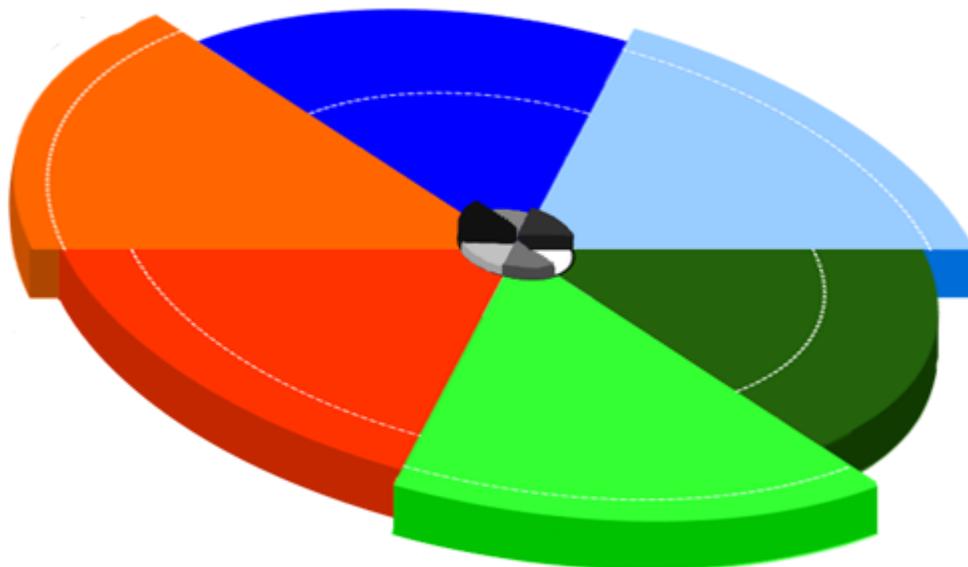


The University of North Carolina at Chapel Hill

Gillings School of Global Public Health

ToxPi standalone GUI



Version 1.1

TOXICOLOGICAL PRIORITY INDEX GUI

User Manual

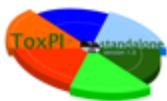
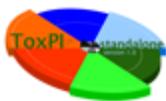


Table of Contents

1. About ToxPi GUI.....	3
2. ToxPi GUI Installation and Requirements.....	4
3. Home screen.....	5
4. Data format.....	7
5. Slices Information Window.....	8
6. Results Window.....	11
6.1 Chemical window.....	14
6.2 Chart Window.....	15
6.2.1 Confidence intervals.....	16
6.2.2 Zooming and Dragging.....	16
6.2.3 Highlighting the Data.....	16
7. Description of ToxPi Bootstrapping Algorithm.....	17





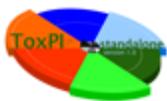
1. About ToxPi GUI

ToxPi GUI (Toxicological Priority Index graphical user interface) is a flexible prioritization support software tool based on the “Endocrine Profiling and Prioritization of Environmental Chemicals Using ToxCast Data” article by David M. Reif et al. (<http://ehsehplp03.niehs.nih.gov/article/info%3Adoi%2F10.1289%2Fehp.1002180>). This tool incorporates a chemical’s bioactivity profiles, inferred toxicity pathways, dose estimates, exposure data, chemical structural descriptors, etc. into easy-to-understand output. The ToxPi GUI was designed to help quickly import, analyze, visualize, highlight and export diverse data.

The output can be used to:

- Calculate the comprehensive toxicity potential of a chemical.
- Provide a visual that represents relative contribution of each data domain
- Enable a new weight-of-evidence framework for diverse prioritization.
- Numerically integrate multiple information domains.
- Incorporate existing knowledge, prioritization schemes and different types of data.
- Provide multivariate of toxicity relative to any set of chemicals.
- Customize components for diverse prioritization tasks.
- Estimate uncertainty in various indices calculated from the underlying data.
- Provide complete transparency in deriving a toxicity score.

ToxPi GUI is written in Java in order to provide cross-platform compatibility and can be used on any computer where Java is installed. The application uses JFreeChart package for ToxPi chart representation, Apache commons library - for CSV files operations and Swing libraries - for design purposes.



ToxPi GUI allows you to extract data from one or multiple files at a time. Once you add a file, you will be able to quickly see a visual representation of the data as well as the prioritization results, allowing you to preview your ToxPi “pies”. When you integrate data from multiple files, you will be able to see the data (or specific “components”) in common among the files for all chemicals or for particular chemicals that you select. You will also see basic statistical analysis for the selected components and will have the option to create custom data types for selected components from your input file.

The results window will show the table of chemical toxicity priority indexes based on the components you selected. You can switch between an extended information table and a simple image table. The chart visualization option allow users to focus on a particular chemical or sets of chemicals, zoom into a specific area of the ToxPi chart, highlight any specific chemicals, and browse from one chemical to another by simply dragging the plot.

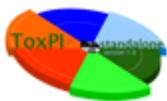
2. ToxPi GUI Installation and Requirements

In order to use ToxPi GUI you need to have Java installed on your machine. You can download Java for free from this website <http://www.java.com/en/> . Starting from version 1.1 users can choose between Windows executable file or Java JAR file. We recommend to use Windows EXE file on Windows OS and JAR on any other OS.

For some big datasets user can face some memory errors. If you see that errors you’ll need to use Java JAR file but set the memory size for the application using the following string in command line: `java -Xmx1500m -jar path_to_jar_file` where *path_to_jar_file* is a system path to the location here you saved **ToxPi v1.1.jar** file.

After you have successfully installed Java you can double click on the **ToxPi.jar** file in order to launch the application. The most recent version of ToxPi GUI is located on the Comptox website <http://comptox.unc.edu/resources.html>.

If you have troubles running the application on Windows OS please verify that your JAVA_HOME system variable is pointing to the installation directory. To do so, click “**Start**” and then right-mouse-click on the “**My Computer**” or “**Computer**” item. Select “**Properties**” in the appeared list. On the next window locate the “**Advanced System Settings**” button and click it. Next locate the “**System Variables**” button and click it. Look inside the “**System Variables**” list for **JAVA_HOME** record. If you cannot find it then click “**New**” button and enter **JAVA_HOME** as



a Variable name and path to your java folder as a value. Usually the path will look like `c:\Program Files\Java\jre6\bin`.

3. Home screen

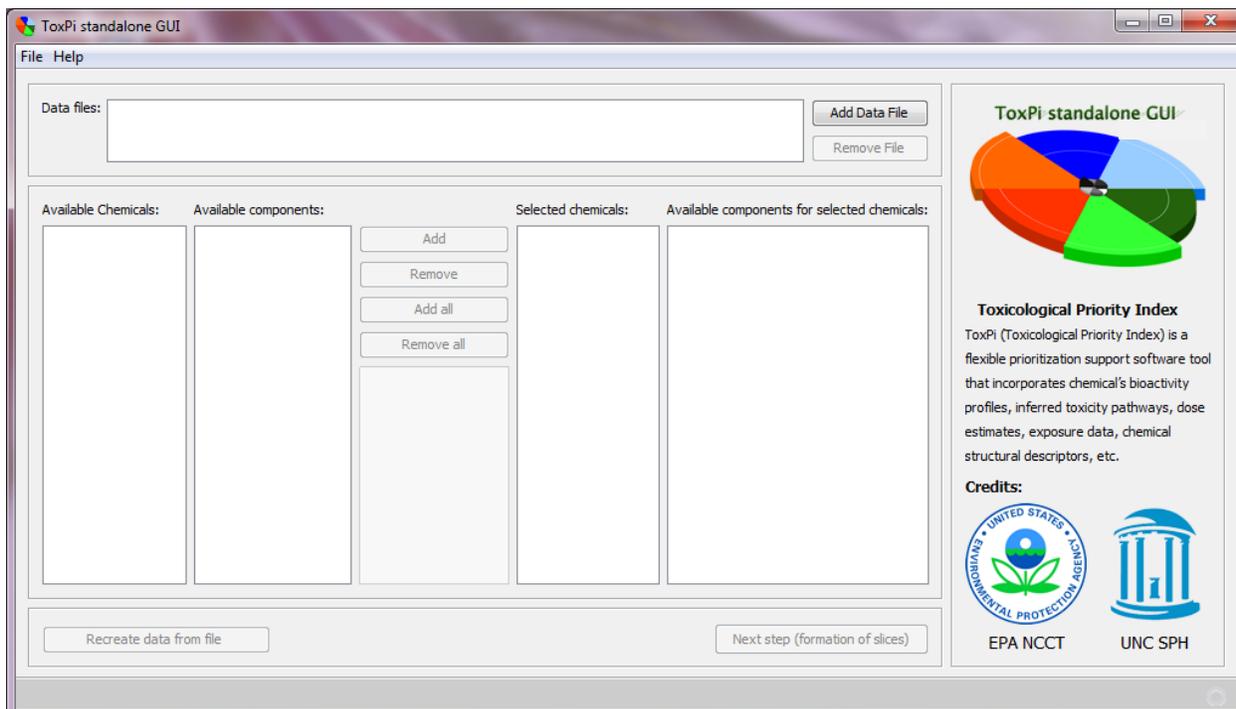
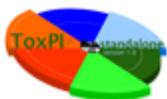


Figure 1. ToxPi GUI home window

Once you have opened the application, you will see the ToxPi GUI main screen, which allows you to select the files and chemicals to work with ([See Figure 1](#)).

To select your input data file, click on the **“Add Data File”** button. The file browsing dialog box should appear. Browse to your data file, select it, and click **“Open”**. (See [Section 4](#) for information on how the input data file should appear.)

For a bigger file, you will see the loading window with the progress bar indicating the amount of data already processed. Please note that all the files will be checked for duplicated



chemicals and components. When there are duplicates, the error window will appear indicating the name of the duplicates so you can easily locate and delete repeated data from your file or resolve the error as-needed.

You can remove a file from the list by selecting it and pressing the **“Remove File”** button.

At this point, you should see the list of the chemicals and list of the components from the file you loaded in the left side of the window under **“Available chemicals”** and **“Available components”**, respectively.

You can add one chemical or multiple chemicals to your analysis by selecting a chemical from the left **“Available Chemicals”** list and pressing the **“Add”** button. If you’d like to add and analyze all available chemicals, click the **“Add all”** button. To remove chemicals from your analysis list, select chemicals from the **“Selected chemicals”** list and click the **“Remove”** button or the **“Remove all”** button to clear the list. (See Figure 2)

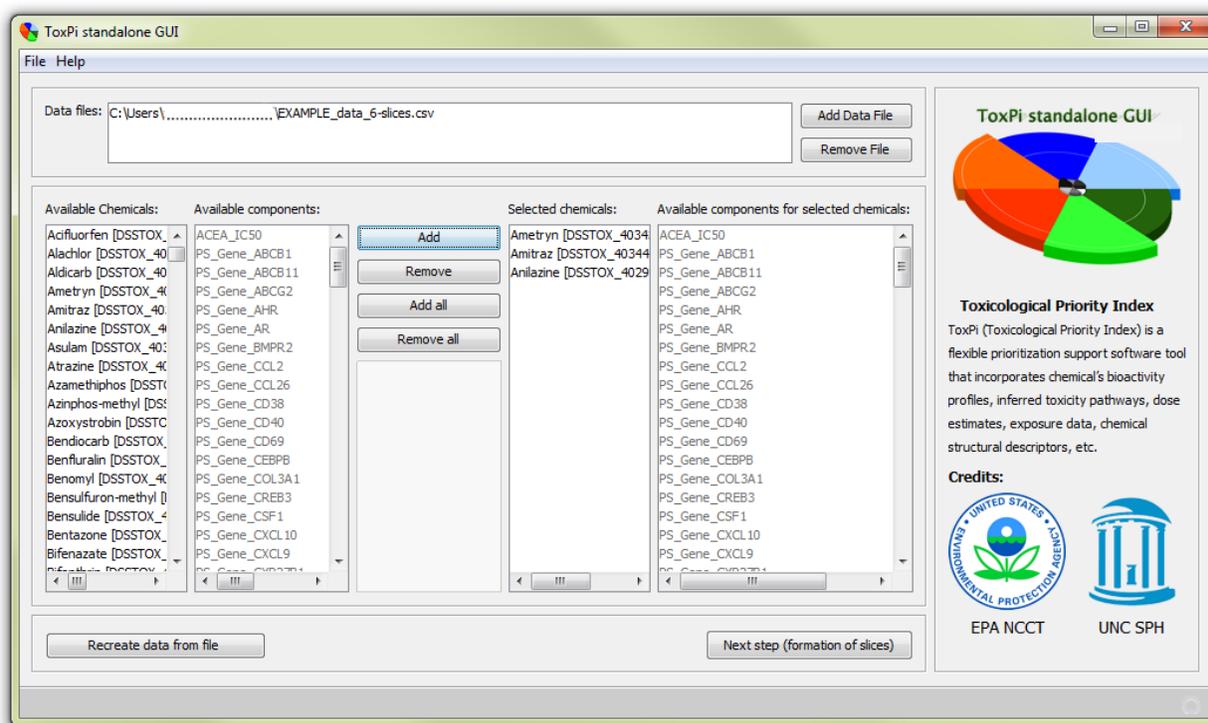
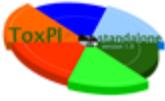


Figure 2. Available chemicals and Selected chemicals



Note: in the case of multiple input files, only common components for all selected files and chemicals will be added to the right **“Available components for selected Chemicals”** list.

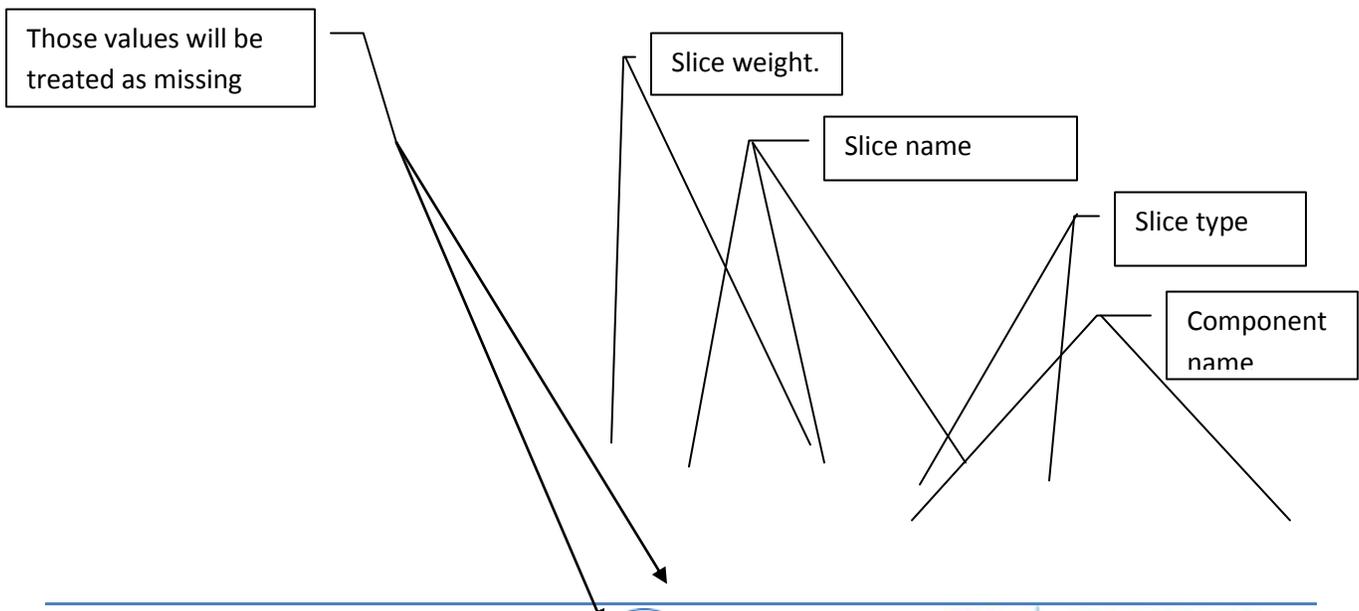
For example: You have added two files, F1 and F2, each containing different chemicals. F1 has data for components C1 and C2 on chemical A and F2 contains data for components C1 and C3 on chemical B. If you select both chemicals A and B, then in the list of “Available components for selected chemicals” you will only see component C1 as it would be the only component for which we have data for both chemicals.

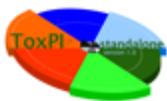
“Recreate data from file” allows you to quickly get to the results of the analysis without selecting any components. The results will be based on the input file and will be an exact visual representation of the input file. This option is available when only one file is added to the list of files to analyze.

“Next step (formation of slices)” proceeds to a window where you will be able to select different combinations of slices, components, types, colors, and calculation methods for your ToxPi analysis.

4. Data format

Data in the input data file should be separated using “,” (comma) delimiter. File extension should be .csv. You can see data format in the example data files provided here: <http://comptox.unc.edu/toxpi.php>





	A	B	C	D	E	F	G	H
1					1	1	1	1
2					Cell electronic sensing	LogPow and MW	LogPow and MW	Steroid Metabolism
3					Assay	ChemProp	ChemProp	Pathway
4					ACEA	QikProp	QikProp	PS_KEGG_0
5	row_order	chemical_source_sid	casrn	chemical_name	ACEA_IC50	logPow_QP	MW_QP	PS_KEGG_Androgen_and_estrogen_metabolism
6	1	DSSTOX_40338	*50594-66-6	Acifluorfen	1000000	3.73	361.661	1000000
7	2	DSSTOX_40339	*15972-60-8	Alachlor	64.30000001	3.154	269.77	184516.032
8	3	DSSTOX_40340	*116-06-3	Aldicarb	NA	1.032	190.26	1000000
9	4	DSSTOX_40343	*834-12-8	Ametryn	1000000	3.401	227.327	16586.61894
10	5	DSSTOX_40344	*33089-61-1	Amitraz	-29.39999997	5.497	293.411	1000000
11	6	DSSTOX_40299	*101-05-3	Anilazine	1000000	3.101	275.524	1000000
12	7	DSSTOX_40374	*3337-71-1	Asulam	1000000	-0.071	230.238	36845.62222
13	8	DSSTOX_40346	*1912-24-9	Atrazine	1000000	2.609	215.685	192525.1417

Figure 3. Input file format

Component
source id

The default value for the slice weight should be 1.

Please note that first 4 (four) columns in the input data file are obligatory as well as first 5 (five) columns. If your file will have different format ToxPi GUI will not accept that file.

Please note that the bootstrapping step will not be executed if each slice contains only one component.

Please note that any values that are below zero or are not numbers will be ignored and treated as a missing data.

5. Slices Information Window

After you click **“Next step (formation of slices)”**, you will see the window where you can control the inputs and information that will be used in the ToxPi analysis.

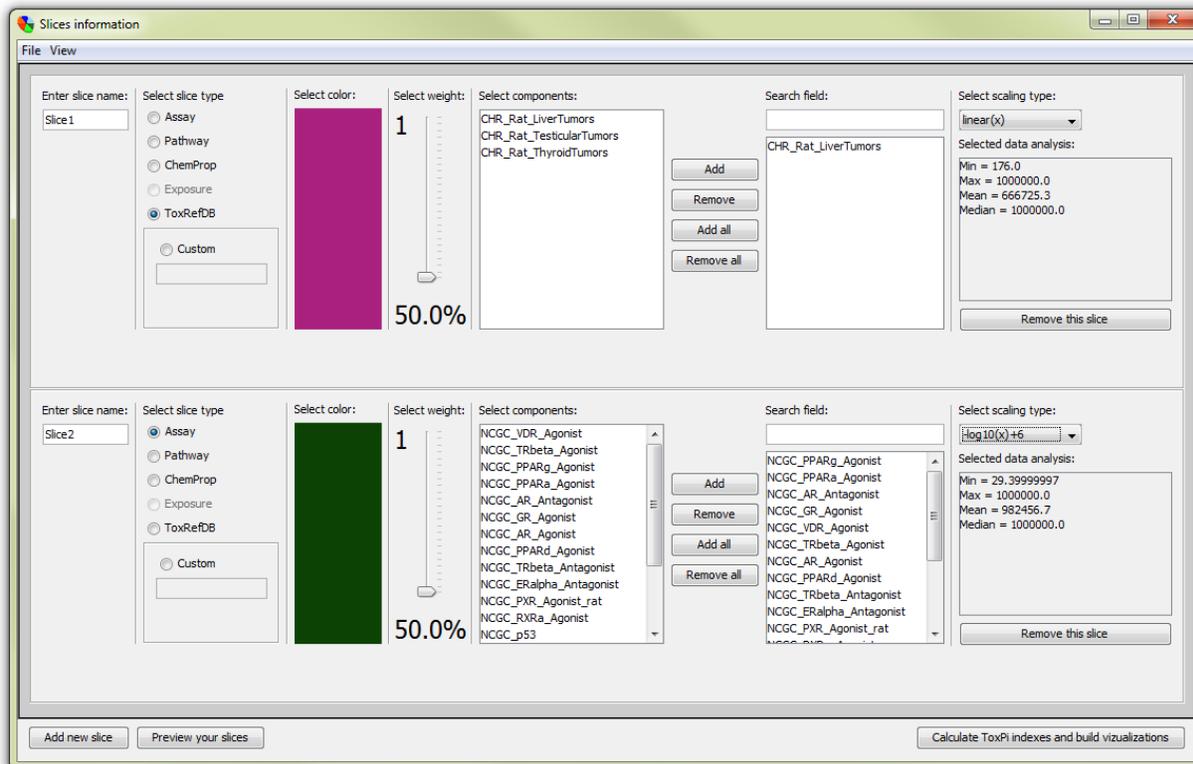
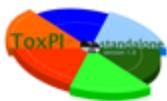
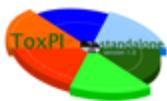


Figure 4. Slices Information Window

In the window shown in Figure 3, you can add a new slice by clicking **“Add new slice”**, preview the slices by clicking **“Preview your slices”**, and start the analysis calculation by clicking **“Calculate ToxPi scores and build visualizations”**.

Initially you will see one slice in the window. You can set some unique name for the slice by rewriting the text in the **“Enter slice name:”** field. In the **“Select slice type”** area you will see the possible types based on the data from your input file. You can create your custom type by selecting **“Custom”** and entering the type name in the field below. You can select the slice color by clicking on the colored rectangle below the **“Select color:”** text. There are predefined color sets for five base types:

- Assay – gradient of green color
- Pathway – gradient of blue color
- ChemProp – gradient of orange color
- Exposure – gradient of gray color



- ToxRefDB – gradient of purple color
- Custom type – gradient of red color

You can set the slice weight in relation to other slices by dragging the slider in the “**Slice weight**” area to a value from 1 to 25. The weight will be displayed in percent ratio and in absolute values, and controls the width of the slice (i.e., slices with higher weights will be wider than slices with lower weights).

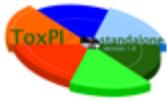
In the “**Select components**” area you will be able to select the components you want to add to a given slice. You can search for some particular components using the “**Search**” field. The matches will be highlighted in the components list.

In the “**Select scaling type**” area you will be able to select a scaling option for the data in the components. There are 5 scaling type allowed:

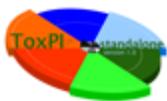
- **$-\log_{10}(x)+6$** – all the values (x) will be normalized using the following formula:
 $-1 * \log_{10}(x) + 6$ (use it when 1,000,000 value indicates no effect)
- **$-\log_{10}(x)+\log_{10}(\max(x))$** - all the values (x) will be normalized using the following formula: $-1 * \log_{10}(x) + \log_{10}(\max(X))$
- **hit count** – all the values (x) will be treated as binary data, meaning that all non-zero values will be replaced with ones, and all zeros will remain zeros (use it when you have only binary data in the input files)
- **$-\ln(x)$** – all the values (x) will be normalized using the following formula
 $-1 * \ln(x)$
- **sqrt(x)** – all the values (x) will be normalized using the square root value of the original value
- **linear(x)** – no normalization will be applied

You will see some basic statistical analysis of the included components in the “**Selected data analysis**” field.

You can remove the slice by clicking “**Remove this slice**”.



6. Results Window



Click **“Calculate ToxPi indexes and build visualizations”** to see the output of your selections above. When all the necessary calculations are completed, the ToxPi GUI results window will appear ([Figure 5](#) below).

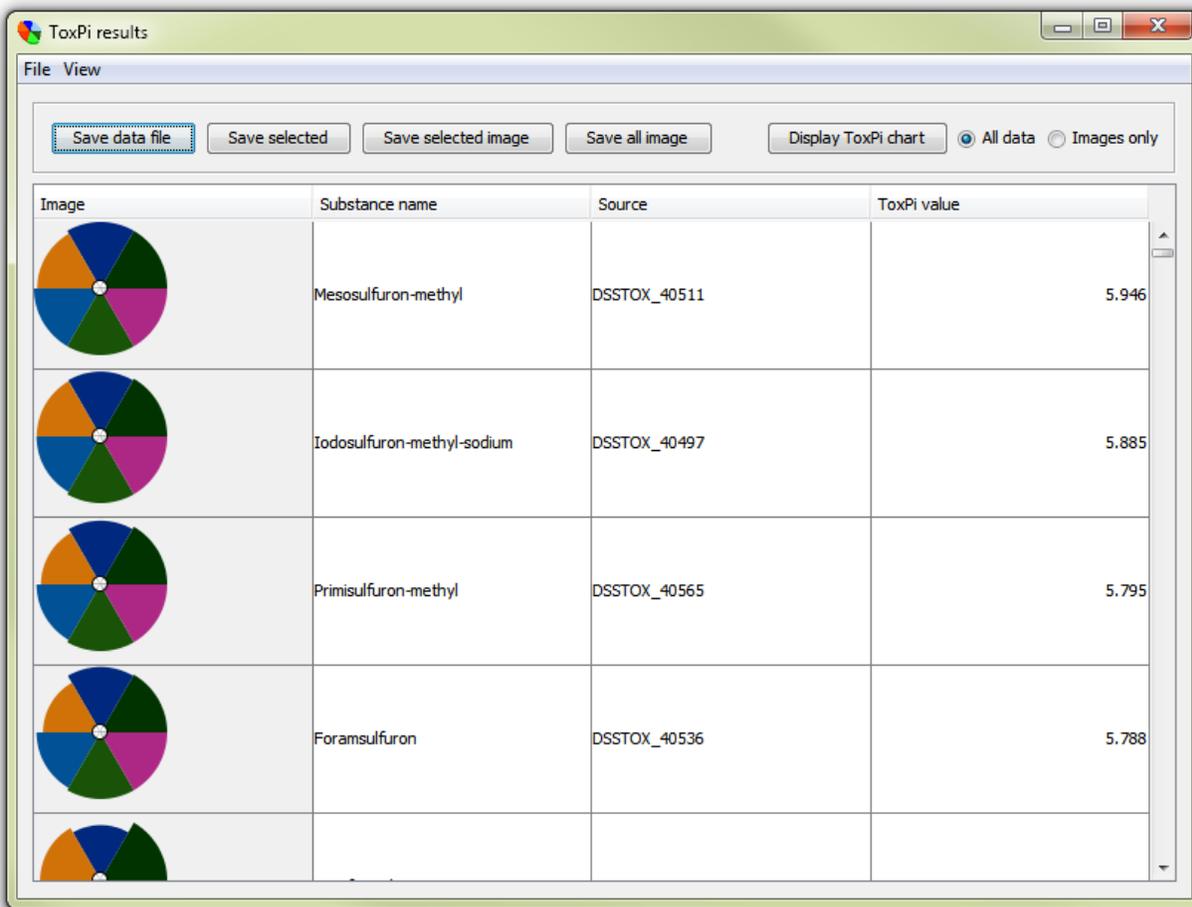
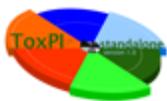


Figure 5. Results Window

The results window will initially display the table of all the chemicals used for the calculations. The table will contain the ToxPi pie image, Chemical name, Source ID, and calculated ToxPi score for each chemical.

Above the table you will see three buttons and two radio controls. The three buttons are:

- **Save data file** – This will save your data. You will be asked to select or enter the file name where the current data will be saved, and later, the saved file can be

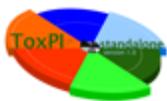


used to recreate your current results using the **“Recreate data from file”** button on the home screen.

- **Save selected** – This will save only selected chemicals into a file.
- **Save Selected image** - This will save only images for selected chemicals into one file.
- **Save all image** – This will save all the chemical images into one file, sorted according to the ToxPi score.
- **Display ToxPi chart** – This will display the ToxPi chart in a new window.

Radio controls allow you to switch between different views of your results data. Extended information will be displayed when **“All data”** is selected, while only images will be displayed if **“Images only”** is selected.

If you click on a ToxPi image on either table (extended or images only), a new window will appear containing detailed information about the selected chemical (See [Figure 6](#)).



6.1 Chemical Window

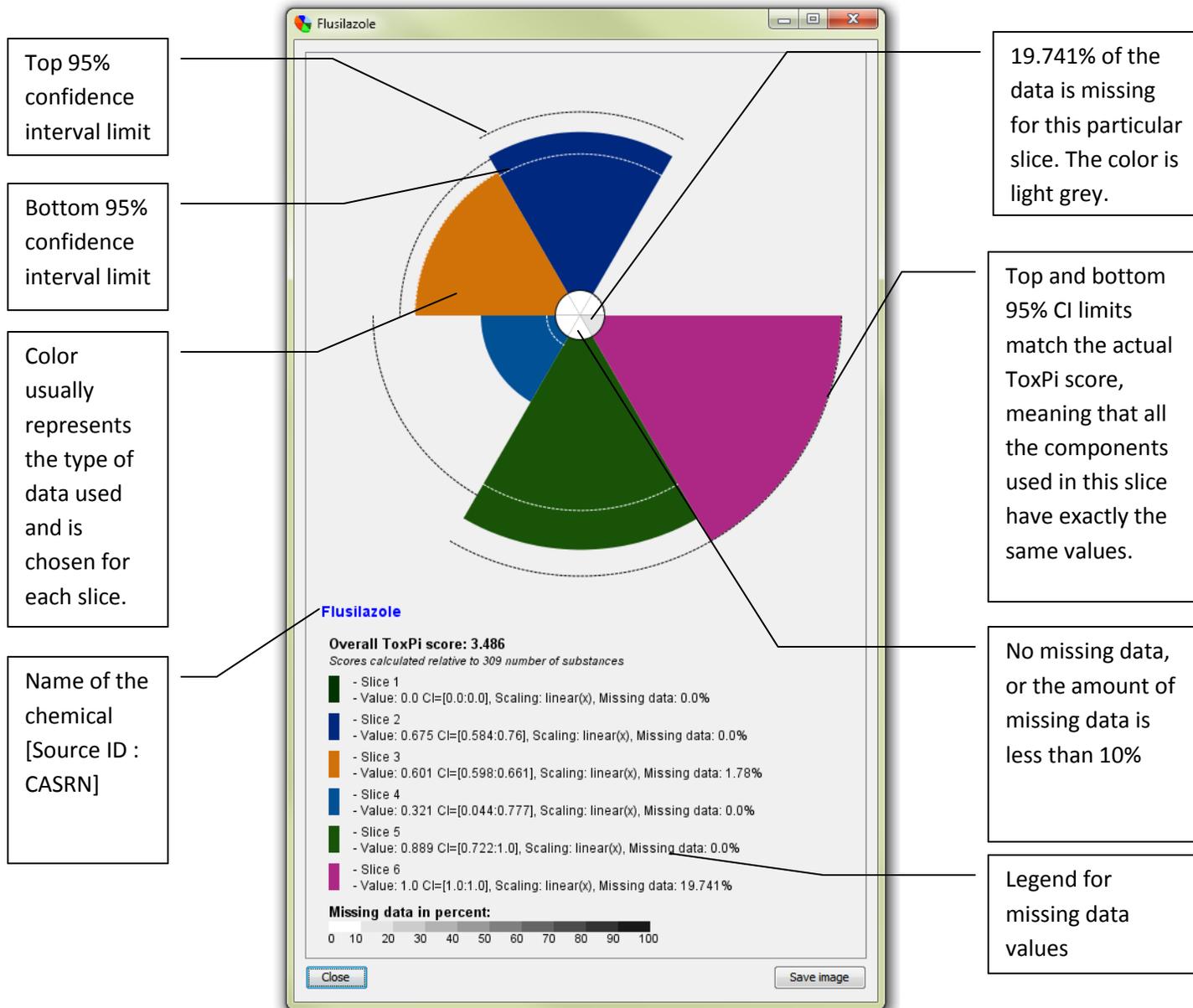
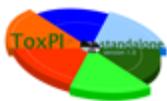


Figure 6. Chemical Information Window

The chemical information window will contain a ToxPi pie image with 95% confidence intervals (CI) highlighted for each slice (upper limit – black dashed line, lower limit – white dashed line). The core of the image represents the amount of missing data, denoted by the color (the lighter the core, the lower the amount of missing values). You can open multiple Chemical information windows simultaneously.

You can open multiple chemical information windows simultaneously.



You can save your image by clicking **“Save image”** and specifying the file name. The image will be saved in PNG format with a transparent background and publication ready quality.

You can close the window by clicking the **“Close”** button.

6.2 Chart Window

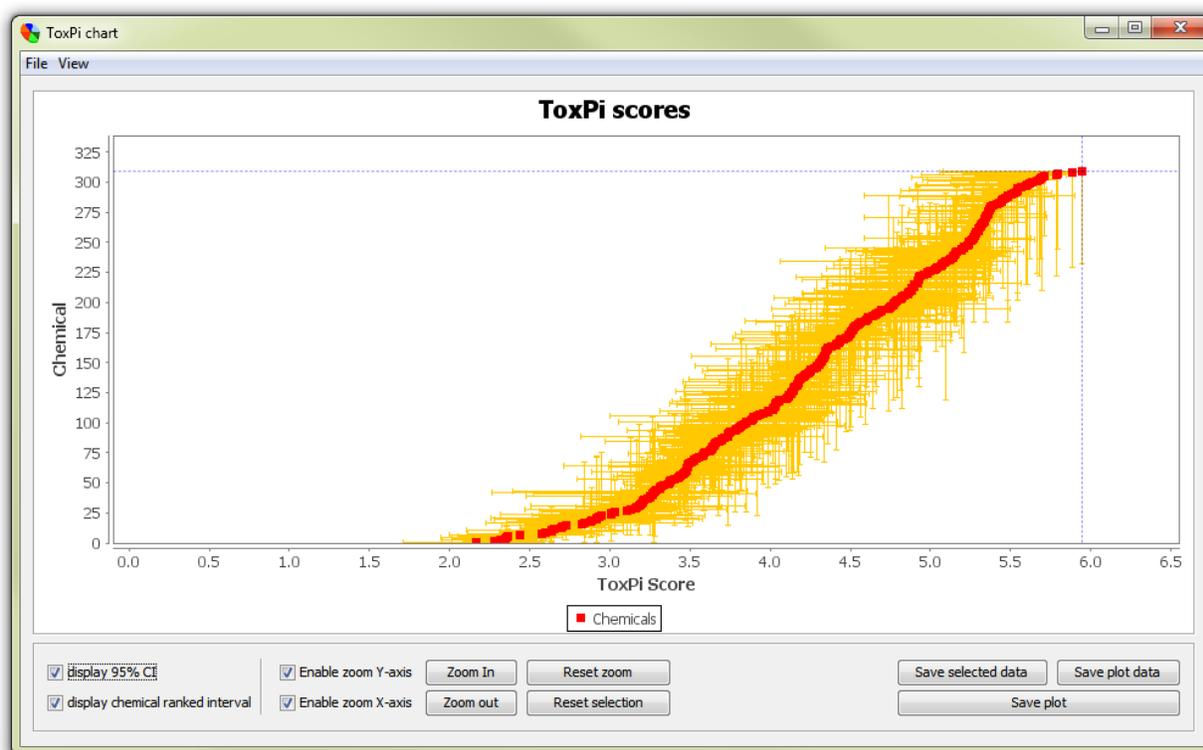
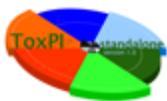


Figure 7. ToxPi Chart Window

The ToxPi chart window will present a chart of the chemicals, based on the ToxPi scores on the X-axis and ranked chemicals on the Y-axis. The actual calculated ToxPi score for each chemical will be represented as red dots.

You can save the plot image by clicking the **“Save plot”** button. The image saved will reflect the contents of the window at the time of saving.



Two additional buttons - “Save selected data” and “Save plot data” allow one to save the actual information (ToxPi Score, 95% confidence intervals and rank intervals) about all or only selected chemicals.

6.2.1 Confidence intervals

The horizontal orange bars, along the X-axis, indicate the bootstrapped 95% confidence interval for a particular chemical. The vertical orange bars, along the Y-axis, indicate the ranking interval for a specific chemical.

You can show or hide the confidence intervals and/or the ranking intervals by checking or unchecking the “**Display 95% CI**” and “**Display chemical ranked interval**” check boxes.

6.2.2 Zooming and Dragging

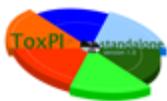
In order to zoom in or out on the plot, you are encouraged to use your mouse or track pad scroll wheel. To move the plot to the area you are interested in, click your left mouse button and drag.

You can disable zoom for either axis by unchecking the “**Enable zoom X-axis**” or “**Enable zoom Y-axis**” checkboxes. To reset the zoom and return to the initial view, click on the “**Reset zoom**” button.

6.2.3 Highlighting the Data

In order to highlight specific data, simply click on the red dot representing the chemical you are interested in. The highlighted chemical should become blue and will be highlighted in the results table as well. You can also select chemicals in your results table and then see the selected chemicals highlighted blue on the chart plot.

To unselect chemicals click on your selections and the dot should change color from blue to red. Alternatively one can use “**Reset selection**” button to clear all selections.



7. Description of ToxPi Bootstrapping Algorithm.

Confidence intervals for the ToxPi score of each chemical are generated by bootstrapping the values within each slice. The assay values within each slice are sampled with replacement, and these resampled values are used to calculate a bootstrap ToxPi statistic. This bootstrap ToxPi statistic is calculated exactly as the original statistic, but on the resampled data. This process is repeated 1000 times, and these 1000 or more bootstrap statistics are used to assess the stability of the estimated ToxPi score for the chemical. In particular, a 95% confidence interval for the ToxPi score is generated in the standard way: the lower bound is given by the 2.5 percentile in the bootstrap statistics and the upper bound is given by the 97.5 percentile of the bootstrap statistics.

Intuitively, the width of the confidence interval for a chemical depends on the amount of variability within each slice. The ToxPi score will have a narrow confidence interval if the assay values within a slice are very similar and a wide confidence interval if the assay values within a slice are very different. The bootstrapping approach does not make any assumption regarding the distribution of the data values, and should give appropriate confidence intervals in most contexts. However, we caution that confidence intervals generated using the approach described here may be unreliable in the following situations:

- If the measurements within each slice are on dramatically different scales.
- If there are just a small number (1-3) of assays within each slice.
- If there are extreme values (outliers) in the data.