

OECD QSAR Toolbox v.3.1

Example for predicting Repeated dose
toxicity of 2,3-dimethylaniline

Outlook

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

Background

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

Outlook

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

Objectives

This presentation demonstrates a number of functionalities of the Toolbox:

- Identify analogues for a target chemical.
- Retrieve experimental results available for those analogues.
- Fill data gaps by read across.

Outlook

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

The exercise

- In this exercise we will predict the repeated dose toxicity of **2,3-dimethylaniline CAS 87-59-2**
- Define initial category of similar analogues based on US-EPA New chemical categories.
- Gather available experimental data for the target chemical and identified analogues
- Apply read across prediction based on analogue approach

Outlook

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

Workflow

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

Outlook

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

Chemical Input Overview

- This module provides the user with several means of entering the chemical of interest or the target chemical.
- Since all subsequent functions are based on chemical structure, the goal here is to make sure the molecular structure assigned to the target chemical is the correct one.

Chemical Input

Ways of Entering a Chemical

User Alternatives for Chemical ID:

A. Single target chemical

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

B. Group of chemicals

- User List/Inventory
- Specialized Databases

Getting Started

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

Chemical Input Screen

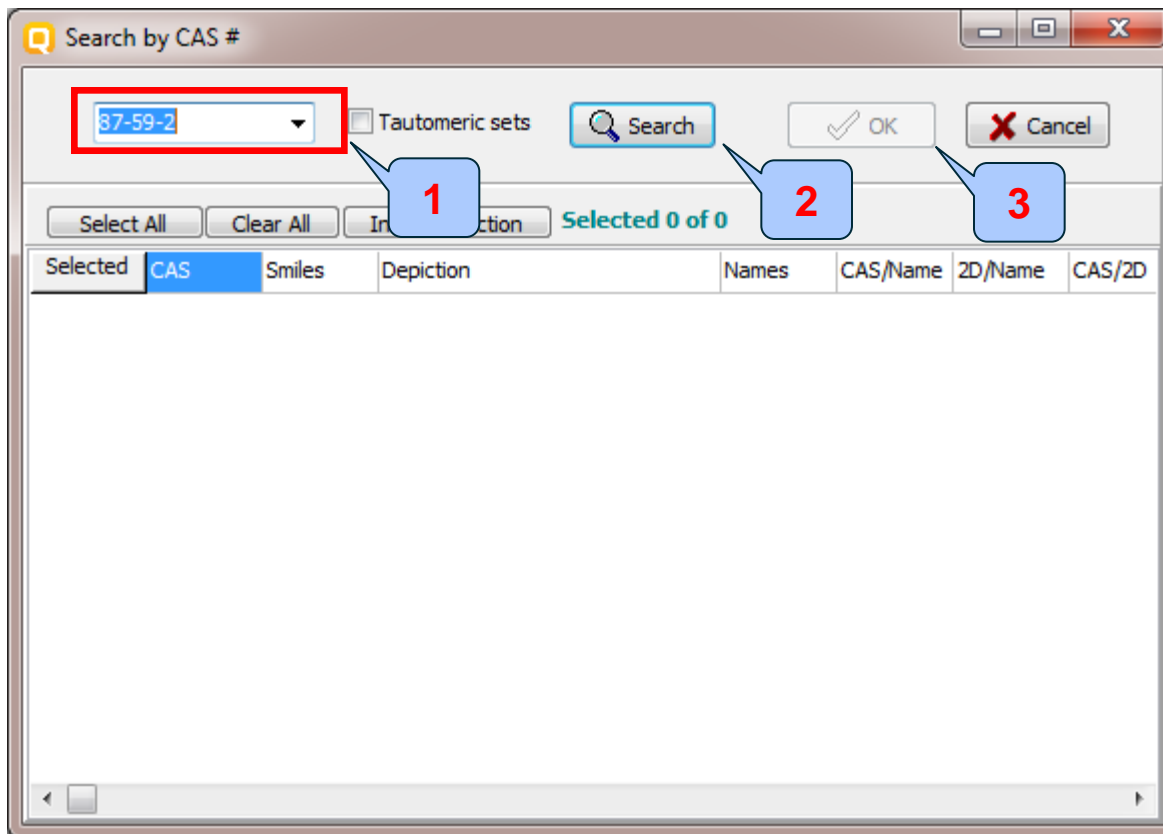
Input target chemical by CAS#

The screenshot displays the QSAR Toolbox 3.1.0.21 software interface. The top menu bar includes options for Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. Below this, a secondary menu bar offers actions like New, Open, Close, Save, and CAS#. The 'CAS#' button is highlighted with a red rectangular box, and a blue callout bubble with the number '1' points to it. The main workspace is divided into a left sidebar with a 'Documents' panel and a central area with a 'Structure' view and a 'filter endpoint tree...' panel. The 'filter endpoint tree...' panel lists categories such as Substance Identity, Physical Chemical Properties, Environmental Fate and Trans..., Ecotoxicological Information, and Human Health Hazards. At the bottom left, there are buttons for '... select filter type ...', 'Create', and 'Apply'.

1. Click on CAS#

Chemical Input Screen

Enter CAS# 87-59-2



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

Chemical Input

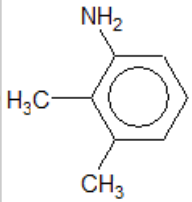
Target chemical identity

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

Search by CAS #

87-59-2 Tautomeric sets

Select All Clear All Invert Selection Selected 1 of 1

Selected	CAS	Smiles	Depiction	Names	CAS/Name	2D/Name	CAS/2D
1. Yes	87-59-2	c1(N)c(C)c(C)cc1		1: 2,3-dimethyl-4-aminobenzene 2: dimethyl-4-aminobenzene 3: 2,3-xylidene-4-aminobenzene 4: 2,3-dimethyl-4-aminobenzene 5: 2,3-dimethyl-4-aminobenzene 6: 2,3-dimethyl-4-aminobenzene 7: benzene 8: 118	2: Low Quality 1: Aquatic 2: Biocidal 3: High Quality 1: Bactericidal 2: Biocidal 3: DSS 4: ECH 5: ECH 6: Geraniol 7: MET 8: REA 9: US 4: Low Quality	2: Low Quality 1: Aquatic 2: Biocidal 3: High Quality 1: Bactericidal 2: Biocidal 3: Geraniol 4: REA 5: ECH 6: DSS 7: MET 8: US 9: ECH 4: Low Quality	2: 3: 4: 5: 6: 7: 8: 9: 10: 11: 12: 13: 14: 15: 16:

Chemical Input

Target chemical identity

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

Chemical Input

Target chemical identity

The screenshot displays the QSAR Toolbox 3.1.0.21 interface. The main window shows the 'Input' tab with a chemical structure of 2,3-dimethylaniline (2,3-dimethylbenzenamine) highlighted in a red circle. The structure is shown as a benzene ring with an amino group (NH₂) at the top position and two methyl groups (H₃C and CH₃) at the 2 and 3 positions.

The 'Substance Identity' section is expanded, showing the following information:

- CAS Number: 87-59-2
- EC Number: 201-755-0
- Einecs Number: 2017550
- Chemical Name: 2,3-dimethylaniline, 2,3-dimethylbenzenamine, 2,3-xylidine, 2,3-dimethyl-phenylamine, 2,3-dimethylbenzenamine, 2,3-dimethyl-aniline, benzenamine, 2,3-dimethyl-118
- Structural Formula: c1(N)c(C)c(C)ccc1

The 'Chemical List' section is also visible, showing the chemical structure and the SMILES string: c1(N)c(C)c(C)ccc1.

Outlook

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

Profiling Overview

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

Profiling

Side-Bar to Profiling

- For most of the profilers, background information can be retrieved by highlighting one of the profilers (for example, Repeated dose toxicity (HESS) and clicking on “View” (see next screen shot).

Profiling Side-Bar to Profiling

1. Highlight the profiler

2. Click View

List with categories

Textual description

Anilines (Hemolytic anemia with methemoglobinemia) Rank A

1. Toxicity Information

The toxicant of methemoglobinemia induced by anilines is considered to be N-hydroxyl anilines that are metabolites of anilines in the liver^{1,2}. The hemolytic anemia induced by anilines is considered to be related to the oxidation of erythrocytes by N-hydroxyl anilines^{3, 4}.

- 1) Anilines are metabolized in hepatocytes by oxidases such as P450 to N-hydroxyl anilines.
- 2) N-hydroxyl anilines react with hemoglobin (Hgb) in erythrocytes to produce nitrosoaniline and methemoglobin (Met-Hgb). The resulting increase in the concentration of Met-Hgb is observed in hematological examination.
- 3) Erythrocytes are degenerated (peroxidation of lipid membrane etc.) by reactive oxygen species (ROS) produced in the above reaction³.
- 4) Phagocytosis of degenerate erythrocytes, mainly in the spleen, results in hemolysis⁴.
- 5) The result is: decrease in red blood cells (RBC), decrease in Hgb, decreased hematocrit (Hct) and increase in reticulocytes (Ret) observed upon hematological examination in RDT test. In addition, pigmentation of hemosiderin and congestion are observed in the spleen on histopathological examination⁵.
- 6) As a compensatory response to anemia, extramedullary hematopoiesis (mainly in the spleen) is observed on histopathological examination⁴.

Diagram:

The diagram illustrates the metabolic pathway of aniline. It starts with aniline (NH₂) in the liver, which is converted to N-hydroxyl aniline (NHOH). This intermediate then reacts with hemoglobin (Hb) to form methemoglobin (Met-Hb) and reactive oxygen species (ROS). The ROS leads to the peroxidation of the lipid membrane of erythrocytes, causing phagocytosis of damaged erythrocytes and subsequent hemolysis. The final result is a decrease in RBC, HGB, and HTC, and an increase in Reticulo and Met-Hb.

1. Highlight the profiler
2. Click View

Profiling Side-Bar to Profiling

1. Highlight the profiler

2. Click View

1. Highlight the profiler
2. Click View

Profiling

Side-Bar to Profiling results

QSAR Toolbox 3.1.0.21 [Document_1]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Profiling Profiling Schemes

Apply New View Delete

Profiling methods

Select All Unselect All Invert About

- Bioaccumulation – metabolism alerts
- Bioaccumulation – metabolism half-lives
- Biodegradation Fragments (BioWIN MITI)
- Carcinogenicity (genotox and nongenotox) alerts by ISS
- DNA alerts for AMES, MN and CA 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 skin sensitization by OASIS v.1
- rHER Expert System ver. 1 - USEPA
- Skin irritation/corrosion Exclusion rules by BfR
- Skin irritation/corrosion Inclusion rules by BfR

Empiric

- Chemical elements
- Groups of elements
- Lipinski Rule Oasis
- Organic functional groups
- Organic functional groups (nested)
- Organic functional groups (US EPA)
- Organic functional groups, Norbert Haider (checkmod)

Metabolism/Transformations

Select All Unselect All Invert About

- Documented
 - Observed Mammalian metabolism
 - Observed Microbial metabolism
 - Observed Rat In vivo metabolism
 - Observed Rat Liver S9 metabolism
- Simulated
 - Autoxidation simulator
 - Autoxidation simulator (alkaline medium)
 - Dissociation simulator
 - Hydrolysis simulator (acidic)
 - Hydrolysis simulator (basic)
 - Hydrolysis simulator (neutral)
 - Microbial metabolism simulator
 - Rat liver S9 metabolism simulator
 - Skin metabolism simulator

Filter endpoint tree... 1 [target]

Structure

Chemical structure: Cc1cccc(N)c1

Predefined

- General Mechanistic
 - Biodeg BioHC half-life (Bio... Not calculated
 - Biodeg primary (Biowin 4) days - weeks
 - Biodeg probability (Biowin 1) Biodegrades Fast
 - Biodeg probability (Biowin 2) Biodegrades Fast
 - Biodeg probability (Biowin 5) Does NOT Biodegrade Fast
 - Biodeg probability (Biowin 6) Does NOT Biodegrade Fast
 - Biodeg probability (Biowin 7) Does NOT Biodegrade Fast
 - Biodeg ultimate (Biowin 3) weeks - months
- DNA binding by OASIS v.1.1
 - Radical
 - Radical >> ROS formation ...
 - Radical >> ROS formation ...
 - SN1
 - SN1 >> Nitrenium ion form...
 - SN1 >> Nitrenium ion form...
 - SN1
 - SN1 >> Nitrenium ion form...
 - SN1 >> Nitrenium ion form...
- DNA binding by OECD
 - DPRA Cysteine peptide d... Not possible to classify ac...
 - DPRA Lysine peptide depl... Not possible to classify ac...
 - Estrogen Receptor Binding Weak binder, NH2 group
 - Hydrolysis half-life (Ka, p... Not calculated
 - Hydrolysis half-life (Ka, p... Not calculated
 - Hydrolysis half-life (Kb, p... Not calculated
 - Hydrolysis half-life (Kb, p... Not calculated
 - Organic functional groups... Primary amine
 - Organic functional groups... Primary aromatic amine
 - Organic functional groups... Stable form
- Tautomers unstable
- Toxicological
 - Repeated dose (HES) Anilines (Hemolytic anemi... Anilines (Hepatotoxicity) R...

The target has a potential to interact with DNA according to DNA binding profilers

The target chemical could cause RDT toxicity through two different effects according to RDT profiler

Profiling

Side-Bar to Profiling results

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The main window displays the results for "Anilines (Hemolytic anemia with methemoglobinemia) Rank A".

- Left Sidebar:** Contains "Profiling methods" (e.g., Bioaccumulation, Carcinogenicity) and "Metabolism/Transformations" (e.g., Documented, Simulated).
- Central Window:**
 - Target:** Shows the chemical structure of Aniline (Nc1ccccc1) and a highlighted fragment (Nc1ccc(C)cc1).
 - Boundary Options:** Shows "Metabolism" and a "Fragment" with the SMILES string: c1(N("Exh13")"Exh13")c("Exh14")c("Exh14")c("Exh14")c1"Exh14".
 - Common Fragments:** A table with 8 columns. The first column is highlighted in blue and contains "[Exh13]".
 - Profile Description:** Contains the title "Anilines (Hemolytic anemia with methemoglobinemia) Rank A" and a section "1. Toxicity Information" with text: "The toxicant of methemoglobinemia induced by anilines is considered to be N-hydroxyl anilines that are metabolites of anilines in the liver^{1,2}. The hemolytic anemia induced by anilines is considered to be related to the oxidation of erythrocytes by N-hydroxyl anilines^{3, 4}."
- Right Side:** Two text annotations with red arrows:
 - "Structural boundary of the category" pointing to the SMILES string.
 - "Mechanistic justification of the category" pointing to the toxicity text.
- Bottom:** A callout box with the number "1" points to a cell in the "Metabolism/Transformations" sidebar containing "Anilines (Hemolytic anemia with methemoglobinemia) R".

1. Double click on the cell with profiling result to see why this chemical is classified as Anilines

Outlook

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

Endpoint Overview

- “Endpoint” refers to the electronic process of retrieving the environmental fate, ecotoxicity and toxicity data that are stored in the Toolbox.
- Data gathering can be executed in a global fashion (i.e., collecting all data for all endpoints) or on a more narrowly defined basis (e.g., collecting data for a single or limited number of endpoints).

Endpoint

Case study

- In this example, we limit our data gathering to a single toxicity endpoint: repeated dose toxicity
- In this example, we collect data from the databases containing experimental results for Repeat dose toxicity (Repeated Dose Toxicity (HESS)).
- **Click** on “Endpoint” in the Toolbox workflow.
- **Expand the** “Human Health Hazards” section
- **Click** on the box to select that database.
- **Click** on “Gather data” (see next screen shot).

Endpoint Gather data

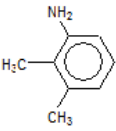

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The 'Endpoint' tab is active. In the 'Databases' list on the left, 'Human Health Hazards' is expanded (callout 1), and 'Repeated Dose Toxicity HESS' is selected (callout 2). The 'Gather' button in the top toolbar is highlighted (callout 3). The right panel shows the 'Filter endpoint tree...' for a target chemical, 2,3-dimethylaniline, with its structure and various identifiers listed.

Structure	1 [target]
Substance Identity	
— CAS Number	87-59-2
— Chemical IDs	EC Number:201-755-0 Einecs Number:2017550 2,3-dimethylaniline dimethylaniline, 2,3- 2,3-xylidine
— Chemical Name	2,3-dimethyl-phenylamine 2,3-dimethylbenzenamine 2,3-dimethyl-aniline benzenamine, 2,3-dimethyl- 118
— Structural Formula	c1(N)c(C)c(C)ccc1
Physical Chemical Properties	
Environmental Fate and Transport	
Ecotoxicological Information	
Human Health Hazards	
Profile	
— Predefined	
— General Mechanistic	
— Endpoint Specific	
— Empiric	
— Toxicological	
— Repeated dose (HESS)	Anilines (Hemolytic anemi... Anilines (Hepatotoxicity) R...

- 1. Expand** the Human Health Hazards section
- 2. Select** database related to the target endpoint: Repeated dose toxicity HESS
- 3. Click** Gather

Endpoint Gather data

Repeated values for: 105 data-points, 22 groups, 1 chemicals

Data points...						
	Endpoint	CAS	Structure	Value	Administration period (...)	
<input checked="" type="checkbox"/>	NOEL	87-59-2		300 mg/kg/day	28	Baso chan
<input checked="" type="checkbox"/>	NOEL	87-59-2		300 mg/kg/day	28	Baso chan
<input checked="" type="checkbox"/>	NOEL	87-59-2		300 mg/kg/day	28	Baso chan
<input checked="" type="checkbox"/>	NOEL	87-59-2		300 mg/kg/day	28	Baso chan
<input checked="" type="checkbox"/>	NOEL	87-59-2		300 mg/kg/day	28	Baso chan
<input checked="" type="checkbox"/>	NOEL	87-59-2		300 mg/kg/day	28	Dilat

Buttons: Select one, Invert, Check All, Uncheck All, **OK**, Cancel

1

QSAR Toolbox 3.1.0.21

860 data points gathered across 1 chemicals.

OK

1. **Select OK.**
2. The message informs you for number of retrieved data points. **Click OK**

Endpoint Gather data

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The 'Endpoint' menu is active, displaying a tree view of various toxicity endpoints. The 'Structure' panel shows the chemical structure of 2,4-dimethylaniline. The data matrix below shows measured data for the target chemical.

Endpoint	Measured Data
(1/76)M	12 mg/kg/day, 12 mg/kg/day, 60 mg/k...
(1/701)M	12 mg/kg/day, 12 mg/kg/day, 12 mg/k...
ACF	

Measured data for the target appeared on data matrix. We will try to reproduce measured data by read-across

(1/76)M: 12 mg/kg/day, 12 mg/kg/day, 60 mg/k...
 (1/701)M: 12 mg/kg/day, 12 mg/kg/day, 12 mg/k...
 ACF

Recap

- In the first module, you have entered the target chemical being sure of the correctness of the structure.
- In the second module, you have profiled the target chemical and found that the target could cause RDT toxicity through two different effects
- In the third module, you have found that there is experimental RDT data for the target structure. We will try to reproduce it using read across analysis
- But before the user can proceed with the "Filling Data Gap" module, he/she should define a category with similar analogues
- **Click** on "Category Definition" to move to the next module.

Outlook

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

Category Definition

Grouping methods

- The different grouping methods allow the user to group chemicals into chemical categories according to different measures of “similarity” so that within a category data gaps can be filled by read-across.
- Detailed information about grouping chemical (Chapter 4) could be found in document “Manual for Getting started” published on OECD website:

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

Basic guidance for category formation and assessment

Suitable categorization phases:

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

Performing categorization:

1. The categorization phases should be applied successively
2. The application order of the phases depend on the specificity of the data gap filling
3. More categories of same Phase could be used in forming categories
4. Some of the phases could be skipped if consistency of category members is reached

Graphical illustration of suitable categorization phases is shown on next slide

Suitable Categorization/Assessment Phases

Phase I. Structure based

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

**Broad grouping
Endpoint Non-specific**

Repeating Phase I due to Multifunctionality of chemicals

Phase II. Mechanism based

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

**Subcategorization
Endpoint Specific**

Metabolism accounted for

Phase III. Eliminating dissimilar chemicals

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

**Subcategorization
Endpoint Specific**

Category Definition

Grouping methods –phase I

Suitable Categorization/Assessment Phases

Phase I. Structure based

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

Broad grouping
Endpoint Non-specific

Phase I categorization in Toolbox

Method	Result
Not categorized	38 analogues are identified
Anilines (Acute toxicity)	10 analogues are identified
Anilines (Hindered)	33 analogues are identified
Aniline	33 analogues are identified
Aryl	33 analogues are identified
Aniline	33 analogues are identified
Structural similarity, Dice ACF, 50%	12 analogues are identified

Category Definition

Grouping methods

- Based on these classifications and basic guidance for grouping chemicals explained on the previous slides the US-EPA (as broader group: 38 analogues) is used for defining initial group of analogues (phase I)
- For refinement of category and eliminating dissimilar chemicals a sequence of endpoint specific and structural profilers are applied (phase II)
 - US-EPA New chemical categories
 - Repeated dose (HESS)
 - Chemical elements
 - Structural similarity

Category Definition

Defining US-EPA New Chemical categories

The screenshot displays the QSAR Toolbox 3.1.0.21 interface. The 'Category Definition' menu item is active. In the 'Grouping methods' list, 'US-EPA New Chemical Categories' is highlighted. The 'Define' button is circled. The 'Filter endpoint tree...' window shows a chemical structure and a list of endpoints, with 'Human Health Hazards' selected. A dialog box titled 'US-EPA New Chemical Categories' is open, showing 'Anilines (Acute toxicity)' selected. Another dialog box titled 'Define category name' is open, showing 'Acute toxicity (US-EPA New Chemical Categories)' as the category name. The 'OK' button in the dialog is circled.


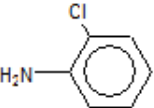

1. **Highlight** the "US-EPA New Chemical Categories"
2. **Click** Define
3. **Click** OK to confirm the defined category for the target chemical
4. **Click** OK

Category Definition

Defining US-EPA New Chemical categories

Repeated values for: 3502 data-points, 793 groups, 766 chemicals

Data points...

	Endpoint	CAS	Structure	Value	Administration period
<input checked="" type="checkbox"/>	LOEL	121-88-0		200 mg/kg/day	91
<input checked="" type="checkbox"/>	LOEL	121-88-0		200 mg/kg/day	91
<input checked="" type="checkbox"/>	NOEL	95-51-2		160 mg/kg/day	93
<input checked="" type="checkbox"/>	NOEL	95-51-2		160 mg/kg/day	93
<input checked="" type="checkbox"/>	NOEL	95-51-2		160 mg/kg/day	93
<input checked="" type="checkbox"/>	NOEL	95-51-2		160 mg/kg/day	93
<input checked="" type="checkbox"/>	NOEL	95-51-2		160 mg/kg/day	93
<input checked="" type="checkbox"/>	NOEL	99-57-0		1E3 mg/kg/day	91
<input checked="" type="checkbox"/>	NOEL	99-57-0		1E3 mg/kg/day	91

Buttons: Select one, Invert, Check All, Uncheck All, **OK**, Cancel

1

1. **Click OK** to retrieve all available experimental data

Category Definition

Defining US-EPA New Chemical categories

The experimental results for the analogues appeared on datamatrix

The screenshot displays the QSAR Toolbox 3.1.0.21 interface. The 'Category Definition' tab is active, showing a datamatrix table with columns for chemical structures and various endpoints. A red box highlights the 'Repeated Dose Toxicity' section, which includes 'LOEL' and 'NOEL' endpoints. A blue callout box with the number '1' points to this section.

Filter endpoint tree...	1 [target]	2	3	4	5	6	7
Structure	<chem>Cc1ccc(N)cc1</chem>	<chem>Nc1ccccc1Cl</chem>	<chem>Nc1ccc2c(c1)C(=O)c3ccccc23</chem>	<chem>O=C(N)Nc1ccc(O)cc1</chem>	<chem>O=C(N)Nc1ccc(O)cc1</chem>	<chem>Nc1ccc2c(c1)C(=O)c3ccccc23</chem>	<chem>Nc1ccc2c(c1)C(=O)c3ccccc23</chem>
Substance Identity							
Physical Chemical Properties							
Environmental Fate and Transport							
Ecotoxicological Information							
Human Health Hazards							
Acute Toxicity							
Carcinogenicity							
Developmental Toxicity / Teratogenicity							
Genetic Toxicity							
Immunotoxicity							
Irritation / Corrosion							
Neurotoxicity							
Repeated Dose Toxicity							
LOEL	(35/1027)M: 12 mg/kg/day, 12 mg/kg/...	M: 10 mg/kg/day, 1...	M: 192 mg/kg/day, ...	M: 125 mg/kg/day, ...	M: 100 mg/kg/day, ...	M: 192 mg/kg/day, ...	M: 1E3 mg/kg/day, ...
NOEL	(38/21339)M: 12 mg/kg/day, 12 mg/kg/...	M: 10 mg/kg/day, 1...	M: 192 mg/kg/day, ...	M: 62.5 mg/kg/day,...	M: 100 mg/kg/day, ...	M: 92.3 mg/kg/day,...	M: 200 mg/kg/day, ...
Sensitisation							
Toxicity to Reproduction							
Toxicokinetics, Metabolism and Distrib...							
Profile							

1. Chemical statistics presenting the number of chemicals and the available experimental data for the two endpoints.

Recap

- In this module, you have defined the category of similar analogues.
- In the next module, you should apply read across in order to fill in data gap
- But before the user can proceed with the “Filling Data Gap” module, he/she should navigate to the target endpoint: In our case we will predict RDT of target for two endpoints: Total NOEL and Total LOEL; Route: Oral (gavage)
- Total NOEL and Total LOEL values coincide with minimal values for all LOELs (NOELs) of the current chemical (more info could be found on next snapshot)
- **Click** on “Data Gap Filling” to move to the next module.

Total LOEL/NOEL

Filter endpoint tree... 1 [target]

CAS 108-69-0 Cc1ccc(N)cc1

Endpoint	Value
LOEL - Total	(1/2) M: 60 mg/kg/day, 60 mg/kg/day
NOEL - Total	(1/2) M: 10 mg/kg/day, 10 mg/kg/day

Minimal value across all LOEL values

Total value coincide with minimal values for all LOELs (10; 60 mg/kg/day)

Minimal value across all NOEL values

Total value coincide with minimal values for all LOELs (10; 60 mg/kg/day)

Now you are ready to continue with next module data gap filling

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Category definition
 - **Data gap filling**

Data Gap Filling Overview

- “Data Gap Filling” module give access to three different data gap filling tools:
 - Read-across
 - Trend analysis
 - (Q)SAR models
- Depending on the situation, the most relevant data gap mechanism should be chosen, taking into account the following considerations:
 - Read-across is the appropriate data-gap filling method for “qualitative” endpoints like skin sensitisation or mutagenicity for which a limited number of results are possible (e.g. positive, negative, equivocal). Furthermore read-across is recommended for “quantitative endpoints” (e.g., 96h-LC50 for fish) in case a low number of analogues with experimental results are identified.
 - Trend analysis is the appropriate data-gap filling method for “quantitative endpoints” (e.g., 96h-LC50 for fish) if a high number of analogues with experimental results are identified.
 - “(Q)SAR models” can be used to fill a data gap if no adequate analogues are found for a target chemical.
- In this example, we use read-across.

Data Gap Filling

Interpreting Read-across

- In this example, all the analogues have repeated dose toxicity data (LOEL and NOEL values)
- Predicted values for the target compound is based on initial group of **Anilines** defined by US-EPA New Chemical categories
- The following subcategorizations are used for filtering the initial group of analogues:
 - US-EPA New chemical categories
 - Repeated dose (HESS)
 - Chemical elements
 - Structural similarity
- Before applying the read across, we should navigate to the target endpoint Total NOEL

See next screen shots

Data Gap Filling

Navigation of endpoint tree: Repeated dose toxicity/NOEL/oral gavage/Total

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The 'Data Gap Filling' tab is active. The endpoint tree on the left is expanded to the 'Total' node. The main window displays a table with 7 columns representing different chemical structures and their toxicity data for various endpoints. The 'Total' node is highlighted in green, and its data is shown in the table. A red callout '1' points to the 'Total' node in the tree, and another red callout '2' points to the '+' sign next to the 'Total' node.

Endpoint	1 [target]	2	3	4	5	6	7
Structure	<chem>Cc1ccc(N)cc1</chem>	<chem>Nc1ccccc1Cl</chem>	<chem>O=C1C=CC(=O)N1</chem>	<chem>Oc1ccc(cc1)[N+](=O)[O-]</chem>	<chem>Oc1ccc(cc1)[N+](=O)[O-]</chem>	<chem>Nc1ccc2c(c1)c(=O)c(N)cc2=O</chem>	<chem>Nc1ccc2c(c1)c(O)cc2</chem>
Substance Identity							
Environmental Fate and Transport							
Human Health Hazards							
Repeated Dose Toxicity							
- LOEL	(34/63) M: 12 mg/kg/day, 60 mg/kg/...	M: 10 mg/kg/day, 1...	M: 192 mg/kg/day, ...	M: 125 mg/kg/day, ...	M: 100 mg/kg/day, ...	M: 192 mg/kg/day, ...	M: 1E3 mg/kg/day
- NOEL							
- Rat							
- Oral (Feed)	(3/6)		M: <192 mg/kg/day...			M: 92.3 mg/kg/day,...	
- Oral (Gavage)							
- Whole Body							
- Total	(32/59) M: 12 mg/kg/day, <12 mg/kg...	M: <10 mg/kg/day, ...		M: 62.5 mg/kg/day,...	M: 100 mg/kg/day, ...		M: 200 mg/kg/day
- Oral (Water Contam.)							
Profile							

1. **Type** "Total" in the filter filed
2. **Expand** the tree to "Total" node by single left click on the "+" sign

Data Gap Filling

Apply read across for Total NOEL

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The top menu bar includes 'Input', 'Profiling', 'Endpoint', 'Category Definition', 'Data Gap Filling', and 'Report'. The left sidebar shows 'Data Gap Filling Method' with 'Read-across' selected. The main window displays a table with columns for different target chemicals and rows for various endpoints. The 'Total' row for 'NOEL' is highlighted, and a specific cell is circled in red.

Total	1 [target]	2	3	4	5	6	7
Structure	<chem>Cc1ccc(N)cc1</chem>	<chem>Nc1ccc(Cl)cc1</chem>	<chem>Nc1ccc2c(c1)C(=O)N2</chem>	<chem>Nc1ccc(O)cc1</chem>	<chem>Nc1ccc(O)cc1</chem>	<chem>Nc1ccc(N)cc1</chem>	<chem>Nc1ccc(N)cc1</chem>
Substance Identity							
Environmental Fate and Transport							
Human Health Hazards							
Repeated Dose Toxicity							
NOEL	(34/63) M: 12 mg/kg/day, 60 mg/kg/...	M: 10 mg/kg/day, 1...	M: 192 mg/kg/day, ...	M: 125 mg/kg/day, ...	M: 100 mg/kg/day, ...	M: 192 mg/kg/day, ...	M: 1E3 mg/kg/day
Rat							
Oral (Feed)	(3/6)		M: <192 mg/kg/day...			M: 92.3 mg/kg/day...	
Oral (Gavage)							
Whole Body							
Total	(32/59) M: 12 mg/kg/day, <12 mg/kg/day, ...	M: <10 mg/kg/day, ...		M: 62.5 mg/kg/day, ...	M: 100 mg/kg/day, ...		M: 200 mg/kg/day
Oral (Water Containing)	(3/6)						
Profile							

1. **Click** on the cell corresponding to "NOEL" total value for the target chemical.
2. **Select** Read-across
3. **Click** Apply

Data Gap Filling

Read-across

- The resulting plot is experimental results of all analogues (Y axis) according to a descriptor (X axis) with the default descriptor of log Kow (see next screen shot).
- The **RED** dot represents predicted results for the target chemical .
- The **BROWN** dots represent the experimental results available for the analogues that are used for the read-across.
- The **BLUE** dot represent the experimental results available for the analogues but not used for read-across.

Data Gap Filling Read-across for NOEL

QSAR Toolbox 3.1.0.21 [Document_1]

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

Human Health Hazards Repeated Dose Toxicity NOEL Rat Oral (Gavage) Whole Body Total

Total	1 [target]	2	4	5	7	8	9
Structure	<chem>Cc1ccc(N)cc1</chem>	<chem>Nc1ccc(Cl)cc1</chem>	<chem>Oc1ccc(N)cc1</chem>	<chem>Oc1ccc(N)cc1</chem>	<chem>Nc1ccc(O)cc1</chem>	<chem>Nc1ccc(O)cc1</chem>	<chem>Nc1ccc(O)cc1</chem>
Total (32/59)	M. 12 mg/kg/day, <12 mg/kg	M. <10 mg/kg/day, ...	M. 62.5 mg/kg/day, ...	M. 100 mg/kg/day, ...	M. 200 mg/kg/day	M. 2 mg/kg/day	M. 10 mg/kg/day

Descriptors Prediction

Read across prediction of NOEL, taking the average from the nearest 5 neighbours, based on 5 values from 5 neighbour chemicals, Observed target value: 12.0 mg/kg/day, Predicted target value: 10.6 mg/kg/day

Accept prediction
Return to matrix

- Select/filter data
- Selection navigation
- Gap filling approach
- Descriptors/data
- Model/(Q)SAR
- Calculation options**
- Data usage
- Prediction approach options
- Use target data for prediction
- Visual options
- Information
- Miscellaneous

Set usage of data per chemical

- All
- Minimal
- Maximal**
- Average
- Median
- Lower median
- Higher median

OK Cancel

Worst case scenario is applied. Follow the steps:

1. **Select** Calculation options
2. **Click** on the "Data usage".
3. **Select** "Maximal"
4. **OK**

*Maximal values are selected because (i.e log/1 of endpoint (LOEL)) correspond to maximal hazard

Data Gap Filling

Subcategorization by US-EPA New Chemical Categories

The screenshot displays the OECD QSAR Toolbox interface. On the left, the 'Subcategorization' window is open, showing a list of predefined categories. The 'US-EPA New Chemical Categories' is selected and highlighted with a red box and a callout '2'. The main window shows a grid of chemical structures and their predicted target values. A scatter plot below the grid shows the relationship between log Kow and NOEL. A callout '1' points to the 'Subcategorize' button in the 'Accept prediction' panel. A callout '3' points to the 'Remove' button in the 'Selected 11 (20/31)' list.

1. Click Subcategorize **2. Select US-EPA New Chemical Categories**
3. Click Remove to eliminate dissimilar chemicals.

Data Gap Filling

Subcategorization by Repeated dose (HESS)

The screenshot displays the 'Subcategorization' window in the QSAR Toolbox. The left sidebar lists various methods, with 'Repeated dose (HESS)' highlighted in a red box and labeled with a '2'. The central workspace shows a table of chemical structures and their corresponding predicted values. The table has columns for chemical structures and predicted values. The predicted values are: (32/59) M. 12 mg/kg/day, <12 mg/kg/day, M. <10 mg/kg/day, M. 200 mg/kg/day, M. 10 mg/kg/day, M. 50 mg/kg/day, 5, M. 2 mg/kg/day, <2, M. <15 mg/kg/day. A scatter plot at the bottom shows the relationship between log Kow (x-axis) and log10(NoEL) (y-axis). The plot is titled 'Read across prediction of NOEL, taking the average from the nearest 5 neighbours, based on 5 values from 5 neighbour chemicals, Observed target value: 12.0 mg/kg/day, Predicted target value: 8.83 mg/kg/day'. The plot shows a positive correlation between log Kow and log10(NoEL). A '3' is placed near the plot. The right sidebar contains a 'Select/filter data' section with a 'Subcategorize' button, which is labeled with a '1'.

1. Click Subcategorize
2. Select Repeated dose (HESS) profiler
3. Click Remove to eliminate dissimilar chemicals.

Data Gap Filling

Subcategorization by Chemical elements

The screenshot displays the QSAR Toolbox software interface. The main window is titled 'Subcategorization' and shows the 'Data Gap Filling' tab. The interface includes a sidebar with various grouping methods, a main window with a table of chemical structures, and a scatter plot of NOEL vs log Kow. Three callouts (1, 2, 3) highlight key steps in the process.

1. Click Subcategorize. This step is highlighted in the 'Accept prediction' panel on the right side of the interface.

2. Select Chemical elements. This step is highlighted in the sidebar on the left side of the interface.

3. Click Remove to eliminate two dissimilar chemicals. This step is highlighted in the 'Selected 4 (6/10)' panel at the bottom of the interface.

1. **Click** Subcategorize.
2. **Select** Chemical elements.
3. **Click** Remove to eliminate two dissimilar chemicals.

Data Gap Filling

Subcategorization by Structural similarity

The screenshot displays the 'Subcategorization' window in the QSAR Toolbox. The interface is divided into several sections:

- Left Sidebar:** Lists various grouping methods under categories like 'Empiric', 'Toxicological', 'Custom', 'Metabolism/Transformations', and 'Simulated'. The 'Structure similarity' method is highlighted with a red box and a callout '2'.
- Top Panel:** Contains navigation buttons for Profiling, Endpoint, Category Definition, Data Gap Filling, and Report. A 'Target' field is set to 'Similar 100%'. Below this, options for 'Differ from target by:' are shown, with 'At least one category' selected.
- Table:** Displays a list of chemical structures and their predicted values. The table has columns for 'Structure', 'Total', and predicted values. The first row shows a target structure with a predicted value of 12 mg/kg/day. Subsequent rows show similar structures with predicted values ranging from 5 to 10 mg/kg/day. A callout '3' points to the 'Remove' button at the bottom of the table.
- Graph:** A scatter plot titled 'Read across prediction of NOEL, taking the average from the nearest 5 neighbours, based on 5 values from 5 neighbour chemicals, Observed target value: 12.0 mg/kg/day, Predicted target value: 8.83 mg/kg/day'. The y-axis is 'NOEL (obs.), log₁₀(mg/kg/day)' and the x-axis is 'log Kow'. Data points are plotted as colored dots.
- Right Panel:** Contains the 'Accept prediction' and 'Return to matrix' sections. The 'Subcategorize' button is highlighted with a callout '1'.

Structural similarity is applied in order to refine the category to the most similar analogues

- 1. Click** Subcategorize.
- 2. Select** Structural similarity (Dice, Atom centered fragments, Atom type; Count H attached; Hybridizations). Select the first two categories (hold Ctrl button)
- 3. Click** Remove to eliminate chemicals with similarity less than 60%

Data Gap Filling

Read across result for Total NOEL

QSAR Toolbox 3.1.0.21 [Document_1]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Filing 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

Human Health Hazards Repeated Dose Toxicity NOEL Rat Oral (Gavage) Whole Body Total

Total	1 [target]	10	11	17
Structure	<chem>Cc1ccc(N)cc1</chem>	<chem>Cc1ccc(N)cc1</chem>	<chem>Cc1ccc(N)cc1</chem>	<chem>Cc1ccc(N)cc1</chem>
Total (32/59)	M: 12 mg/kg/day, <12 mg/kg	M: 50 mg/kg/day, 5	M: 2 mg/kg/day, <2	M: 10 mg/kg/day, 2
Oral (Water Containing) (3/6)				

Descriptors Prediction

Read across prediction of NOEL, taking the average from the nearest 5 neighbours, based on 3 values from 5 neighbour chemicals, Observed target value: 12.0 mg/kg/day, Predicted target value: 10.00 mg/kg/day

NOEL (obs.), log (1 mol/kg/day)

log Kow

Descriptor X: log Kow

Accept prediction
Return to matrix

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

1. Click Accept prediction

2. Click Return to matrix

Data Gap Filling

Result of read-across prediction

QSAR Toolbox 3.1.0.21 [Document_1]

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

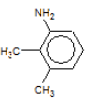






Apply

Data Gap Filling Method

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

Target Endpoint

Human Health Hazards Repeated Dose Toxicity LOEL

Total	1 [target]	2	3	4	5	6	7
Structure							
Substance Identity							
Environmental Fate and Transport							
Human Health Hazards							
Repeated Dose Toxicity							
LOEL	(34/63) M: 12 mg/kg/day, 60 mg/kg/...	M: 10 mg/kg/day, 1...	M: 192 mg/kg/day, ...	M: 125 mg/kg/day, ...	M: 100 mg/kg/day, ...	M: 192 mg/kg/day, ...	M: 1E3 mg/kg/day
NOEL							
Rat							
Oral (Feed)	(3/6)		M: <192 mg/kg/day...			M: 92.3 mg/kg/day...	
Oral (Gavage)							
Whole Body							
Total	(32/60) M: 12 mg/kg/day, <12 mg/kg/... R: 10 mg/kg/day	M: <10 mg/kg/day, ...		M: 62.5 mg/kg/day,...	M: 100 mg/kg/day, ...		M: 200 mg/kg/day
Oral (Water Containing)	(3/6)						
Profile							

1. Read across prediction 10 mg/kg/day coincide with experimental data (12 mg/kg/day)

Data Gap Filling

Apply read-across for Total LOEL

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The 'Data Gap Filling Method' is set to 'Read-across'. A table displays data for seven target chemicals. The 'LOEL' row for the first target chemical is highlighted with a red circle and labeled '1'. The 'Read-across' method is selected in the left sidebar, labeled '2'. The 'Apply' button is circled in red and labeled '3'.

Total	1 [target]	2	3	4	5	6	7
Structure	<chem>Cc1ccc(N)cc1</chem>	<chem>Nc1ccccc1Cl</chem>	<chem>O=C1C=CC(=O)N1</chem>	<chem>Oc1ccc(N)cc1</chem>	<chem>Oc1ccc(N)cc1</chem>	<chem>Nc1ccc(N)cc1</chem>	<chem>Nc1ccc(N)cc1</chem>
Substance Identity							
Environmental Fate and Transport							
Human Health Hazards							
Repeated Dose Toxicity							
LOEL							
Rat							
Oral (Feed)	(3/6)		M: 192 mg/kg/day, ...			M: 192 mg/kg/day, ...	
Oral (Gavage)							
Whole Body							
Total	(28/51) M: 12 mg/kg/day, 60 mg/kg/day, ...	M: 10 mg/kg/day, 1...		M: 125 mg/kg/day, ...	M: 100 mg/kg/day, ...		M: 1E3 mg/kg/day
Oral (Water Containing)	(3/6)						
NOEL	(38/72) M: 12 mg/kg/day, <12 mg/kg/day, ...	M: <10 mg/kg/day, ...	M: <192 mg/kg/day, ...	M: 62.5 mg/kg/day, ...	M: 100 mg/kg/day, ...	M: 92.3 mg/kg/day, ...	M: 200 mg/kg/day
Profile							

1. **Click** on the cell corresponding to "LOEL" total value for the target chemical.
2. **Select** Read-across
3. **Click** Apply

Data Gap Filling Read-across for LOEL

QSAR Toolbox 3.1.0.21 [Document_1]

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

Human Health Hazards Repeated Dose Toxicity LOEL Rat Oral (Gavage) Whole Body Total

Total	1 [target]	2	4	5	7	8	9
Structure							
Total (28/51)	M: 12 mg/kg/day, 60 mg/kg/day	M: 10 mg/kg/day, 1...	M: 125 mg/kg/day, ...	M: 100 mg/kg/day, ...	M: 1E3 mg/kg/day	M: 10 mg/kg/day	M: 60 mg/kg/day

Descriptors Prediction

Read across prediction of LOEL, taking the average from the nearest 5 neighbours, based on 6 values from 6 neighbour chemicals, Observed target value: 26.8 mg/kg/day, Predicted target value: 17.6 mg/kg/day

Accept prediction

Return to matrix

- Select/filter data
- Selection navigation
- Gap filling approach
- Descriptors/data
- Model/(Q)SAR
- Calculation options
 - Data usage **2**
 - Prediction approach option
 - Use target data for prediction
- Visual options
- Information
- Miscellaneous

Set usage of data **3**

- All
- Minimal
- Maximal**
- Average
- Median
- Lower media
- Higher media

4 OK

Worst case scenario is applied. In this respect
2. **Click** on the "Data usage.

*Maximal values are selected because (i.e log/1 of endpoint (LOEL)) correspond to maximal hazard

1. **Select** Calculation options
3. **Select** "Maximal"

4. **OK**

Data Gap Filling

Subcategorization by US-EPA New Chemical Categories

The screenshot displays the QSAR Toolbox 3.1.0.21 interface. The 'Subcategorization' panel on the left shows a list of predefined methods, with 'US-EPA New Chemical Categories' highlighted. The central area shows a table of chemical structures with their corresponding predicted values. The scatter plot below shows the relationship between log Kow and LOEL (obs., log¹⁰ (mg/kg/day)). The right-hand panel shows the 'Accept prediction' options, with 'Subcategorize' highlighted.

1. Click Subcategorize
3. Click Remove to eliminate dissimilar chemicals.

2. Select US-EPA New Chemical Categories

Data Gap Filling

Subcategorization by Repeated dose (HESS)

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The 'Data Gap Filling' tab is active. The main window displays a table of chemical structures and their predicted values. The table has columns for chemical structures and predicted values. The predicted values are: (28/51) M. 12 mg/kg/day, 60 mg/kg/day; M. 10 mg/kg/day, 1; M. 1E3 mg/kg/day; M. 60 mg/kg/day; M. 2 mg/kg/day, 10; M. 15 mg/kg/day, 1; M. 10 mg/kg/day, 1.

The scatter plot shows LOEL (obs.) on the y-axis (ranging from 2.20 to 4.80) and log Kow on the x-axis (ranging from -6.00 to 3.00). The plot title is: "Read across prediction of LOEL, taking the average from the nearest 5 neighbours, based on 6 values from 6 neighbour chemicals, Observed target value: 12.0 mg/kg/day, Predicted target value: 14.1 mg/kg/day".

The right-hand panel contains the 'Accept prediction' and 'Return to matrix' sections. The 'Return to matrix' section has a 'Select/filter data' dropdown menu. The 'Subcategorize' option is selected. The 'Remove' button is highlighted with a callout '3'.

1. Click Subcategorize.
2. Select Repeated dose (HESS)
3. Click Remove to eliminate dissimilar chemicals.

Data Gap Filling

Subcategorization by Chemical elements

The screenshot displays the QSAR Toolbox interface for subcategorization. The 'Subcategorization' panel on the left lists various grouping methods, with 'Chemical elements' highlighted by a red box and a callout '2'. The main data table shows a list of chemicals with their structures and predicted values. The central scatter plot shows the relationship between $\log(\text{LOEL})$ and $\log(\text{Kow})$, with a callout '3' pointing to the 'Remove' button at the bottom of the left panel. The right panel, 'Return to matrix', has a callout '1' pointing to the 'Subcategorize' button.

1. Click Subcategorize

2. Select Chemical elements

3. Click Remove to eliminate dissimilar chemicals.

Data Gap Filling

Subcategorization by Chemical elements

The screenshot displays the QSAR Toolbox 3.1.0.21 interface. The 'Subcategorization' panel on the left shows various grouping methods, with 'Structure similarity' highlighted in a red box (callout 2). The main window shows a table of chemical structures and their predicted values. The table has columns for chemical structures and predicted values (M: 12 mg/kg/day, 60 mg/kg/day, 10, 30 mg/kg/day, 40 mg/kg/day, 5, 20 mg/kg/day, 60 mg/kg/day, 6). Below the table is a scatter plot titled 'Read across prediction of LOEL, taking the average from the nearest 5 neighbours, based on 5 values from 5 neighbour chemicals, Observed target value: 12.0 mg/kg/day, Predicted target value: 19.6 mg/kg/day'. The plot shows LOEL (obs.) on the y-axis (log₁₀(mg/kg/day)) and log Kow on the x-axis. Three callout boxes are present: 1. A blue callout box with the number '1' pointing to the 'Subcategorize' button in the 'Accept prediction' panel. 2. A blue callout box with the number '2' pointing to the 'Structure similarity' option in the 'Grouping methods' list. 3. A blue callout box with the number '3' pointing to the 'Remove' button in the 'Selected 2 (3/5)' list.

Structural similarity is applied in order to refine the category to the most similar analogues

1. **Click** Subcategorize.

2. **Select** Structural similarity (Dice, Atom centered

fragments, Atom type; Count H attached; Hybridizations). Select first two Categories (hold Ctrl button)

3. **Click** Remove to eliminate chemicals with similarity less than 50%

Data Gap Filling

Read across result for Total LOEL

The screenshot displays the QSAR Toolbox 3.1.0.21 interface. The 'Data Gap Filling' tab is active. The central workspace shows a table with columns for chemical structures and their corresponding LOEL values. The table is as follows:

Structure	1 [target]	11	14	17
Structure	<chem>Cc1ccc(N)cc1</chem>	<chem>Cc1ccc(N)cc1</chem>	<chem>Cc1ccc(N)cc1</chem>	<chem>Cc1ccc(N)cc1</chem>
Total (28/51)	M: 12 mg/kg/day, 60 mg/kg/day	M: 2 mg/kg/day, 10 mg/kg/day	M: 30 mg/kg/day	M: 40 mg/kg/day, 50 mg/kg/day
Oral (Water Containing) (3/6)				

Below the table, a scatter plot shows the 'Read across prediction of LOEL, taking the average from the nearest 5 neighbours, based on 3 values from 3 neighbour chemicals, Observed target value: 12.0 mg/kg/day, Predicted target value: 13.9 mg/kg/day'. The x-axis is 'log Kow' and the y-axis is 'LOEL (obs.), log(1/mmol/kg/day)'. A red circle highlights the predicted target value of 13.9 mg/kg/day in the text above the plot.

On the right sidebar, two callouts are present:

- Callout 1 points to the 'Accept prediction' button.
- Callout 2 points to the 'Return to matrix' button.

1. Click Accept prediction

2. Click Return to matrix

Data Gap Filling

Result of read-across prediction

QSAR Toolbox 3.1.0.21 [Document_1]

QSAR TOOLBOX

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

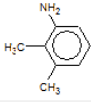
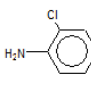
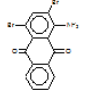
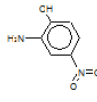
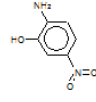
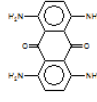
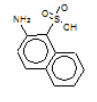
The OECD QSAR Toolbox for Grouping Chemicals into Categories
Developed by LMC, Bulgaria

Data Gap Filling Method

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

Target Endpoint

not defined

Total	1 [target]	2	3	4	5	6	7
Structure							
Substance Identity							
Environmental Fate and Transport							
Human Health Hazards							
Repeated Dose Toxicity							
LOEL							
Rat							
Oral (Feed)	(3/6)			M: 192 mg/kg/day, ...			M: 192 mg/kg/day, ...
Oral (Gavage)							
Whole Body							
Total	(28/52) M: 12 mg/kg/day, 60 mg/kg/day R: 13.9 mg/kg/day	M: 10 mg/kg/day, 1...		M: 125 mg/kg/day, ...	M: 100 mg/kg/day, ...		M: 1E3 mg/kg/day
Oral (Water Containing)	(3/6)						
NOEL	(38/72) M: 12 mg/kg/day, <12 mg/kg/day R: 10 mg/kg/day	M: <10 mg/kg/day, ...	M: <192 mg/kg/day, ...	M: 62.5 mg/kg/day, ...	M: 100 mg/kg/day, ...	M: 92.3 mg/kg/day, ...	M: 200 mg/kg/day
Profile							

1. Read across prediction for LOEL: 13.9 mg/kg/day coincide with experimental LOEL: 12 mg/kg/day

Read across predictions for 2,3 dimethylaniline (CAS 87-59-2) Result

Ultimate prediction:

Total NOEL – 10 mg/kg/day

Total LOEL – 13.9 mg/kg/day

Based on obtained results (for total LOEL and total NOEL) the target chemical is classified as *Category 2* regarding GHS classification ¹

Table 3.9.2: Guidance values to assist in Category 2 classification

Route of exposure	Units	Guidance value range (dose/concentration)
Oral (rat)	mg/kg bw/d	10 - 100
Dermal (rat or rabbit)	mg/kg bw/d	20 - 200
Inhalation (rat) gas	ppm/6h/d	50 - 250
Inhalation (rat) vapour	mg/litre/6h/d	0.2 - 1.0
Inhalation (rat) dust/mist/fume	mg/litre/6h/d	0.02 - 0.2

¹ Globally Harmonized System of Classification and Labeling of Chemicals (GHS):

http://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev04/English/ST-SG-AC10-30-Rev4e.pdf

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
 - Input
 - Profiling
 - Endpoint
 - Category definition
 - Data gap filling
- **Report**

Report

- Remember the report module allows you to generate a report on the predictions performed with the Toolbox. This module contains predefined report templates as well as a template editor with which users can define their own user defined templates. The report can then be printed or saved in different formats.
- Generating the report is shown on next screenshots

Report

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The 'Report' menu item is highlighted in the top bar. On the left, the 'Data Gap Filling Method' sidebar is active, showing 'Target Endpoint' as 'Human Health Hazards Repeated Dose Toxicity LOEL Rat Oral (Gavage) Whole Body Total'. The main workspace displays a tree view of 'Human Health Hazards Repeated Dose Toxicity' with a table of predictions. A context menu is open over a prediction, with 'Report' selected. A dialog box titled 'Select prediction to edit reporting info' is also visible, showing prediction details and a '3' in a blue callout.

Structure	1 [target]	2	3	4	5	6	7
Structure	<chem>Cc1ccc(N)cc1</chem>	<chem>Clc1ccccc1</chem>	<chem>Cc1ccc(N)cc1</chem>				
Substance Identity							
Environmental Fate and Transport							
Human Health Hazards							
Repeated Dose Toxicity							
LOEL							
Rat							
Oral (Feed)							
Oral (Gavage)							
Whole Body							
Total	(23/44) M: 12 mg/kg/day, 60 mg/kg/day R: 19.6 mg/kg/day		M: 10 mg/kg/day, 1... M: 125 mg/k				
Oral (Water Containing) (3/6)							
NOEL	(33/64) M: 12 mg/kg/day, <12 mg/kg/day R: 10 mg/kg/day		5 mg/k				
Profile							

1. **Select** prediction
2. **Right Click** and **Select Report**
3. **Select** the prediction for which you want to generate the report

Report

The screenshot shows the QSAR Toolbox 3.1.0.21 interface. The main window displays a report titled "Prediction of NOEL for 2,3-dimethylaniline". The report content includes the text "QSAR Toolbox prediction for single chemical" which is circled in red. A callout box with the number "1" points to the report content. The left sidebar shows available report templates, with "QSAR Toolbox Prediction Report (TPRF v.3.1)" selected under the "Standard (predefined)" category. The top menu bar includes options like Input, Profiling, Endpoint, Category Definition, Data Gap Filling, and Report.

1. Report for NOEL

Report

QSAR Toolbox 3.1.0.21 [Document_2]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Repository

Register Unregister Update Clone Design

Available data to report

- Predictions
- (Q)SARs
- Categories

Available report templates

- Standard (predefined)
 - QSAR Model Reporting Format (QMRF v.3.1)
 - QSAR Toolbox Category Report (CCRF v.3.1)
 - QSAR Toolbox Prediction Report (TPRF v.3.1)
- Custom (user defined)
 - Editable copy of QSAR Model Reporting Format (QMRF v.3.1)
 - Editable copy of QSAR Toolbox Category Report (CCRF v.3.1)
 - Editable copy of QSAR Toolbox Prediction Report (TPRF v.3.1)

show only relevant templates

Prediction [1]

Prediction of NOEL for 2,3-dimethylaniline 9 / 20

Human Health Hazards#Repeated Dose Toxicity	12.0 mg/kg/day
---	----------------

h. Additional data eliminations (not determined by domain):
Not available

i. Predicted value (model result):
10.00 mg/kg/day

j. Predicted value (comments):
Not provided by the user *manually editable field*

4.3. Applicability domain (OECD Principle 3):
The target chemical FALLS within applicability domain (see Section 3.1.b for detailed description of the domain)

1 2 3 4 5 6 7 8 9 10

Diagram showing applicability domain nodes (1-10) with 'AND' and 'NOT' labels. Nodes 3, 4, and 7 are marked 'NOT' with a red 'X'. Nodes 1, 2, 5, 6, 8, 9, 10 are marked 'AND' with a green checkmark.

1. Predicted value

2. Applicability domain

Report

QSAR Toolbox 3.1.0.21 [Document_1]

QSAR TOOLBOX

Input Profiling Endpoint Category Definition Data Gap Filling Report

Repository

Register Unregister Update Clone Design

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

Available data to report

- Predictions
- (Q)SARs
- Categories

Available report templates

Standard (predefined)

- QSAR Model Reporting Format (QMRF v.3.1)
- QSAR Toolbox Category Report (CCRF v.3.1)
- QSAR Toolbox Prediction Report (TPRF v.3.1)

Custom (user defined)

- Editable copy of QSAR Model Reporting Forma
- Editable copy of QSAR Toolbox Category Rept
- Editable copy of QSAR Toolbox Prediction Rep

Prediction [2]

Prediction of NOEL for 2,3-dimethylaniline 9 / 23

Appendix 1 - Category members

QSAR Toolbox prediction based on read-across

Prediction of NOEL for 2,3-dimethylaniline

APPENDIX 1 - Category members

From 26 category members, 10 chemicals are reported in more detail, the remaining category members are listed in a table with basic information

1. Cat. member No.1:

1.1. CAS number:
95-64-7

1.2. Other regulatory numbers:
Not reported

1.3. Chemical name(s):
benzenamine, 3,4-dimethyl-
3,4-dimethylaniline (3,4-xylidine)

used for read-across

1. Additional information for category members