

OECD QSAR Toolbox v.3.4

Step-by-step example of how to evaluate an ad-hoc category of aliphatic amines and to predict an ecotoxicological endpoint

Background

- This is a step-by-step presentation designed to take you through the workflow of the Toolbox for evaluating of an ad-hoc category.
- You will learn several new functionalities which we be repeated to assure a consistent category is defined.
- It is assumed that you now have some experience in using the Toolbox so there will be multiple key strokes between screen shots.

Outlook

- Background
- **Category evaluation - overview**

Category Evaluation

Two phases of evaluation process

Phase I. Evaluation of Category & Data Gap Filling:

- Investigating the structural consistency of an ad hoc category (e.g., a category submitted by an industry consortium to a regulatory assessment program).
- Implementation of an ad hoc category building and data gap filling.

Phase II. Extension of Category & Data Gap Filling:

- Search for other analogues which are consistent with the submitted category.
- Data gap filling using new data matrix.

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - ***Case study***

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Case Study

- The submission consists of a category with **19 aliphatic amines**.
- The predicted ecotoxicological endpoint EC50, 48h, *D.magna* of **2-Butanamine(CAS 13952-84-6)** will be reviewed.

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Case study*
 - **Workflow**

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow

The following input workflow is used:

- Input the file for submission as an user list.(In this case input file Aliphatic amines.smi*)
- Evaluate the category applying the following profiling schemes:
 - US-EPA New Chemical Categories
 - Aquatic toxicity classification by ECOSAR
 - Aquatic toxicity MOA of action
 - Organic functional groups (nested)

Aliphatic amines.smi* - file is packed in the zip file with all power point tutorials posted on the LMC website.

Also it is placed in the example folder of the Toolbox installation.

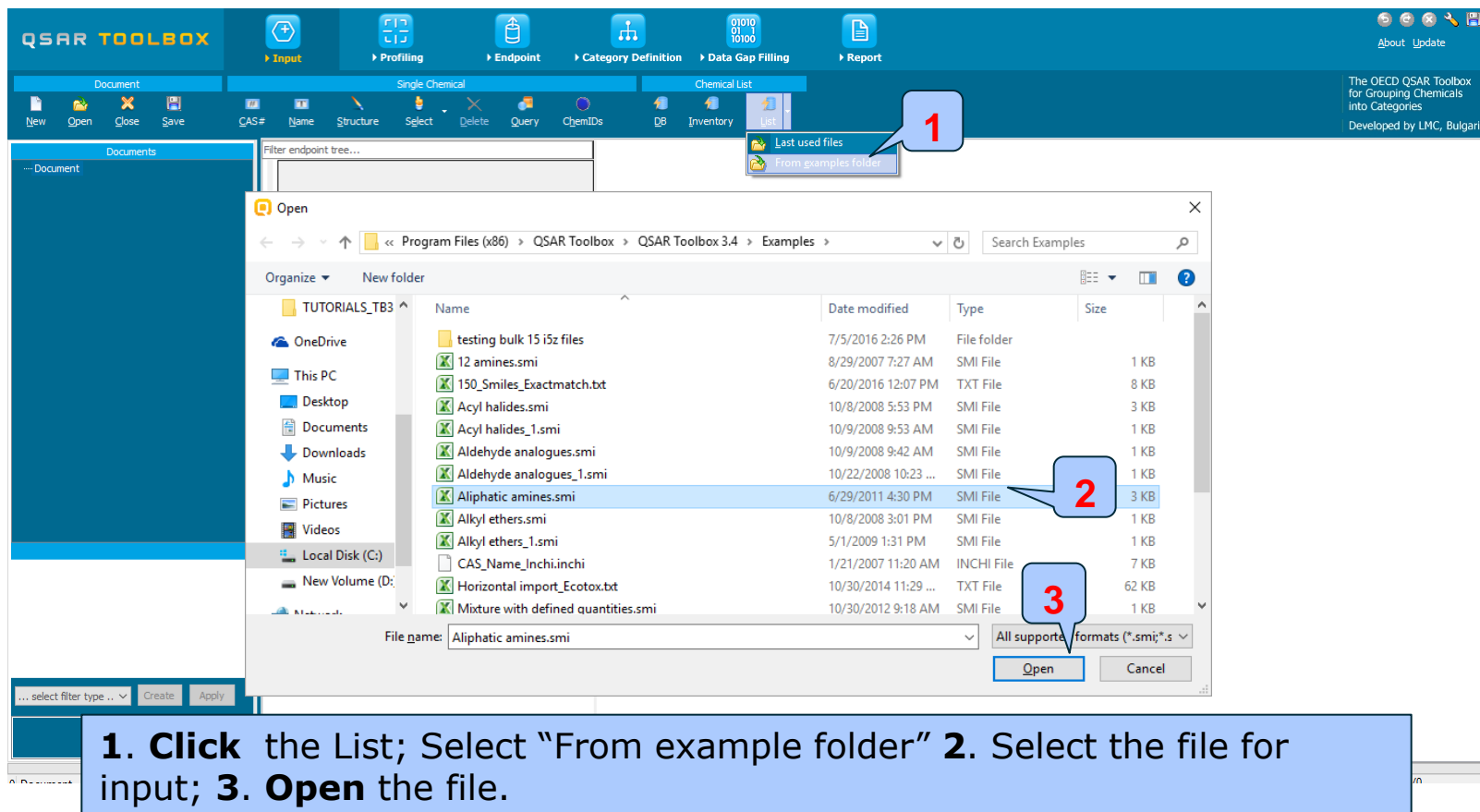
Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Case study*
 - ***Workflow***
 - ***Input***

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Input



The screenshot shows the QSAR Toolbox software interface. The 'Open' dialog box is open, displaying a list of files in the 'Examples' folder. The files listed are:

Name	Date modified	Type	Size
testing bulk 15 i5z files	7/5/2016 2:26 PM	File folder	
12 amines.smi	8/29/2007 7:27 AM	SMI File	1 KB
150_Smiles_Exactmatch.txt	6/20/2016 12:07 PM	TXT File	8 KB
Acyl halides.smi	10/8/2008 5:53 PM	SMI File	3 KB
Acyl halides_1.smi	10/9/2008 9:53 AM	SMI File	1 KB
Aldehyde analogues.smi	10/9/2008 9:42 AM	SMI File	1 KB
Aldehyde analogues_1.smi	10/22/2008 10:23 ...	SMI File	1 KB
Aliphatic amines.smi	6/29/2011 4:30 PM	SMI File	3 KB
Alkyl ethers.smi	10/8/2008 3:01 PM	SMI File	1 KB
Alkyl ethers_1.smi	5/1/2009 1:31 PM	SMI File	1 KB
CAS_Name_Inchi.inchi	1/21/2007 11:20 AM	INCHI File	7 KB
Horizontal import_Ecotox.txt	10/30/2014 11:29 ...	TXT File	62 KB
Mixture with defined quantities.smi	10/30/2012 9:18 AM	SMI File	1 KB

The 'Open' dialog box shows the file name 'Aliphatic amines.smi' and the file type 'All supported formats (*.smi;*.s)'. The 'Open' button is highlighted.

1. Click the List; Select "From example folder" **2. Select** the file for input; **3. Open** the file.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

***Workflow/*Input**

- You have now inserted your chemical list into the system.
- **Click** on the box next to "Substance Identity"; this displays the chemical identification information (see next screen shot).

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Input

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Document Single Chemical Chemical List

New Open Close Save CAS# Name Structure Select Delete Query ChemIDs DB Inventory List

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

Documents

Document

> [set][Cust]Chemicals file: Aliphatic amines.am

Filter endpoint tree... 1 [target] 2 [target] 3 [target] 4 [target] 5 [target] 6 [target] 7 [target] 8 [target] 9 [target]

Structure

1

Substance Identity

CAS Number	109-73-9	74-89-5	75-04-7	75-31-0	75-64-9	78-96-6	104-75-6	108-91-8	110-62-3
Chemical IDs	EINECS:2036992	EINECS:2008200	EINECS:2008347	EINECS:2008609	EINECS:2008881	EINECS:2011627	EINECS:2032338	EINECS:2036290	EINECS:2036290
Chemical Name	butylamine n-butylamine 1-butanamine 1-amino-butane 1-butylamine butan-1-amine	methylamine amino-methane methanamine monomethylamine methylamine, in aq...	ethylamine monoethylamine ethanamine amino-ethane ethyl amine	isopropylamine 2-amino-propane 2-amino-propane 2-propanamine propan-2-amine 2-propylamine mono-isopropylamine isopropyl amine	(tert)butylamine 2-propanamine, 2-... tert-butylamine 1-butylamine 2-methylpropan-2-...	1-amino-2-propanol 2-propanol, 1-amino- 1-amino-propan-2-ol monoisopropanola...	2-ethylhexylamine 1-hexanamine, 2-ethyl- 2-ethylhexan-1-amine hexylamine, 2-ethyl- 2-ethyl-1-hexanamine	cyclohexanamine cyclohexylamine, hcl cyclohexylamine aminocyclohexane cyclohexylamine (...)	1-cyclohexylamine n-cyclohexylamine 1-aminocyclohexane 1-cyclohexylamine (...)
Molecular Formula	C4H11N	CH5N	C2H7N	C3H9N	C4H11N	C3H9NO	C8H19N	C6H13N	C6H13N
Structural Formula	CCCCN	CN	CCN	CC(C)N	CC(C)CN	CC(O)CN	CCCC(C)CN	NC1CCCCC1	NC1CCCCC1
Physical Chemical Properties									
Environmental Fate and Transport									
Ecotoxicological Information									
Human Health Hazards									

CCCCN

select filter type... Create Apply

1. Click on "Substance identity"

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Case study*
 - **Workflow**
 - *Input*
 - **Profiling**

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

- The first step of the category evaluation according to Phase I is the consistency check of the category which starts with evaluation of the **robustness with respect to structural functionalities**. The following schemes could be used for this purpose:
 - US-EPA New Chemical Categories
 - Aquatic toxicity classification by ECOSAR
 - Aquatic toxicity MOA of action
 - Organic functional groups (nested)
- **Select "profiling methods"** by clicking on the boxes before the names of the profilers and **Click "Apply"**. Before selecting the profiling methods unselect all (see next screen shot).

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

1

2

Structure	1 [target]	2 [target]	3 [target]	4 [target]	5 [target]	6 [target]	7 [target]	8 [target]	9
Structure	<chem>CCCCN</chem>	<chem>CH5N</chem>	<chem>C2H7N</chem>	<chem>CC(C)N</chem>	<chem>C4H11N</chem>	<chem>C3H9NO</chem>	<chem>C8H19N</chem>	<chem>C6H13N</chem>	
Substance Identity									
- CAS Number	109-73-9	74-89-5	75-04-7	75-31-0	75-64-9	78-96-6	104-75-6	108-91-8	11
- Chemical IDs	EINECS:2036992	EINECS:2008200	EINECS:2008347	EINECS:2008609	EINECS:2008881	EINECS:2011627	EINECS:2032338	EINECS:2036290	11
- Chemical Name	butylamine n-butylamine 1-butylamine 1-amino-butane 1-butylamine butan-1-amine	methylamine amino-methane methanamine monomethylamine methylamine, in aq...	ethylamine monoethylamine ethanamine amino-ethane ethyl amine	isopropylamine 2-amino-propane 2-propanamine propan-2-amine 2-propylamine mono-isopropylamine isopropyl amine	(tert)butylamine 2-propanamine, 2-... tert-butylamine t-butylamine 2-methylpropan-2-... tert-butylamine (2-...	1-amino-2-propanol 2-propanol, 1-amino- 1-aminopropan-2-ol monoisopropanola...	2-ethylhexylamine 1-hexanamine, 2-e... 2-ethylhexan-1-amine 2-ethylhexyl amine hexylamine, 2-ethyl- 2-ethyl-1-hexanamine	cyclohexylamine cyclohexylamine hcl n-cyclohexylamine aminocyclohexane cyclohexylamine (... oc	
- Molecular Formula	C4H11N	CH5N	C2H7N	C3H9N	C4H11N	C3H9NO	C8H19N	C6H13N	Ci
- Structural Formula	CCCCN	CN	CCN	CC(C)N	CC(C)C)N	CC(O)CN	CCCC(C)CN	NC1CCCCC1	Ci
Physical Chemical Properties									
Environmental Fate and Transport									
Ecotoxicological Information									
Human Health Hazards									

19 Document

1. Go to profiling; 2. Select Unselect All.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

QSAR Toolbox 3.4.0.17 [Document]

Profiling methods

Substance Type

- USEPA New Chemical Categories

General Mechanistic Endpoint Specific

- Acute aquatic toxic. by OASIS
- Acute toxicity classification by ECOSAR
- Biodegradation - metabolism alerts
- Biodegradation - metabolism half-lives
- Biodegradation fragments (BioWIN MIT)
- Cardiotoxicity (genotox and nongenotoxic)
- DART scheme v.1.0
- DNA alerts for CA and MIT by OASIS
- DNA alerts for Ames and MIT by OASIS
- Eye irritation/corrosion Exclusion rules
- Eye irritation/corrosion Inclusion rules
- In vitro mutagenicity (Ames test) alerts
- In vivo mutagenicity (Micronucleus) alerts
- Keratinocyte gene expression
- Oncologic Primary Classification
- Protein binding alerts for Chromosomal
- Protein binding alerts for skin sensitization
- Respiratory sensitisation
- Retinoid Acid Receptor Binding
- RTE Expert System ver. 1 - USEPA
- Skin irritation/corrosion Exclusion rules
- Skin irritation/corrosion Inclusion rules

Empiric

- Chemical elements
- Groups of elements
- Lipinski Rule Oasis
- Organic Functional groups
- Organic Functional groups (nested)
- Organic functional groups (US EPA)

Metabolism/Transformations

Documented

Structure	1 [target]	2 [target]	3 [target]	4 [target]	5 [target]	6 [target]	7 [target]	8 [target]	9 [target]
Structure	<chem>CCCCN</chem>	<chem>CN</chem>	<chem>CCN</chem>	<chem>CC(C)N</chem>	<chem>CC(C)(C)N</chem>	<chem>CC(O)CN</chem>	<chem>CCCC(C)CN</chem>	<chem>NC1CCCCC1</chem>	
Substance Identity									
CAS Number	109-73-9	74-89-5	75-04-7	75-31-0	75-64-9	78-96-6	104-75-6	108-91-8	
Chemical IDs	EINECS:2036992	EINECS:2008200	EINECS:2008347	EINECS:2008609	EINECS:2008881	EINECS:2011627	EINECS:2032338	EINECS:2036290	
Chemical Name	butylamine n-butylamine 1-butanamine 1-amino-butane 1-butylamine butan-1-amine	methylamine amino-methane methanamine monomethylamine methylamine, in aq...	ethylamine monoethylamine ethanamine amino-ethane ethyl amine	isopropylamine 2-amino-propane 2-amino-propane 2-propanamine propan-2-amine 2-propylamine mono-isopropylamine isopropyl amine	(tert)butylamine 2-propanamine, 2-... tert-butylamine t-butylamine 2-methylpropan-2-... tert-butylamine (2-...	1-amino-2-propanol 2-propanol, 1-amino- 1-amino-propan-2-ol monoisopropanola...	2-ethylhexylamine 1-hexanamine, 2-e... 2-ethylhexan-1-amine 2-ethylhexyl amine hexylamine, 2-ethyl- 2-ethyl-1-hexanamine	cyclohexylamine, hcl cyclohexylamine aminocyclohexane cyclohexylamine (...)	
Molecular Formula	C4H11N	CH5N	C2H7N	C3H9N	C4H11N	C3H9NO	C8H19N	C6H13N	
Structural Formula	CCCCN	CN	CCN	CC(C)N	CC(C)(C)N	CC(O)CN	CCCC(C)CN	NC1CCCCC1	
Physical Chemical Properties									
Environmental Fate and Transport									
Ecotoxicological Information									
Human Health Hazards									

1. Check the USEPA, MOA of action, ECOSAR and OFG (nested) profilers; 2. Click Apply.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

- The actual profiling will take several seconds depending on the number and type of selected profilers.
- The results of profiling automatically appeared as a dropdown box under the target chemical.
- The result from profile statistics, can be seen by **right clicking** in the space above the profiler in the endpoint tree and **select Profile statistics** from the dropdown menu (see next screen shot).

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Profiling Profiling Schemes

Apply New View Delete

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

Filter endpoint tree...

Structure

Substance Identity
Physical Chemical Properties
Environmental Fate and Transport
Ecotoxicological Information
Human Health Hazards
Profile

Predefined
US-EPA New Chemical Categories
Endpoint Specific
Acute aquatic toxicity MOA by OASIS
Aquatic toxicity classification by ECOSAR
Empiric

Organic Functional groups (nested)

1

2

1 [target]	2 [target]	3 [target]	4 [target]	5 [target]	6 [target]	7 [target]	8 [target]	9 [target]
<chem>CCCCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>
Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines
Hide								
Show hidden								
Collapse all								
Sort (targets priority)								
Sort								
Function...								
Export CAS list								
Export								
Copy path								
Profile Statistic								

Metabolism/Transformations

Documented

1. Right click over the row with US-EPA profiler; 2. Select Profile Statistics.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to US-EPA

Profile statistic

Select profiles Report Stop

Stack mode Stacked to 100% Group by category

Profiles in use: US-EPA New Chemical Categories

#	Category	Count	%
1	Aliphatic Amines	19	100.00

Chemical groups

All chemicals

19 Save to smi Print... Add in new doc Add as target list Remove

1 CAS# 109-73-9 <chem>CCCCN</chem>	2 CAS# 74-89-5 <chem>CCN</chem>	3 CAS# 75-04-7 <chem>CCN</chem>	4 CAS# 75-31-0 <chem>CC(C)N</chem>	5 CAS# 75-64-9 <chem>CC(C)(C)N</chem>
6 CAS# 78-96-6 <chem>CC(O)N</chem>	7 CAS# 104-75-6 <chem>CCCCN</chem>	8 CAS# 108-91-8 <chem>C1CCNCC1</chem>	9 CAS# 111-86-4 <chem>CCCCCN</chem>	10 CAS# 1761-71-3 <chem>C1CCNCC1</chem>
11 CAS# 5332-73-0 <chem>CCCCCN</chem>	12 CAS# 13952-84-6 <chem>CC(C)N</chem>	13 CAS# 111-68-2 <chem>CCCCCN</chem>	14 CAS# 111-26-2 <chem>CCCCCN</chem>	15 CAS# 102-82-9 <chem>CCCCCN</chem>

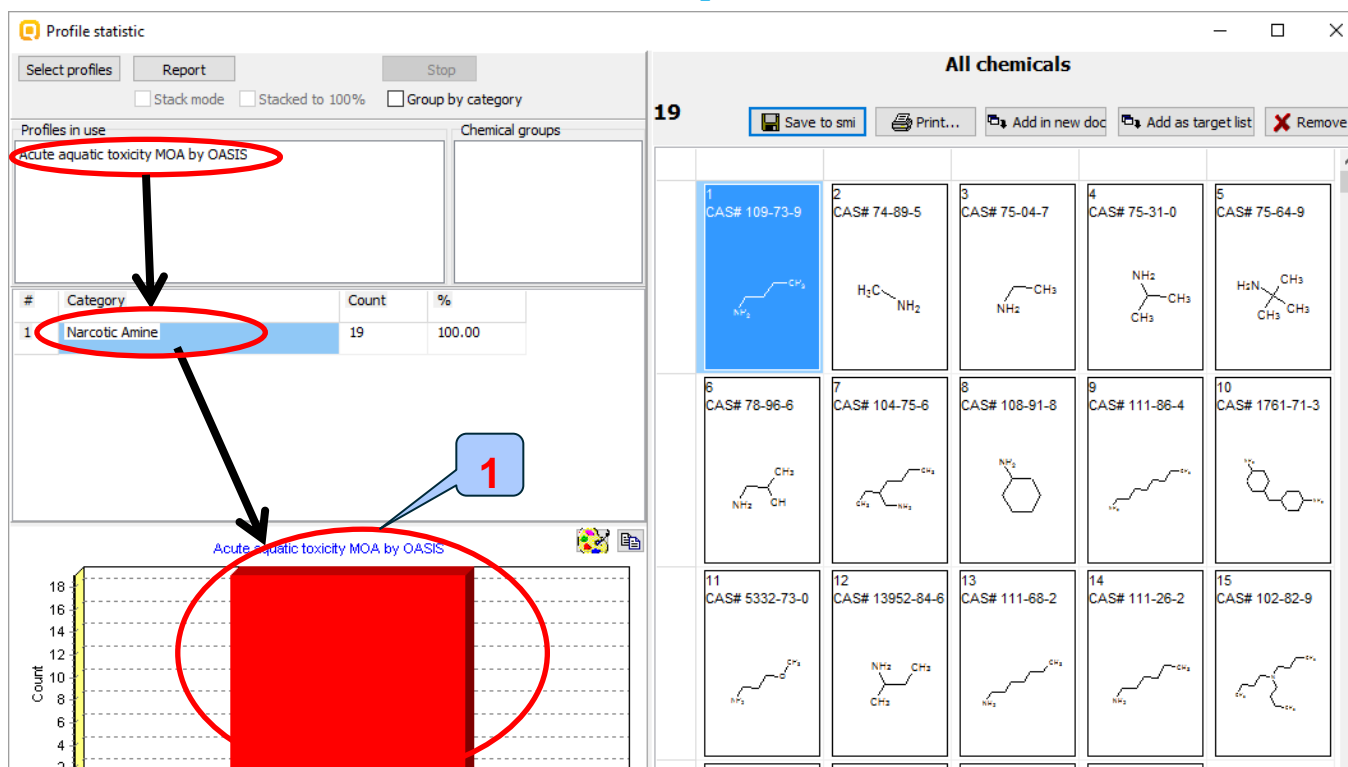
Count

1. In this case all 19 chemicals are Aliphatic amines according to US-EPA profiler.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to Aquatic toxicity MOA by OASIS



1. In this case all 19 chemicals are Narcotic amines according to MOA by OASIS profiling scheme

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to ECOSAR

Profile statistic

Select profiles Report Stop

Stack mode Stacked to 100% Group by category

Profiles in use: **Aquatic toxicity classification by ECOSAR**

Chemical groups

#	Category	Count	%
1	Aliphatic Amines	19	100.00

All chemicals

19 Save to smi Print... Add in new doc Add as target list Remove

1 CAS# 109-73-9
2 CAS# 74-89-5
3 CAS# 75-04-7
4 CAS# 75-31-0
5 CAS# 75-64-9
6 CAS# 78-96-6
7 CAS# 104-75-6
8 CAS# 108-91-8
9 CAS# 111-86-4
10 CAS# 1761-71-3
11 CAS# 5332-73-0
12 CAS# 13952-84-6
13 CAS# 111-68-2
14 CAS# 111-26-2
15 CAS# 102-82-9

Count

18
16
14
12
10
8
6
4
2

1. In this case all 19 chemicals are Aliphatic amines according to ECOSAR scheme

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Workflow/Profiling/Statistics according to OFG(nested)

The screenshot displays the 'Profile statistic' window with the following data:

#	Category	Count	%
1	Alcohol+Aliphatic Amine, primary	1	5.26
2	Aliphatic Amine, primary	10	52.63
3	Aliphatic Amine, primary+Alkane, branched	1	5.26
4	Aliphatic Amine, primary+Alkane, branched	1	5.26
5	Aliphatic Amine, primary+Cydoalkane+Over	1	5.26
6	Aliphatic Amine, primary+Ether	1	5.26

The bar chart below the table shows the count for each category, with the 'Aliphatic Amine, primary' bar reaching a count of 10. A blue callout box with the number '1' points to this bar.

The 'All chemicals' window on the right shows a list of 15 chemicals with their CAS numbers and chemical structures. The first chemical (CAS# 108-91-8) is highlighted in blue.

1. Aliphatic amine fragment is available.

Phase I: Evaluation of Category & Data Gap Filling

Step 1. Investigating the structural consistency of an ad hoc category

Recap

- **Chemicals are defined as:**
 - Aliphatic amines (broader category than primary amines)
 - US-EPA categories, ECOSAR classification
 - The statistics of organic functional groups provide detailed alert description of all 19 structures. However, all 19 chemicals have aliphatic amines fragment.
- **It could be concluded that the category is consistent with respect to structural functionalities (chemicals are empirically similar).**

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Step 2: Investigating the applicability domain of an ad hoc category.*

Phase I: Evaluation of Category & Data Gap Filling

Step 2. Investigating the applicability domain of an ad hoc category

- Boundaries of structural functionalities
 - Aliphatic amines
- Parametric boundaries
 - log Kow (from 0.64 to 7.71)
 - Molecular weight (from 31 to 269 Da)
 - Water solubility (from 0.48×10^{-1} to 1×10^6 mg/l)

You are now ready to extract the 2D and/or 3D parameters (see next screen shot).

Phase I: Evaluation of Category & Data Gap Filling


Step 2. Investigating the applicability domain of an ad hoc category

Extracting 2D and 3D parameters

1. Right click above endpoint tree; 2. Select Show hidden.

Phase I: Evaluation of Category & Data Gap Filling

Step 2. Investigating the applicability domain of an ad hoc category Extracting 2D and 3D parameters

1. Double click on the box  to open the nodes with 2D/3D parameters.

1 [target]	2 [target]	3 [target]	4 [target]	5 [target]	6 [target]	7 [target]	8 [target]
<chem>CCCCN</chem>	<chem>CCN</chem>	<chem>CCN</chem>	<chem>CN(C)C</chem>	<chem>CN(C)(C)C</chem>	<chem>CC(O)N</chem>	<chem>CCCCN</chem>	<chem>CCCCN</chem>
Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines
Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine	Narcotic Amine
Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines	Aliphatic Amines
Aliphatic Amine, pr...	Aliphatic Amine, pr...	Aliphatic Amine, pr...	Aliphatic Amine, pr...	Aliphatic Amine, pr...	Aliphatic Amine, pr...	Aliphatic Amine, pr...	Aliphatic Amine, pr...

Phase I: Evaluation of Category & Data Gap Filling

Step 2. Investigating the applicability domain of an ad hoc category

Extracting 2D and 3D parameters

The screenshot shows the QSAR Toolbox 3.4.0.17 interface. The 'Calculate structure' dialog box is open, displaying 'Current operation: MOPAC calculations Precise' and 'Conf. # 1/1 Start MOPAC (Precise): 15:12:34'. A context menu is open over the '2D' parameter tree, with the option 'Calculate/Extract all 2D parameters for all chemicals' selected. Callouts 1, 2, 3, and 4 highlight specific actions: 1 points to the right-click, 2 points to the selected menu item, 3 points to the 'Calculate all parameters for all chemicals' option, and 4 points to the 'Cancel' button in the dialog box.

! If some of the parameters (in case of new chemicals) are not calculated then the message (4) appears. If you want to stop the calculation process, click the **Cancel** button.

- 1. Right click** above the parameter in the endpoint tree; **2.** The user can calculate all parameters or extract current parameter for all chemicals in the current row. **3.** In this exercise calculate all (2D) parameters.

Phase I: Evaluation of Category & Data Gap Filling

Step 2. Investigating the applicability domain of an ad hoc category

Extracting 2D and 3D parameters - results

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Document Single Chemical Chemical List

New Open Close Save CAS# Name Structure Select Delete Query ChemIDs DB Inventory List

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

Documents

Document > [set][Cust]Chemicals file: Aliphatic amines.sm

Filter endpoint tree...

Structure

Substance Identity

Parameters

2D

- (Q) Acidic pKa (Chemaxon)
- (Q) Basic pKa (Chemaxon)
- Acidic Ionization at pH 1
- Acidic Ionization at pH 4
- Acidic Ionization at pH 7.4
- Acidic Ionization at pH 9
- BAF
- BAF (lower trophic)
- BAF (mid trophic)
- BAF (upper trophic)
- BAF (upper trophic, biotransformation rate is zero)
- Basic Ionization at pH 1
- Basic Ionization at pH 4
- Basic Ionization at pH 7.4
- Basic Ionization at pH 9
- BCF
- BCF (lower trophic)
- BCF (mid trophic)
- BCF (upper trophic)
- BCF (upper trophic, biotransformation rate is zero)
- Bio Half-Life
- Biodeg probability (Biowin 1)

CCCCN

Structure

select filter type... Create Apply

	1 [target]	2 [target]	3 [target]	4 [target]	5 [target]	6 [target]	7 [target]
Structure							
(Q) Acidic pKa (Chemaxon)	Can not calculate	Can not calculate	Can not calculate	Can not calculate	Can not calculate	15.3	Can not calc
(Q) Basic pKa (Chemaxon)	10.2	10.1	10.2	10.4	10.7	9.6	10.2
Acidic Ionization at pH 1	Can not calculate	Can not calculate	Can not calculate	Can not calculate	Can not calculate	4.96E-13 %	Can not calc
Acidic Ionization at pH 4	Can not calculate	Can not calculate	Can not calculate	Can not calculate	Can not calculate	4.96E-10 %	Can not calc
Acidic Ionization at pH 7.4	Can not calculate	Can not calculate	Can not calculate	Can not calculate	Can not calculate	1.25E-6 %	Can not calc
Acidic Ionization at pH 9	Can not calculate	Can not calculate	Can not calculate	Can not calculate	Can not calculate	4.96E-5 %	Can not calc
BAF	0.25 log(L/kg wet)	-0.04 log(L/kg wet)	-0.02 log(L/kg wet)	0.02 log(L/kg wet)	0.05 log(L/kg wet)	-0.05 log(L/kg wet)	1.75 log(L/kg wet)
BAF (lower trophic)	0.168 log(L/kg wet)	-0.02 log(L/kg wet)	-0.008 log(L/kg wet)	0.018 log(L/kg wet)	0.035 log(L/kg wet)	-0.025 log(L/kg wet)	1.57 log(L/kg wet)
BAF (mid trophic)	0.186 log(L/kg wet)	-0.023 log(L/kg wet)	-0.01 log(L/kg wet)	0.019 log(L/kg wet)	0.038 log(L/kg wet)	-0.028 log(L/kg wet)	1.62 log(L/kg wet)
BAF (upper trophic)	0.25 log(L/kg wet)	-0.039 log(L/kg wet)	-0.02 log(L/kg wet)	0.023 log(L/kg wet)	0.052 log(L/kg wet)	-0.046 log(L/kg wet)	1.75 log(L/kg wet)
BAF (upper trophic, biotransformation rate is zero)	0.28 log(L/kg wet)	-0.035 log(L/kg wet)	-0.012 log(L/kg wet)	0.037 log(L/kg wet)	0.066 log(L/kg wet)	-0.043 log(L/kg wet)	1.89 log(L/kg wet)
Basic Ionization at pH 1	100 %	100 %	100 %	100 %	100 %	100 %	100 %
Basic Ionization at pH 4	100 %	100 %	100 %	100 %	100 %	100 %	100 %
Basic Ionization at pH 7.4	99.8 %	99.8 %	99.9 %	99.9 %	99.9 %	99.4 %	99.8 %
Basic Ionization at pH 9	94.1 %	92.3 %	94.4 %	96.4 %	97.8 %	79.9 %	94.3 %
BCF	0.5 log(L/kg wet)	0.5 log(L/kg wet)	0.5 log(L/kg wet)	0.5 log(L/kg wet)	0.5 log(L/kg wet)	0.5 log(L/kg wet)	1.53 log(L/kg wet)
BCF (lower trophic)	0.168 log(L/kg wet)	-0.02 log(L/kg wet)	-0.008 log(L/kg wet)	0.018 log(L/kg wet)	0.035 log(L/kg wet)	-0.025 log(L/kg wet)	1.57 log(L/kg wet)
BCF (mid trophic)	0.186 log(L/kg wet)	-0.023 log(L/kg wet)	-0.01 log(L/kg wet)	0.019 log(L/kg wet)	0.038 log(L/kg wet)	-0.028 log(L/kg wet)	1.62 log(L/kg wet)
BCF (upper trophic)	0.25 log(L/kg wet)	-0.039 log(L/kg wet)	-0.02 log(L/kg wet)	0.023 log(L/kg wet)	0.052 log(L/kg wet)	-0.046 log(L/kg wet)	1.75 log(L/kg wet)
BCF (upper trophic, biotransformation rate is zero)	0.277 log(L/kg wet)	-0.035 log(L/kg wet)	-0.012 log(L/kg wet)	0.036 log(L/kg wet)	0.065 log(L/kg wet)	-0.043 log(L/kg wet)	1.85 log(L/kg wet)
Bio Half-Life	0.215 Days	0.0724 Days	0.0962 Days	0.125 Days	0.175 Days	0.025 Days	0.723 Days
Biodeg probability (Biowin 1)	0.975	0.887	0.88	0.873	0.683	1.02	0.948
Biodeg probability (Biowin 1)	0.800	0.875	0.87	0.861	0.700	0.800	0.800

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Step 2: Investigating the applicability domain of an ad hoc category.*
 - *Step 3: Reading data for the analogues*

Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data

- Next, the Data matrix is constructed by extracting available experimental results for all 19 members of the category.
- Based on preceding category evaluation, no outliers have been identified violating the structural and mechanistic consistency of the category.

Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data

The screenshot shows the QSAR Toolbox 3.4.0.17 interface. The left sidebar is titled 'Database' and contains a tree view of categories. The 'Ecotoxicological Information' section is expanded, and several sub-items are checked, including 'Aquatic ECOTOC', 'Aquatic Japan MoE', 'Aquatic OASIS', 'ECHA CHEM', and 'ECOTOX'. A red circle highlights this section, with a callout '1' pointing to it. The main window shows a 'Filter endpoint tree...' dialog box with 'Structure' selected. Below this, there are seven tabs labeled '1 [target]' through '7 [target]'. A 'Read data?' dialog box is open in the center, with 'All endpoints' selected and 'from Tautomers' checked. A callout '3' points to the 'OK' button. The background shows a table of chemical structures and their corresponding categories, such as 'Aliphatic Amines' and 'Narcotic Amine'.

1. Select databases related to aquatic toxicity; **2. Click** Gather; **3. Click** OK.

Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data

Due to the overlap between the Toolbox databases same data for intersecting chemicals could be found simultaneously in more than one database. The data redundancy is identified and the user has the opportunity to select either a single data value or all data values.

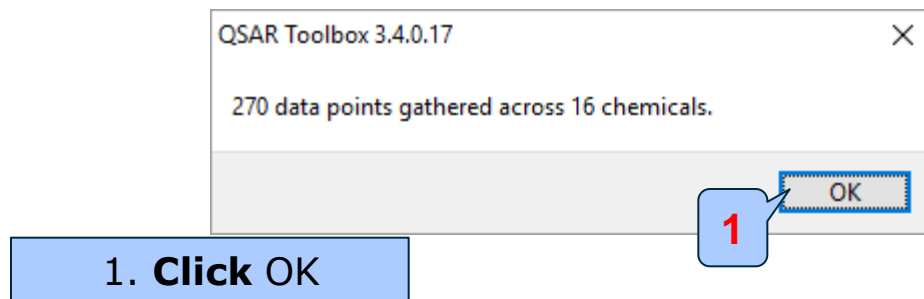
Data points...	Endpoint	CAS	Structure	Value	additional_com
<input checked="" type="checkbox"/>	LOEC	104-75-6	<chem>CCCC</chem>	360 micrograms per liter	
<input type="checkbox"/>	LOEC	104-75-6	<chem>CCCC</chem>	360 micrograms per liter	
<input checked="" type="checkbox"/>	LC50	109-73-9	<chem>CCCCN</chem>	2.68E5 micrograms per liter	
<input type="checkbox"/>	LC50	109-73-9	<chem>CCCCN</chem>	2.68E5 micrograms per liter	Affected fish lost schooling behavior and swam near the tank surface. They were hyperactive, had increased respiration, and lost equilibrium.

1. Click Select one; 2. Click OK
3. The indication for repeating values

Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data

The system automatically gives indication for the number of gather experimental data points



Phase I: Evaluation of Category & Data Gap Filling

Step 3. Reading data.

Inserting data into the data matrix

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Data Import Export Delete Tautomerize

Gather Import IUCLID5 Export IUCLID5 Database Inventory Database

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

Databases: Select All Unselect All Invert About

- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
 - Aquatic ECETOC
 - Aquatic Japan MoE
 - Aquatic OASIS
 - ECHA CHEM
 - ECOTOX
- Human Health Hazards

Inventories: Select All Unselect All Invert About

- Canada DSL
- COSING
- DSSTOX
- ECHA PR
- ENHES
- HPVC OECD
- METI Japan
- NICNAS
- REACH ECB
- RIFM_Inventory
- TSCA
- US EPA Challenge Program

Filter endpoint tree...

Structure

Substance Identity








Parameters

Physical Chemical Properties

Environmental Fate and Transport

Ecotoxicological Information

- Aquatic Toxicity (16/259)
 - Sediment Toxicity
 - Terrestrial Toxicity (5/11)
- Human Health Hazards
- Profile

	1 [target]	2 [target]	3 [target]	4 [target]	5 [target]	6 [target]	7 [target]
Structure							
Aquatic Toxicity	M: 139 mg/L, 237 ...	M: 938 mg/L, 0.12...	M: 226 mg/L, 226 ...	M: 448 mg/L, 1E3 ...	M: 136 mg/L, 581 ...	M: 2.54E3 mg/L, 2...	M: 15.9 mg/L, 7
Terrestrial Toxicity	M: 578 mg/L	M: 0.466 g	M: 1.17 g				

19 Document 1/0/0

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Step 2: Investigating the applicability domain of an ad hoc category.*
 - *Step 3: Reading data for the analogues*
 - *Step 4: Data gap filling for 2-Butanamine*

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine

1. Right click above the target chemicals. 2. Select Remove all as targets. 3. Go to Data Gap Filling

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine

QSAR Toolbox 3.4.0.17 [Document]

Document | Single Chemical | Chemical List

Filter endpoint tree...

Structure

Substance Identity

Parameters

Physical Chemical Properties

Environmental Fate and Transport

Ecotoxicological Information (16/270)

Human Health Hazards

Profile

10 11 12 13 14 15 16

M: 278 mg/L, 2

M: 56.9... M: 8(6.1;10.8) mg/... M: 178 mg/L, 16.6

Focus

- Add as target
- Set AOP target
- Add to Study Pad
- Select all as targets
- Remove all as targets
- Edit and add target
- Add target
- Add in category
- Delete
- Delete all except current
- Save to SMI file (DayLight format)
- Save to SMI file
- Print structures
- Export data for targets
- Export CAS list
- Query tool matrix Ctrl+F3

CCCCN

CCCCN

1. Scroll with the cursor and find out the target chemical (2-Butanamine)(#12)
2. Right click above the target chemical ;
3. Select Add as target;
4. Go to Data Gap Filling

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine

Navigate to the target endpoint: Intoxication, D.magna, EC50, 48h

1 Type Daphnia in the filter field; **2**. Navigate to target endpoint by opening the nodes and **3**. **Highlight** the gap which will be filled in; **4**. **Select** Trend analysis; **5**. **Click** Apply.

	10	11	12 [target]	13	14	15	16
Structure							
Substance Identity							
Parameters							
Ecotoxicological Information							
Aquatic Toxicity							
Behavior							
Immobilisation							
Intoxication							
EC50							
24 h							
48 h							
Animalia							
Arthropoda (Invertebrates)							
Branchiopoda (Branchiopods)							
Daphnia magna							
IC50							
LC0							
LC100							
LC50							
NOEL							
Mortality							

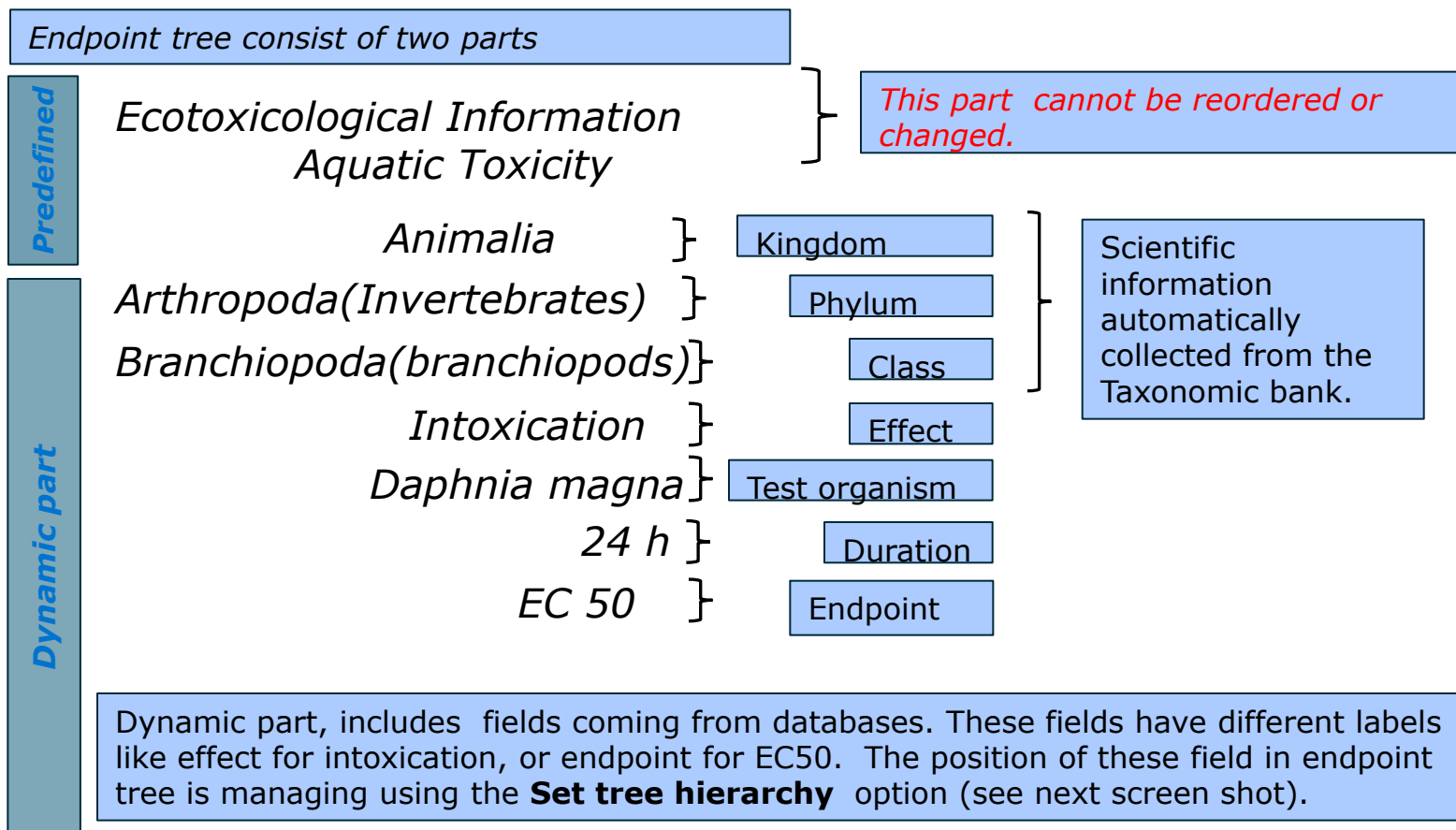
Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Step 2: Investigating the applicability domain of an ad hoc category.*
 - *Step 3: Reading data for the analogues*
 - *Step 4: Data gap filling for 2-Butanamine*
 - *Step 4.1*: How to reorder the endpoint tree*

Phase I: Evaluation of Category & Data Gap Filling

Step 4.1*. How to reorder the endpoint tree

* Detailed information describing how to reproduce the order of endpoint tree shown in the current presentation



Phase I: Evaluation of Category & Data Gap Filling

Step 4.1. How to reorder the endpoint tree

The screenshot shows the QSAR Toolbox interface with the 'Data Gap Filling' tab active. The left sidebar contains 'Data Gap Filling Method' (Read-across, Trend analysis, (Q)SAR models) and 'Target Endpoint' (Ecotoxicological Information Aquatic Toxicity Behavior). The main area displays a table of results for various endpoints across different chemical structures. A context menu is open over the endpoint tree, with three callouts: 1 points to a blue triangle indicating the tree is ordered; 2 points to a right-click action; 3 points to the 'Set tree hierarchy...' option in the menu.

1. The little blue triangle is indication that the endpoint tree below is ordered; To see order of nodes of the tree; **2.** Right click above endpoint tree near this triangle; **3.** Select Set tree hierarchy.

Phase I: Evaluation of Category & Data Gap Filling

Step 4.1. How to reorder the endpoint tree

The diagram illustrates the process of reordering an endpoint tree. It consists of three main parts:

- Conceptual Endpoint Tree:** A tree structure where nodes are labeled with italicized text and corresponding field names in blue boxes. The nodes are:
 - Intoxication* (Effect)
 - Daphnia magna* (Test organism)
 - 24 h* (Duration)
 - EC50* (Endpoint)
- Software Screenshot:** A window titled "Ecotoxicological Information#Aquatic Toxicity" showing a list of "Metabolic variables" on the left and "Sub-nodes" on the right. The "Metabolic variables" list includes: Assay, Experimental test system, Metabolic activation, Organ, Route, Route of administration, Sexual maturation (offspring), Strain, StudyType, Test guideline, Test type, Tissue, Type of genotoxicity, and Type of method. The "Sub-nodes" list includes: Effect, Endpoint, Duration, and Test organisms (species). A red circle highlights the "Metabolic variables" list, and a black arrow points from this circle to the "Sub-nodes" list.
- Annotations:**
 - A blue box on the left states: "Left panel contains most usable labels of fields coming from databases." with an arrow pointing to the "Metabolic variables" list.
 - A blue box on the right states: "The endpoint tree is organized following the order shown in right panel" with an arrow pointing to the "Sub-nodes" list.

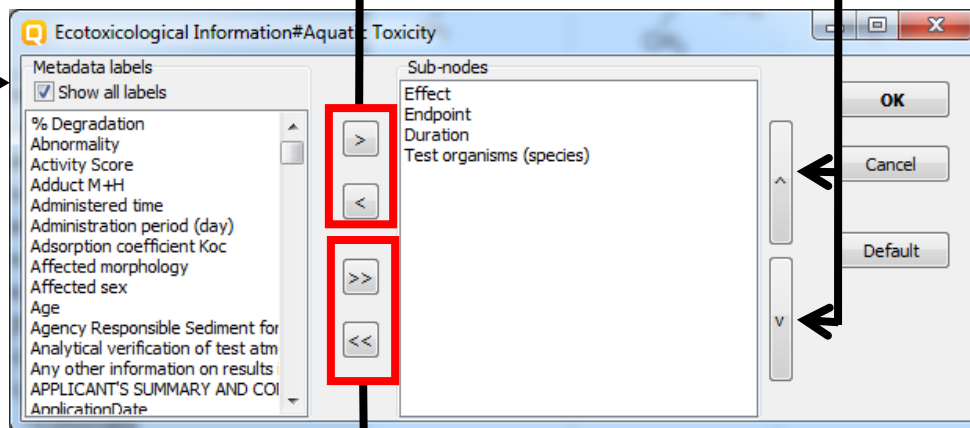
Phase I: Evaluation of Category & Data Gap Filling

Step 4.1. How to reorder the endpoint tree

All available fields coming from databases could be shown when the box **Show all labels** is selected.

To remove or change some of the fields which are visualizing on the tree, highlight the field from left panel and move it to the right panel using the **Right** or **Left** button.

Reordering of the fields is possible using the **Up** and **Down** buttons.



To move all fields form one panel to another use **Right** and **Left** button.

Outlook

- Background
- Category evaluation – overview
- **Phase I. Evaluation of Category & Data Gap Filling**
 - *Step 1: Investigating the structural consistency of an ad hoc category.*
 - *Step 2: Investigating the applicability domain of an ad hoc category.*
 - *Step 3: Reading data for the analogues*
 - *Step 4: Data gap filling for 2-Butanamine*
 - *Back to Trend analysis*

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

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

Data Gap Filling Method

- Read-across
- Trend analysis
- (Q)SAR models

Target Endpoint

Ecotoxicological Information Aquatic Toxicity Intoxication EC50 48 h Animalia Arthropoda Branchiopoda Daphnia magna

Daphnia

Structure

Daphnia magna (11/13)

M: >100 mg/L M: 163(147.180) m... M: 198(167.231) m... M: 48.6 mg/L M: 1.9(1.5,2.4) mg/L M: 9.4(6.9,12)

Descriptors Prediction Adequacy Cumul. freq. Statistics Residuals

Trend analysis prediction of EC50, making a linear approximation, based on 11 values from 11 analogue chemicals, Observed target value: N/A, Predicted target value: 64.9 mg/L

Model equation: $EC50 = +2.66 + 0.516 * \log Kow$

Descriptor X: log Kow

Accept prediction

Return to matrix

- Select/filter data
- Selection navigation
- Gap filling approach
- Descriptors/data
- Model/(Q)SAR
- Calculation options
- Visual options
- Information
- Miscellaneous

19 Document Create prediction by gap filling 0/1 1/1/0

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine/Subcategorize by OFG(nested)

The screenshot shows the 'Subcategorization' window in the OECD QSAR Toolbox. The main window displays a scatter plot of log Kow vs EC50 (obs., log1(mol/L)) with a red trend line. The plot title is 'Trend analysis prediction of EC50, making a linear approximation, based on 11 values from 11 analogue chemicals, Observed target value: N/A, Predicted target value: 64.9 mg/L, Model equation: EC50 = +2.66 +0.516 * log Kow'. Several data points are highlighted with callouts showing chemical structures: 2-Butanamine, N,N-Dimethyl-2-butanamine, and N,N-Diethyl-2-butanamine. A blue box at the bottom contains the text '1. Remove dissimilar chemicals'.

1. Remove dissimilar chemicals

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine/Predicted result

The screenshot shows the 'Data Gap Filling' step in the OECD QSAR Toolbox. The main window displays a trend analysis plot of log Kow vs log EC50. The plot includes the model equation: $EC50 = +2.34 + 0.822 * \log Kow$. A callout box labeled '1' points to the predicted result for 2-Butanamine, which is 79.0 mg/L. The plot also shows the observed target value: N/A, and the predicted target value: 79.0 mg/L. The interface includes a list of chemical structures and their predicted EC50 values.

Chemical Structure	M	EC50 (mg/L)
<chem>CCN</chem>	163	(147-180) mg/L
<chem>CCCCN</chem>	19	(15.2-24) mg/L
<chem>CCCCCN</chem>	9.4	(6.9-12) mg/L
<chem>CCCCCCN</chem>	8.6	(7-11) mg/L
<chem>CCCCCN</chem>	56	(27-170) mg/L
<chem>CCCCCN</chem>	0.58	(0.39-0.79) mg/L

1. Predicted result is 79.0 mg/l

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine/Cumulative frequency

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Filling Apply

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

Data Gap Filling Method

- Read-across
- Trend analysis
- (Q)SAR models

Target Endpoint

Ecotoxicological Information Aquatic Toxicity Intoxication EC50 48 h Animalia Arthropoda Branchiopoda Daphnia magna

Daphnia

Structure

Daphnia magna (7/8)

CCN
CCCCN
CCCCCN
CCCCCCN
CCCCCN
CCCCCCN

M: 163(147,180) m... M: 1.9(1.5,2.4) mg/L M: 9.4(6.9,12) mg/L M: 8.6(7,11) mg/L M: 56(27,170) mg/L M: 0.58(0.39,0.79) ...

Descriptors Prediction Adequacy Cumul. freq. Statistics Residuals

95% of Residuals = < 0.243, log(1/mol/L)

Cumulative frequency, %

Residuals, Y - Y_{calc}

Accept prediction

Return to matrix

- Select/filter data
 - Subcategorize
 - Mark chemicals by WS
 - Mark chemicals by descriptor value
 - Mark outlier points
 - Filter points by test conditions
 - Mark focused chemical
 - Mark focused points
- Selection navigation
- Gap filling approach
 - Descriptors/data
 - Model/(Q)SAR
 - Calculation options
 - Visual options
 - Information
 - Miscellaneous

1. Cumulative frequency is less then 0.24 log units

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine/Statistics

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

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

Filling
Apply

Data Gap Filling Method

- Read-across
- Trend analysis
- (Q)SAR models

Target endpoint

Ecotoxicological Information Aquatic Toxicity Intoxication EC50 48 h Animalia Arthropoda Branchiopoda Daphnia magna

Daphnia

Structure

Daphnia magna

1

CCCCN CCCCCN CCCCCCN CCCCCN CCCCCN CCCCCN

M: 163(147,180) m... M: 1.9(1.5,2.4) mg/L M: 9.4(6.9,12) mg/L M: 8.6(7,11) mg/L M: 56(27,170) mg/L M: 0.56(0.39,0.79) ...

Descriptors	Prediction	Adequacy	Cumul. freq.	Statistics	Residuals
Statistical characteristics		TA model			
Number of data points, (N)				7	
Coefficient of determination, (R2)				0.974	
Adjusted coefficient of determination, (R2adj)				0.969	
Coefficient of determination -leave one out-, (Q2)				0.949	
Coefficient of correlation for external set, (r2)				-	
Sum of squared residuals, (SSR)				0.220	
Standard deviation of residuals, (sN)				-	
Sample standard deviation of residuals, (s)				0.210	
Fisher function, (F)				189	
Fisher threshold for statistical significance, (Fa)				10.6	
b0		Intercept			
- model descriptor				2.34	
- coeff. value				± 0.34	
- coeff. range				Yes	
- significance				0.377 (vs b1)	
- max. covariation					
b1		log Kow			
- model descriptor				0.822	
- coeff. value				± 0.154	
- coeff. range				Yes	
- significance				0.377 (vs b0)	
- max. covariation					

Accept prediction

Return to matrix

- Select/filter data
 - Subcategorize
 - Mark chemicals by WS
 - Mark chemicals by descriptor value
 - Mark outlier points
 - Filter points by test conditions
 - Mark focused chemical
 - Mark focused points
- Selection navigation
- Gap filling approach
- Descriptors/data
- Model/(Q)SAR
- Calculation options
- Visual options
- Information
- Miscellaneous

Phase I: Evaluation of Category & Data Gap Filling

Step 4. Data Gap Filling for 2-Butanamine/Interpretation of the result

- The structurally similar analogs across category of aliphatic amines is used for data gap filling
- Subcategorization by Organic functional groups(nested) is applied
- The prediction based on the defined category is acceptable.
- The predicted value based on predefined category of aliphatic amines is 79.0 mg/l

Outlook

- Background
- Category evaluation – overview
 - Phase I. Evaluation of Category & Data Gap Filling
 - **Phase II. Extension of the Category& Data gap Filling**

Phase II: Extension Category & Data gap filling

- The extension of the category is performed by using Phase II of the category evaluation process (*Extension of Category & Data Gap Filling*). Other analogues are searched in the Toolbox, which are structurally and mechanistically consistent with the predefined category.
- The structural analogues could be defined using ECOSAR grouping

Outlook

- Background
- Category evaluation – overview
 - Phase I. Evaluation of Category & Data Gap Filling
 - **Phase II. Extension of the Category& Data gap Filling**
 - Step 1: Category definition

Phase II: Extension Category & Data gap filling

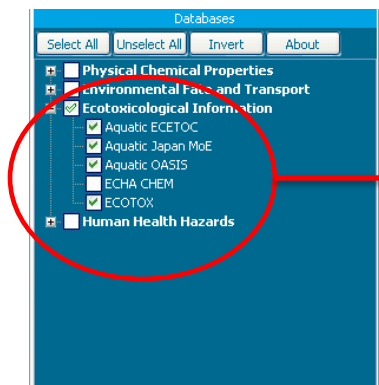
Step 1: Category definition

- The ECOSAR strict category is used to define a broader category used in further analysis.
- According to ECOSAR categorization **a list of 302 aliphatic amines have been selected** for which data are available in the Toolbox (identified in the Toolbox databases).
- The same endpoint: EC 50 48h *D.magna* will be predicted as with the predefined category

Phase II: Extension Category & Data gap filling

Step 1: Category definition

- Define ECOSAR category
- Before defining the category, the following databases related to the predicted endpoint are selected:



Aquatic ECETOC
 Aquatic Japan MoE
 Aquatic OASIS
 ECOTOX

Phase II: Extension Category & Data gap filling

Step 1: Category definition/Defining ECOSAR (strict)

1. Select the target chemical; 2. Highlight "Aquatic toxicity classification by ECOSAR"; 3. Click Define; 4. Select Strict. 5. Click OK

Phase II: Extension Category & Data gap filling

Step 1: Category definition/Defining ECOSAR (strict)

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint **Category Definition** Data Gap Filling Report

Define Define with metabolism Subcategorize Combine Clustering Delete Delete All

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

Grouping methods

- Protein binding by OASIS v1.4
- Protein binding by OECD
- Protein binding potency
- Superfragments
- Toxic hazard classification by Cramer (extension)
- Toxic hazard classification by Cramer (original)
- Ultimate biodeg

Endpoint Specific

- Acute aquatic toxicity classification by Verhaar (Modified)
- Acute aquatic toxicity MOA by OASIS
- Aquatic toxicity classification by ECOSAR
- Bioaccumulation - metabolism alerts
- Bioaccumulation - metabolism half-lives
- Biodegradation fragments (BioWIN MTTI)
- Carcinogenicity (genotox and nongenotox) alerts by ISS
- DART scheme v.1.0
- DNA alerts for Ames by OASIS v. 1.4
- DNA alerts for CA and MNT by OASIS v. 1.1
- Eye irritation/corrosion Exclusion rules by BfR
- Eye irritation/corrosion Inclusion rules by BfR
- in vitro mutagenicity (Ames test) alerts by ISS
- in vivo mutagenicity (Micronucleus) alerts by ISS
- Keratinocyte gene expression
- Oncologic Primary Classification
- Protein binding alerts for Chromosomal aberration by OECD
- Protein binding alerts for skin sensitization by OASIS v. 1.4
- Respiratory sensitisation
- Retinoic Acid Receptor Binding
- rTER Expert System ver. 1 - USEPA
- Skin irritation/corrosion Exclusion rules by BfR
- Skin irritation/corrosion Inclusion rules by BfR

Empiric

Defined Categories

- Document
- [3] Subcategorized in GP: Organic Functional groups (hes)

Daphnia

Structure

Substance Identity

Parameters

Ecotoxicological Information

- Aquatic Toxicity
 - Behavior
 - (4/12) M: 9.5 mg/L, 18 m... M: 49 mg/L, 120 m...
 - Immobilisation
 - (1/1) M: 36 mg/L
 - Intoxication
 - EC50
 - (12/14) M: 48.6 m... M: 1.9(1.5;2.4) mg/L
 - 24 h
 - T: 79(20;311) mg/L M: 9.4(6...
 - 48 h
 - M: 58(47,72) mg/L
 - Animalia
 - Arthropoda (Invertebrates)
 - Branchiopoda (Branchiopods)
 - Daphnia magna
 - (2/2) M: 58(47,72) mg/L

EC50 (2/2) M: 58(47,72) mg/L

LC0 (4/4) M: 0.62 mg/L M: 50 mg/L

LC100 (4/4) M: 5 mg/L M: 120 mg/L

LC50 (4/4) M: 2.2 mg/L M: 80 mg/L

NOEL (1/1)

Mortality (2/2) M: 58.4 mg/L

Photoinduced Toxicity

Reproduction (1/2) M: 1.6 mg/L, 3.9 ...

Profile

Define category name

Category name (302 chemicals) **Strict (Aquatic toxicity classification by ECOSAR)**

OK Cancel

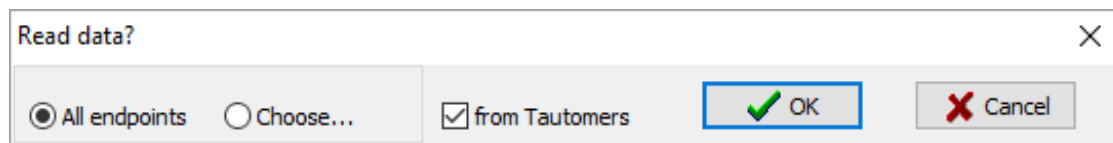
1

1. Click OK to confirm the name of the category.

Phase II: Extension Category & Data gap filling

Step 1: Category definition/Reading data

- The Toolbox will now retrieve those chemicals that have the same ECOSAR functionality as the target compound.
- The Toolbox automatically request the user to select the endpoint that should be retrieved.
- The user can either select the specific endpoint or by default choose to retrieve data on all endpoints (see below).



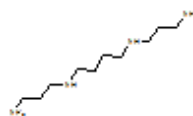
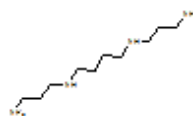
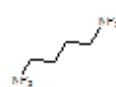
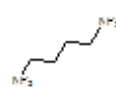
- In this example, as only databases are selected that contain information for aquatic toxicity endpoint, both options give the same results.

Phase II: Extension Category & Data gap filling

Step 1: Category definition/Reading data

Repeated values for: 1406 data-points, 422 groups, 168 chemicals

Data points...

	Endpoint	CAS	Structure	Value	additional_comments
<input checked="" type="checkbox"/>	NOEC	71-44-3		1 milliMolar (millimoles per liter)	SOURCE/COLLECT FROM VERAVAL COAST,
<input checked="" type="checkbox"/>	NOEC	71-44-3		1 milliMolar (millimoles per liter)	SOURCE/COLLECT FROM VERAVAL COAST,
<input checked="" type="checkbox"/>	NOEC	110-60-1		1 milliMolar (millimoles per liter)	SOURCE/COLLECT FROM VERAVAL COAST,
<input checked="" type="checkbox"/>	NOEC	110-60-1		1 milliMolar (millimoles per liter)	SOURCE/COLLECT FROM VERAVAL COAST,

1

Select one

Invert

Check All

Uncheck All

2

OK

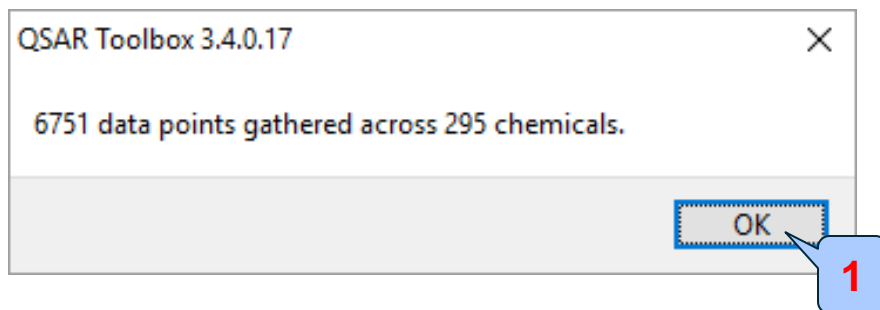
Cancel

1. Click Select one and then; **2. Click** OK.

Phase II: Extension Category & Data gap filling

Step 1: Category definition/Reading data

The system automatically gives indication for the number of gather experimental data points



1. **Click OK**

Outlook

- Background
- Category evaluation – overview
 - Phase I. Evaluation of Category & Data Gap Filling
 - **Phase II. Extension of the Category& Data gap Filling**
 - Step 1: Category definition
 - Step 2: Navigate to the target endpoint

Phase II: Extension Category & Data gap filling

Step 2: Navigate to the target endpoint/EC 50, 48h, D.magna

The screenshot shows the QSAR Toolbox 3.4.0.17 interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Data Gap Filling' menu item is highlighted with a red callout '3'. The left sidebar shows a tree view of 'Databases' and 'Inventories'. The 'Databases' tree is expanded to 'Ecotoxicological Information' and 'Human Health Hazards'. The 'Inventories' list includes 'Canada DSL', 'COSING', 'DSSTOX', 'ECHA PR', 'EINECS', 'HPVC OECD', 'METI Japan', 'NICNAS', 'REACH ECB', 'RIFM_Inventory', 'TSCA', and 'US EPA Challenge Program'. The main window displays a table of data for 'Daphnia' endpoints. The table has columns for '1 [target]', '2', '3', '4', '5', '6', and '7'. The 'EC50' endpoint is expanded to show '24 h', '48 h', and '21 Days'. The '48 h' endpoint is further expanded to show 'Animalia', 'Arthropoda (Invertebrates)', 'Branchiopoda (Branchiopoda)', 'Ceriodaphnia dubia', 'Daphnia magna', and 'Daphnia pulex'. The 'Daphnia magna' row is highlighted in blue, with a red callout '2' pointing to the gap in the 'EC50' column. The 'Daphnia magna' row shows 'M: 5.76(4.34;7.55) ... M: 70.1(51.1;99.3) ... M: 73.7(62.6;91.2) ... M: 8(6'.

1. Navigate to target endpoint by **opening the nodes**; **2.** Highlight the gap which will be filled in; **3.** Move to Data Gap Filling

Outlook

- Background
- Category evaluation – overview
 - Phase I. Evaluation of Category & Data Gap Filling
 - **Phase II. Extension of the Category& Data gap Filling**
 - Step 1: Category definition
 - Step 2: Navigate to the target endpoint
 - Step 3: Data Gap Filling

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling

QSAR Toolbox 3.4.0.17 [Document]

Input Profiling Endpoint Category Definition **Data Gap Filling** Report

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

1 Read-across Trend analysis (Q)SAR models

2 Apply

Structure

Target	1 [target]	2	3	4	5	6	7
Structure	<chem>CC(N)C</chem>	<chem>CCCCCCCCCCCCCCCC</chem>	<chem>c1ccc(cc1)C(=O)c2ccccc2</chem>	<chem>CC1CCCCC1</chem>	<chem>C1CCNCC1</chem>	<chem>CCCC(O)C</chem>	
Aquatic Toxicity							
Behavior	(13/41)						
Biochemistry	(2/4)						
Development	(1/4)						
Enzyme(s)	(2/5)						
Feeding Behavior	(1/1)						
Genetics	(1/2)						
Growth	(8/34)						
Immobilisation	(24/24)		M: 10 mg/L		M: 8 mg/L		
Intoxication							
EC50							
24 h	(41/44)	M: >180 mg/L		M: 26.9(19.7;40.4) ...	M: 192(148;272) m...	M: 188(153;241) m...	M: 19...
48 h							
Animalia							
Arthropoda (Invertebrates)							
Branchiopoda (Branchiopods)							
Ceriodaphnia dubia	(4/4)						
Daphnia magna	(67/78)			M: 5.76(4.34;7.55) ...	M: 70.1(51.1;99.3) ...	M: 73.7(62.6;91.2) ...	M: 8(6...
Daphnia pulex	(1/2)						
52 h	(1/1)						
96 h	(1/1)						
21 Days	(1/1)						
LC50	(7/7)						
LC0	(10/10)						
LC100	(10/10)						

302 Aliphatic Amines Strict (Aquatic toxicity classification by ECOSAR)

1. Select Trend analysis; 2. Click Apply

1/0/0

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

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

Filling

Apply

Data Gap Filling Method

- Read-across
- Trend analysis
- (Q)SAR models**

Target Endpoint

Ecotoxicological Information Aquatic Toxicity Intoxication EC50 48 h Animalia Arthropoda Branchiopoda Daphnia magna

Daphnia

1 [target] 4 5 6 7 8 10

Structure

Daphnia magna (67/78)

M: 5.76(4.34,7.55) M: 70.1(51.1,99.3) M: 73.7(62.6,91.2) M: 8(6.1,10.8) mg/L M: 38.5(30.49,4) m M: >1

Descriptors Prediction Adequacy Cumul. freq. Statistics Residuals

Trend analysis prediction of EC50, making a linear approximation, based on 67 values from 67 analogue chemicals, Observed target value: N/A, Predicted target value: 7.29 mg/L, Model equation: $EC50 = +3.78 + 0.297 * \log Kow$

EC50 (obs.), log1(mg/L)

log Kow

Descriptor X: log Kow

Accept prediction

Return to matrix

- Select/filter data
- Selection navigation
- Gap filling approach
- Descriptors/data
- Model/(Q)SAR
- Calculation options
- Visual options
- Information
- Miscellaneous

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/Subcategorize by Water solubility

QSAR Toolbox 3.4.0.17 [Document]

Input Profiling Endpoint Category Definition **Data Gap Filling** Report

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

Data Gap Filling Method

- Read-across
- Trend analysis
- (Q)SAR models**

Target Endpoint

Ecotoxicological Information Aquatic Toxicity Intoxication EC50
48 h Animalia Arthropoda Branchiopoda Daphnia magna

Daphnia 1 [target] 4 5 6 7 8 10

Structure

Daphnia magna (67/78) M: 5.76(4.34.7.55) M: 70.1(51.1.99.3) M: 73.7(62.6.91.2) M: 8(6.1.10.8) mg/L M: 38.5(30.49.4) m M: >1

Descriptors Prediction Adequacy Cumul. freq. Statistics Residuals

Trend analysis prediction of EC50, making a linear approximation, based on 67 values from 67 analogue chemicals, Observed target value: N/A, Predicted target value: 7.29 mg/L

Model equation: $EC50 = +3.78 + 0.297 * \log Kow$

Descriptor X: log Kow

Accept prediction 1

Return to matrix

Select/filter data

- Subcategorize
- Mark chemicals by WS**
- Mark chemicals by descriptor value
- Mark outlier points
- Filter points by test conditions
- Mark focused chemical
- Mark focused points

Selection navigation

- Gap filling approach
- Descriptors/data
- Model/(Q)SAR
- Calculation options
- Visual options

Select a WS descriptor:

- Name
- Exp Water Solubility
- Water Solubility mg/L
- Water Solubility (fragments) log(mol/L)
- WS Multicase log(mol/m3)

2 3 4

1. Open Select/Filter data panel; **2. Click** Mark chemical by WS; **3. Select** Water solubility (fragments); **4. Click** OK

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/Subcategorize by Water solubility

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

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

Data Gap Filling Method

- Read-across
- Trend analysis
- (Q)SAR models**

Target Endpoint

Ecotoxicological Information Aquatic Toxicity Intoxication EC50
48 h Animalia Arthropoda Branchiopoda Daphnia magna

Daphnia

1 [target] 4 5 6 7 8 10

Structure

Daphnia magna (67/78)

M: 5.76(4.34;7.55) M: 70.1(51.1;99.3) M: 73.7(62.6;91.2) M: 8(6.1;10.8) mg/L M: 38.5(30.49;4 m M: >1

Descriptors Prediction Adequacy Cumul. freq. Statistics Residuals

Information

Marked 9 point(s) for 9 chemical(s)

OK

EC50 obs., log(f time/L)

log Kow

Descriptor X: log Kow

Accept prediction

Return to matrix

Select/filter data

- Subcategorize
- Mark chemicals by WS**
- Mark chemicals by descriptor value
- Mark outlier points
- Filter points by test conditions
- Mark focused chemical
- Mark focused points
- Invert existing marks
- Remove marked chemicals/points
- Clear existing marks

Selection navigation

- Gap filling approach
- Descriptors/data
- Model/(Q)SAR
- Calculation options
- Visual options
- Information
- Miscellaneous

1. Nine chemicals are marked in green; 2. Click OK; 3. Remove marked chemicals

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/Subcategorize by OFG(nested)

The screenshot displays the 'Subcategorization' window in the OECD QSAR Toolbox. The 'Grouping methods' list on the left includes 'Organic functional groups (nested)' which is highlighted with a red box labeled '2'. The main workspace shows a chemical structure and a table of analogues for *Daphnia magna* with columns for predicted values and observed values. A 'Trend analysis prediction of EC50' plot is shown, with the model equation $EC50 = +3.75 + 0.289 * \log Kow$. The right sidebar has 'Subcategorize' selected, highlighted with a red box labeled '1'. At the bottom of the 'Subcategorization' window, the 'Remove' button is highlighted with a red box labeled '3'.

1. Click Subcategorize;
2. Select Organic functional groups(nested);
3. Remove dissimilar chemicals

302 Aliphatic Amines Strict (Aquatic toxicity classification)

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/Subcategorize by Lipinski rules

The screenshot displays the QSAR Toolbox interface during the subcategorization phase. The left sidebar lists various grouping methods, with 'Lipinski Rule Oasis' selected under the 'Empiric' category. The top navigation bar shows the 'Data Gap Filling' step. A dialog box titled '4 structures from: Not bioavailable' is open, showing four chemical structures with their CAS numbers and Lipinski Rule Oasis status. A red circle highlights four points on the plot that are outliers from the main trend line. Numbered callouts 1-4 indicate specific actions: 1. Selecting Lipinski rules in the sidebar; 2. Double-clicking on the 'Not bioavailable' category in the dialog; 3. Closing the dialog box; 4. Removing dissimilar chemicals from the plot.

Chemicals with very long chain could be removed from the category due to their non-bioavailability. **1. Select** Lipinski rules; **2. Double click** to see "Not- bioavailable" chemicals; **3. Close** the appeared window; **4. Remove** dissimilar chemicals

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/Prediction result

The screenshot shows the 'Data Gap Filling' step in the OECD QSAR Toolbox. The main window displays a trend analysis plot of EC50 (log10 mg/L) vs log Kow. The plot shows 11 data points and a linear regression line. A callout box labeled '1' points to the predicted target value of 77.8 mg/L. The interface includes a sidebar with grouping methods, a top navigation bar, and a right-hand panel with prediction options.

Trend analysis prediction of EC50, making a linear approximation, based on 11 values from 11 analogue chemicals, Observed target value: N/A, Predicted target value: 77.8 mg/L, Model equation: $EC50 = +2.34 + 0.838 * \log Kow$

1

302 Aliphatic Amines Strict (Aquatic toxicity classification by ECOSAR)

Prediction result is 77.8 mg/l.

Phase II: Extension Category & Data gap filling

Step 3: Data Gap Filling/Accept the Prediction result

QSAR Toolbox 3.4.0.17 [Document]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

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

Data Gap Filling Method

- Read-across
- Trend analysis
- (Q)SAR models**

Target Endpoint

Ecotoxicological Information Aquatic Toxicity Intoxication EC50 48 h Animalia Arthropoda Branchiopoda Daphnia magna

Daphnia

Structure

Daphnia magna (12/13)

M: 1.6(1.2,2) mg/L M: 0.026(0.018,0.0) M: 0.58(0.39,0.79) M: 0.5(0.4,0.61) mg/L M: 163(14) M: 56(27,170) mg/L

Descriptors Prediction Adequacy Cumul. freq. Statistics Residuals

Information

The current prediction was accepted

OK

Accept prediction

Return to matrix

- Select/filter data
 - Subcategorize
 - Mark chemicals by WS
 - Mark chemicals by descriptor value
 - Mark outlier points
 - Filter points by test conditions
 - Mark focused chemical
 - Mark focused points
- Selection navigation
- Gap filling approach
 - Read-across
 - Trend analysis**
 - Descriptors/data
 - Model/(Q)SAR
 - Calculation options
 - Visual options
 - Information
 - Miscellaneous

1. Click on Accept prediction.

2. Click OK; 3. Click Return to the matrix

Outlook

- Background
- Category evaluation – overview
 - Phase I. Evaluation of Category & Data Gap Filling
 - Phase II. Extension of the Category& Data gap Filling
- Save the prediction result

Saving the prediction result

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

Saving the prediction result

The screenshot shows the QSAR Toolbox 3.4.0.17 interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The 'Input' section is active, showing a 'Documents' sidebar on the left with a list of files. The main workspace displays chemical structures and a table of prediction results. A 'Save As' dialog box is open, showing the file path 'This PC > New Volume (D:) > Ksenia > CURRENT WORK > NEW TB 3.3.6', a list of files, and the 'File name' field set to 'Tutorial 7'. The 'Save as type' is 'Toolbox work file (*.tbw)'. The 'Save' button is highlighted.

1. Go to Input section
2. Click on Save button
3. Define name of the file;
4. Click Save button

Open saved file

The screenshot shows the QSAR Toolbox 3.4.0.17 interface. The 'Document' menu is open, and the 'Open' option is highlighted. A 'Select file' dialog box is open, showing a list of files in the 'NEW TB 3.3.6' folder. The file 'Tutorial 7.tbw' is selected. The 'Open' button in the dialog box is highlighted.

Name	Date modified	Type	Size
Tutorial 7.tbw	7/13/2016 4:25 PM	TBW File	523 KB
For testing Report Tutorial 11.tbw	7/16/2016 1:18 PM	TBW File	109 KB
Tutorial 19.tbw	7/16/2016 6:00 PM	TBW File	102 KB
Tutorial 22_Filter by Ames data.tbw	7/16/2016 6:45 PM	TBW File	89 KB
test save doc.tbw	7/19/2016 3:58 PM	TBW File	317 KB
Test edit prediction info.tbw	7/1/2016 2:27 PM	TBW File	494 KB
File with Filter by Ames.tbw	7/1/2016 8:55 AM	TBW File	15 KB
Not accurate message for Nasko.tbw	6/29/2016 12:59 PM	TBW File	176 KB
Test edit predction info.tbw	6/24/2016 6:41 PM	TBW File	235 KB
Test nasko.tbw	6/24/2016 2:54 PM	TBW File	165 KB
File with edit prediction info.tbw	6/22/2016 10:59 AM	TBW File	226 KB
File for Ames 13013177.tbw	6/22/2016 9:31 AM	TBW File	217 KB
File for RDT.tbw	6/22/2016 8:37 AM	TBW File	1,250 KB

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

Open saved file

The screenshot displays the QSAR Toolbox 3.4.0.17 interface. The main window shows a list of chemical structures and their associated properties. A dialog box titled "Information" is open, displaying the message "The file was executed successfully" with an "OK" button. A blue callout box with the number "1" points to the "OK" button. The interface includes a menu bar with options like "New", "Open", "Close", "Save", and a toolbar with various chemical analysis tools. The left sidebar shows a document tree with "Tutorial 7.tbw" selected. The bottom of the window features a filter endpoint tree and a search bar.

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