diethyl_Sulfate	<chem>CCOS(=O)(=O)OCC</chem>	human health hazards#Genetic Toxicity	Gene mutation I	Positive
56273	METHYL_METHANESULFONATE	<chem>COS(C)(=O)=O</chem>	human health hazards#Genetic Toxicity	Gene mutation I	Positive

Cancel < Back Next > Finish

Important: The endpoint tree path should point to a leaf of the predefined endpoint tree. For more information check chapter 4 **Endpoint tree path** in this document.

Select : DATA Main Value/Scale value

File import wizard step 2 of 3

File layout: ☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES

Endpoint tree path: **Data** (2)

Define value: ☐ Scale ☐ Low qualifier ☐ Upper value ☒ Mean value/Scale value (3) ☐ Mean qualifier ☐ Low value ☐ Unit ☐ Undefined

Metadata: Water alkalinity, Water flow, Water Hardness, Water media type, Water type, WaterType, Weekly exposure (days/ week), Wet_Dry, Year, Undefined

Endpoint tree path: Genetic Toxicity

Endpoint tree path	Scale	Data Mean value/Scale value	Data Unit	Endpoint	Type of genotoxicity	Type of method	Test type
Endpoint tree path#Genetic Toxicity	Data.Scale	Data.Mean value/Scale value (1)	Data.Unit	Gene mutation	Gene mutation	in vitro	bacterial revers
Human health hazards#Genetic Toxicity	Gene mutation I	Positive		Gene mutation	Gene mutation	in vitro	bacterial revers
Human health hazards#Genetic Toxicity	Gene mutation I	Positive		Gene mutation	Gene mutation	in vitro	bacterial revers
Human health hazards#Genetic Toxicity	Gene mutation I	Positive		Gene mutation	Gene mutation	in vitro	bacterial revers
Human health hazards#Genetic Toxicity	Gene mutation I	Positive		Gene mutation	Gene mutation	in vitro	bacterial revers
Human health hazards#Genetic Toxicity	Gene mutation I	Positive		Gene mutation	Gene mutation	in vitro	bacterial revers

Cancel < Back Next > Finish

Select : DATA Unit

File import wizard step 2 of 3

File layout: ☐ Vertical ☒ Horizontal ☒ I have a header row [What do different layouts mean?](#)

☐ Private data

Define new region

SMILES

Endpoint tree path: **Data** (2)

Define value: ☐ Scale ☐ Low qualifier ☐ Upper value ☒ Mean value/Scale value (3) ☐ Mean qualifier ☐ Low value ☐ Unit ☐ Undefined

Metadata: Water alkalinity, Water flow, Water Hardness, Water media type, Water type, WaterType, Weekly exposure (days/ week), Wet_Dry, Year, Undefined

Endpoint tree path: Genetic Toxicity

Endpoint tree path	Scale	Data Mean value/Scale value	Data Unit	Endpoint	Type of genotoxicity	Type of method	Test type
Endpoint tree path#Genetic Toxicity	Data.Scale	Data.Mean value/Scale value (1)	Data.Unit	Gene mutation	Gene mutation	in vitro	bacterial revers
Human health hazards#Genetic Toxicity	Gene mutation I	Positive		Gene mutation	Gene mutation	in vitro	bacterial revers
Human health hazards#Genetic Toxicity	Gene mutation I	Positive		Gene mutation	Gene mutation	in vitro	bacterial revers
Human health hazards#Genetic Toxicity	Gene mutation I	Positive		Gene mutation	Gene mutation	in vitro	bacterial revers
Human health hazards#Genetic Toxicity	Gene mutation I	Positive		Gene mutation	Gene mutation	in vitro	bacterial revers
Human health hazards#Genetic Toxicity	Gene mutation I	Positive		Gene mutation	Gene mutation	in vitro	bacterial revers

Cancel < Back Next > Finish



I. The type of the column is specified by clicking on the column and

then clicking its type (CAS/Chemical name/SMILES) from the list box in the **Define new region** panel or selecting a metadata field label from the list box in the **Metadata** panel. To remove designations click a column and then click on **Undefined** from the list box.

- II. All fields defined with the **Define new region** panel are color-distinguished from the fields of the **Metadata** panel.

Scroll bar

File import wizard step 2 of 3

File layout:
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value:
☐ Scale
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low qualifier
☐ Low value
☐ Upper qualifier
☐ Upper value
☒ Unit
☐ Undefined

Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
☐ is value add

Endpoint	Type of genotoxicity	Type of method	Test type	Test organisms (species)	Strain	Species/strain	Metabolic activation
Gene mutation	Gene mutation	in vitro	bacterial reverse mutation assay (e.g. Ames)	Salmonella typhimurium	TA 100		without S9
Gene mutation	Gene mutation	in vitro	bacterial reverse mutation assay (e.g. Ames)	Salmonella typhimurium	TA 100		without S9
Gene mutation	Gene mutation	in vitro	bacterial reverse mutation assay (e.g. Ames)	Salmonella typhimurium	TA 100		without S9
Gene mutation	Gene mutation	in vitro	bacterial reverse mutation assay (e.g. Ames)	Salmonella typhimurium	TA 100		without S9
Gene mutation	Gene mutation	in vitro	bacterial reverse mutation assay (e.g. Ames)	Salmonella typhimurium	TA 100		without S9

Cancel < Back Next > Finish

Scroll the bar to review data column designations. Click on "Next" button

File import wizard step 2 of 3

File layout:
☐ Vertical
☒ Horizontal
☒ I have a header row
[What do different layouts mean?](#)

☐ Private data

Define new region

SMILES
 Endpoint tree path
 Data
 Undefined

Define value:
☐ Scale
☐ Mean qualifier
☐ Mean value/Scale value
☐ Low qualifier
☐ Low value
☐ Upper qualifier
☐ Upper value
☒ Unit
☐ Undefined

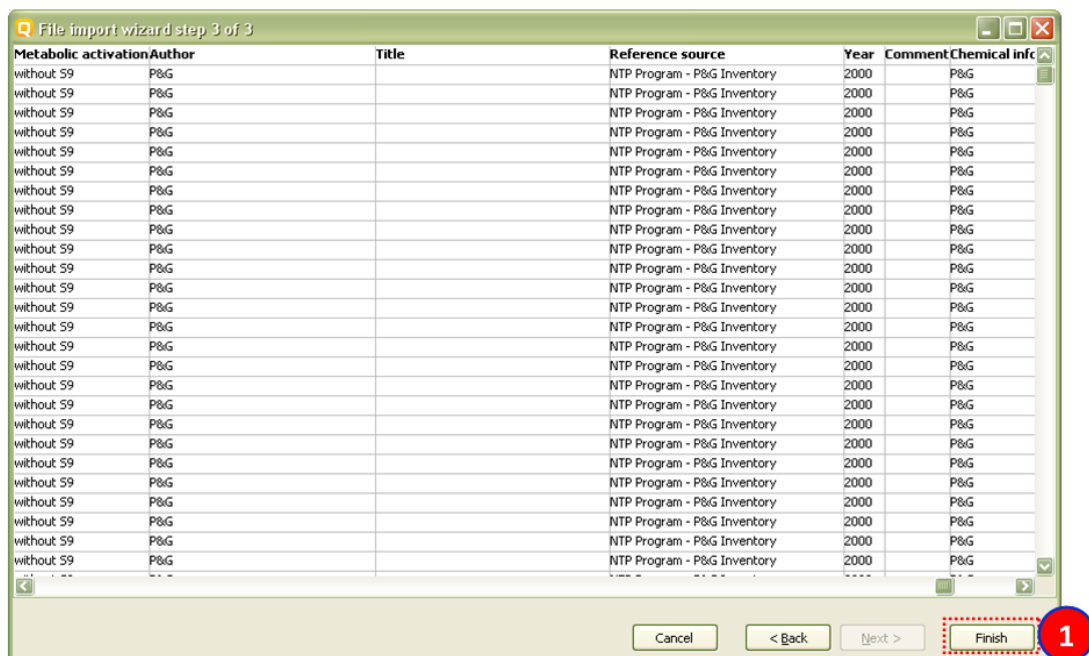
Scales editor
 Select one or more columns from the grid below and then assign the selection to a region by choosing from the lefthand side radio buttons.

Metadata
☐ is value add

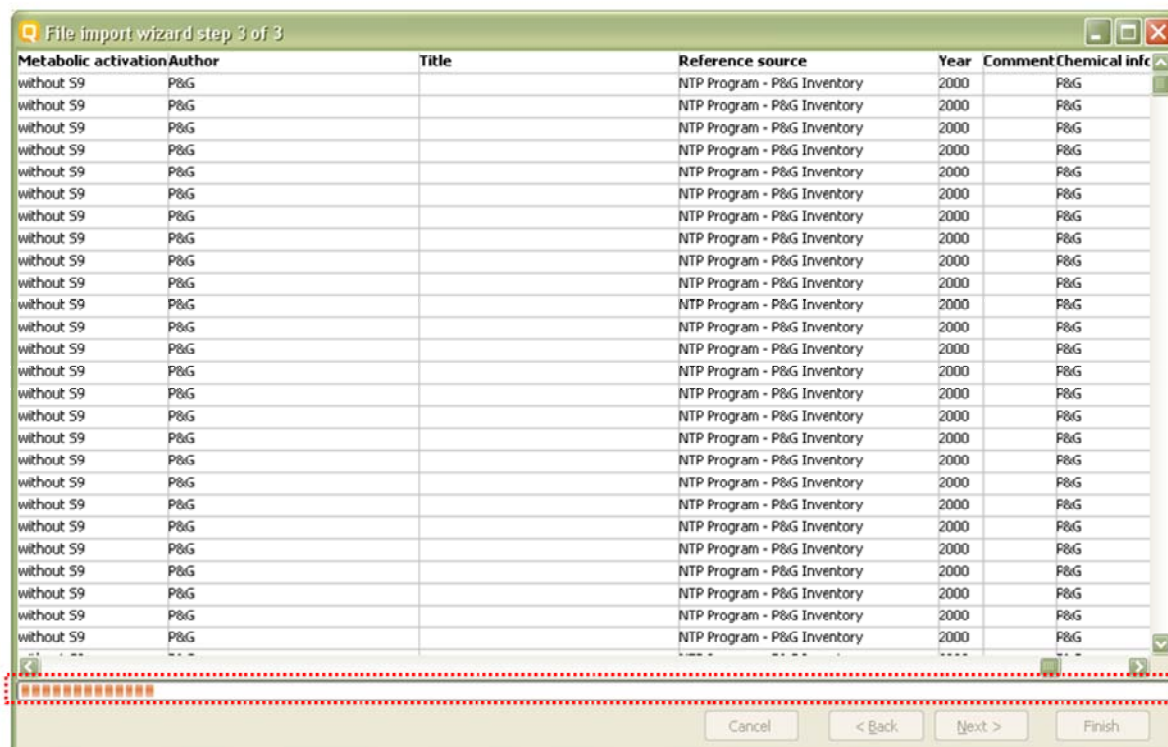
Author	Title	Reference source	Year	Comment	Chemical info
P&G		NTP Program - P&G Inventory	2000		P&G
P&G		NTP Program - P&G Inventory	2000		P&G
P&G		NTP Program - P&G Inventory	2000		P&G
P&G		NTP Program - P&G Inventory	2000		P&G
P&G		NTP Program - P&G Inventory	2000		P&G

Cancel < Back Next > Finish

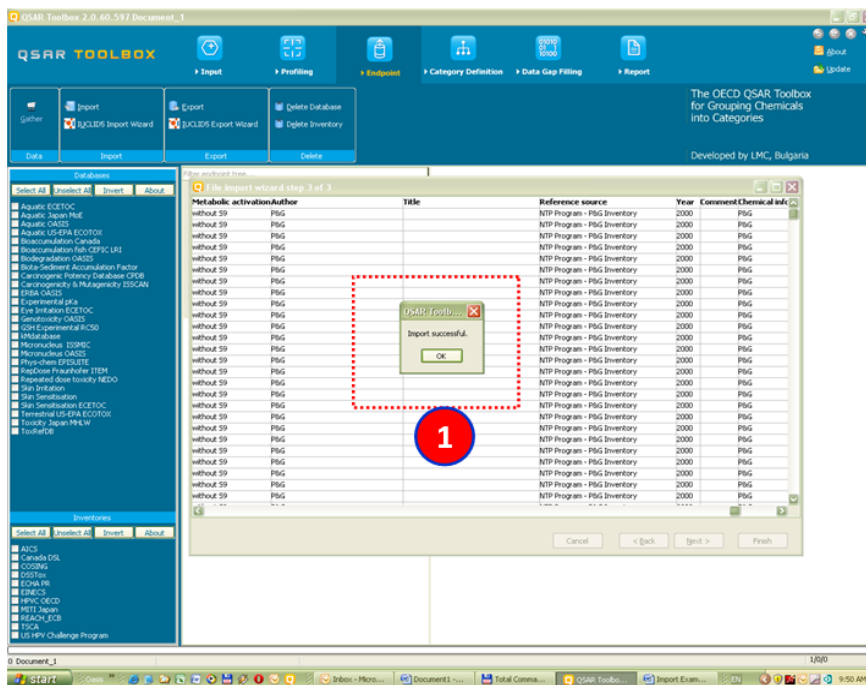
Click on "Finish" button



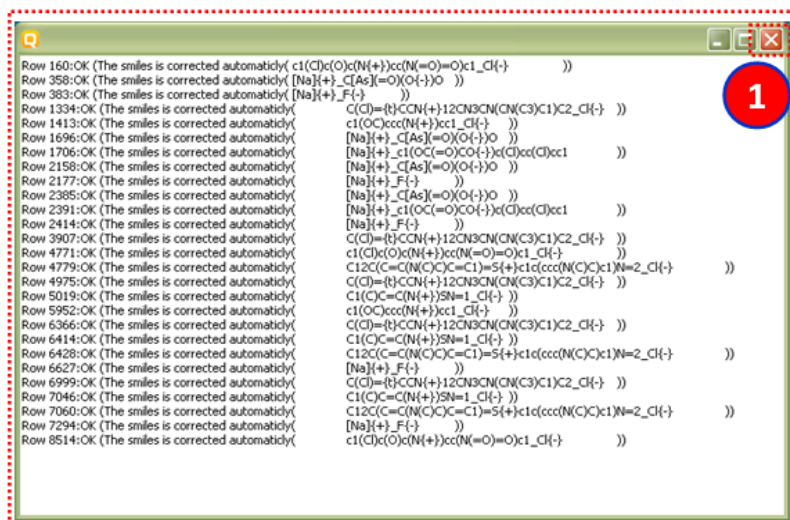
Database is being imported: Progress bar.



Database is imported successfully. Click "OK".



Error message log file



OECD

2, rue André Pascal
75775 Paris Cedex 16
France

Tel.: +33 1 45 24 82 00

Fax: +33 1 45 24 85 00

ehscont@oecd.org