QSAR TOOLBOX

The OECD QSAR Toolbox
for Grouping Chemicals
into Categories

OECD QSAR Toolbox v.4.1

Example illustrating endpoint vs. endpoint correlation for apical endpoints

Outlook

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

Background

This presentation is designed to introduce the user with:

- Illustration of different types endpoint vs. endpoint correlations using:
  - LLNA and GPMT skin sensitization data
  - DPRA and LLNA skin sensitization data
  - Skin sensitization and Ames mutagenicity data

Outlook

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

Objectives

This presentation demonstrates a number of functionalities of the Toolbox:

- Illustration of endpoint vs. endpoint correlations using different type endpoint data

Outlook

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

The exercise

- Illustration of different endpoint data correlations:
  - ➢ LLNA vs. GPMT skin sensitization data
  - ➢ DPRA (reactivity) vs. LLNA (skin sensitization) data
  - ➢ GPMT (skin sensitization) vs. Ames mutagenicity data

Outlook

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

Workflow

- **The Toolbox has six modules which are typically used in a workflow:**
  - Chemical Input
  - Profiling
  - Endpoints
  - Category Definition
  - Filling Data Gaps
  - Report
- **In this example we will use the modules in a different order, tailored to the aims of the example.**

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - **Correlation of data - background**

Correlation of endpoint data
Background

- This functionality introduce the user with opportunity to analyze correlations between selected gap filling endpoint (endpoint used for prediction) and other endpoint data.
- It is applicable for correlation analysis of data presented in ordinary, interval or ratio scale.
- If correlated data are measured in interval or ratio scale they are transformed in ordinary scale and the strength of the correlation is estimated by Spearman correlation coefficient.
- Basically, this functionality provides a correlation between target endpoint (this is the initial endpoint selected by the user) displayed on ordinate axis (Y-axis) and other endpoint data displayed on abscissa (X-axis).

Correlation of endpoint data
Spearman coefficient factor

- Spearman’s rank correlation coefficient is a nonparametric rank statistic proposed by Charles Spearman as a measure of the strength of an association between two variables. It assesses how well the relationship between two variables can be described using a monotonic function.
- Spearman correlation coefficient could be used for exploring the covary between:
  - two ranked variables
  - one measurement variable and one ranked variable (in this case, the measurement variable need to be to converted to ranks)
- Spearman correlation varies from -1 to +1 and the interpretation of the coefficient factor is provided below:
  - 0.00 – 0.19 – very weak correlation
  - 0.20 – 0.39 – weak correlation
  - 0.40 – 0.59 – moderate correlation
  - 0.60 – 0.79 – strong correlation
  - 0.80 – 1.0 – very strong

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - Correlation of data – background
  - **Types endpoint correlations**

Types endpoint correlations

Types endpoint correlations are as follows:

- Continuous vs. continuous*
- Categorical vs. categorical:
  - ✓Categorical vs. categorical
  - ✓Categorized continuous vs. categorical
  - ✓Categorized continuous vs. categorized continuous*

*Both type correlation is not illustrated in this presentations. They are presented in "Tutorial_4_TB 4.1_llustrating endpoint vs. endpoint correlation using ToxCast data"

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - Correlation of data – background
  - **Types endpoint correlations**
    - Categorical vs. categorical

Types endpoint correlations
Categorical vs. categorical

- The aim of this type correlation is to illustrate how categorical type data correlates each other.

- Categorical type data is the statistical data type consisting of categorical variables or of data that has been converted into that form. Such data is binary Ames data (dichotomic type): positive, negative or polytomic type data such as GPMT data: strong, weak and negative.

- Two examples illustrating this type correlation will be demonstrated:
  - Example 1: Correlation of two types skin sensitization data
    - LLNA (Positive, Negative) vs. GPMT (Weakly positive, Strongly positive, Negative)
  - Example 2: Correlation of skin sensitization and Ames mutagenicity data
    - LLNA (Negative, Weakly positive, Strongly positive) vs. AMES (Positive, Equivocal, Negative)

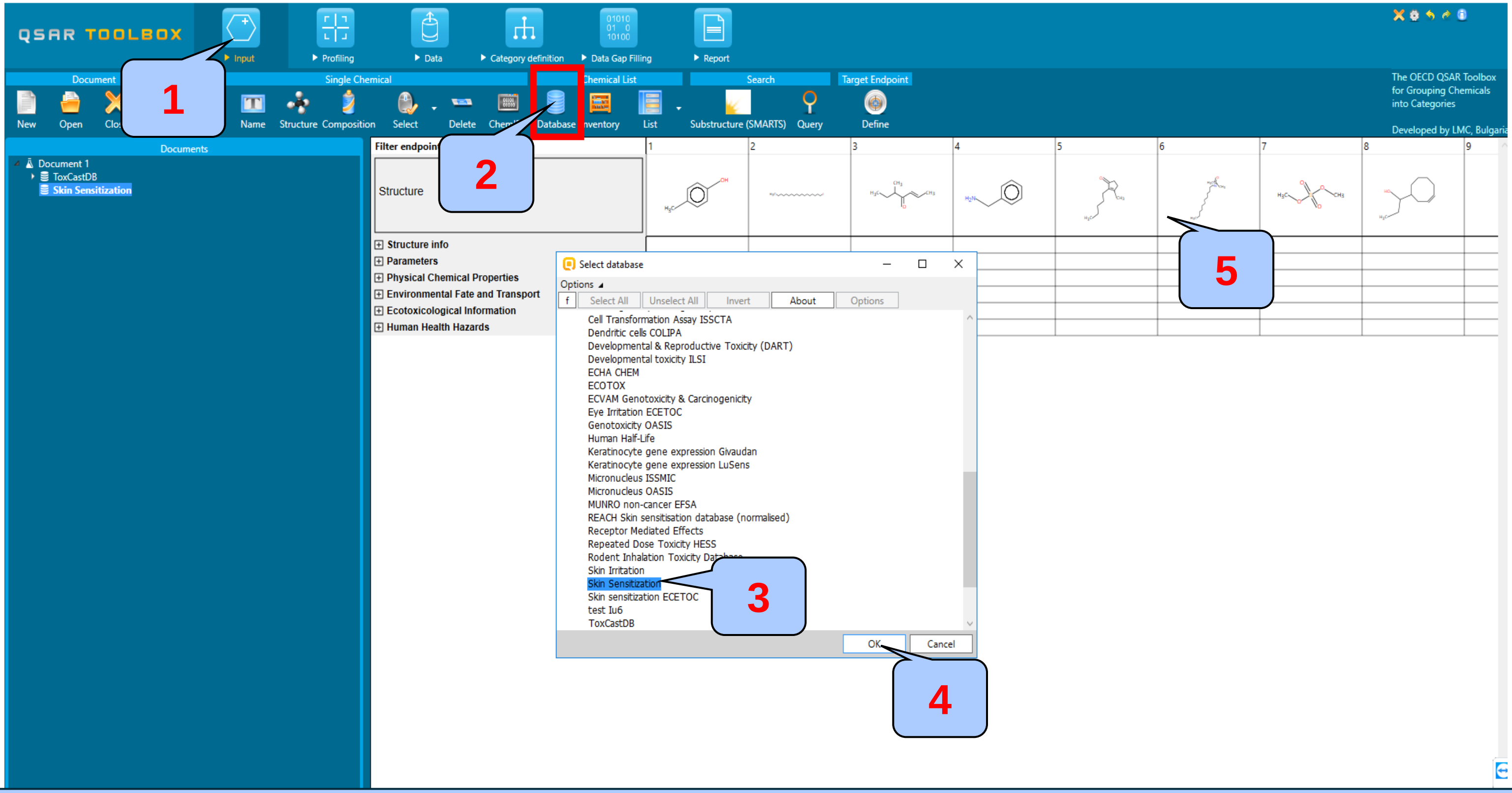
- Step by step workflow is presented on next few slides. Summary of the workflow steps are provided below:
    - *Load Skin sensitization database (step 1)*
    - *Gather experimental data (step 2)*
    - *Define target endpoint (step 3)*
    - *Enter Gap filling (step 4)*
    - *Perform correlation between endpoints (step 5).*

Types endpoint correlations

Categorical vs. categorical

Load Skin sensitization database – step 1

Example 1: Correlation of LLNA and GPMT data



Cell Transformation Assay ISSCTA
Dendritic cells COLIPA
Developmental & Reproductive Toxicity (DART)
Developmental toxicity ILSI
ECHA CHEM
ECOTOX
ECVAM Genotoxicity & Carcinogenicity
Eye Irritation ECETOC
Genotoxicity OASIS
Human Half-Life
Keratinocyte gene expression Givaudan
Keratinocyte gene expression LuSens
Micronucleus ISSMIC
Micronucleus OASIS
MUNRO non-cancer EFSA
REACH Skin sensitisation database (normalised)
Receptor Mediated Effects
Repeated Dose Toxicity HESS
Rodent Inhalation Toxicity Database
Skin Irritation
Skin Sensitization
Skin sensitization ECETOC
test Iu6
ToxCastDB

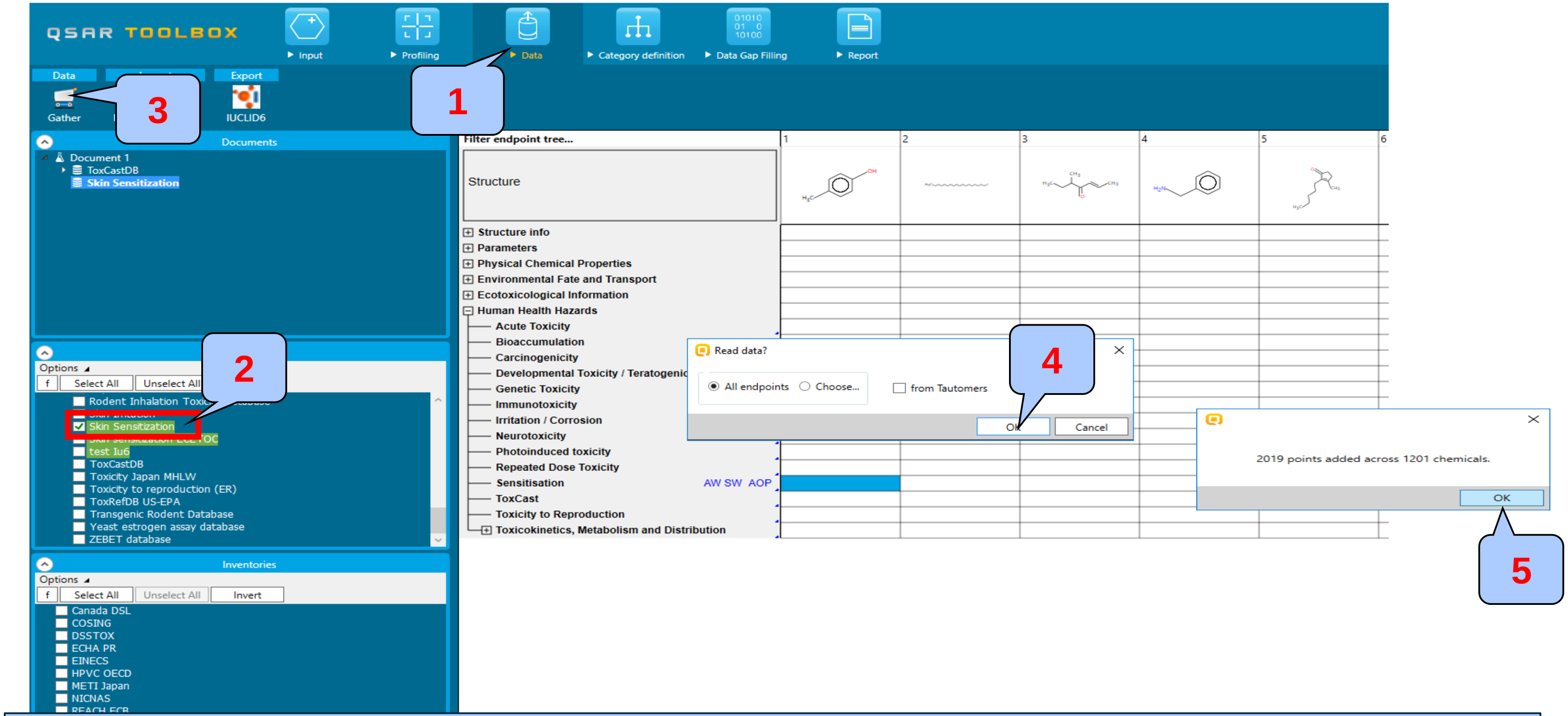
1. **Go** to “Input”; 2. **Click** “Database” button; 3. **Select** “Skin sensitization” database;
4. **Click** OK; 5. The chemicals from database have been loaded on datamatrix;

Types endpoint correlations

Categorical vs. categorical

Gather experimental data – step 2

Example 1: Correlation of LLNA and GPMT data



2019 points added across 1201 chemicals.

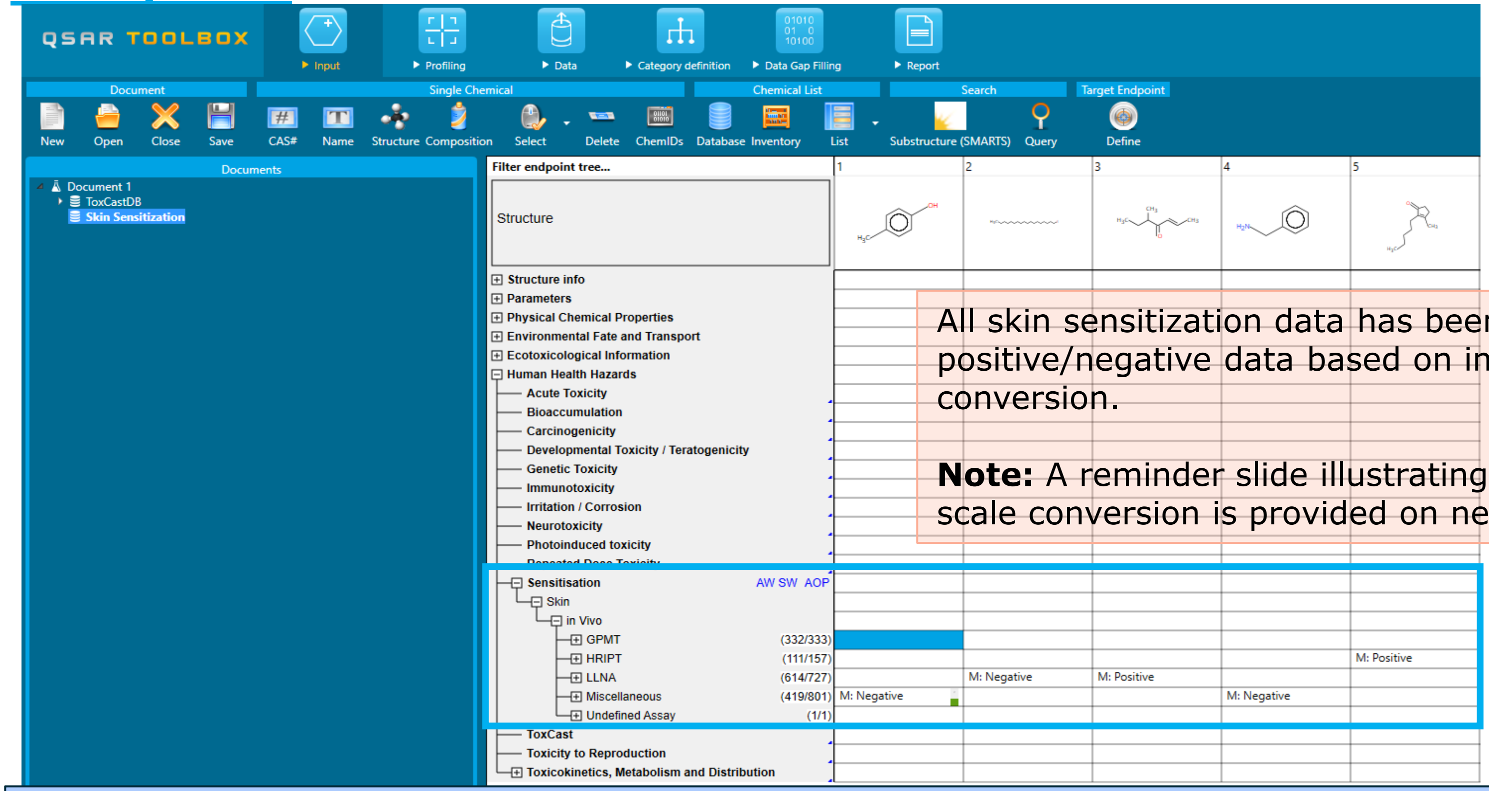
1. Go to “Data”; 2. **Select** “Skin sensitization”; 3. **Click** “Gather” 4. **Click** “OK”
5. Click “OK”;

Types endpoint correlations

Categorical vs. categorical

Gather experimental data – step 2

Example 1: Correlation of LLNA and GPMT data



Filter endpoint tree...	1	2	3	4	5
Structure					
Structure info					
Parameters					
Physical Chemical Properties					
Environmental Fate and Transport					
Ecotoxicological Information					
Human Health Hazards					
Acute Toxicity					
Bioaccumulation					
Carcinogenicity					
Developmental Toxicity / Teratogenicity					
Genetic Toxicity					
Immunotoxicity					
Irritation / Corrosion					
Neurotoxicity					
Photoinduced toxicity					
Repeated Dose Toxicity					
Sensitisation					
Skin					
in Vivo					
GPMT (332/333)					
HRIPT (111/157)					M: Positive
LLNA (614/727)		M: Negative	M: Positive		
Miscellaneous (419/801)	M: Negative			M: Negative	
Undefined Assay (1/1)					
ToxCast					
Toxicity to Reproduction					
Toxicokinetics, Metabolism and Distribution					

All skin sensitization data has been converted into positive/negative data based on implemented scale conversion.

Note: A reminder slide illustrating what is scale and scale conversion is provided on next click.

1. Skin sensitization data appeared on data matrix.
2. Data associated with different type assay (e.g LLNA, GPMT, HRIPT) are distributed in separate nodes

What is “scale” and “scale conversion” ?

Reminder slide

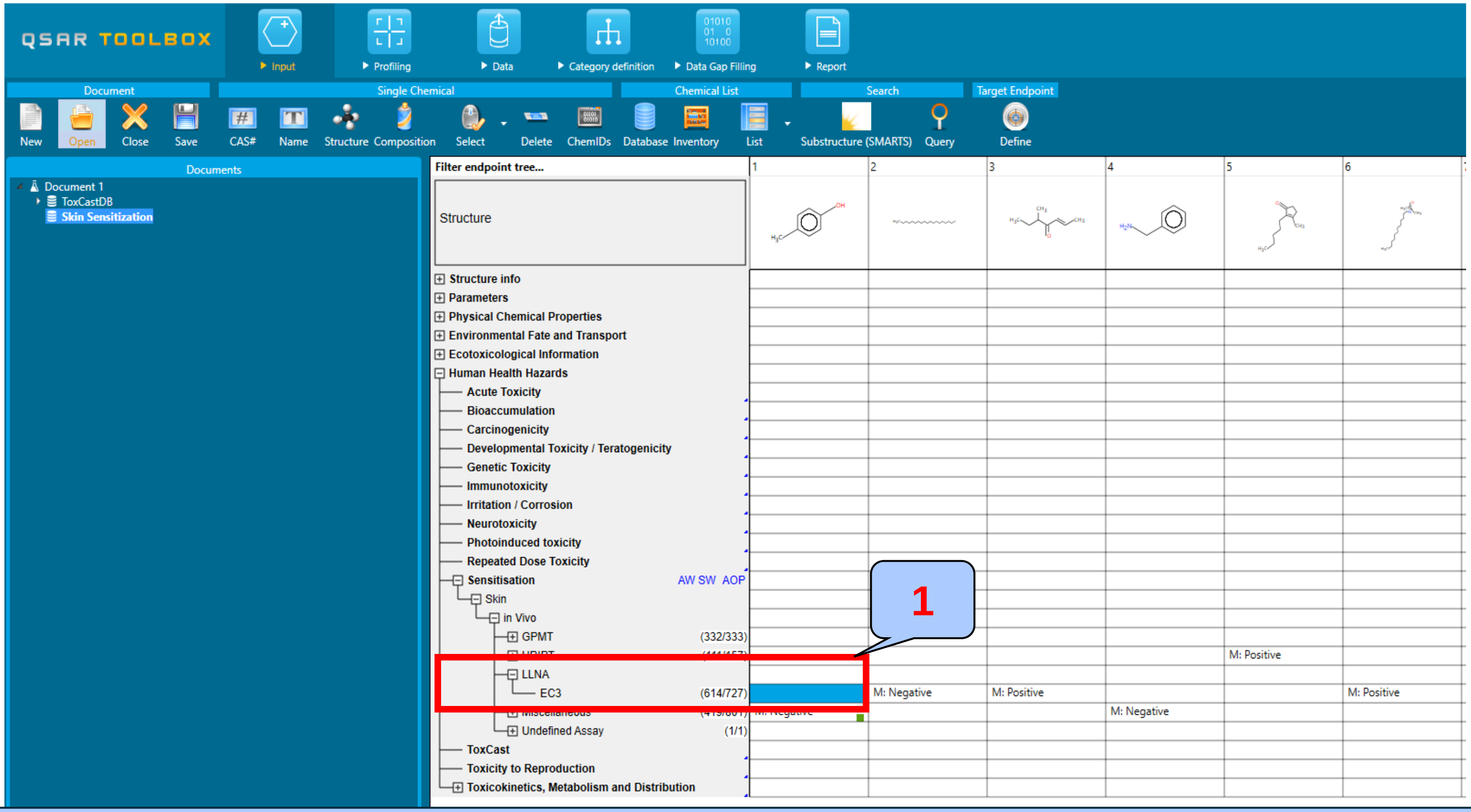
- Skin sensitisation as an example is a “qualitative” endpoint for which the results are presented with categorical type of data (for example: positive; negative; weak sensitizer; strong sensitizer, etc).
- Skin sensitisation potential of the chemicals came from different authors coded with different names (for example: data from John Moores University of Liverpool are: *Strongly sensitizing, Moderately sensitizing etc*.; data from European centre for Ecotoxicology and Toxicology of chemicals are: *Positive, Negative, and Equivocal*).
- The main purpose of the scales is to unify all data available in the Toolbox databases for a certain endpoint.
- “Scale conversion” is the TB instrument to create conversions between scales. More reasonable is to convert more informative to less informative scale.
- The default scale for Skin Sensitisation data is “Skin Sensitisation ECETOC”. It converts all skin sensitization data into: Positive and Negative. This allows skin sensitization data to be used as much as possible for gap filling purposes.

Types endpoint correlations

Categorical vs. categorical

Define target endpoint – step 3

Example 1: Correlation of LLNA and GPMT data



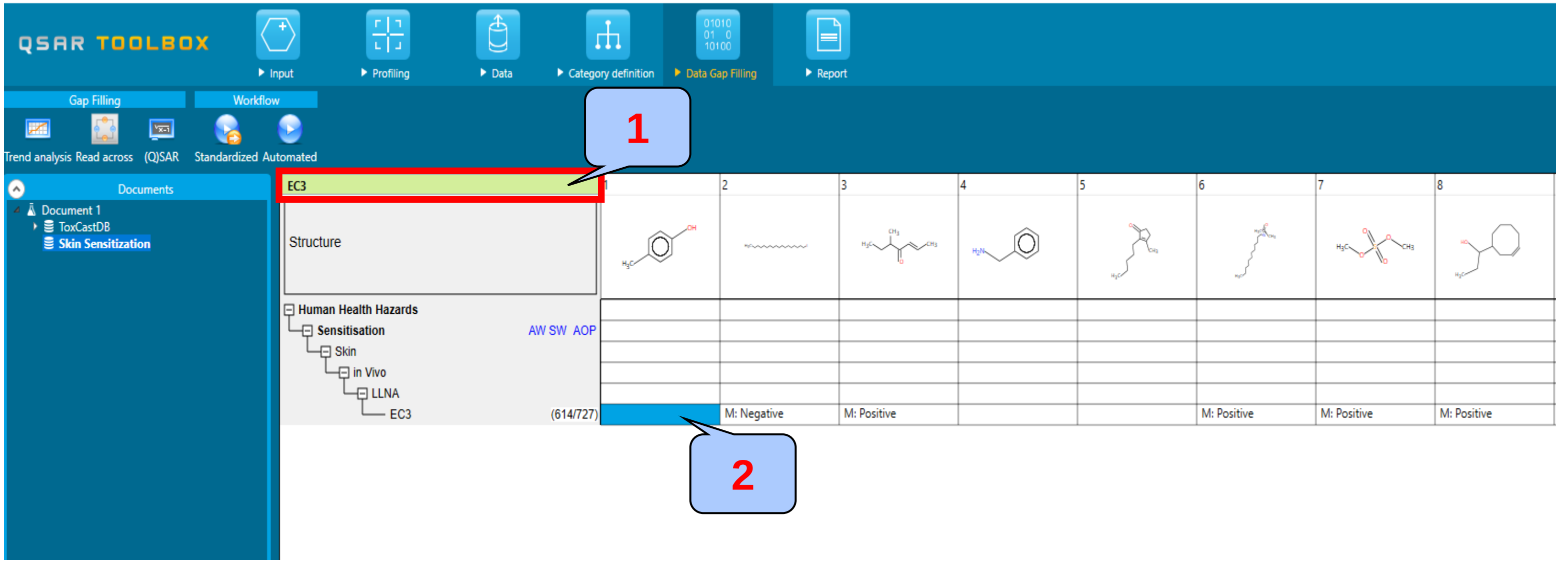
The target endpoint is EC3 data associated with LLNA assay.
1. **Click** on the cell associated with target endpoint;

Types endpoint correlations

Categorical vs. categorical

Define target endpoint – step 3

Example 1: Correlation of LLNA and GPMT data



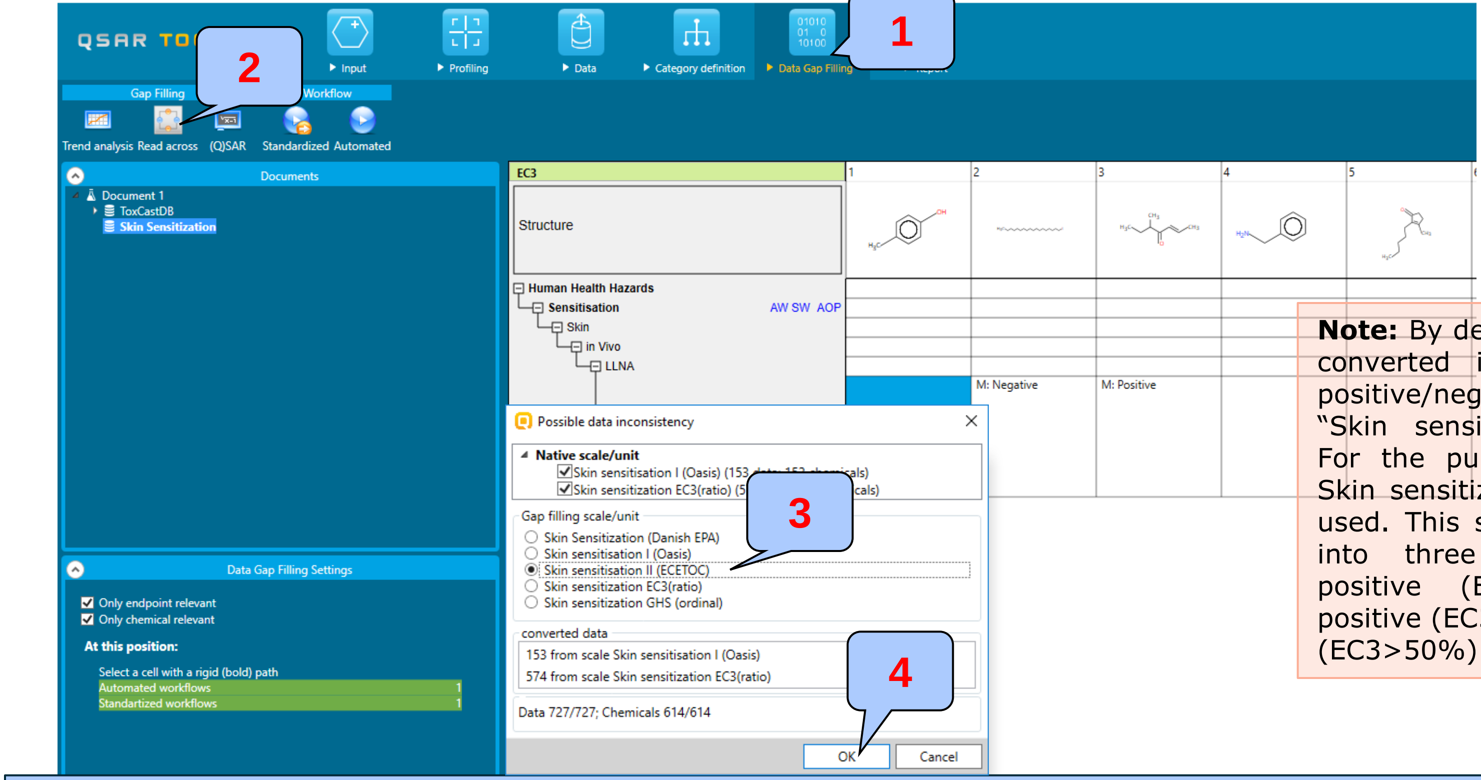
1. **Insert** type "EC3" data associated with LLNA assay in the filter box, then **press** "Enter" and automatically opening the tree to the target endpoint;
2. **Click** on the cell associated with target endpoint;

Types endpoint correlations

Categorical vs. categorical

Enter Gap filling – step 4

Example 1: Correlation of LLNA and GPMT data



Data Gap Filling Settings:

☑ Only endpoint relevant
☑ Only chemical relevant

At this position:

Select a cell with a rigid (bold) path
Automated workflows 1
Standartized workflows 1

Possible data inconsistency

Native scale/unit

☑ Skin sensitisation I (Oasis) (153 data; 153 chemicals)
☑ Skin sensitization EC3(ratio) (5 cals)

Gap filling scale/unit

○ Skin Sensitization (Danish EPA)
○ Skin sensitisation I (Oasis)
● Skin sensitisation II (ECETOC)
○ Skin sensitization EC3(ratio)
○ Skin sensitization GHS (ordinal)

converted data

153 from scale Skin sensitisation I (Oasis)
574 from scale Skin sensitization EC3(ratio)

Data 727/727; Chemicals 614/614

OK Cancel

Note: By default EC3 data has been converted into binary categories: positive/negative based on scale "Skin sensitization II (ECETOC)". For the purpose of this exercise, Skin sensitization I (OASIS) will be used. This scale converts EC3 data into three categories: Strongly positive (EC3 0-10%), Weakly positive (EC3 10-50%) and Negative (EC3>50%).

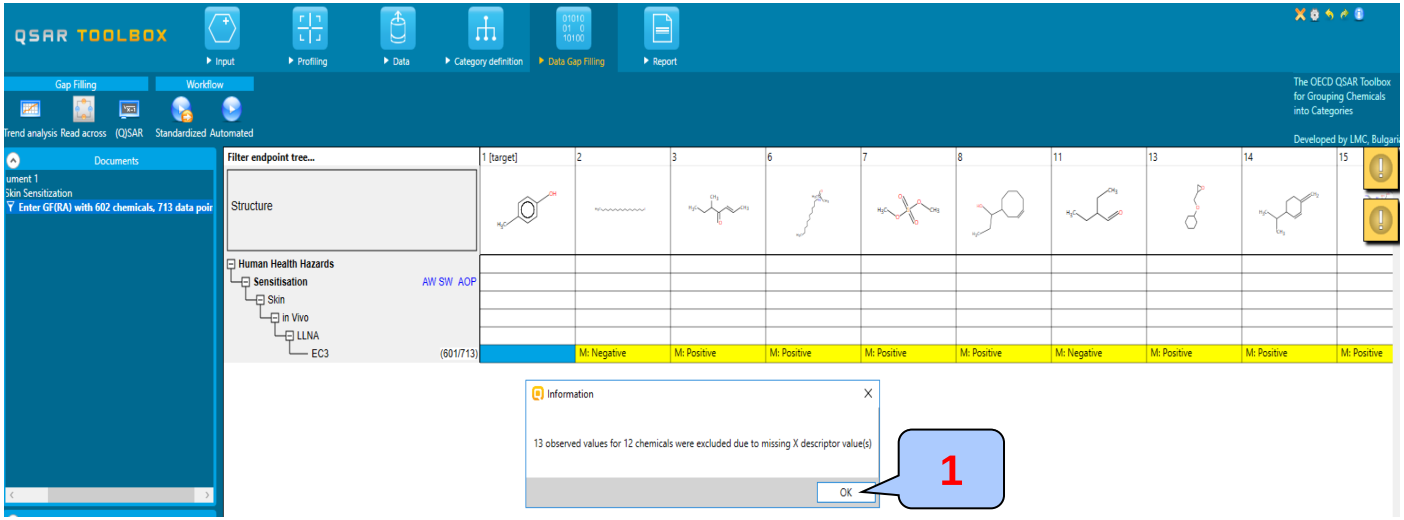
Enter Gap filling and apply read across. Read across is applied because a categorical type data is analyzed. Follow the steps:
1. **Go** to "Data Gap filling"; 2. **Select** "Read-across"; 3. **Select** "Skin sensitization II (ECETOC)" scale (see Note); 4. **Click** "OK";

Types endpoint correlations

Categorical vs. categorical

Enter Gap filling – step 4

Example 1: Correlation of LLNA and GPMT data



13 observed values for 12 chemicals were excluded due to missing X descriptor value(s)

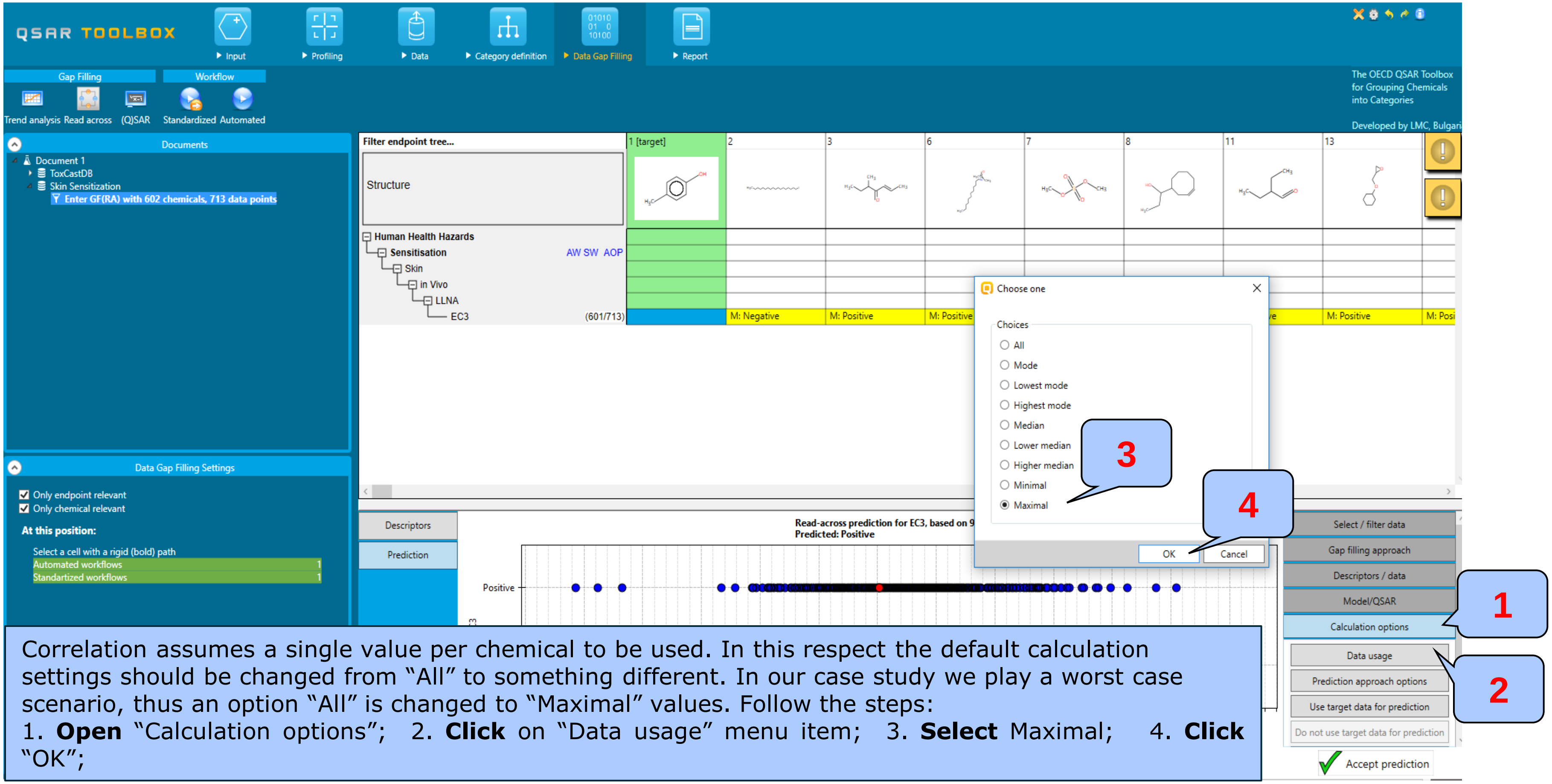
The message informing the user for how many chemicals with experimental data are excluded from gap filling due to missing X-descriptor value appeared 1. **Click** “OK”;

Types endpoint correlations

Categorical vs. categorical

Perform correlation between LLNA and GPMT data– step 5

Example 1: Correlation of LLNA and GPMT data



Choose one

Choices

- ○ All
- ○ Mode
- ○ Lowest mode
- ○ Highest mode
- ○ Median
- ○ Lower median
- ○ Higher median
- ○ Minimal
- ● Maximal

OK Cancel

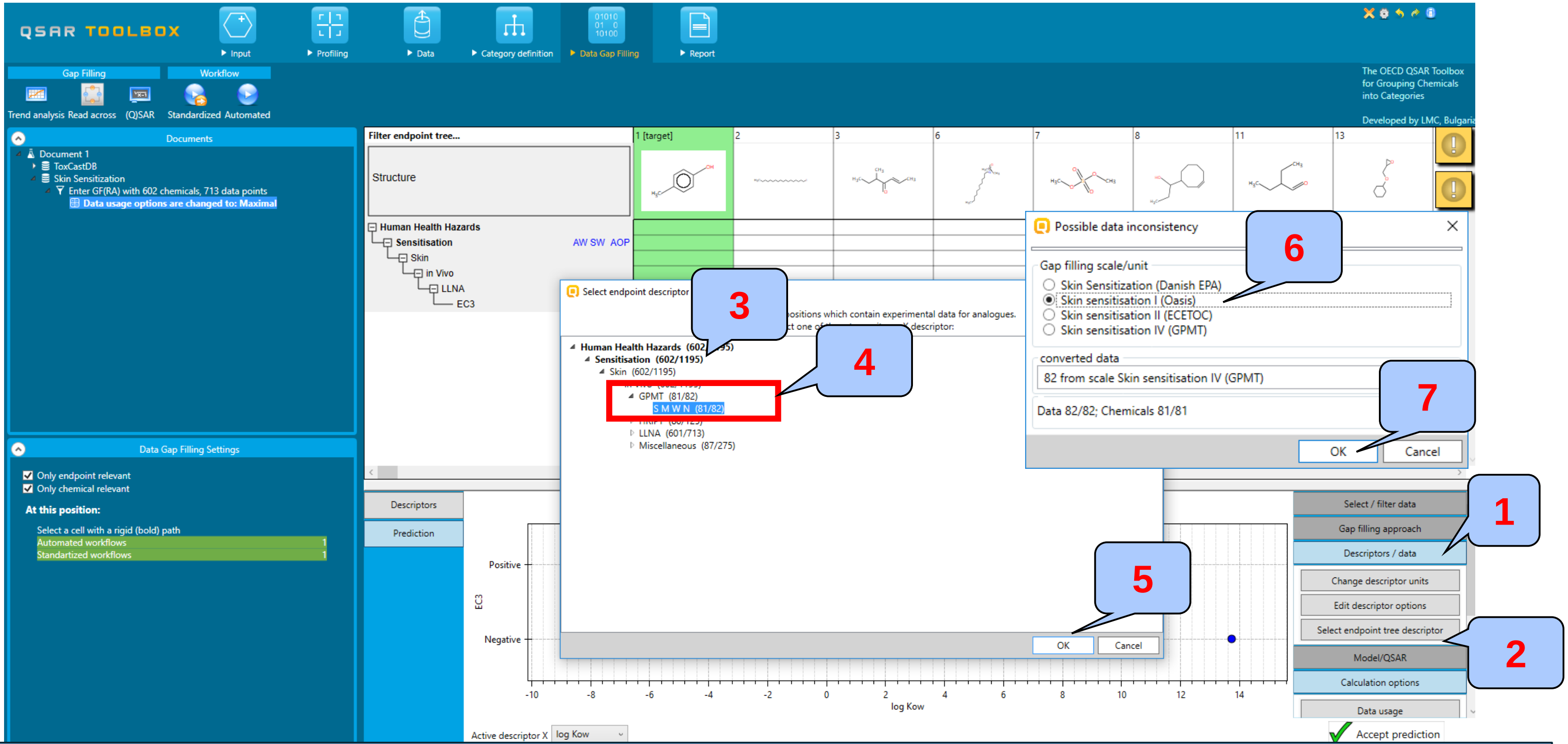
Correlation assumes a single value per chemical to be used. In this respect the default calculation settings should be changed from “All” to something different. In our case study we play a worst case scenario, thus an option “All” is changed to “Maximal” values. Follow the steps:
1. **Open** “Calculation options”; 2. **Click** on “Data usage” menu item; 3. **Select** Maximal; 4. **Click** “OK”;

Types endpoint correlations

Categorical vs. categorical

Perform correlation between LLNA and GPMT data– step 5

Example 1: Correlation of LLNA and GPMT data



Document 1
- ToxCastDB
- Skin Sensitization
  - Enter GF(RA) with 602 chemicals, 713 data points
    - **Data usage options are changed to: Maximal**

Human Health Hazards
- Sensitisation
  - Skin
    - in Vivo
      - LLNA
        - EC3

Data Gap Filling Settings

☑ Only endpoint relevant
☑ Only chemical relevant

At this position:

Select a cell with a rigid (bold) path

Automated workflows	1
Standartized workflows	1

Active descriptor X: log Kow

Select endpoint descriptor

...ositions which contain experimental data for analogues. ...ct one of ... X descriptor:

- **Human Health Hazards (602/...95)**
  - **Sensitisation (602/1195)**
    - Skin (602/1195)
      - in Vivo (602/1195)
        - GPMT (81/82)
          - S M W N (81/82)
        - HRIPT (80/125)
        - LLNA (601/713)
        - Miscellaneous (87/275)

OK Cancel

Possible data inconsistency

Gap filling scale/unit
○ Skin Sensitization (Danish EPA)
● Skin sensitisation I (Oasis)
○ Skin sensitisation II (ECETOC)
○ Skin sensitisation IV (GPMT)

converted data
82 from scale Skin sensitisation IV (GPMT)

Data 82/82; Chemicals 81/81

OK Cancel

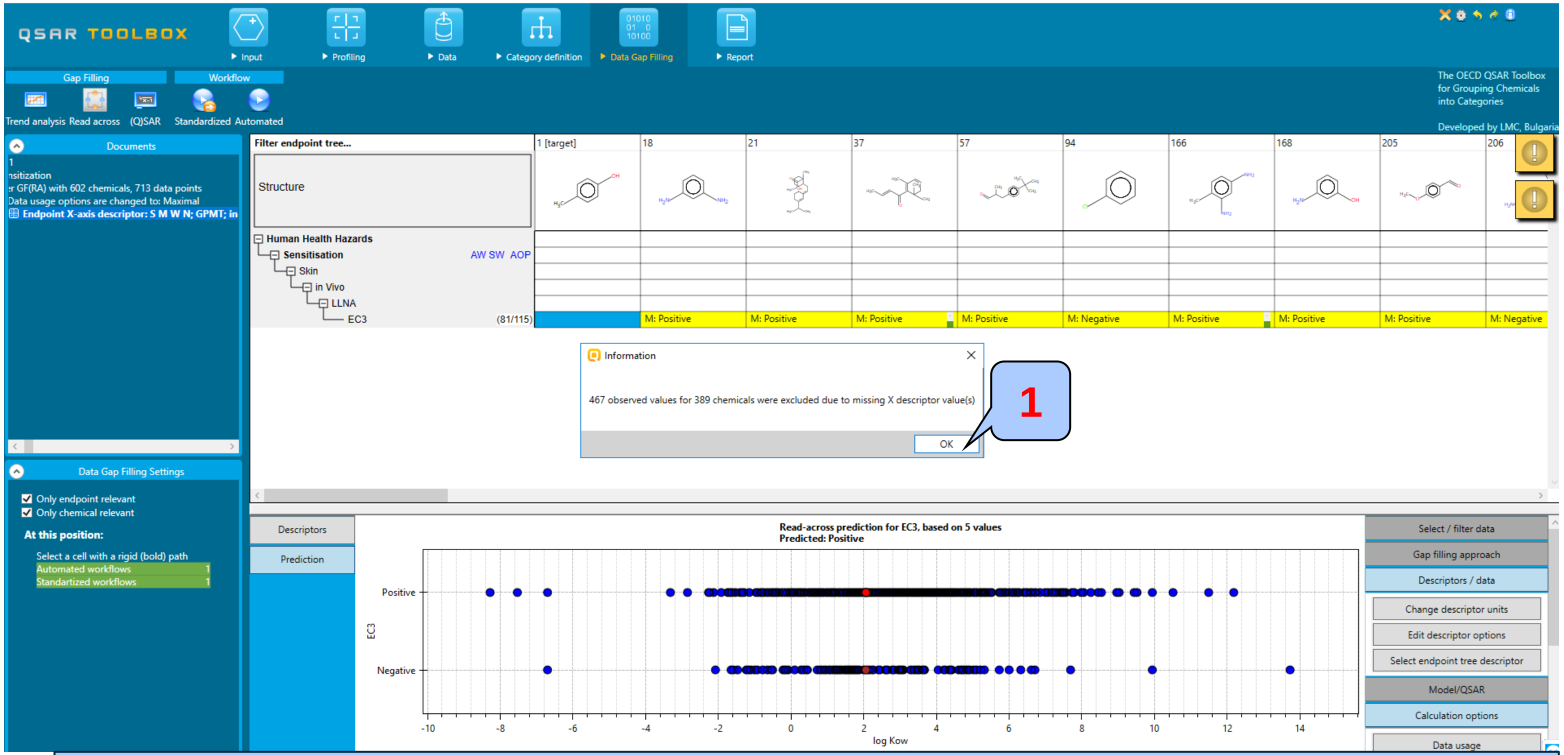
1. Open Descriptor/data tab; 2. Click on Select endpoint tree descriptor; 3. **Open** nodes under “Sensitization” node; 4. **Select** second endpoint, which will be placed on X-axis circled in red box: SMWN; 5. **Click** “OK” button; 6. **Select** Scale I OASIS 7. **Click** OK

Types endpoint correlations

Categorical vs. categorical

Perform correlation between LLNA and GPMT data– step 5

Example 1: Correlation of LLNA and GPMT data



1
nsitization
r GF(RA) with 602 chemicals, 713 data points
Data usage options are changed to: Maximal
Endpoint X-axis descriptor: S M W N; GPMT; in

☑ Only endpoint relevant
☑ Only chemical relevant

At this position:

Select a cell with a rigid (bold) path
Automated workflows 1
Standardized workflows 1

Information

467 observed values for 389 chemicals were excluded due to missing X descriptor value(s)

OK

Read-across prediction for EC3, based on 5 values
Predicted: Positive

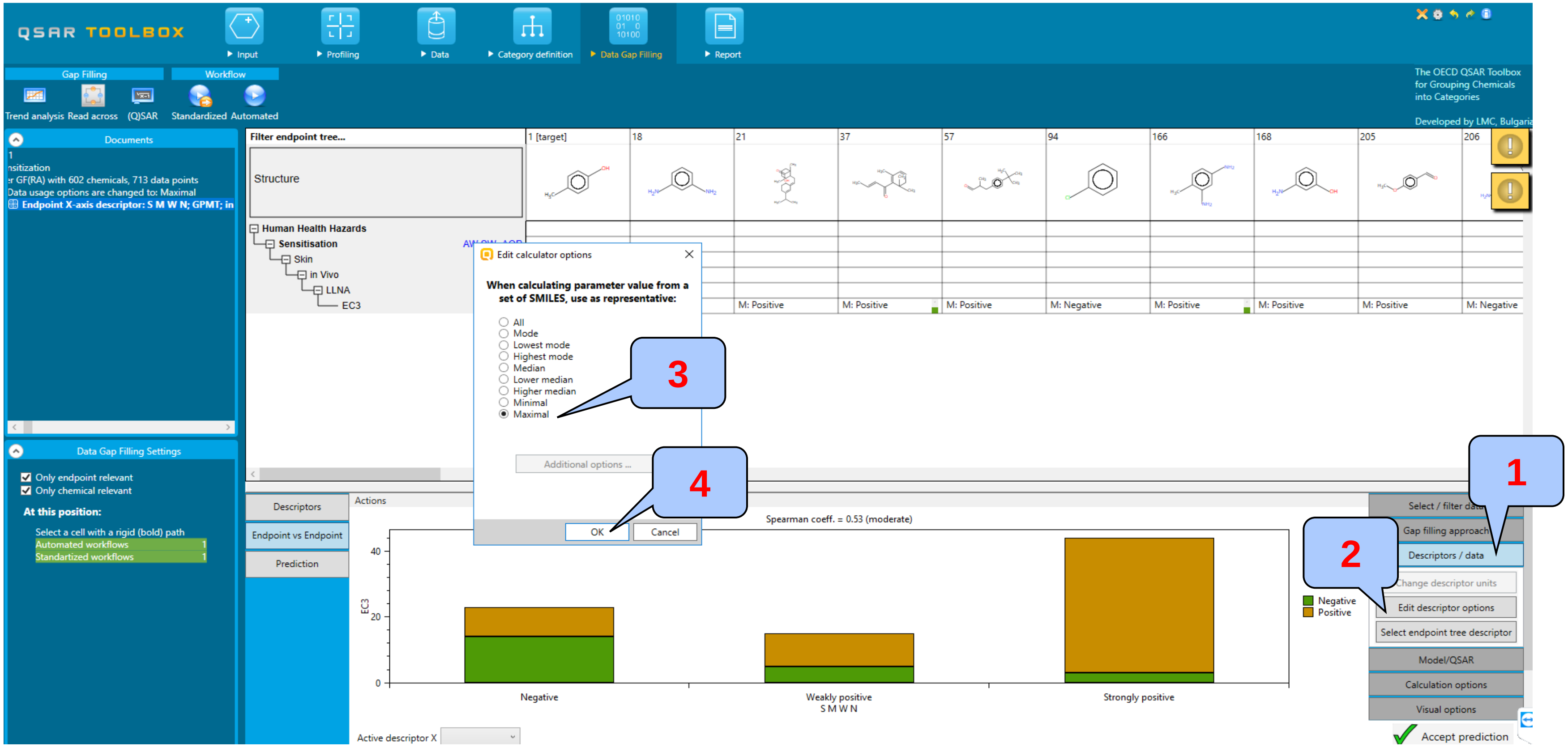
The message inform the user for how many chemicals will be excluded from correlation due to missing data for SMWN endpoint appears. This will not affect the value of correlation coefficient 1. **Click** “OK”;

Types endpoint correlations

Categorical vs. categorical

Perform correlation between LLNA and GPMT data– step 5

Example 1: Correlation of LLNA and GPMT data



1
nsitization
r GF(RA) with 602 chemicals, 713 data points
Data usage options are changed to: Maximal
Endpoint X-axis descriptor: S M W N; GPMT; in

☑ Only endpoint relevant
☑ Only chemical relevant

At this position:

Select a cell with a rigid (bold) path
Automated workflows 1
Standardized workflows 1

Edit calculator options

When calculating parameter value from a set of SMILES, use as representative:

○ All
○ Mode
○ Lowest mode
○ Highest mode
○ Median
○ Lower median
○ Higher median
○ Minimal
● Maximal

Additional options ...

OK Cancel

Spearman coeff. = 0.53 (moderate)

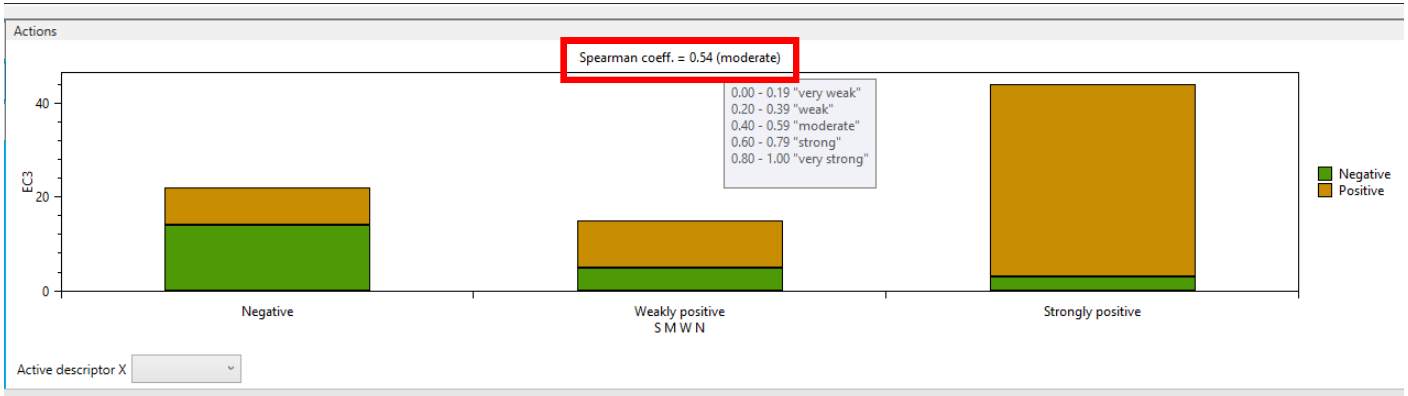
1.**Select** “Descriptor/ data”; 2. **Click** “Edit descriptor options” ; 3.**Select** “Maximal”; 4. **Click** OK

Types endpoint correlations

Categorical vs. categorical

Interpretation of correlation results (LLNA vs. GPMT)

Example 1: Correlation of LLNA and GPMT data



- Correlation analysis between two categorical type skin sensitization data (LLNA and GPMT) shows moderate endpoint correlation (Spearman coefficient is 0.54).

Types endpoint correlations
Categorical vs. categorical

- The second example illustrating categorical vs. categorical type correlation is:
  - Example 2: Correlation of Skin sensitization and Ames mutagenicity data
    - LLNA (Negative, Weakly positive, Strongly positive)
    - AMES (Positive, Equivocal, Negative)

- Step by step workflow is presented on next few slides. Summary of the workflow steps are provided below:
    - *Load Skin sensitization database (step 1) – skipped, because this database is already loaded on data matrix*
    - *Gather experimental data (step 2)*
    - *Define target endpoint (step 3)*
    - *Enter Gap filling (step 4)*
    - *Perform correlation between endpoints (step 5)*

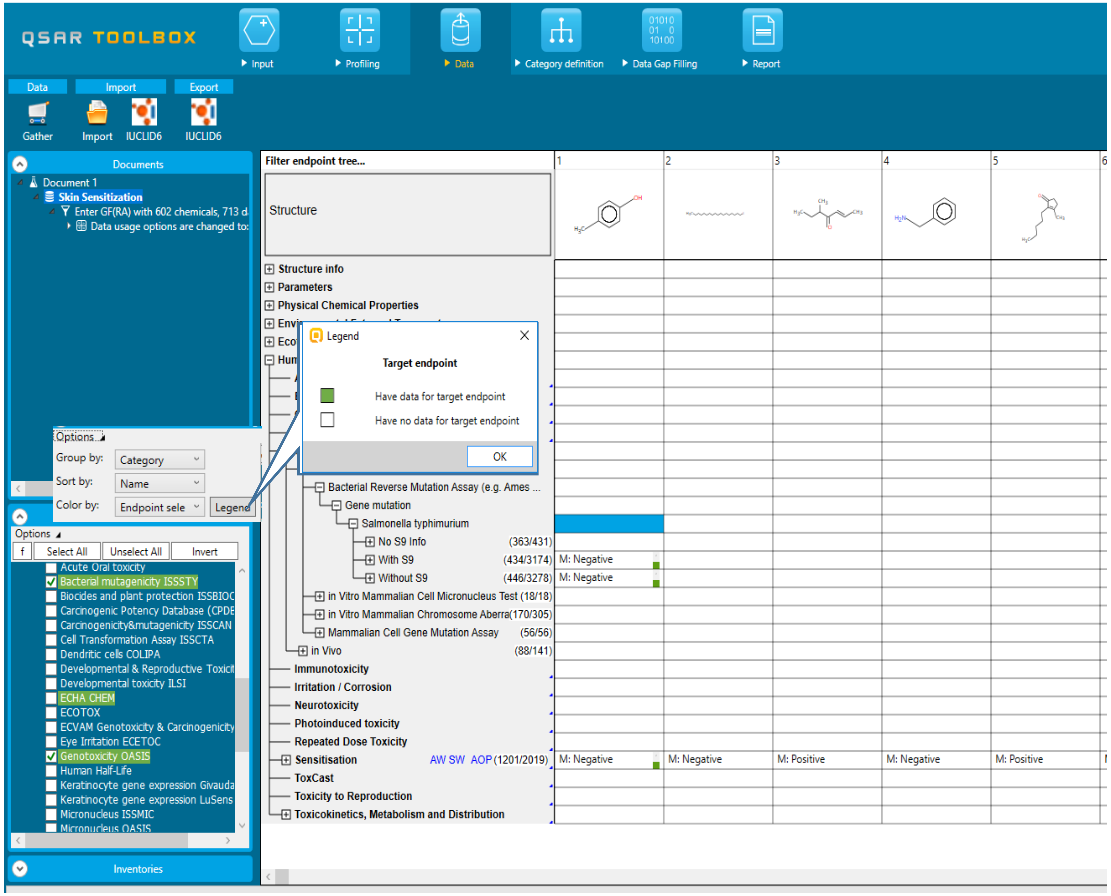
Types endpoint correlations

Categorical vs. categorical

Gather experimental data – step 2

Sidebar of database relevancy

Once the endpoint is selected, the relevant databases are highlighted.



Document 1
- Skin Sensitization
  - Enter GF(RA) with 602 chemicals, 713 d
    - Data usage options are changed to:

Options
Group by: Category
Sort by: Name
Color by: Endpoint sele | Legend

Legend

Target endpoint

🟩 Have data for target endpoint

□ Have no data for target endpoint

OK

Options
f | Select All | Unselect All | Invert

- □ Acute Oral toxicity
- ☑ Bacterial mutagenicity ISSSTY
- □ Biocides and plant protection ISSBIOC
- □ Carcinogenic Potency Database (CPDE
- □ Carcinogenicity&mutagenicity ISSCAN
- □ Cell Transformation Assay ISSCTA
- □ Dendritic cells COLIPA
- □ Developmental & Reproductive Toxicit
- □ Developmental toxicity ILSI
- □ ECHA CHEM
- □ ECOTOX
- □ ECVAM Genotoxicity & Carcinogenicity
- □ Eye Irritation ECETOC
- ☑ Genotoxicity OASIS
- □ Human Half-Life
- □ Keratinocyte gene expression Givauda
- □ Keratinocyte gene expression LuSens
- □ Micronucleus ISSMIC
- □ Micronucleus OASIS

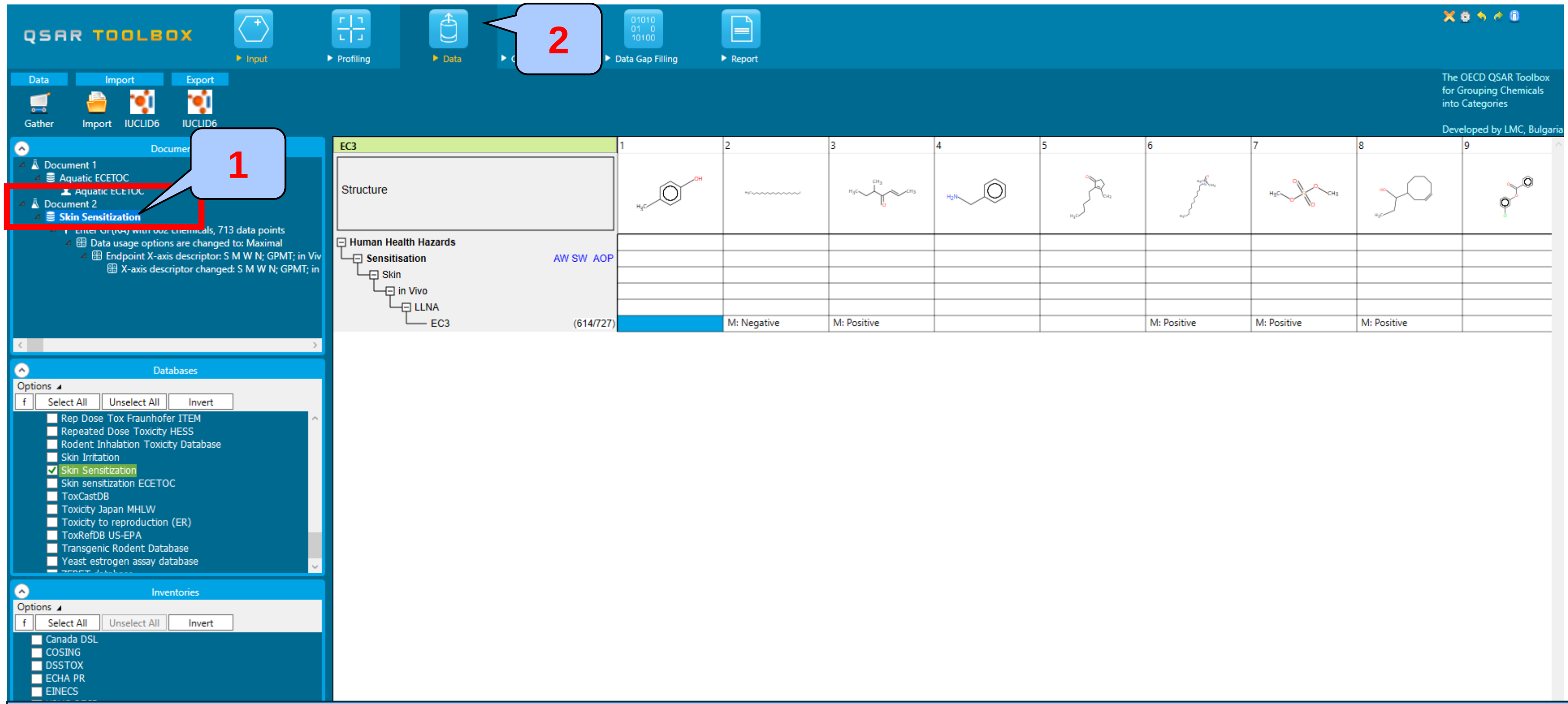
```
⊞ Structure info
⊞ Parameters
⊞ Physical Chemical Properties
⊞ Envi
⊞ Eco
⊟ Hum
      ⊟ Bacterial Reverse Mutation Assay (e.g. Ames ...
         ⊟ Gene mutation
            ⊟ Salmonella typhimurium
               ⊞ No S9 Info                          (363/431)
               ⊞ With S9                            (434/3174)  M: Negative
               ⊞ Without S9                         (446/3278)  M: Negative
      ⊞ in Vitro Mammalian Cell Micronucleus Test   (18/18)
      ⊞ in Vitro Mammalian Chromosome Aberra       (170/305)
      ⊞ Mammalian Cell Gene Mutation Assay          (56/56)
   ⊞ in Vivo                                        (88/141)
   Immunotoxicity
   Irritation / Corrosion
   Neurotoxicity
   Photoinduced toxicity
   Repeated Dose Toxicity
   ⊞ Sensitisation           AW SW  AOP (1201/2019)  M: Negative  M: Negative  M: Positive  M: Negative  M: Positive
   ToxCast
   Toxicity to Reproduction
   ⊞ Toxicokinetics, Metabolism and Distribution
```

Types endpoint correlations

Categorical vs. categorical

Gather experimental data – step 2

Example 2: Correlation of LLNA and AMES data



Documents

- Document 1
  - Aquatic ECETOC
    - Aquatic ECETOC
- Document 2
  - Skin Sensitization
    - Enter GP(RA) with 602 chemicals, 713 data points
      - Data usage options are changed to: Maximal
        - Endpoint X-axis descriptor: S M W N; GPMT; in Viv
          - X-axis descriptor changed: S M W N; GPMT; in

Databases

Options

f | Select All | Unselect All | Invert

☐ Rep Dose Tox Fraunhofer ITEM
☐ Repeated Dose Toxicity HESS
☐ Rodent Inhalation Toxicity Database
☐ Skin Irritation
☑ Skin Sensitization
☐ Skin sensitization ECETOC
☐ ToxCastDB
☐ Toxicity Japan MHLW
☐ Toxicity to reproduction (ER)
☐ ToxRefDB US-EPA
☐ Transgenic Rodent Database
☐ Yeast estrogen assay database

Inventories

Options

f | Select All | Unselect All | Invert

☐ Canada DSL
☐ COSING
☐ DSSTOX
☐ ECHA PR
☐ EINECS

EC3	1	2	3	4	5	6	7	8	9
Structure									
Human Health Hazards									
Sensitisation AW SW AOP									
Skin									
in Vivo									
LLNA									
EC3 (614/727)		M: Negative	M: Positive			M: Positive	M: Positive	M: Positive	

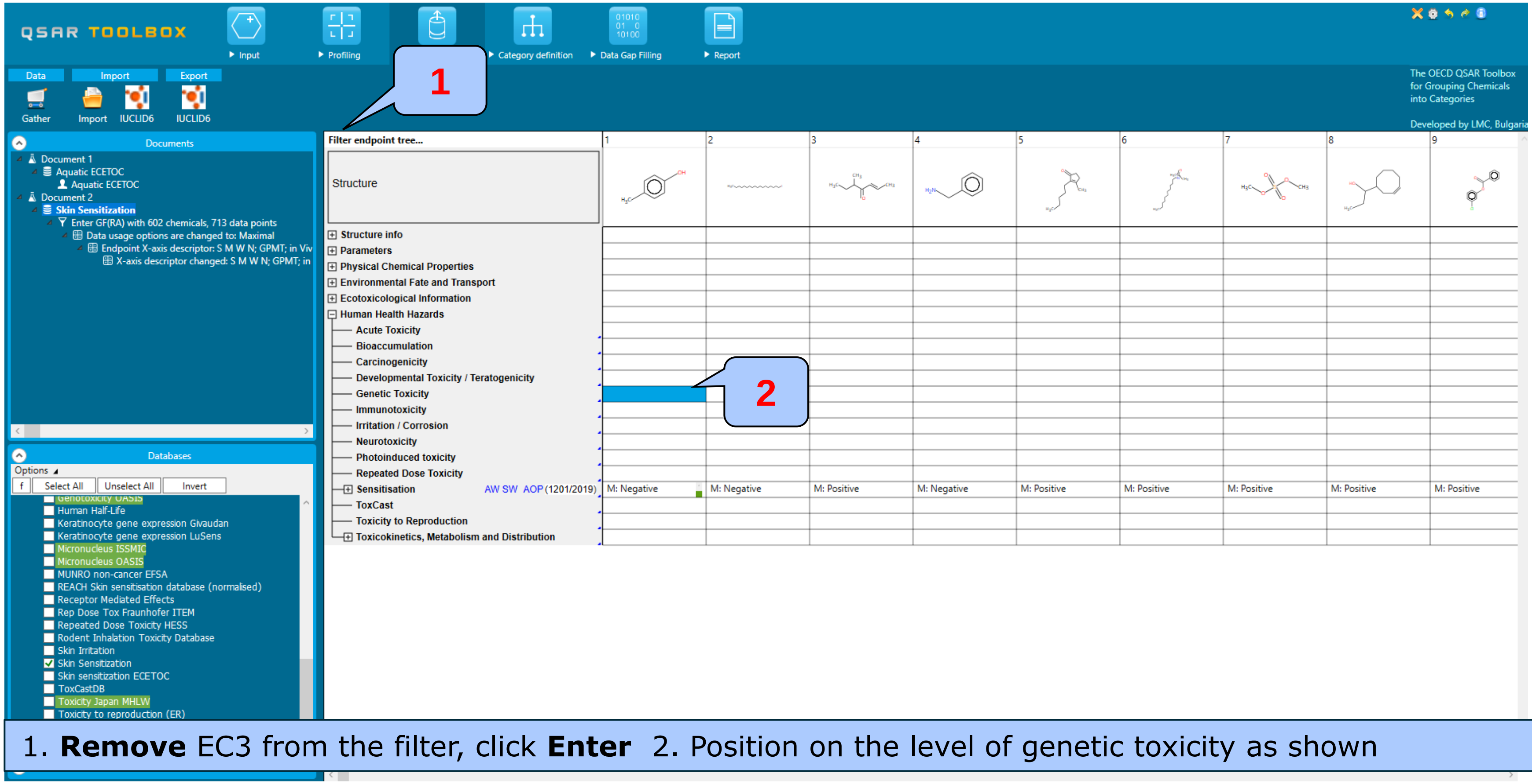
1. In order to start with next example please **click** on the level Skin sensitization from document tree 2. **Click** on Data tab

Types endpoint correlations

Categorical vs. categorical

Gather experimental data – step 2

Example 2: Correlation of LLNA and AMES data


QSAR TOOLBOX
Input
Profiling
Category definition
Data Gap Filling
Report
Data
Import
Export
Gather
Import
IUCLID6
IUCLID6
The OECD QSAR Toolbox for Grouping Chemicals into Categories
Developed by LMC, Bulgaria
1
2
Documents
Document 1
Aquatic ECETOC
Aquatic ECETOC
Document 2
Skin Sensitization
Enter GF(RA) with 602 chemicals, 713 data points
Data usage options are changed to: Maximal
Endpoint X-axis descriptor: S M W N; GPMT; in Viv
X-axis descriptor changed: S M W N; GPMT; in
Databases
Options
Select All
Unselect All
Invert
Genotoxicity OASIS
Human Half-Life
Keratinocyte gene expression Givaudan
Keratinocyte gene expression LuSens
Micronucleus ISSMIC
Micronucleus OASIS
MUNRO non-cancer EFSA
REACH Skin sensitisation database (normalised)
Receptor Mediated Effects
Rep Dose Tox Fraunhofer ITEM
Repeated Dose Toxicity HESS
Rodent Inhalation Toxicity Database
Skin Irritation
☑ Skin Sensitization
Skin sensitization ECETOC
ToxCastDB
Toxicity Japan MHLW
Toxicity to reproduction (ER)
Filter endpoint tree...
Structure
Structure info
Parameters
Physical Chemical Properties
Environmental Fate and Transport
Ecotoxicological Information
Human Health Hazards
Acute Toxicity
Bioaccumulation
Carcinogenicity
Developmental Toxicity / Teratogenicity
Genetic Toxicity
Immunotoxicity
Irritation / Corrosion
Neurotoxicity
Photoinduced toxicity
Repeated Dose Toxicity
Sensitisation AW SW AOP (1201/2019)
ToxCast
Toxicity to Reproduction
Toxicokinetics, Metabolism and Distribution
1 2 3 4 5 6 7 8 9
M: Negative
M: Negative
M: Positive
M: Negative
M: Positive
M: Positive
M: Positive
M: Positive
M: Positive

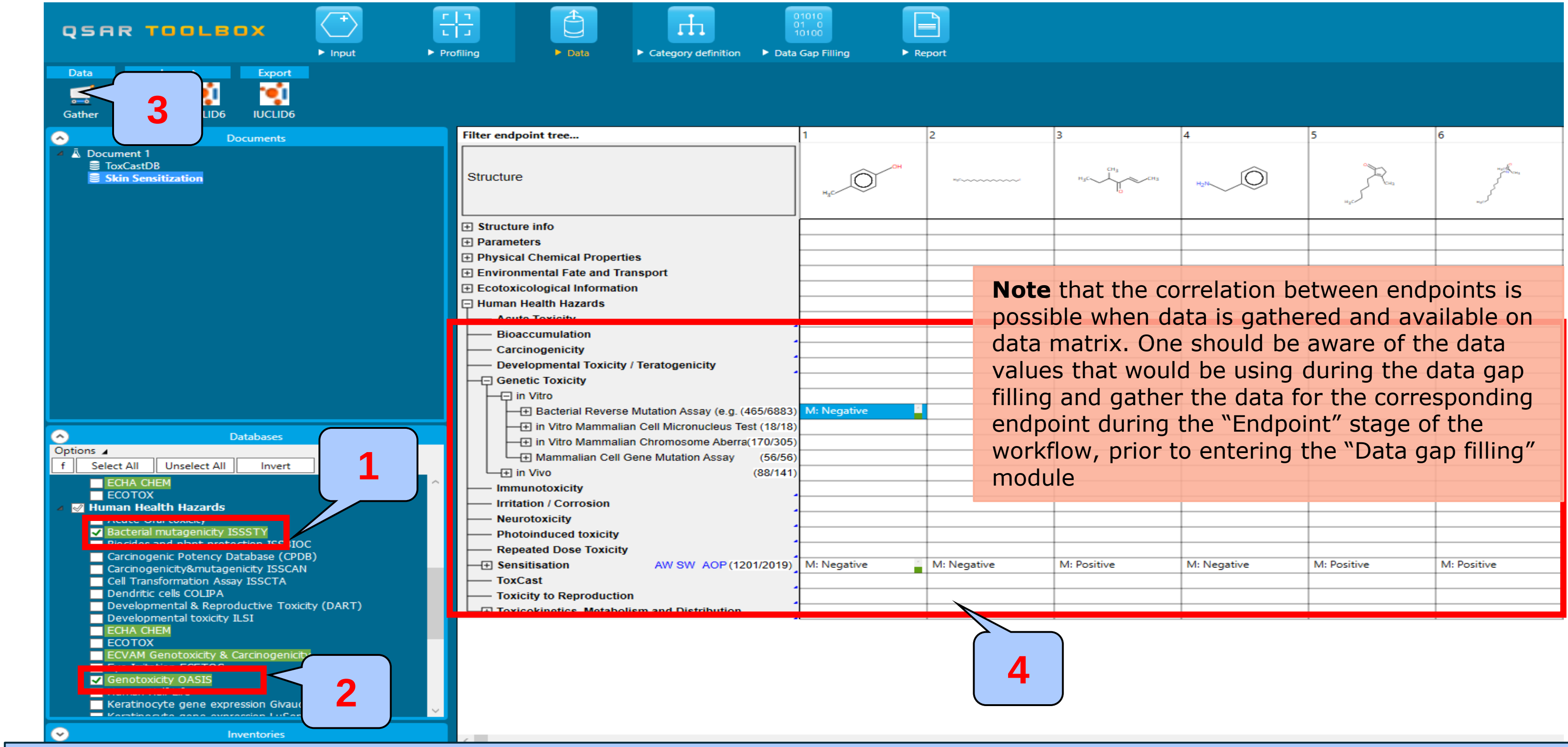
1. **Remove** EC3 from the filter, click **Enter** 2. Position on the level of genetic toxicity as shown

Types endpoint correlations

Categorical vs. categorical

Gather experimental data – step 2

Example 2: Correlation of LLNA and AMES data


QSAR TOOLBOX
Input
Profiling
Data
Category definition
Data Gap Filling
Report
Data
Export
Gather
LID6
IUCLID6
3
Documents
Document 1
ToxCastDB
Skin Sensitization
Databases
Options
f
Select All
Unselect All
Invert
ECHA CHEM
ECOTOX
Human Health Hazards
Acute Oral toxicity
☑ Bacterial mutagenicity ISSSTY
Biocides and plant protection ISSBIOC
Carcinogenic Potency Database (CPDB)
Carcinogenicity&mutagenicity ISSCAN
Cell Transformation Assay ISSCTA
Dendritic cells COLIPA
Developmental & Reproductive Toxicity (DART)
Developmental toxicity ILSI
ECHA CHEM
ECOTOX
ECVAM Genotoxicity & Carcinogenicity
Eye Irritation ECETOC
☑ Genotoxicity OASIS
Human Half-Life
Keratinocyte gene expression Givaudan
Keratinocyte gene expression LuSens
Inventories
1
2
Filter endpoint tree...
Structure
Structure info
Parameters
Physical Chemical Properties
Environmental Fate and Transport
Ecotoxicological Information
Human Health Hazards
Acute Toxicity
Bioaccumulation
Carcinogenicity
Developmental Toxicity / Teratogenicity
Genetic Toxicity
in Vitro
Bacterial Reverse Mutation Assay (e.g. (465/6883)
in Vitro Mammalian Cell Micronucleus Test (18/18)
in Vitro Mammalian Chromosome Aberra(170/305)
Mammalian Cell Gene Mutation Assay (56/56)
in Vivo (88/141)
Immunotoxicity
Irritation / Corrosion
Neurotoxicity
Photoinduced toxicity
Repeated Dose Toxicity
Sensitisation AW SW AOP (1201/2019)
ToxCast
Toxicity to Reproduction
Toxicokinetics, Metabolism and Distribution
1
2
3
4
5
6
M: Negative
M: Negative
M: Negative
M: Positive
M: Negative
M: Positive
M: Positive
4

Note that the correlation between endpoints is possible when data is gathered and available on data matrix. One should be aware of the data values that would be using during the data gap filling and gather the data for the corresponding endpoint during the “Endpoint” stage of the workflow, prior to entering the “Data gap filling” module

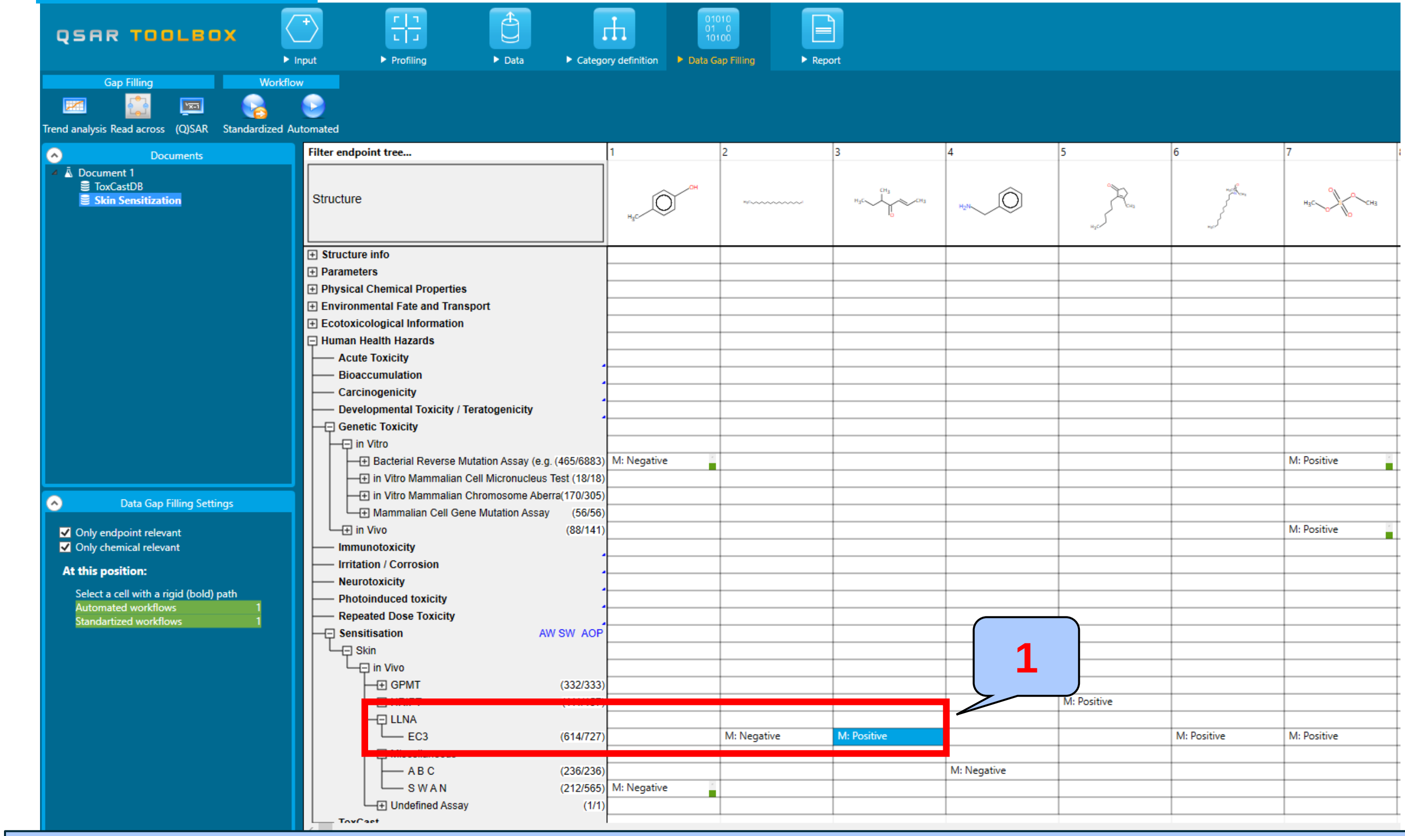
1. **Select** the databases including Ames data (green highlighted). Do not check ECHA Chem database.
2. Skin sensitization DB is already selected; 3. **Click** “Gather” 4. The data appeared on datamatrix;

Types endpoint correlations

Categorical vs. categorical

Define target endpoint – step 3

Example 2: Correlation of LLNA and AMES data



☑ Only endpoint relevant
☑ Only chemical relevant

At this position:

Select a cell with a rigid (bold) path

Automated workflows	1
Standartized workflows	1

Filter endpoint tree...

		1	2	3	4	5	6	7
Structure								
⊞ Structure info								
⊞ Parameters								
⊞ Physical Chemical Properties								
⊞ Environmental Fate and Transport								
⊞ Ecotoxicological Information								
⊟ Human Health Hazards								
Acute Toxicity								
Bioaccumulation								
Carcinogenicity								
Developmental Toxicity / Teratogenicity								
⊟ Genetic Toxicity								
⊟ in Vitro								
⊞ Bacterial Reverse Mutation Assay (e.g.	(465/6883)	M: Negative						M: Positive
⊞ in Vitro Mammalian Cell Micronucleus Test	(18/18)							
⊞ in Vitro Mammalian Chromosome Aberra	(170/305)							
⊞ Mammalian Cell Gene Mutation Assay	(56/56)							
⊞ in Vivo	(88/141)							M: Positive
Immunotoxicity								
Irritation / Corrosion								
Neurotoxicity								
Photoinduced toxicity								
Repeated Dose Toxicity								
⊟ Sensitisation	AW SW AOP							
⊟ Skin								
⊟ in Vivo								
⊞ GPMT	(332/333)							
[illegible]	[illegible]					M: Positive		
⊟ LLNA								
EC3	(614/727)		M: Negative	M: Positive			M: Positive	M: Positive
[illegible]								
A B C	(236/236)				M: Negative			
S W A N	(212/565)	M: Negative						
⊞ Undefined Assay	(1/1)							
ToxCast								

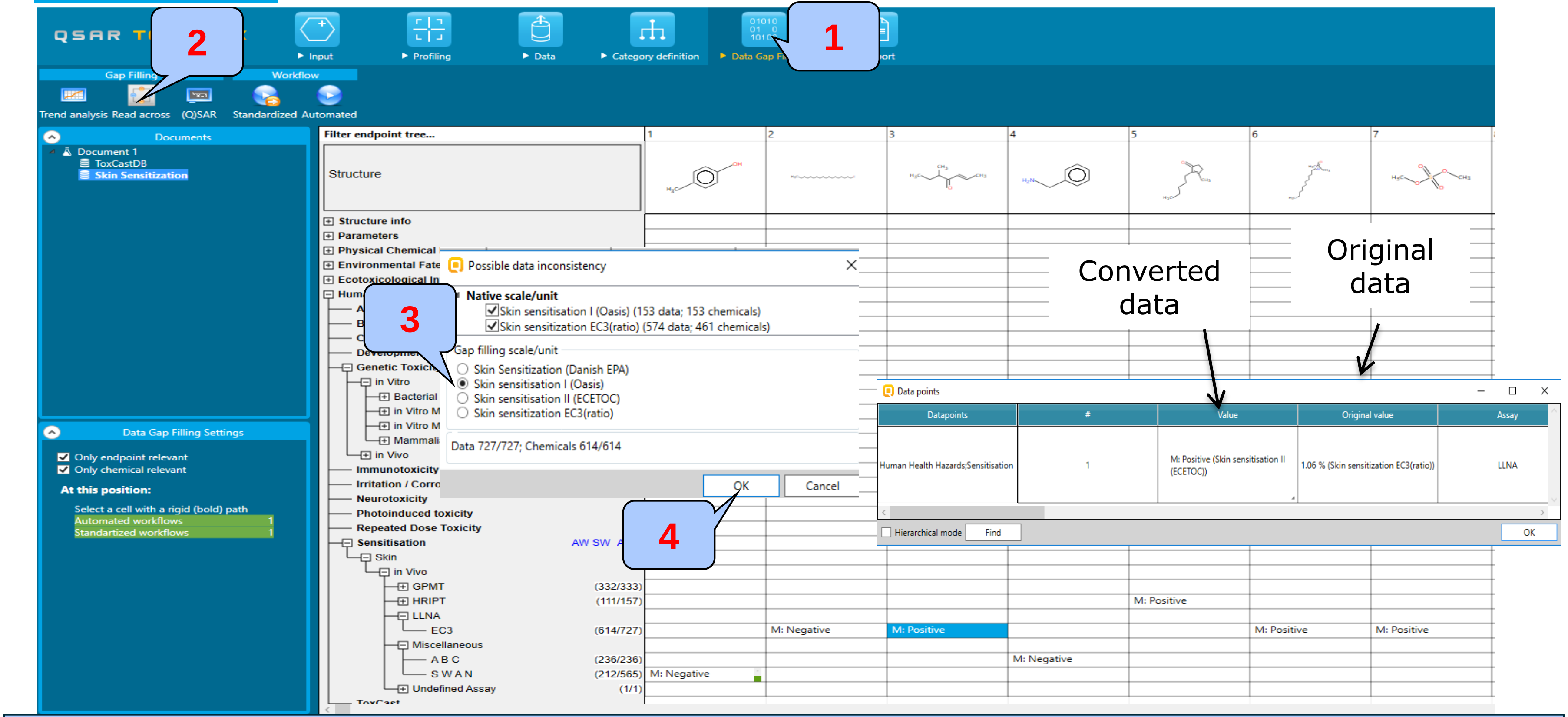
The target endpoint is skin sensitization/in vivo/LLNA/EC3;
1. **Click** on the cell associated with target endpoint;

Types endpoint correlations

Categorical vs. categorical

Enter Gap filling – step 4

Example 2: Correlation of LLNA and AMES data



☑ Only endpoint relevant
☑ Only chemical relevant

At this position:

Select a cell with a rigid (bold) path
Automated workflows 1
Standartized workflows 1

Possible data inconsistency

Native scale/unit
☑ Skin sensitisation I (Oasis) (153 data; 153 chemicals)
☑ Skin sensitization EC3(ratio) (574 data; 461 chemicals)

Gap filling scale/unit
○ Skin Sensitization (Danish EPA)
● Skin sensitisation I (Oasis)
○ Skin sensitisation II (ECETOC)
○ Skin sensitization EC3(ratio)

Data 727/727; Chemicals 614/614

OK Cancel

Data points

Datapoints	#	Value	Original value	Assay
Human Health Hazards;Sensitisation	1	M: Positive (Skin sensitisation II (ECETOC))	1.06 % (Skin sensitization EC3(ratio))	LLNA

□ Hierarchical mode Find OK

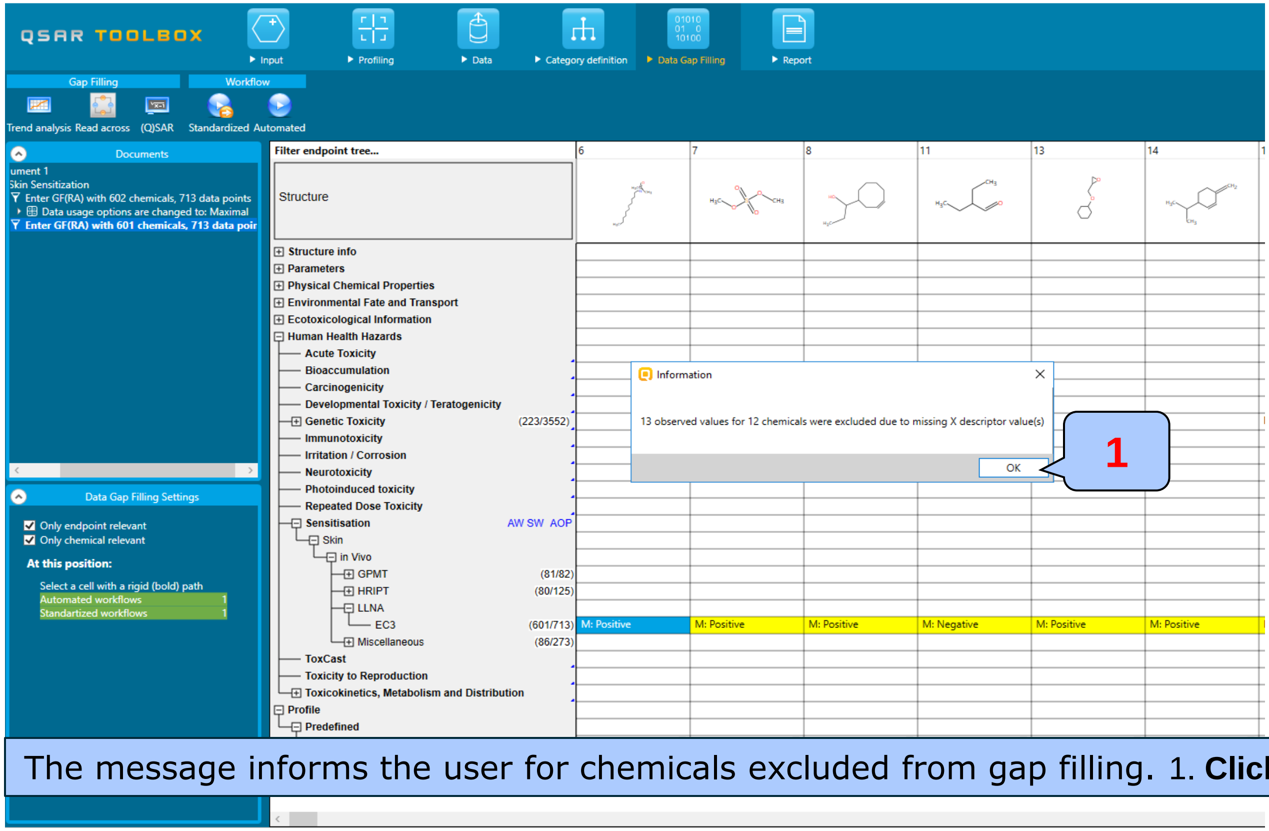
Enter Gap filling applying read across. Read across is applied because a categorical type data is analyzed.
1. **Go** to "Data Gap filling"; 2. **Select** "Read-across"; 3. **Check** "Skin sensitization I (OASIS)" scale;
5. **Click** "OK"

Types endpoint correlations

Categorical vs. categorical

Enter Gap filling – step 4

Example 2: Correlation of LLNA and AMES data



ument 1
Skin Sensitization
Enter GF(RA) with 602 chemicals, 713 data points
Data usage options are changed to: Maximal
Enter GF(RA) with 601 chemicals, 713 data poir

☑ Only endpoint relevant
☑ Only chemical relevant

At this position:

Select a cell with a rigid (bold) path

Automated workflows	1
Standartized workflows	1

- Structure info
- Parameters
- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
- Human Health Hazards
  - Acute Toxicity
  - Bioaccumulation
  - Carcinogenicity
  - Developmental Toxicity / Teratogenicity
  - Genetic Toxicity (223/3552)
  - Immunotoxicity
  - Irritation / Corrosion
  - Neurotoxicity
  - Photoinduced toxicity
  - Repeated Dose Toxicity
  - Sensitisation AW SW AOP
    - Skin
      - in Vivo
        - GPMT (81/82)
        - HRIPT (80/125)
        - LLNA
          - EC3 (601/713)
        - Miscellaneous (86/273)
  - ToxCast
  - Toxicity to Reproduction
  - Toxicokinetics, Metabolism and Distribution
- Profile
  - Predefined

6	7	8	11	13	14
M: Positive	M: Positive	M: Positive	M: Negative	M: Positive	M: Positive

13 observed values for 12 chemicals were excluded due to missing X descriptor value(s)

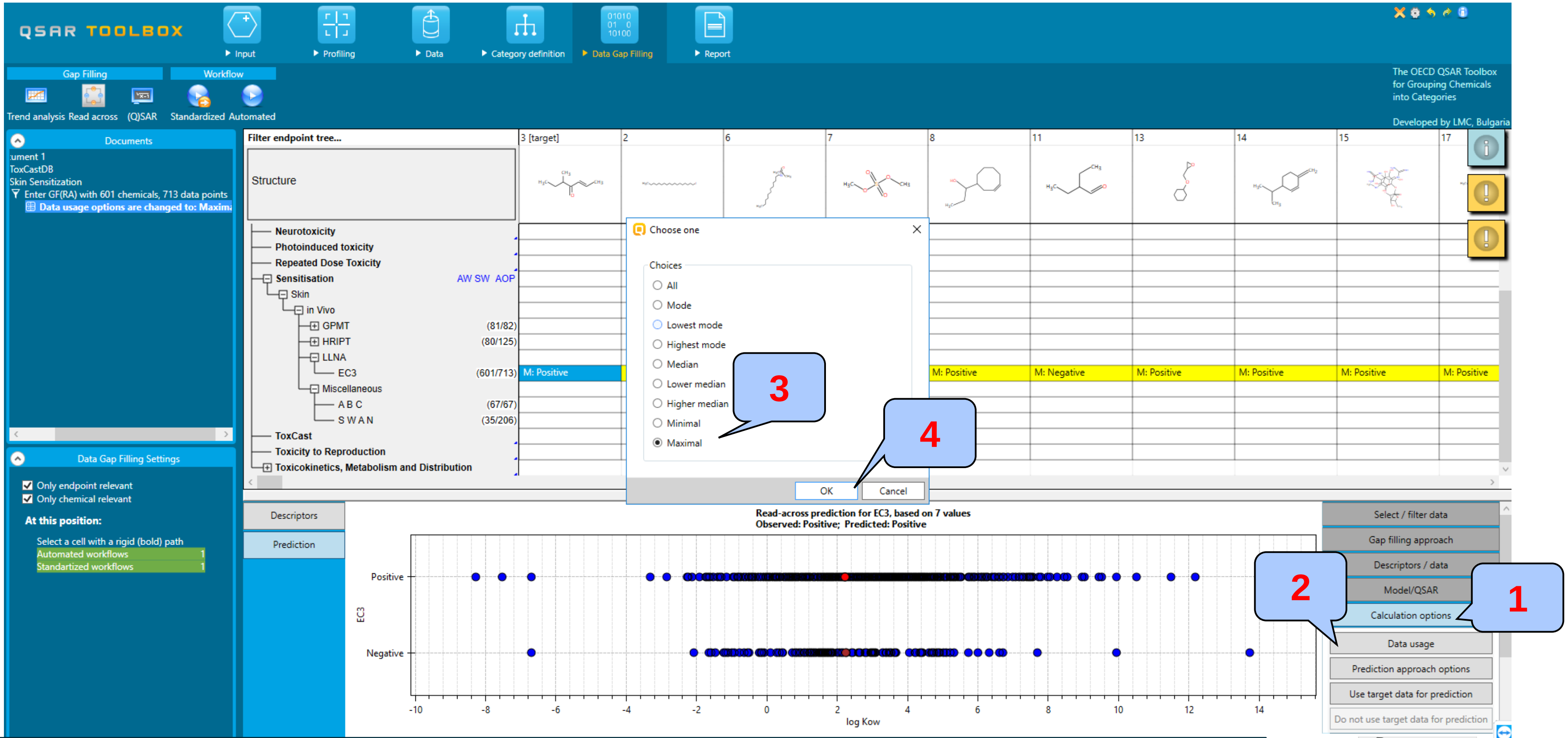
The message informs the user for chemicals excluded from gap filling. 1. **Click** “OK”;

Types endpoint correlations

Categorical vs. categorical

Perform correlation between GPMT and AMES data – step 5

Example 2: Correlation of LLNA and AMES data



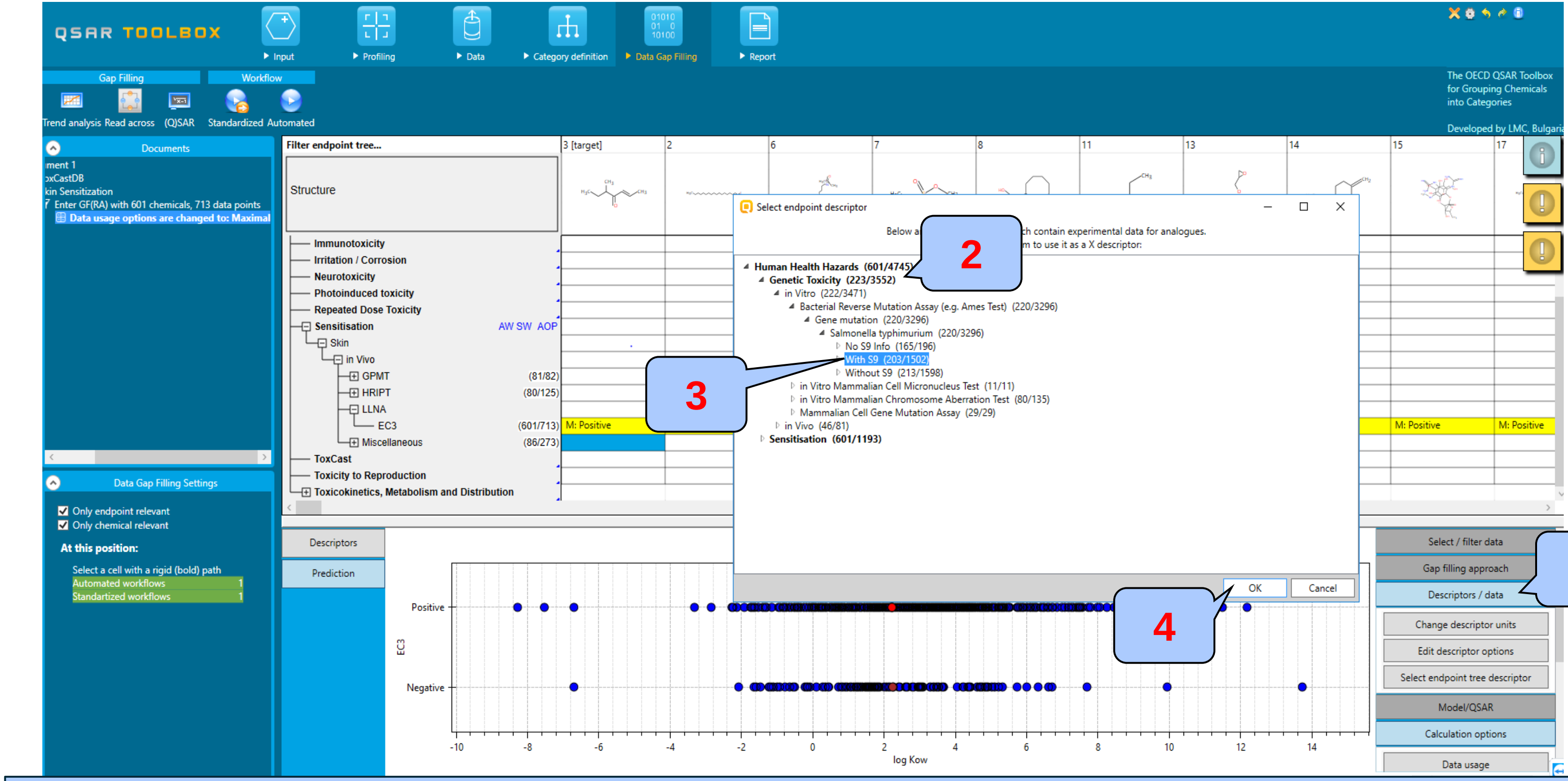
1. **Open** "Calculation" options. 2. **Click** on "Data usage" 3. **Select** "Maximal"   4. **Click** "OK" (refer to slide 62 for details)

Types endpoint correlations

Categorical vs. categorical

Perform correlation between GPMT and AMES data – step 5

Example 2: Correlation of LLNA and AMES data



ment 1
xCastDB
kin Sensitization
7 Enter GF(RA) with 601 chemicals, 713 data points
Data usage options are changed to: Maximal

☑ Only endpoint relevant
☑ Only chemical relevant

At this position:

Select a cell with a rigid (bold) path
Automated workflows 1
Standartized workflows 1

Immunotoxicity
Irritation / Corrosion
Neurotoxicity
Photoinduced toxicity
Repeated Dose Toxicity
Sensitisation AW SW AOP
Skin
in Vivo
GPMT (81/82)
HRIPT (80/125)
LLNA
EC3 (601/713)
Miscellaneous (86/273)
ToxCast
Toxicity to Reproduction
Toxicokinetics, Metabolism and Distribution

Select endpoint descriptor

Below a ch contain experimental data for analogues.
m to use it as a X descriptor:

Human Health Hazards (601/4745)
Genetic Toxicity (223/3552)
in Vitro (222/3471)
Bacterial Reverse Mutation Assay (e.g. Ames Test) (220/3296)
Gene mutation (220/3296)
Salmonella typhimurium (220/3296)
No S9 Info (165/196)
With S9 (203/1502)
Without S9 (213/1598)
in Vitro Mammalian Cell Micronucleus Test (11/11)
in Vitro Mammalian Chromosome Aberration Test (80/135)
Mammalian Cell Gene Mutation Assay (29/29)
in Vivo (46/81)
Sensitisation (601/1193)

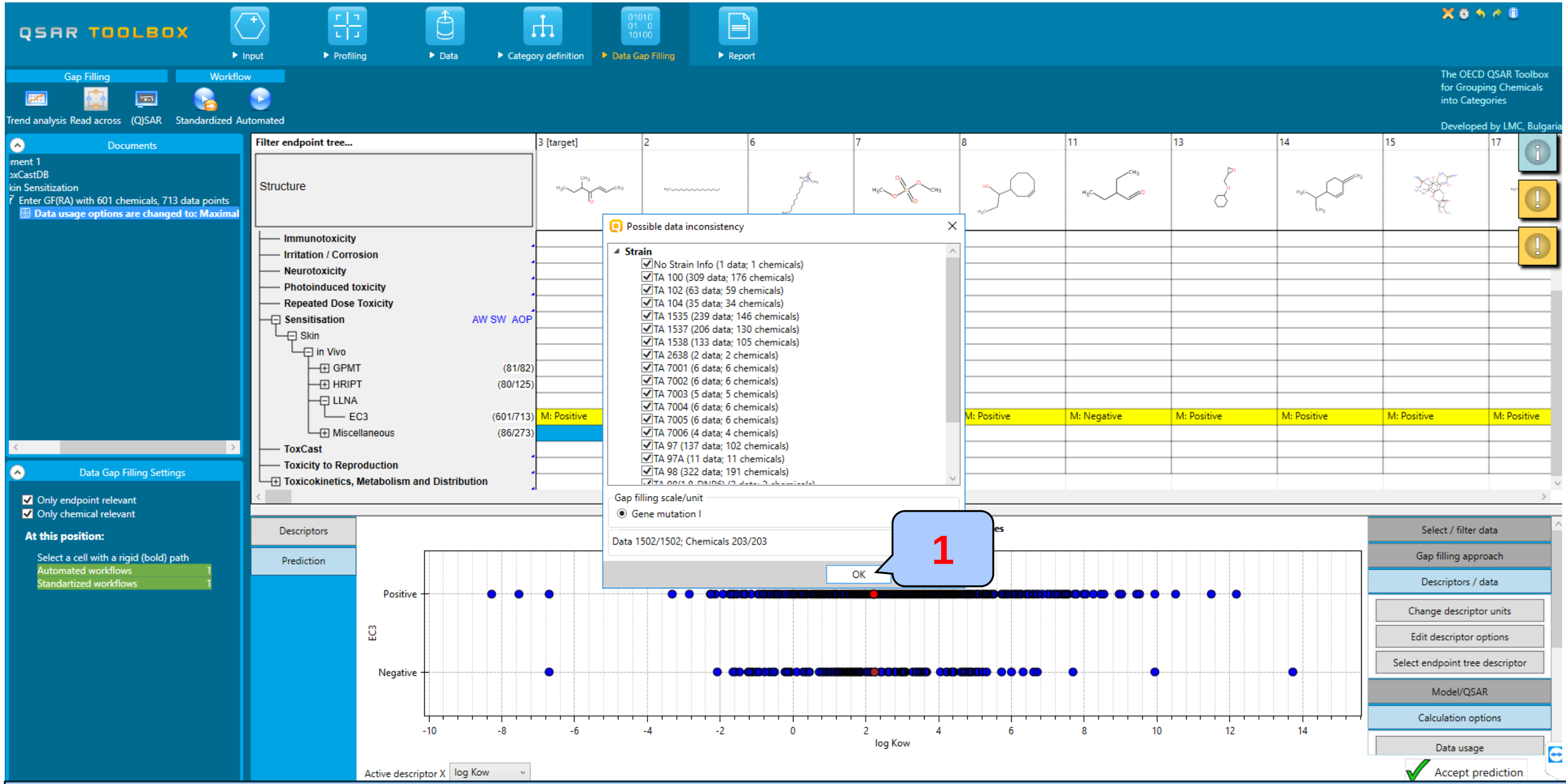
1. **Open** Select/descriptors data/ Select endpoint tree descriptor; 2. **Open** nodes under “Genetic Toxicity” node; 3. **Select** “With S9” under In Vitro|Bacterial Reverse Mutation Assay (e.g. Ames Test)|Gene Mutation| Salmonella typhimurium; ; 4. **Click** “OK” button;

Types endpoint correlations

Categorical vs. categorical

Perform correlation between GPMT and AMES data – step 5

Example 2: Correlation of LLNA and AMES data



Possible data inconsistency

- ▲ **Strain**
  - ☑ No Strain Info (1 data; 1 chemicals)
  - ☑ TA 100 (309 data; 176 chemicals)
  - ☑ TA 102 (63 data; 59 chemicals)
  - ☑ TA 104 (35 data; 34 chemicals)
  - ☑ TA 1535 (239 data; 146 chemicals)
  - ☑ TA 1537 (206 data; 130 chemicals)
  - ☑ TA 1538 (133 data; 105 chemicals)
  - ☑ TA 2638 (2 data; 2 chemicals)
  - ☑ TA 7001 (6 data; 6 chemicals)
  - ☑ TA 7002 (6 data; 6 chemicals)
  - ☑ TA 7003 (5 data; 5 chemicals)
  - ☑ TA 7004 (6 data; 6 chemicals)
  - ☑ TA 7005 (6 data; 6 chemicals)
  - ☑ TA 7006 (4 data; 4 chemicals)
  - ☑ TA 97 (137 data; 102 chemicals)
  - ☑ TA 97A (11 data; 11 chemicals)
  - ☑ TA 98 (322 data; 191 chemicals)
  - ☑ TA 98(1,8-DNP6) (2 data; 2 chemicals)

Gap filling scale/unit

● Gene mutation I

Data 1502/1502; Chemicals 203/203

OK

1

Documents

ment 1
oxCastDB
kin Sensitization
7 Enter GF(RA) with 601 chemicals, 713 data points
Data usage options are changed to: Maximal

Data Gap Filling Settings

☑ Only endpoint relevant
☑ Only chemical relevant

At this position:

Select a cell with a rigid (bold) path
Automated workflows 1
Standartized workflows 1

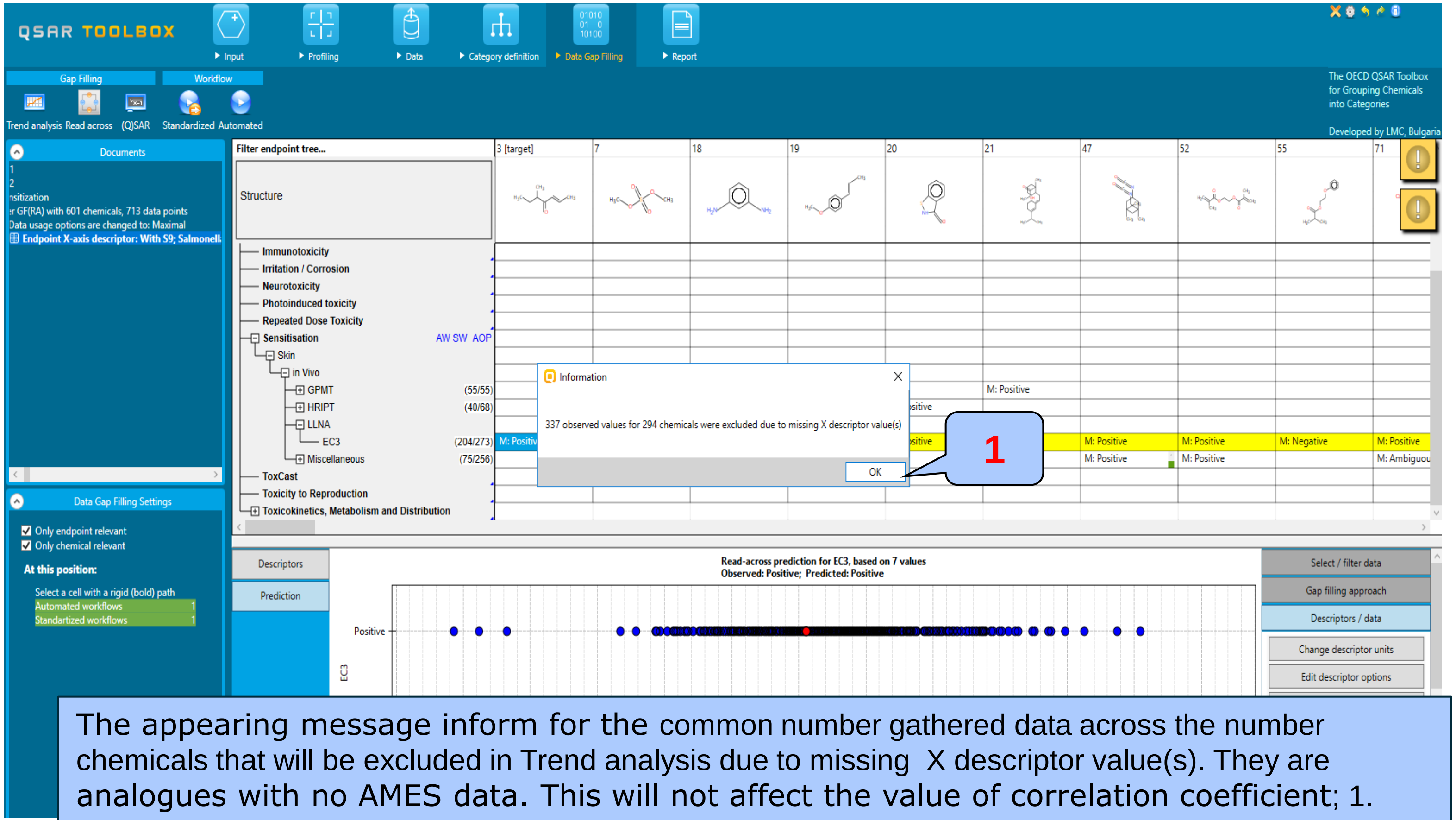
Possible data inconsistency window is appearing. **Click** OK.

Types endpoint correlations

Categorical vs. categorical

Perform correlation between GPMT and AMES data – step 5

Example 2: Correlation of LLNA and AMES data



Information

337 observed values for 294 chemicals were excluded due to missing X descriptor value(s)

OK

Read-across prediction for EC3, based on 7 values
Observed: Positive; Predicted: Positive

Data usage options are changed to: Maximal

Endpoint X-axis descriptor: With S9; Salmonell

☑ Only endpoint relevant
☑ Only chemical relevant

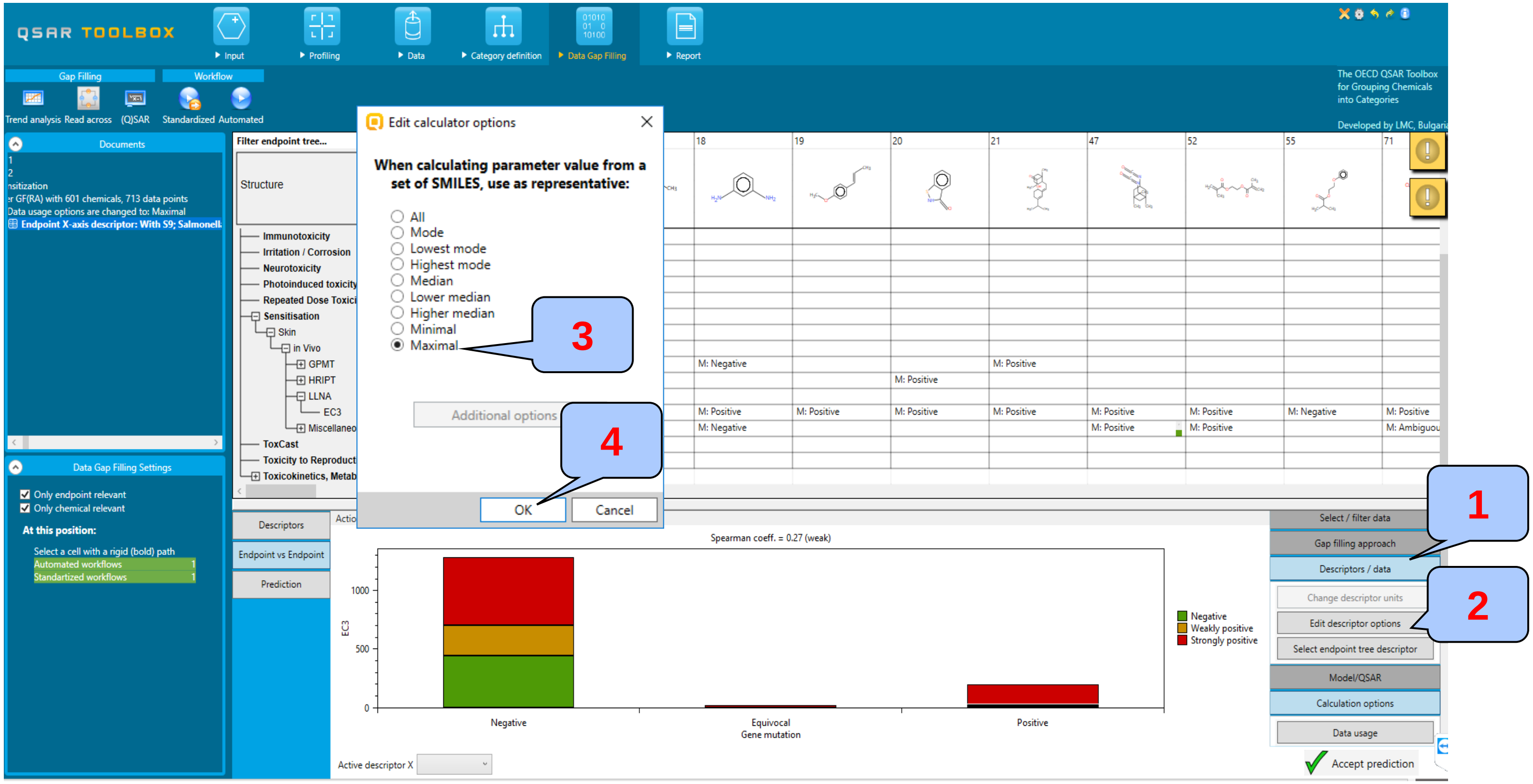
The appearing message inform for the common number gathered data across the number chemicals that will be excluded in Trend analysis due to missing X descriptor value(s). They are analogues with no AMES data. This will not affect the value of correlation coefficient; 1. **Click** “OK”;

Types endpoint correlations

Categorical vs. categorical

Perform correlation between GPMT and AMES data – step 5

Example 2: Correlation of GPMT and AMES data



Edit calculator options

When calculating parameter value from a set of SMILES, use as representative:

- ○ All
- ○ Mode
- ○ Lowest mode
- ○ Highest mode
- ○ Median
- ○ Lower median
- ○ Higher median
- ○ Minimal
- ● Maximal

Additional options

OK Cancel

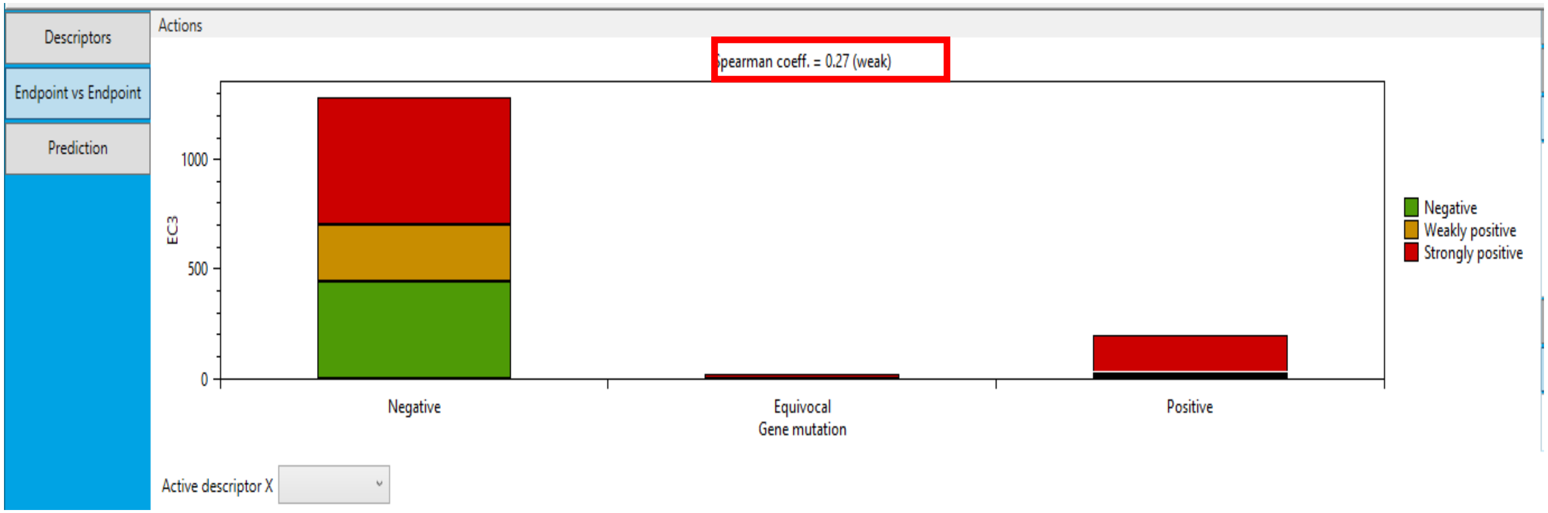
Spearman coeff. = 0.27 (weak)

1.**Select** "Descriptor/ data"; 2. **Click** "Edit descriptor options" ; 3.**Select** "Maximal" (worst case); 4. **Click** OK

Types endpoint correlations

Categorical vs. categorical

Interpretation of correlation results (GPMT vs. AMES)



Correlation analysis between two categorical type data: GPMT and AMES shows weak correlation between two endpoints (Spearman coefficient is 0.3).

Outlook

- Background
- Objectives
- The exercise
- **Workflow**
  - Correlation of data – background
  - **Types endpoint correlations**
    - Categorical vs. categorical
    - Categorized continuous vs. categorical

Types endpoint correlations
Categorized continuous vs. categorical

- The aim of this type correlation is to illustrate how categorized continuous and categorical type data correlates each other.

- Categorized continuous data is the continuous type data (e.g LC50 or AC50, EC3, %) converted into categories.

- In this example we will illustrated how DPRA ratio data (%) correlates with LLNA data:
  - DPRA (ratio data expressed in % and converted in categories)
  - LLNA (categorical type: Strongly positive, Weakly positive, Negative)

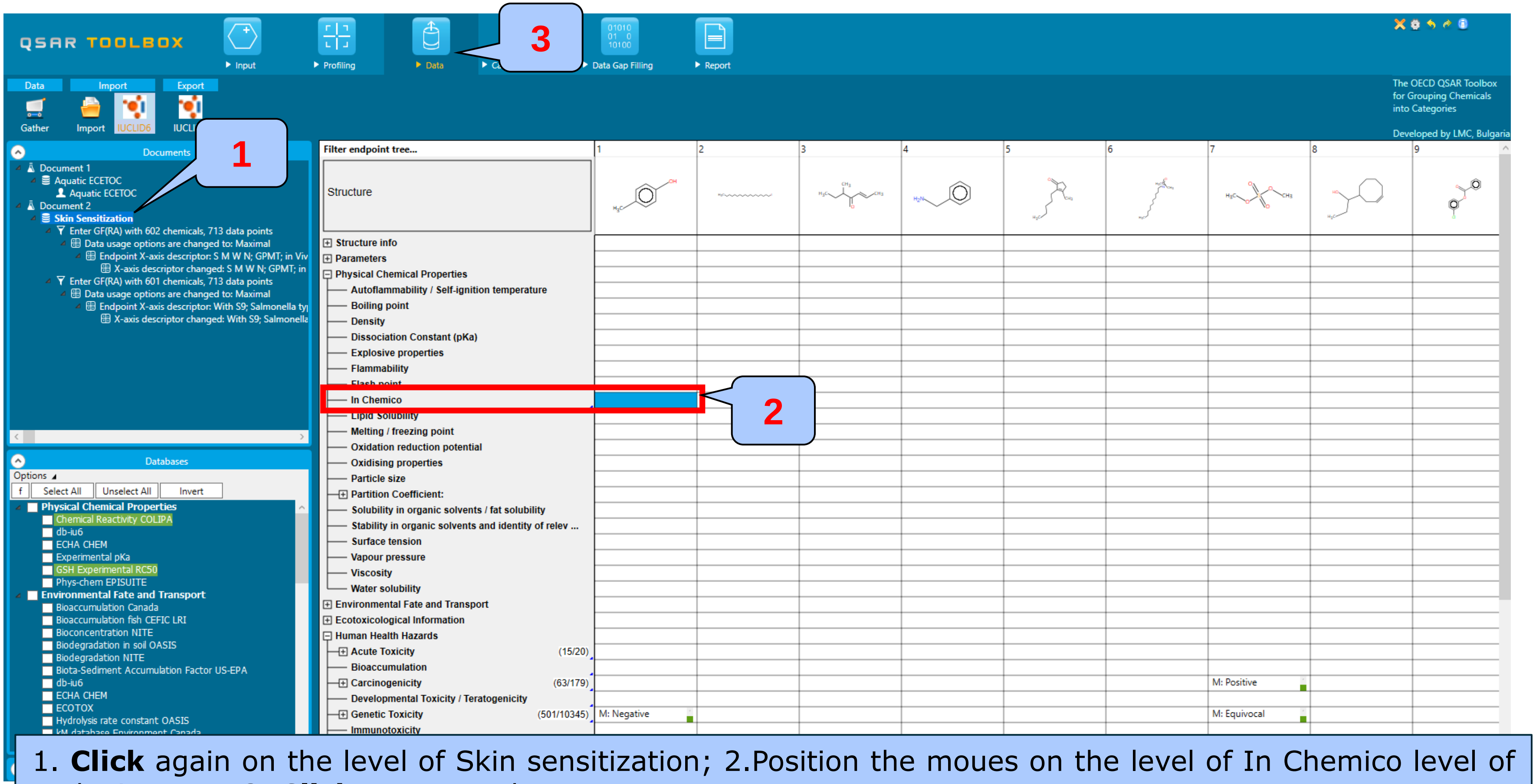
- Step by step workflow is presented on next few slides. Summary of the workflow steps are provided below:
  - *Load Skin sensitization database (step 1) – skipped, because this database has been already loaded on data matrix*
  - *Gather experimental data (step 2)*
  - *Define target endpoint (step 3)*
  - *Enter Gap filling (step 4)*
  - *Perform correlation between endpoints (step 5).*

Types endpoint correlations

Categorized continuous vs. categorical

Gather experimental data – step 2

Example: Correlation of DPRA (%) and LLNA (Strongly positive, Weakly positive, Negative) data



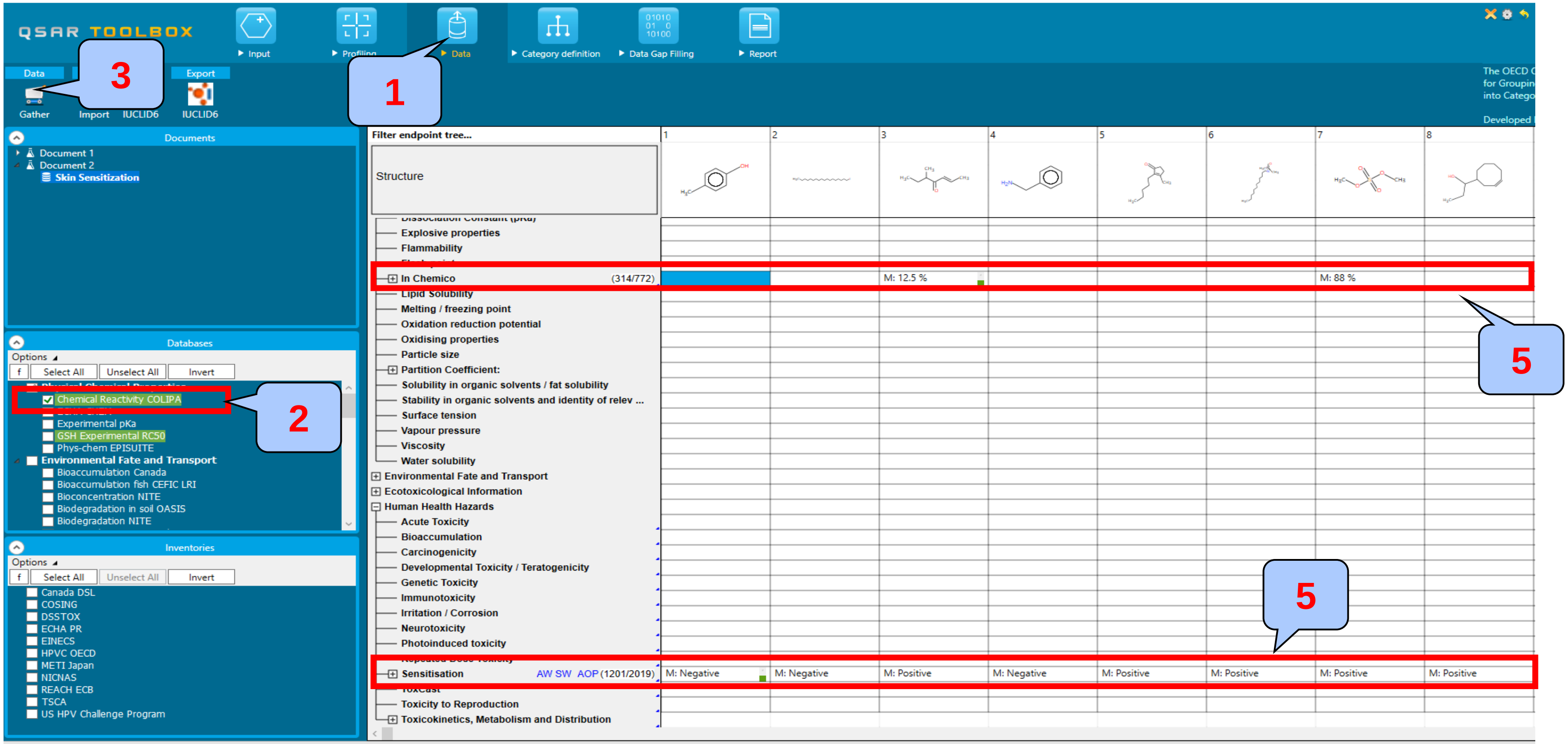
1. **Click** again on the level of Skin sensitization; 2.Position the moues on the level of In Chemico level of endpoint tree; 3. **Click** on Data tab

Types endpoint correlations

Categorized continuous vs. categorical

Gather experimental data – step 2

Example: Correlation of DPRA (%) and LLNA (Strongly positive, Weakly positive, Negative) data



Documents
- Document 1
- Document 2
  - Skin Sensitization

Databases

Options

f | Select All | Unselect All | Invert

- Physical Chemical Properties
  - ☑ Chemical Reactivity COLIPA
  - □ ECHA CHEM
  - □ Experimental pKa
  - □ GSH Experimental RC50
  - □ Phys-chem EPISUITE
- □ Environmental Fate and Transport
  - □ Bioaccumulation Canada
  - □ Bioaccumulation fish CEFIC LRI
  - □ Bioconcentration NITE
  - □ Biodegradation in soil OASIS
  - □ Biodegradation NITE

Inventories

Options

f | Select All | Unselect All | Invert

- □ Canada DSL
- □ COSING
- □ DSSTOX
- □ ECHA PR
- □ EINECS
- □ HPVC OECD
- □ METI Japan
- □ NICNAS
- □ REACH ECB
- □ TSCA
- □ US HPV Challenge Program

Filter endpoint tree...	1	2	3	4	5	6	7	8
Structure								
Dissociation Constant (pKa)								
Explosive properties								
Flammability								
Flash point								
⊞ In Chemico (314/772)			M: 12.5 %				M: 88 %	
Lipid Solubility								
Melting / freezing point								
Oxidation reduction potential								
Oxidising properties								
Particle size								
⊞ Partition Coefficient:								
Solubility in organic solvents / fat solubility								
Stability in organic solvents and identity of relev ...								
Surface tension								
Vapour pressure								
Viscosity								
Water solubility								
⊞ Environmental Fate and Transport								
⊞ Ecotoxicological Information								
⊟ Human Health Hazards								
Acute Toxicity								
Bioaccumulation								
Carcinogenicity								
Developmental Toxicity / Teratogenicity								
Genetic Toxicity								
Immunotoxicity								
Irritation / Corrosion								
Neurotoxicity								
Photoinduced toxicity								
Repeated Dose Toxicity								
⊞ Sensitisation AW SW AOP (1201/2019)	M: Negative	M: Negative	M: Positive	M: Negative	M: Positive	M: Positive	M: Positive	M: Positive
ToxCast								
Toxicity to Reproduction								
⊞ Toxicokinetics, Metabolism and Distribution								

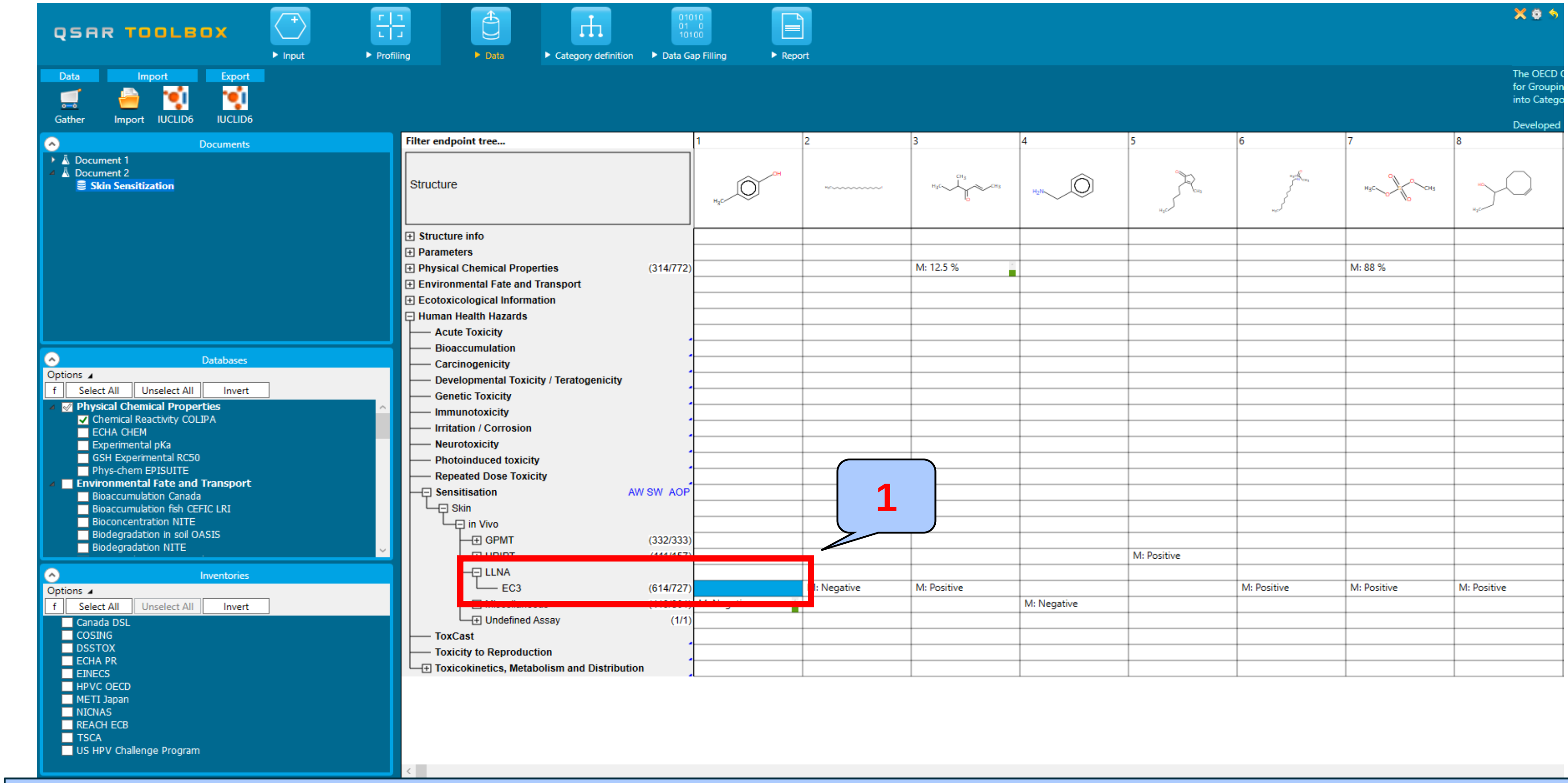
1. **Go** to "Data"; 2. **Select** "Chemical reactivity COLIPA" database. Skin sensitization DB is already selected; 3. **Click** "Gather" button; 5. The data appeared on datamatrix;

Types endpoint correlations

Categorized continuous vs. categorical

Define target endpoint – step 3

Example: Correlation of DPRA (%) and LLNA (Strongly positive, Weakly positive, Negative) data



Document 1
Document 2
Skin Sensitization

Databases:
☑ Physical Chemical Properties
☑ Chemical Reactivity COLIPA
□ ECHA CHEM
□ Experimental pKa
□ GSH Experimental RC50
□ Phys-chem EPISUITE
□ Environmental Fate and Transport
□ Bioaccumulation Canada
□ Bioaccumulation fish CEFIC LRI
□ Bioconcentration NITE
□ Biodegradation in soil OASIS
□ Biodegradation NITE

Inventories:
□ Canada DSL
□ COSING
□ DSSTOX
□ ECHA PR
□ EINECS
□ HPVC OECD
□ METI Japan
□ NICNAS
□ REACH ECB
□ TSCA
□ US HPV Challenge Program

Endpoint tree:
- Structure info
- Parameters
- Physical Chemical Properties (314/772) — 3: M: 12.5 %; 7: M: 88 %
- Environmental Fate and Transport
- Ecotoxicological Information
- Human Health Hazards
  - Acute Toxicity
  - Bioaccumulation
  - Carcinogenicity
  - Developmental Toxicity / Teratogenicity
  - Genetic Toxicity
  - Immunotoxicity
  - Irritation / Corrosion
  - Neurotoxicity
  - Photoinduced toxicity
  - Repeated Dose Toxicity
  - Sensitisation AW SW AOP
    - Skin
      - in Vivo
        - GPMT (332/333)
        - HRIPT (111/157) — 5: M: Positive
        - LLNA
          - EC3 (614/727) — 2: M: Negative; 3: M: Positive; 6: M: Positive; 7: M: Positive; 8: M: Positive
        - Miscellaneous (113/301) — 1: M: Negative; 4: M: Negative
        - Undefined Assay (1/1)
  - ToxCast
  - Toxicity to Reproduction
  - Toxicokinetics, Metabolism and Distribution

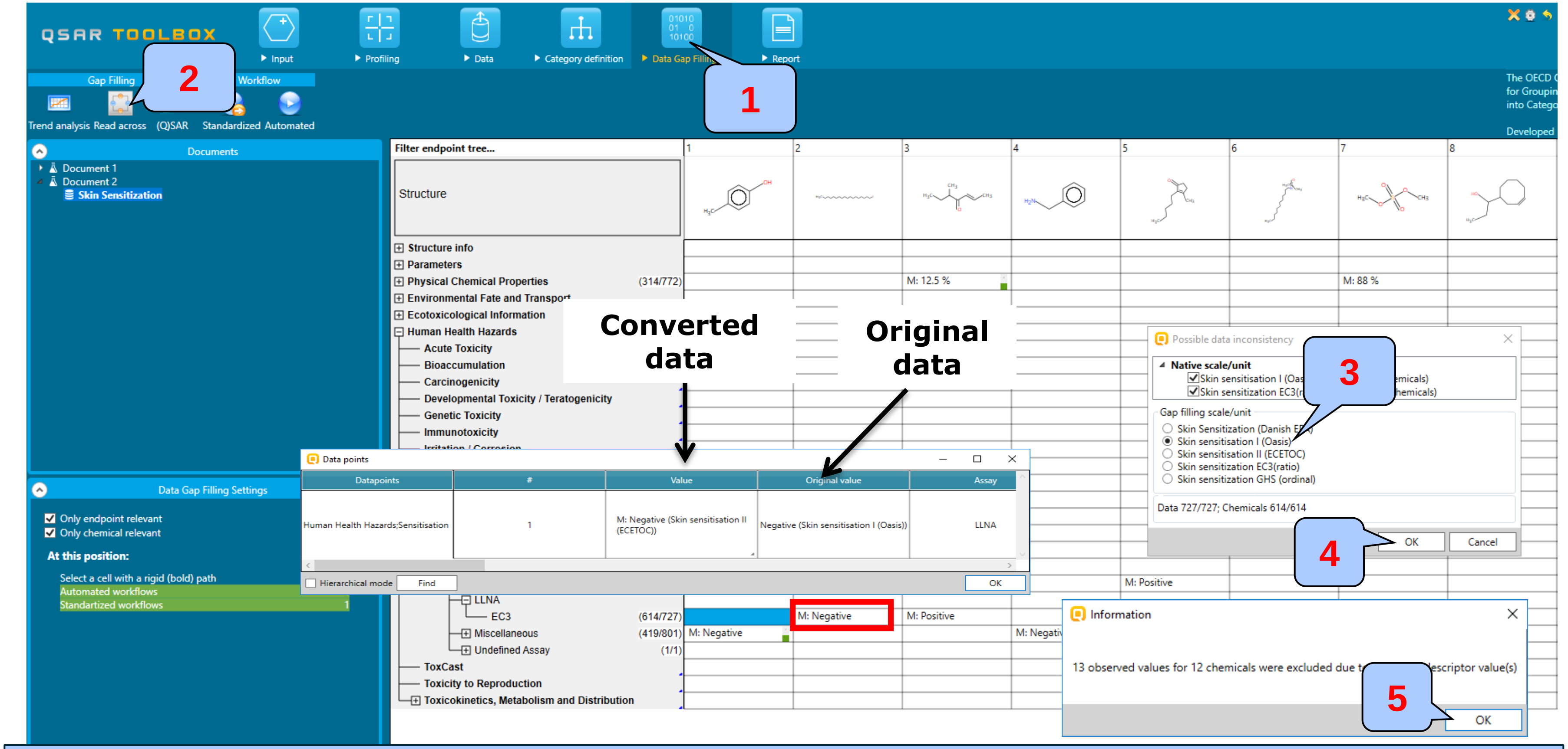
The target endpoint is EC3;
1. **Click** on the cell associated with target endpoint and target chemical;

Types endpoint correlations

Categorized continuous vs. categorical

Enter Gap filling – step 4

Example: Correlation of DPRA (%) and LLNA (Strongly positive, Weakly positive, Negative) data



Data points

Datapoints	#	Value	Original value	Assay
Human Health Hazards;Sensitisation	1	M: Negative (Skin sensitisation II (ECETOC))	Negative (Skin sensitisation I (Oasis))	LLNA

□ Hierarchical mode   Find   OK

Possible data inconsistency

Native scale/unit
☑ Skin sensitisation I (Oas ... emicals)
☑ Skin sensitization EC3(r ... hemicals)

Gap filling scale/unit
○ Skin Sensitization (Danish EPA)
● Skin sensitisation I (Oasis)
○ Skin sensitisation II (ECETOC)
○ Skin sensitization EC3(ratio)
○ Skin sensitization GHS (ordinal)

Data 727/727; Chemicals 614/614

OK   Cancel

Information

13 observed values for 12 chemicals were excluded due t ... escriptor value(s)

OK

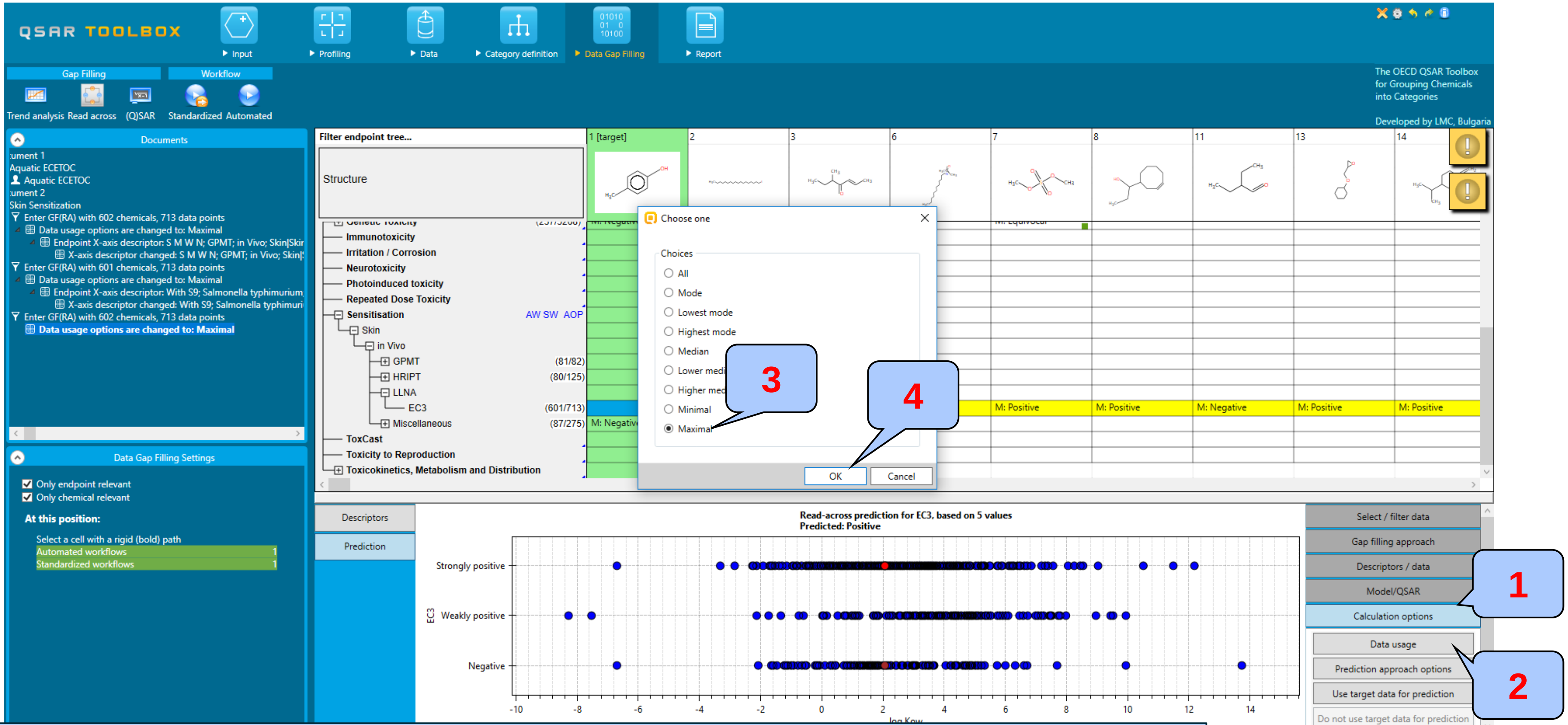
Enter Gap filling and apply read across. Read across is applied because a categorical type data is analyzed.
1. **Go** to “Data Gap filling”; 2. **Select** “Read-across”; 3. **Select** “Skin sensitization I (OASIS)” scale ; 4-5. **Click** “OK”;

Types endpoint correlations

Categorized continuous vs. categorical

Perform correlation between DPRA and LLNA data – step 5

Example: Correlation of DPRA (%) and LLNA (Strongly positive, Weakly positive, Negative) data



Filter endpoint tree...

Structure

Immunotoxicity
Irritation / Corrosion
Neurotoxicity
Photoinduced toxicity
Repeated Dose Toxicity
Sensitisation AW SW AOP
Skin
in Vivo
GPMT (81/82)
HRIPT (80/125)
LLNA
EC3 (601/713)
Miscellaneous (87/275)
ToxCast
Toxicity to Reproduction
Toxicokinetics, Metabolism and Distribution

Documents

Skin Sensitization
Enter GF(RA) with 602 chemicals, 713 data points
Data usage options are changed to: Maximal
Enter GF(RA) with 601 chemicals, 713 data points
Data usage options are changed to: Maximal
Enter GF(RA) with 602 chemicals, 713 data points
Data usage options are changed to: Maximal

Data Gap Filling Settings

☑ Only endpoint relevant
☑ Only chemical relevant

At this position:

Select a cell with a rigid (bold) path
Automated workflows 1
Standardized workflows 1

Choose one

Choices
○ All
○ Mode
○ Lowest mode
○ Highest mode
○ Median
○ Lower median
○ Higher median
○ Minimal
● Maximal

OK Cancel

M: Positive M: Positive M: Negative M: Positive M: Positive

Read-across prediction for EC3, based on 5 values
Predicted: Positive

Select / filter data
Gap filling approach
Descriptors / data
Model/QSAR
Calculation options
Data usage
Prediction approach options
Use target data for prediction
Do not use target data for prediction
Accept prediction

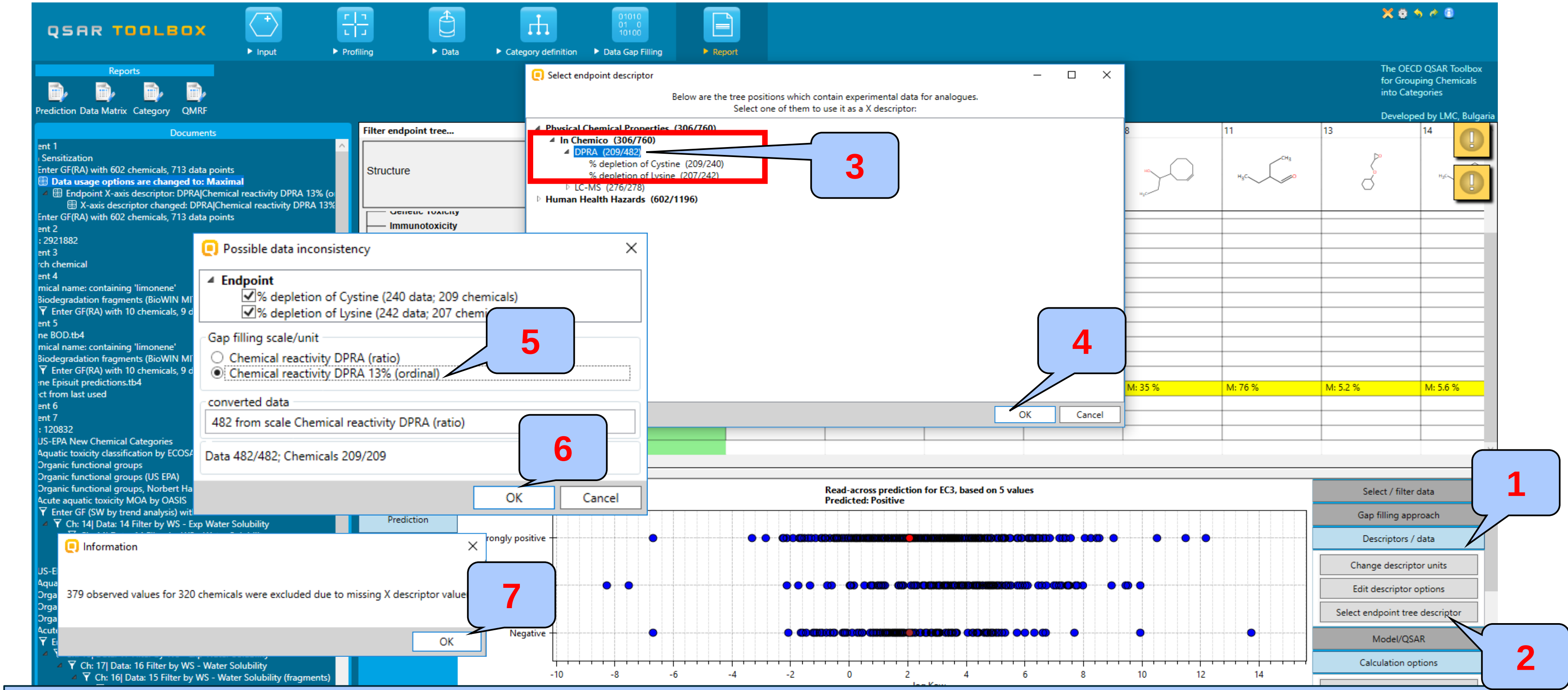
1. **Open** “Calculation options”; 2. **Click** on “Data usage”; 3. **Select** “Maximal”
4. **Click** “OK”

Types endpoint correlations

Categorized continuous vs. categorical

Perform correlation between DPRA and LLNA data – step 5

Example: Correlation of DPRA (%) and LLNA (Strongly positive, Weakly positive, Negative) data



Select endpoint descriptor

Below are the tree positions which contain experimental data for analogues.
Select one of them to use it as a X descriptor:

- Physical Chemical Properties (306/760)
  - In Chemico (306/760)
    - DPRA (209/482)
      - % depletion of Cystine (209/240)
      - % depletion of Lysine (207/242)
    - LC-MS (276/278)
- Human Health Hazards (602/1196)

Possible data inconsistency

Endpoint
☑ % depletion of Cystine (240 data; 209 chemicals)
☑ % depletion of Lysine (242 data; 207 chemi

Gap filling scale/unit
○ Chemical reactivity DPRA (ratio)
● Chemical reactivity DPRA 13% (ordinal)

converted data
482 from scale Chemical reactivity DPRA (ratio)

Data 482/482; Chemicals 209/209

Information

379 observed values for 320 chemicals were excluded due to missing X descriptor value

Read-across prediction for EC3, based on 5 values
Predicted: Positive

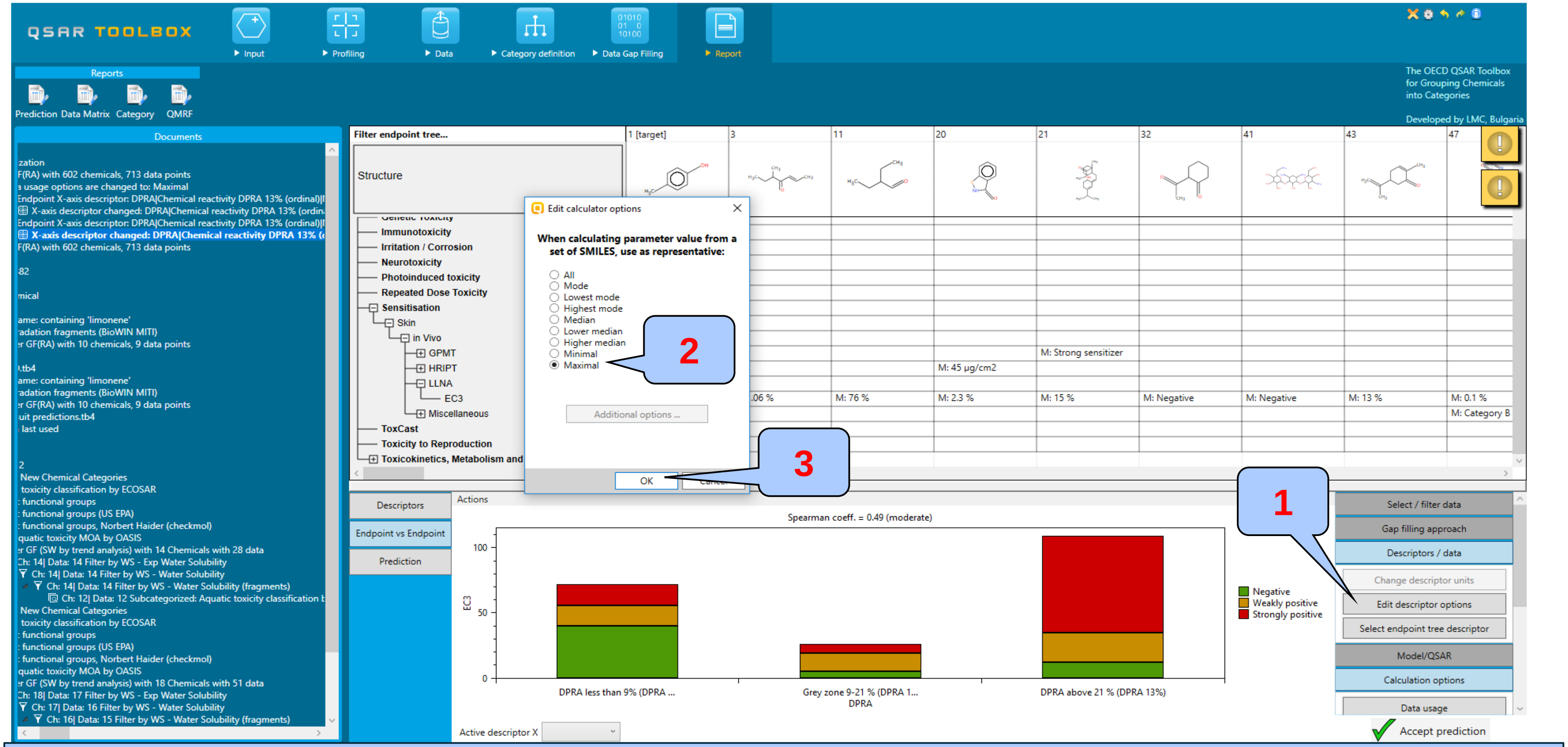
1. **Open Descriptors/data** 2. **Click** on "Select endpoint tree descriptor" button 3. **Click** on the endpoint tree on the level of "DPRA". In this case we mixed DPRA lysine and Cysteine data 4. **Click** on OK 5. **Select** Chemical reactivity DPRA 13% (ordinal) scale 6. **Click** OK 7. **Click** OK on the appeared message

Types endpoint correlations

Categorized continuous vs. categorical

Perform correlation between DPRA and LLNA data – step 5

Example: Correlation of DPRA (%) and LLNA (Strongly positive, Weakly positive, Negative) data



Edit calculator options

When calculating parameter value from a set of SMILES, use as representative:

○ All
○ Mode
○ Lowest mode
○ Highest mode
○ Median
○ Lower median
○ Higher median
○ Minimal
● Maximal

Spearman coeff. = 0.49 (moderate)

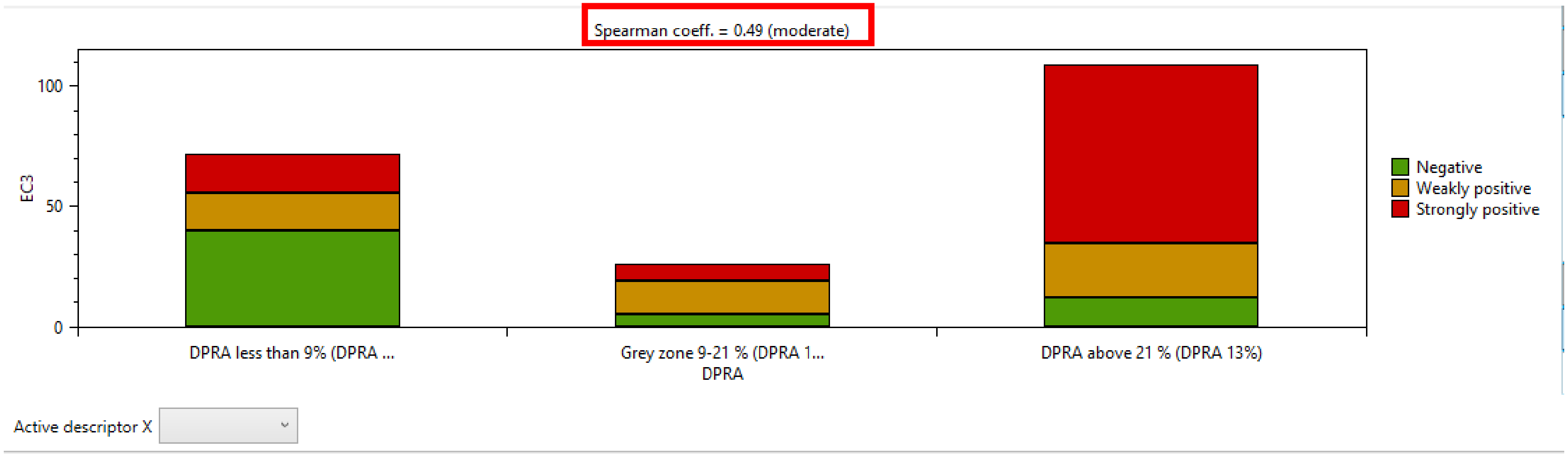
zation
F(RA) with 602 chemicals, 713 data points
usage options are changed to: Maximal
Endpoint X-axis descriptor: DPRA|Chemical reactivity DPRA 13% (ordinal)
X-axis descriptor changed: DPRA|Chemical reactivity DPRA 13% (ordin
Endpoint X-axis descriptor: DPRA|Chemical reactivity DPRA 13% (ordinal)
X-axis descriptor changed: DPRA|Chemical reactivity DPRA 13% (
F(RA) with 602 chemicals, 713 data points
82
mical
ame: containing 'limonene'
adation fragments (BioWIN MITI)
r GF(RA) with 10 chemicals, 9 data points
.tb4
ame: containing 'limonene'
adation fragments (BioWIN MITI)
r GF(RA) with 10 chemicals, 9 data points
uit predictions.tb4
last used
2
New Chemical Categories
toxicity classification by ECOSAR
functional groups
functional groups (US EPA)
functional groups, Norbert Haider (checkmol)
quatic toxicity MOA by OASIS
r GF (SW by trend analysis) with 14 Chemicals with 28 data
Ch: 14| Data: 14 Filter by WS - Exp Water Solubility
Ch: 14| Data: 14 Filter by WS - Water Solubility
Ch: 14| Data: 14 Filter by WS - Water Solubility (fragments)
Ch: 12| Data: 12 Subcategorized: Aquatic toxicity classification b
New Chemical Categories
toxicity classification by ECOSAR
functional groups
functional groups (US EPA)
functional groups, Norbert Haider (checkmol)
quatic toxicity MOA by OASIS
r GF (SW by trend analysis) with 18 Chemicals with 51 data
Ch: 18| Data: 17 Filter by WS - Exp Water Solubility
Ch: 17| Data: 16 Filter by WS - Water Solubility
Ch: 16| Data: 15 Filter by WS - Water Solubility (fragments)

1. **Open** Edit descriptor options 2. **Select** maximal values (worst case) 3. **Click** OK

Types endpoint correlations

Categorized continuous vs. categorical

Interpretation of correlation results (DPRA vs. LLNA)


Spearman coeff. = 0.49 (moderate)

100
50
0
EC3
DPRA less than 9% (DPRA ...
Grey zone 9-21 % (DPRA 1...
DPRA above 21 % (DPRA 13%)
DPRA
🟩 Negative
🟧 Weakly positive
🟥 Strongly positive

Active descriptor X

- In this example we have correlate continues DPRA (%) data distributed into 3 bins (shown below) and categorical LLNA data (Strongly positive, Weakly positive, Negative)
  - Less than 9%
  - Grey zone 9 – 21%
  - Above 21%

- The high value of Spearman coefficient (0.49) shows moderate correlation between DPRA and LLNA data

Summary

- Different type correlations have been illustrated in this tutorial based on type of endpoint data:
  - Categorical vs. categorical:
  - Categorized continuous vs. categorical
- Correlation analysis has been evaluated by Spearman coefficient
- Moderate endpoint correlations have been obtained for 2 out of 3 illustrated examples.