



# ImprovRisk

**v.2.0.0**

## **A model for Dietary Risk Assessment**

Description and User Manual

**July 2019**

**Nicosia, Cyprus**



## **Preface**

This ImproRisk user manual corresponds with the updated version of ImproRisk (v.2.0.0) released in July 2019.

In this new version of the model, two major updates have been implemented.

### *Exposure calculation at FoodEx1 Level 3*

The user is now able to load occurrence data both in Level 2 and in Level 3 of the foodEx1 food characterization system. For that, a new occurrence template is provided where the risk assessor can provide aggregated occurrence values for up to 1000 level-3 food items.

Moreover, there is a new pivot table in Contribution Worksheet that allows the user to calculate the contribution of each level-3 food item to the total intake of the population and the contribution within Level 2. The contribution is currently calculated only at the MB occurrence scenario.

### *Weighting Coefficients*

The ImproRisk version 2.0.0 accommodates weighting coefficients to adjust the sample for non-representativeness within the population. Exposure statistics e.g. mean exposure and standard error will now be adjusted to reflect the population from which the sample was taken from.

We thank all the users of the model, for their valuable feedback

## Table of Contents

Introduction.....	5
Benefits .....	6
Ownership and intellectual rights .....	6
Contact.....	6
Methodology.....	7
Launching the ImproRisk .....	8
WORKSHEETS of the ImproRisk model .....	9
Worksheet <b>Subjects</b> .....	9
Worksheet <b>Foodlist</b> .....	10
Worksheet <b>Consumption</b> .....	11
Refined Exposure.....	11
Level 3 categorisation.....	12
Worksheet <b>OccurrenceL2</b> .....	16
Worksheet <b>MeanConsumption</b> .....	18
Worksheet <b>Contribution</b> .....	20
Level 1 contribution.....	20
Level 2 contribution.....	23
Level 3 contribution.....	24
Worksheet <b>Exposure</b> .....	26
Worksheet <b>ExposureMB-LB-UB</b> .....	28
Summary statistics.....	28
Frequency distribution table .....	30
Worksheet <b>GENDER</b> .....	32
Comparison between FEMALES and MALES.....	32
Worksheet <b>AREA</b> .....	35
Worksheet <b>POP_CLASS</b> .....	37
Worksheet <b>DrillDown</b> .....	38
Worksheet <b>COMPREHENSIVE</b> .....	39
Worksheet <b>LOG</b> .....	42
Updating the Occurrence and Consumption .....	43
Occurrence.....	43

Level2.....	43
Level 3 .....	44
Updating.....	45
Consumption .....	47
Worksheet "Subjects" .....	47
Worksheet "Consumption" .....	48
Weighting coefficients .....	51
Calculating weights.....	52
Example - Cyprus.....	52
Subjects and Consumption template.....	54
General Information.....	56
Automatic recalculation .....	56
Locked workbook and worksheet.....	57
Notes within the worksheet.....	58
Named Ranges in Excel .....	58
Statistical Terms.....	59
Cumulative Frequency .....	59
Cumulative percentage .....	59
Quartiles.....	59
Probability Distribution (histogram).....	60
Cumulative Distribution .....	60
Decimal and Thousand separator .....	62
Pop Up Menu for quick navigation across worksheets.....	64
References .....	65

## Introduction

The ImproRisk model is a tool built in MS Excel, for conducting dietary risk assessment analysis.

It combines Food Consumption dataset with Occurrence (of chemical substances) dataset and calculates the exposure rates for the population.

ImproRisk supports Food Consumption datasets at individual level. Meaning, one can load food consumption data that came out of a food survey. All the individuals are used, and the consumption data needs to be coded according to EFSA food classification FoodEx1 at level 4 along with the amounts consumed. The food consumption dataset may contain demographic characteristics of the population and especially **must contain** the individual's own body weight so that the exposure per kilogram of body weight (Kg/ b.w.) is calculated.

Calculation at each food consumption occasion is performed. Every individual, at each food consumption occasion, is exposed to a chemical substance. The exposure at that individual food consumption occasion is calculated by matching the amount of consumption of the food and occurrence of a chemical substance. Individual's body weight is considered. Then, by aggregating the results, this leads to the following:

- A) Exposure rate for the population. Determine the rate of exposure of the population to a given reference value (i.e. 80% of the population is below the reference value). This can be either weekly, or daily reference value according to the chemical compound and EFSA guidelines.
- B) Derivation of probability and cumulative distributions of the exposures. This will enable the risk manager to have a clearer view on how the exposure is distributed over the population of interest.
- C) Explore the effect of demographic characteristics on the exposure. For example, one can compare the exposure towards gender or age group of the population.

The above can be derived for all Lower Bound (LB), Middle Bound (MB), and Upper Bound (UB) scenarios of the occurrence dataset and the exposure is calculated at the FoodEx1 level 2.

The occurrence dataset needs to be in FoodEx1 Level 2 or FoodEx1 Level 3, and the Consumption dataset must be in FoodEx1 level 4.

## Benefits

- I) The calculation is performed at the individual level; thus it enables the estimation of exposure distribution and waves the biased effect of the mean consumption.
- II) Not a closed box model. All calculations are there to inspect. The formulas and all the methodology are transparent so the model results can be validated easily.
- III) A straightforward and user-friendly model.
- IV) Free to use.

## Ownership and intellectual rights

ImproRisk is owned by the State General Laboratory of Cyprus (SGL) and was developed by the private company Improvast. SGL reserves the property rights. The model is free to use, but any unauthorised modification is prohibited under the Cyprus law.

## Contact

This manual was prepared by Improvast. Should you find any errors or have any comments/suggestions please contact [info@improvast.com](mailto:info@improvast.com).

Additionally, should you have any suggestions for further development or encounter bugs please report them to [gstavroulakis@sgl.moh.gov.cy](mailto:gstavroulakis@sgl.moh.gov.cy) or [info@improvast.com](mailto:info@improvast.com).

We welcome all feedback from all interested parties.

## Methodology

The model considers the mean occurrence (LB, MB, & UB scenario) and uses the consumption data at an individual level.

The calculated individual exposure can be later be used to cross tabulate by Gender, Age group and Area.

The model is ready to calculate the intake levels of any substance under study by substituting the occurrence data via an automated VBA code (Visual Basic for Applications). Moreover, the model enables the user to easily install a new Consumption Database.

The Estimated Daily Intake (EDI) of chemical Substance is calculated as follows:

- a) For every single food consumption occasion, at any day, for any food category, the following formula is applied:

$$\frac{\text{Consumption (grams of food)} \times \text{Occurrence (mg/Kg of food)}}{\text{Body Weight (Kg)}} = \text{Exposure (}\mu\text{g/Kg b.w.)}$$

*b.w. = body weight*

- b) The body weight is each individual's body weight. The outcome of the equation is the exposure (in  $\mu\text{g/Kg b.w.}$ ) for each instance the subject consumes each and every specific food.

- c) The Exposure:

The Exposure for each instance, calculated in step a), is summed for all days and for every food category in FoodxLevel2 (160 categories).

The EDI is calculated by taking the average exposure throughout the number of days of the study.

The result is the EDI in  $\mu\text{g/Kg b.w.}$  for each individual in the food survey.

Note that 3 types of EDI's are calculated at LB, MB and UB.

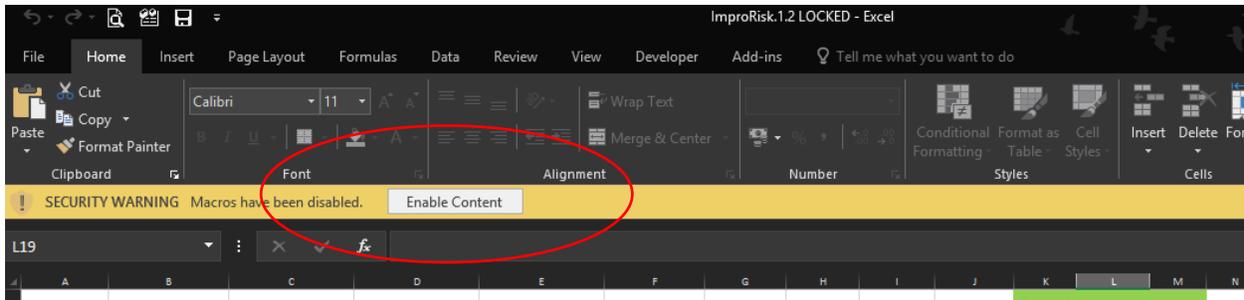
This way the exposure rate is established for each individual at an Optimistic (LB), an "Average" (MB) and a Pessimistic (UB) scenario.

# Launching the ImproRisk

Open the file **ImproRisk.xlsx**

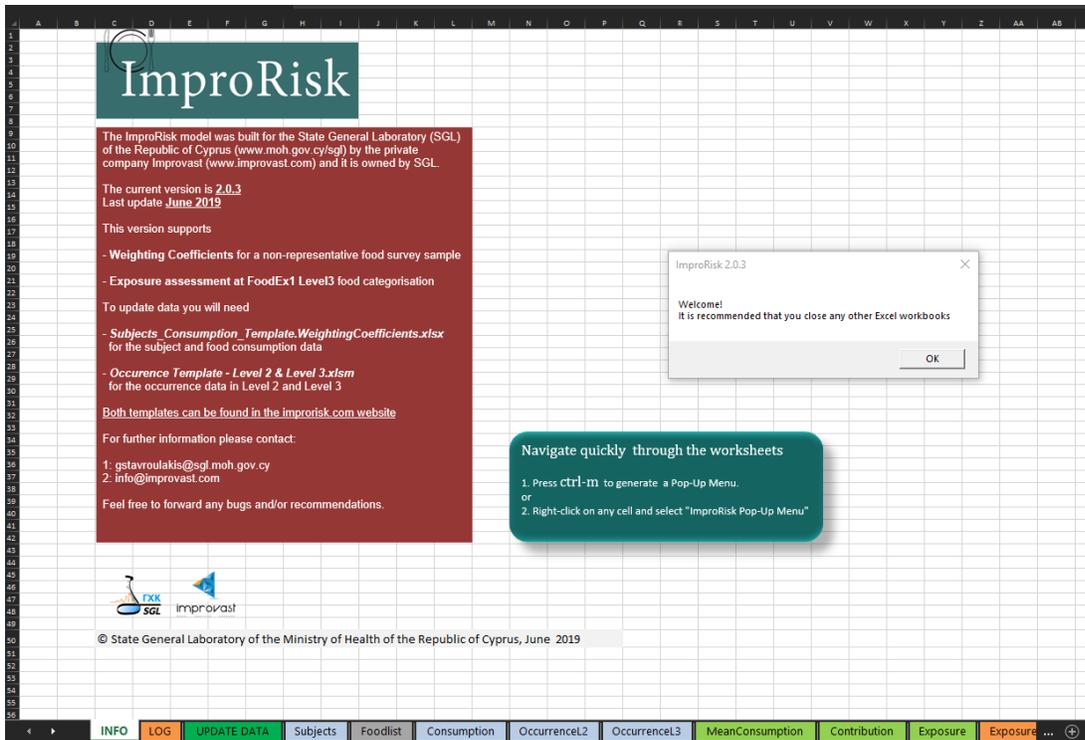
If a security warning is issued (Figure 1) that macros have been disabled, the user should click on the Enable Content button. The Macros are the programming code in VBA (Visual Basic for Applications) which is included in Excel and must be activated.

Figure 1: Security warning for enabling macros in Excel



The user will get a welcoming message informing you to close all other Excel workbooks (Figure 2). This will reduce the processing speed and increase the efficiency of the model.

Figure 2: Opening ImproRisk



# WORKSHEETS of the ImproRisk model

An overview of each worksheet's functionality in the model reporting.

## Worksheet **Subjects**

*Press CTRL-Home to navigate to the TOP LEFT cell (A1)*

The worksheet Subjects tabulates the sample population data to be used (Figure 3). Each line represents an individual. Across the worksheet the user can see the following:

The subject ID, Gender, Age, Weight, Area, Population Class (POP\_CLASS) and WCOEFF

The WCOEFF is the weight coefficient of each subject to generalise the exposure to the total population.

The Sum of Weight Coefficients is the total size of the population from which the that the subjects were drawn from. In the example (Figure 3) we sampled 272 subjects from a population of size 551.204. The *Sample Size* is calculated automatically via the VBA code when the *Consumption and Subjects* dataset is installed.

Figure 3: The Subjects worksheet

GENDER	AGE	WEIGHT	AREA	POP_CLASS	WCOEFF	Sample Size	Sum of the Weight Coefficients (WCOEFF)		
MALE	24,25	76,7	Lemesos	Adults	1993	272	551.204	<p><b>NOTE 1:</b> This worksheet contains the participants' demographic characteristics.</p> <p>When new survey (consumption data) is available, do not alter this sheet! It is highly recommended that first the new data are inserted in the "Subjects_Consumption_Template.xlsx" that is provided, and then use the "UPDATE DATA" worksheet.</p> <p>Sample Size is calculated automatically when the consumption data are updated through the hidden VBA code.</p> <p>Population Size is calculated when summing the weighting coefficients (WCOEFF) in column G.</p>	<p><b>NOTE 2:</b> <b>POP_CLASS - AGE</b> Infants* 3-11 months Toddlers* 1-2 years Other children* 3-9 years Adolescents* 10-17 years Adults* 18-64 years Elderly* 65-74 years Very elderly* &gt;75  *EFSA Journal 2014;12(12):3944</p>
MALE	32,76	60,1	Lemesos	Adults	1993				
MALE	41,31	64,7	Pafos	Adults	1993				
FEMALE	44,53	72	Pafos	Adults	2059				
FEMALE	21,53	82	Pafos	Adults	2059				
MALE	23,88	56	Pafos	Adults	1993				
FEMALE	63,44	92,8	Pafos	Adults	2059				
FEMALE	49,14	57	Lemesos	Adults	2059				
MALE	62,9	89,6	Lemesos	Adults	1993				
MALE	31,37	94,6	Lemesos	Adults	1993				
FEMALE	44,77	138	Pafos	Adults	2059				
MALE	64,59	91	Pafos	Adults	1993				
MALE	24,6	99	Pafos	Adults	1993				
FEMALE	20,8	49,5	Lemesos	Adults	2059				
FEMALE	34,06	67	Pafos	Adults	2059				
MALE	60,64	109	Pafos	Adults	1993				
MALE	32,47	67	Pafos	Adults	1993				
FEMALE	64,79	75,5	Lemesos	Adults	2059				
FEMALE	21,04	57	Lemesos	Adults	2059				
FEMALE	47,3	75,6	Lemesos	Adults	2059				
MALE	58,11	75	Lemesos	Adults	1993				
MALE	53,04	78,8	Lemesos	Adults	1993				
MALE	64,43	90,7	Lemesos	Adults	1993				
MALE	65,89	94,7	Lemesos	Adults	1993				
FEMALE	53,46	48,9	Lemesos	Adults	2059				
FEMALE	64,06	64,9	Lemesos	Adults	2059				
FEMALE	63,43	76,7	Lemesos	Adults	2059				
FEMALE	24,87	92,2	Lemesos	Adults	2059				
FEMALE	31,03	85	Lemesos	Adults	2059				
MALE	44,94	99	Pafos	Adults	1993				
FEMALE	31,94	51,5	Pafos	Adults	2059				
MALE	30,79	79	Pafos	Adults	1993				
FEMALE	37,5	75,6	Pafos	Adults	2059				
FEMALE	39,28	68,1	Lefkosa	Adults	2059				
FEMALE	30,79	88,9	Lemesos	Adults	2059				

# Worksheet Foodlist

Press CTRL-Home to navigate to the TOP LEFT cell (A1)

This worksheet contains the hierarchy of the food classification FoodEx1 from Level 1 to Level 4 as indicated by EFSA (Figure 4). The name of the food in FoodEx1 L4 is in the Column D ("FOODEX\_L4\_name"). The first three columns are the corresponding names of the food, in Level 1 (Column A) up to Level 3 (Column C).

Press CTRL-Down Key to navigate to the END of the worksheet.

Press CTRL-Home to navigate back to the beginning

Figure 4: The Foodlist worksheet

FOODEX_L1_name	FOODEX_L2_name	FOODEX_L3_name	FOODEX_L4_name	NOTE 3:
Grains and grain-based products	<p>Here there are 1897 rows of data.</p> <p>These 1897 rows represent the food categorisation on Level 4. The name of the food in Foodex L4 is in the Column D ("FOODEX_L4_name").</p> <p>The first three columns are the corresponding names of the food, in Level 1 (Column A) up to Level 3 (Column C).</p> <p>For more info please see EFSA 2011; 9(3):1970</p> <p><a href="http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/1970.pdf">http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/1970.pdf</a></p>			
Grains and grain-based products	Grains for human consumption	Grains for human consumption	Grains for human consumption	
Grains and grain-based products	Grains for human consumption	Wheat grain	Wheat grain	
Grains and grain-based products	Grains for human consumption	Wheat grain	Wheat germ	
Grains and grain-based products	Grains for human consumption	Wheat grain	Wheat grain, Durum	
Grains and grain-based products	Grains for human consumption	Wheat grain	Wheat grain, soft	
Grains and grain-based products	Grains for human consumption	Wheat grain	Bulgur wheat	
Grains and grain-based products	Grains for human consumption	Barley grain	Barley grain	
Grains and grain-based products	Grains for human consumption	Barley grain	Barley grain, whole	
Grains and grain-based products	Grains for human consumption	Barley grain	Barley, pearled	
Grains and grain-based products	Grains for human consumption	Corn grain	Corn grain	
Grains and grain-based products	Grains for human consumption	Rye grain	Rye grain	
Grains and grain-based products	Grains for human consumption	Spelt grain	Spelt grain	
Grains and grain-based products	Grains for human consumption	Spelt grain	Spelt grain, rippen	
Grains and grain-based products	Grains for human consumption	Spelt grain	Spelt grain, unripped	
Grains and grain-based products	Grains for human consumption	Buckwheat grain	Buckwheat grain	
Grains and grain-based products	Grains for human consumption	Millet grain	Millet grain	
Grains and grain-based products	Grains for human consumption	Oats, grain	Oats, grain	
Grains and grain-based products	Grains for human consumption	Rice	Rice	
Grains and grain-based products	Grains for human consumption	Rice	Rice, brown	
Grains and grain-based products	Grains for human consumption	Rice	Rice, long-grain	
Grains and grain-based products	Grains for human consumption	Rice	Rice, mixed	
Grains and grain-based products	Grains for human consumption	Rice	Rice, parboiled	
Grains and grain-based products	Grains for human consumption	Rice	Rice, red	
Grains and grain-based products	Grains for human consumption	Rice	Rice, white	
Grains and grain-based products	Grains for human consumption	Rice	Rice, wild	
Grains and grain-based products	Grains for human consumption	Other grains	Other grains	
Grains and grain-based products	Grains for human consumption	Other grains	Mixture of grains	
Grains and grain-based products	Grains for human consumption	Other grains	Sorghum grain	
Grains and grain-based products	Grains for human consumption	Other grains	Enkorn	
Grains and grain-based products	Grains for human consumption	Other grains	Emmer	
Grains and grain-based products	Grain milling products	Grain milling products	Grain milling products	
Grains and grain-based products	Grain milling products	Wheat milling products	Wheat milling products	
Grains and grain-based products	Grain milling products	Wheat milling products	Wheat flour, brown	
Grains and grain-based products	Grain milling products	Wheat milling products	Wheat flour, Durum	
Grains and grain-based products	Grain milling products	Wheat milling products	Wheat flour, white	
Grains and grain-based products	Grain milling products	Wheat milling products	Wheat flour, wholemeal	
Grains and grain-based products	Grain milling products	Wheat milling products	Graham flour	
Grains and grain-based products	Grain milling products	Wheat milling products	Wheat flour, gluten free	
Grains and grain-based products	Grain milling products	Wheat milling products	Couscous	
Grains and grain-based products	Grain milling products	Wheat milling products	Wheat bran	
Grains and grain-based products	Grain milling products	Wheat milling products	Wheat groats	
Grains and grain-based products	Grain milling products	Wheat milling products	Wheat semolina, Durum	
Grains and grain-based products	Grain milling products	Wheat milling products	Wheat semolina, soft wheat	

## Worksheet Consumption

*Press CTRL-Home to navigate to the TOP LEFT cell (A1)*

The first five columns (A, B, C, D, E, coloured in light BLUE background colour) contain the uploaded consumption dataset (Figure 5).

*Press CTRL-Down Key*

These are the **food consumption occasions** within the sample, i.e. the *SUBJECTID* of the person who consumed the food, the sequential *DAY* of consumption in the food survey (what day the individual consumed it), the *amount of food* (in grams) and the *FoodEx1 Level 4 name* of the food. Subjects consumed different and various amounts and type of foods at each day, so the same *SUBJECTID* and *DAY* will appear many times.

*Press CTRL-Home to navigate back to the beginning*

Columns "F" and "G" (*grey coloured*) translate the FoodEx1 level4 name into level 1 and level2 respectively. This is the main use of the Worksheet **Foodlist**.

Columns "H", "I" and "J" hold the occurrence of the chemical in the LB, MB, and UB scenario respectively. These values are obtained (automatically) from the **Worksheet Occurrence** that will be presented later in this user manual.

Columns "K", "L", "M", "N", "O", "P", hold the demographic characteristics of the subject of that particular food consumption occasion; Gender, Area, Pop Class, Age, Weight and the Weight Coefficient (WCOEFF).

Columns "Q", "R", and "S" hold the (unweighted) calculated exposure from that food consumption occasion (Figure 7). Note that the "WCOEFF" column DOES NOT participate in the calculations of these columns.

Column "T" holds the level 3 food name for that consumption occasion

Columns "U", "V", and "W" calculate the exposure at level 3 of that consumption occasion. This is calculated only and if the user provided occurrence values at the level 3 and are in the **Worksheet OccurrenceL3**. The empty cells (designated with the dash "-") are food consumption occasions where no occurrence values at level-3 are provided for that food item. These columns are essential when the exposure at the individual level is calculated in the Worksheet **Exposure**.

## Refined Exposure

Columns X, Y and Z calculate the "refined exposure" at each food consumption occasion (Figure 8). The formulas check whether an exposure at level-3 is present

and take that value as the exposure of that food consumption occasion. If not, then the exposure at level-2 is considered. This highlights the need to additionally provide aggregated occurrence data at level-2.

Figure 5 shows an example of calculating the exposure of individuals in the food survey, where both Level 2 and Level 3 occurrence values are provided.

Subject 1001 had two food consumption instances. In the first occasion, the exposure is 0.0036 and it is calculated at the level-2 occurrence because no occurrence at level-3 was provided. The second food consumption occasion incurs two alternative exposures since both Level 2 and Level 3 occurrence values were provided. The total exposure for that individual, is then calculated as  $0.0026 + 0.0039 = 0.0065$  (the exposure at Level 2 at the first occasion + the exposure at level-3 of the second occasion). The exposure at level-2 of the second occasion is not considered since we have a more accurate estimated (in Level 3) of occurrence, hence exposure.

Figure 5 Example - Calculation of refined exposure with occurrence at level-3

Food consumption occasions			EXPOSURE at the individual level	
SUBJECTID	Exposure at Level2	Exposure at Level3	ID	EXPOSURE
1001	0,0026	-	1001	0,0065
1001	0,0022	0,0039	1002	0,0184
1002	-	0,0069	1003	0,0675
1002	0,0115	-		
1003	0,0065	-		
1003	0,0610	-		

### Level 3 categorisation

Column AA re-organizes the food names at level-3 (Figure 9). The formulas check whether an exposure at level-3 is present, and if yes, then the food name at level-3 is considered. If not, then it uses the level-2 name of the food. In summary, it captures the food name at level-3 and at the same time aggregates all the food consumption occasions to "Other" where no exposure at level-3 is provided. This column is essential for the contribution of food items at level-3 compared to the total intake and to the intake with each level-2 (see Level 3 Contribution section)



Figure 8 Consumption worksheet – Refined exposure at Level 3

Figure 9 Consumption worksheet – Level 3 categorisation

**IMPORTANT NOTE 1:**

1. Calculation of exposure at each food consumption occasion:

The excel built in formula IFERROR() is used when the multiplication is performed.

In case where an invalid Food name is used for a food consumption occasion, then no consumption value for this consumption occasion will be found, thus an error is created in the columns of Exposure (P, Q, and R). This way, if an invalid name is used, then the model assumes that there is no food consumption occasion (leaving the exposure blank) thus no exposure for that particular food consumption

occasion, and the aggregating statistics calculations (sheet Exposure) are performed without any interruption.

2. The occurrence of each food consumption occasion is retrieved in columns H, I, J. When the value is empty, this means that no mean occurrence for that food category is determined in the "Occurrence" sheet.

Solution:

By using the filters provided by Excel, any errors can be identified easily e.g. consumption name not included at the food level 2 or occurrence data values do not exist in columns H, J and I.

### **IMPORTANT NOTE 2:**

Whenever there is a need to update the consumption database, it must be done with a template that is provided (Subjects\_Consumption\_Template.WeightingCoefficients.xlsx). Go to "UPDATE DATA" worksheet and click on the button for updating Consumption data. The user will be asked to locate and select the file in the local disk and replacement of both the Subject information and Consumption information will be performed automatically.

Note that the template name is not of importance. Name the resulting file appropriately for easier reference. What is of importance, it is the name of the two worksheets within the template that need to be named "Subjects" for the subject information and "Consumption" for the Consumption information.

See instructions in the Chapter Updating the Occurrence and Consumption

## Worksheet OccurrenceL2

This worksheet contains the concentration data (occurrences) in FoodEx1 level 2.

**Note that ImproRisk needs the occurrence to be measured in mg/ Kg of food**

Figure 10: Worksheet OccurrenceL2

A		B		C		D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Chemical Substance		Pb		NOTE 4:															
Substance Category		Contaminant		This worksheet contains the concentration data (occurrences) in FoodEx1 at Level 2.															
Reference value (µg/Kg b.w.)		0,63		Note(!) that ImproRisk needs the occurrences to be expressed in mg/ Kg of food															
Type of Reference value		Benchmark Dose Level (BMDL)																	
Type		DAILY																	
				mg/kg food															
				min			mean			median			P95						
FoodEx1_name	FoodEx2_name	No of Samples	LB	MB	UB	Occur_Mean LB	Occur_Mean MB	Occur_Mean UB	LB	MB	UB	LB	MB	UB	LB	MB	UB		
Grains and grain-based products	Grains and grain-based products	9937	-	-	-	0,025	0,029	0,033	-	-	-	0,105	0,105	0,105					
	Grains for human consumption	4936	-	-	-	0,027000	0,031000	0,035000	-	-	-	0,109	0,109	0,109					
	Grain milling products	1842	-	-	-	0,028000	0,029000	0,033000	-	-	-	0,117	0,117	0,117					
	Bread and rolls	1192	-	-	-	0,025000	0,029000	0,033000	-	-	-	0,098	0,098	0,098					
	Pasta (Raw)	381	-	-	-	0,021	0,024	0,027	-	-	-	0,089	0,089	0,091					
	Breakfast cereals	789	-	-	-	0,019000	0,025000	0,030000	-	-	-	0,080	0,080	0,090					
	Fine bakery wares	675	-	-	-	0,021	0,025	0,028	-	-	-	0,080	0,080	0,086					
Vegetables and vegetable products (including fungi)	Vegetables and vegetable products (including fungi)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
	Root vegetables	2350	-	-	-	0,018	0,019	0,021	-	-	-	0,060	0,060	0,060					
	Bulb vegetables	799	-	-	-	0,029	0,031	0,033	-	-	-	0,061	0,064	0,069					
	Fruiting vegetables	2891	-	-	-	0,007	0,011	0,014	-	-	-	0,039	0,045	0,050					
	Brassica vegetables	1977	-	-	-	0,010	0,013	0,016	-	-	-	0,050	0,050	0,050					
	Leaf vegetables	3122	-	-	-	0,037	0,041	0,044	-	-	-	0,101	0,101	0,130					
	Legume vegetables	77	-	-	-	0,022	0,026	0,030	-	-	-	0,081	0,081	0,081					
	Stem vegetables (Fresh)	1267	-	-	-	0,017	0,021	0,025	-	-	-	0,059	0,075	0,078					
	Sugar plants	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	Sea weeds	66	-	-	-	2,663	2,677	2,692	-	-	-	1,100	1,1	1,100					
	Tea and herbs for infusions (Solid)	473	-	-	-	0,353	0,355	0,358	-	-	-	1,530	1,530	1,530					
	Cocoa beans and cocoa products	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	Coffee beans and coffee products (Solid)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	Coffee imitates (Solid)	16	-	-	-	0,223	0,224	0,224	-	-	-	-	-	-	-	-	-	-	
Vegetable products	427	-	-	-	0,030	0,030	0,031	-	-	-	0,105	0,105	0,105						
Fungi, cultivated	1722	-	-	-	0,053	0,057	0,060	-	-	-	0,194	0,194	0,194						
Fungi, wild, edible	786	-	-	-	0,478	0,491	0,503	-	-	-	2,700	-	-						
Starchy roots and tubers	Other starchy roots and tubers	984	-	-	-	0,016	0,017	0,018	-	-	-	0,060	0,060	0,060					
	Potatoes and potatoes products	1370	-	-	-	0,016	0,019	0,022	-	-	-	0,060	0,064	0,066					
	Starchy roots and tubers	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
Legumes, nuts and oilseeds	Legumes, beans, dried	1395	-	-	-	0,030000	0,034000	0,038000	-	-	-	0,122	0,122	0,124					
	Legumes, beans, green, without pods	610	-	-	-	0,013000	0,016000	0,020000	-	-	-	0,064	0,064	0,072					
	Legumes, nuts and oilseeds	4071	-	-	-	0,030000	0,034000	0,038000	-	-	-	0,079	0,079	0,079					
	Oilseeds	1036	-	-	-	0,043000	0,046000	0,048000	-	-	-	0,171	0,171	0,171					
	Other seeds	19	-	-	-	0,080000	0,087000	0,093000	-	-	-	-	-	-					

## Worksheet OccurrenceL3

This worksheet contains the concentration data (occurrences) in FoodEx1 level 3.

**Note that ImproRisk needs the occurrence to be measured in mg/ Kg of food**

Figure 11 Worksheet OccurrenceL3

Chemical Substance		Pb		NOTE 5: This worksheet contains the concentration data (occurrences) in FoodEx1 at Level 3. Note(!) that ImproRisk needs the occurrences to be expressed in mg/ Kg of food												
Substance Category		Contaminant														
Reference value (µg/Kg b.w.)		0,63														
Type of Reference value		Benchmark Dose Level (BMDL)														
Type		DAILY														
				mg/kg food									P95			
				min			mean			median						
				No of Samples	LB	MB	UB	L3_Occur_Mean_LB	L3_Occur_Mean_MB	L3_Occur_Mean_UB	LB	MB	UB	LB	MB	UB
Level 1	Level 2	Level 3														
Alcoholic beverages	Alcoholic mixed drinks	Alcoholic mixed drinks						0,0235	0,0235	0,0235						
Milk and dairy products	Cheese	Cheese, processed, with ham						0,0453	0,0453	0,0453						
Herbs, spices and condiments	Seasoning or extracts	Meat extract						0,058	0,058	0,058						
Non-alcoholic beverages (excepting milk)	Cocoa beverage	Hot chocolate						0,0453	0,0453	0,0453						

### IMPORTANT NOTE:

In the case where the values need to be updated, this can be done with a template that is provided (Occurrence Template - Level 2 & Level 3.xlsm). Go to UPDATE DATA worksheet and click **Select file to update "Occurrence Data"**. ImproRisk will ask the user to locate and select the file in the local disk and the replacement of the values will be performed automatically.

Note that the template name is not of any importance. Name the occurrence workbook appropriately for easier reference. What is of importance, it is the name of the worksheet that needs to "Sheet1".

See instructions in the Chapter Updating the Occurrence and Consumption.

The two Occurrence worksheets are protected (i.e. no changes can be made), however, there is the option of un-protecting the worksheets (Go to Review->Unprotect Sheet in the Protect section) and making changes in the values. After modifying any cell in these two worksheets, ImproRisk detects the changes and asks the user for permission to recalculate. Nevertheless, it is highly recommended to always use the occurrence template to make any changes in the occurrence values and then loading it to ImproRisk.

## Worksheet MeanConsumption

In this worksheet, calculations are performed in order to estimate the average chronic consumption (Figure 12).

### **TABLES in the worksheet:**

**TABLE 1:** First, the average consumption (throughout the survey period) **for each individual** and for each food category in Level 1, is calculated. The *SubjectID* (individuals) is on the vertical axis and the 20 food categories (Level 1) on the horizontal axis. Note that this calculation considers the subject's own number of consumption days (e.g. 2 or 3) which may differ from subject to subject. The average value for each subject is then considered when calculating the average consumption<sup>1</sup>.

**TABLE 2:** The average consumption for:

- Consumers only, is calculated by taking the average of those who consumed the food.
- Population based, is calculated by taking the average of all individuals irrespective of their zero consumption.

Methodology notes:

A. There is a column "Number of Consumers in the Population" in Table 2, which shows the projected number of consumers in the population. This number is simply the summation of the weight coefficients of the sample participants that consumed each food at Level 1. The formula used is:

$$\sum_{i=1}^k w_i,$$

where  $k$  is the number of subjects who consumed the food, and  $w_i$  is the respective weight coefficient of the  $i_{th}$  consumer.

B. The "Consumer based (gr)" calculations (next column) uses the following formula:

$$\frac{\sum_{i=1}^k y_i w_i}{\sum_{i=1}^k w_i},$$

where  $k$  is the number of subjects who consumed the food,  $y_i$  is the mean consumption of the  $i_{th}$  consumer, and  $w_i$  is the respective weight coefficient of the  $i_{th}$  consumer.

Essentially, the formula calculates the total weighted (i.e. multiplied by the weight coefficient) mean daily food consumption for each subject, and divides by the number of consumers in the population.

---

<sup>1</sup> EFSA 2011;9(3):2097

C. The "Population Based (gr)" divides over the whole population (i.e. Sum Of Weights). The formula used is:

$$\frac{\sum_{i=1}^k y_i w_i}{\sum_{i=1}^n w_i}$$

where k is the number of subjects that consumed the food,  $y_i$  is the mean consumption of the  $i_{th}$  consumer,  $w_i$  is the respective weight coefficient of the  $i_{th}$  consumer, and n is the total number of participants in the food survey.

Figure 12 Worksheet MeanConsumption - Calculation of average consumption. Consumer based and Population based

TABLE 1: Mean Daily consumption for each subject in the food survey. By FoodEx Level 1 category																				TABLE 1: Table 1 contains the mean consumption (throughout the survey period) for each individual and for each FoodEx category at Level 1	TABLE 2: Mean daily consumption for each FoodEx L1 category, over the whole N-day period over the total sample				
Subject	Products for special nutritional use	Drinking water (water without any additives)	Alcoholic beverages	Eggs and egg products	Food for infants and young children	Animal and vegetable fats and oils	Snacks, desserts, and other foods	Non-alcoholic beverages (excluding milk based beverages)	Fruit and fruit products	Sugar and confectionery	Herbs, spices and vegetable oils	Fruit and vegetable products	Legumes, nuts and seeds	Starchy roots and tubers	Milk and dairy products	Meat and meat products	Vegetables and other leafy products	Composites and other products	Fish and other seafood	Grains and grain-based products	FoodEx Level 1 Category	Number of Consumers in the sample	Number of Consumers in the Population	Consumer based (gr)	Population based (gr)
70000	0.0	778.0	1030.0	0.0	0.0	39.0	0.0	215.0	0.0	40.0	0.1	128.0	0.0	138.0	480.0	166.5	88.0	44.5	105.5	192.0	Products for special nutritional use	5	10 031	60.4	1.1
70001	0.0	272.0	0.0	0.0	0.0	18.5	0.0	335.0	0.0	60.0	19.6	67.0	21.5	235.5	235.5	195.0	58.0	0.0	0.0	165.0	Drinking water (water without any additives)	272	551 204	1266.3	1266.3
70012	0.0	800.0	0.0	0.0	0.0	53.0	0.0	0.0	0.0	18.0	302.0	82.3	262.0	30.0	245.0	172.0	0.0	0.0	0.0	45.0	Alcoholic beverages	103	207 325	234.7	88.3
70013	0.0	1425.0	20.0	0.0	0.0	51.5	0.0	590.0	132.5	0.0	7.1	24.0	0.0	172.5	202.5	24.0	196.5	0.0	60.0	104.5	Eggs and egg products	129	260 925	24.5	11.5
70015	0.0	1450.0	0.0	70.0	0.0	4.0	0.0	175.0	3.0	8.0	10.5	41.5	0.0	125.0	209.5	73.0	0.0	0.0	100.0	0.0	Food for infants and small children	2	4 052	62.7	0.5
70019	22.5	750.0	40.0	9.0	0.0	28.0	32.5	495.0	0.0	22.5	16.3	185.0	0.0	188.5	94.0	177.5	43.1	5.5	0.0	127.0	Animal and vegetable fats and oils	265	536 857	26.7	25.0
70020	0.0	758.0	0.0	0.0	0.0	39.0	0.0	0.0	51.5	0.0	8.6	23.0	38.0	109.5	62.5	174.0	222.0	0.0	60.0	148.0	Snacks, desserts, and other foods	62	125 876	60.6	13.8
70034	0.0	1005.0	0.0	0.0	0.0	34.0	60.0	75.0	0.0	22.5	6.5	2.5	0.0	0.0	385.0	7.0	121.5	0.0	75.0	185.5	Non-alcoholic beverages (excluding milk)	209	422 873	525.4	249.7
70035	0.0	1625.0	465.0	0.0	0.0	33.0	0.0	0.0	400.0	0.0	28.0	10.0	0.0	149.5	222.5	280.0	469.0	7.0	148.0	52.5	Fruit and fruit products	210	425 856	142.8	110.3
70036	0.0	1377.0	0.0	0.0	0.0	60.0	0.0	120.0	0.0	15.0	0.0	4.0	0.0	253.5	290.0	281.0	407.0	0.0	148.0	0.0	Sugar and confectionery	145	293 869	18.8	9.0
70037	0.0	1575.0	0.0	0.0	0.0	29.0	0.0	0.0	205.0	0.0	1.0	7.5	0.0	68.0	150.0	281.5	885.5	0.0	160.0	0.0	Herbs, spices and condiments	241	488 365	18.2	14.3
70038	0.0	1425.0	150.0	0.0	0.0	53.0	0.0	165.0	410.0	5.0	0.0	0.0	47.5	215.5	120.0	151.0	392.0	0.0	0.0	151.5	Fruit and vegetable juices	181	366 409	76.2	50.6
70039	0.0	1000.0	750.0	24.0	0.0	78.5	0.0	250.0	3.0	6.0	34.1	14.0	57.0	0.0	151.5	126.0	242.0	0.0	0.0	249.5	Legumes, nuts and oilseeds	134	271 418	37.4	18.4
70054	0.0	1500.0	0.0	0.0	0.0	26.5	30.0	165.0	0.0	5.0	0.0	155.0	0.0	293.0	310.0	318.5	33.0	0.0	0.0	225.0	Starchy roots and tubers	186	376 440	131.8	90.0
70056	0.0	2325.0	165.0	0.0	0.0	7.5	0.0	0.0	375.0	0.0	2.5	7.5	0.0	0.0	132.5	199.5	460.0	0.0	0.0	151.5	Milk and dairy products	266	271 418	37.4	18.4
70057	0.0	1724.5	165.0	0.0	0.0	62.5	0.0	0.0	310.0	0.0	0.5	105.5	53.0	112.0	182.5	249.5	607.0	0.0	0.0	247.0	Meat and meat products (including edible)	253	512 479	175.2	162.9
70058	0.0	1770.0	0.0	70.0	0.0	33.0	0.0	465.0	0.0	7.5	50.0	45.0	47.5	172.5	60.0	149.5	282.0	0.0	0.0	115.0	Vegetables and vegetable products (incl)	271	549 211	211.6	210.8
70062	0.0	1117.0	0.0	0.0	0.0	47.0	0.0	0.0	127.5	0.0	0.0	2.5	0.0	368.5	320.0	136.5	289.5	0.0	0.0	410.0	Composite food (including frozen product)	65	131 525	32.1	7.7
70063	0.0	1375.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.5	300.0	0.0	0.0	289.5	25.0	52.5	0.0	0.0	345.5	Fish and other seafood (including amphip)	94	190 048	97.8	33.7
70064	0.0	1953.0	0.0	10.5	0.0	21.5	0.0	170.0	225.0	0.0	0.6	2.5	0.0	78.5	281.0	119.5	154.0	0.0	0.0	175.5	Grains and grain-based products	269	545 159	170.1	168.3
70065	0.0	1212.0	10.5	0.0	0.0	32.5	0.0	150.0	610.0	5.0	6.6	0.5	0.0	93.5	60.0	341.0	223.5	31.5	0.0	445.0					
70071	0.0	986.0	11.0	5.0	0.0	33.0	0.0	0.0	450.0	0.0	2.7	67.0	0.0	0.0	389.0	344.0	313.5	0.0	0.0	37.5	87.0				
70090	0.0	787.5	302.5	6.5	0.0	39.0	0.0	166.5	105.0	25.0	1.1	0.0	0.0	87.5	103.0	236.5	223.0	90.0	0.0	66.0					
70093	0.0	1502.0	0.0	7.0	0.0	47.0	0.0	22.5	420.0	0.0	3.0	0.0	0.0	0.0	431.0	176.0	530.0	0.0	0.0	159.5					

## Worksheet Contribution

The worksheet Contribution presents the calculated aggregated exposure by each FoodEx1 level 1 category, FoodEx1 level 2 category and FoodEx1 level 3. The aggregation is the summation of the exposure for each food consumption occasion.

Eventually, the percent contribution (%) (for LB, MB and UB scenario) is calculated of each food category to the total exposure of the entire sample.

### Level 1 contribution

**Table 3:** Total Exposure ( $\mu\text{g}/\text{Kg}$  of body weight) over the total population broken down by FoodEx1 L1 name.

It is noticeable that the exposure values are significantly high since the weighting coefficients ("WCOEFF") values are included in the formulas. The exposure calculated here, is the estimated exposure projected for the entire population.

**Table 4:** Contribution of each food category (Level 1) to the Total Exposure ( $\mu\text{g}/\text{Kg}$  of body weight) over the total population.

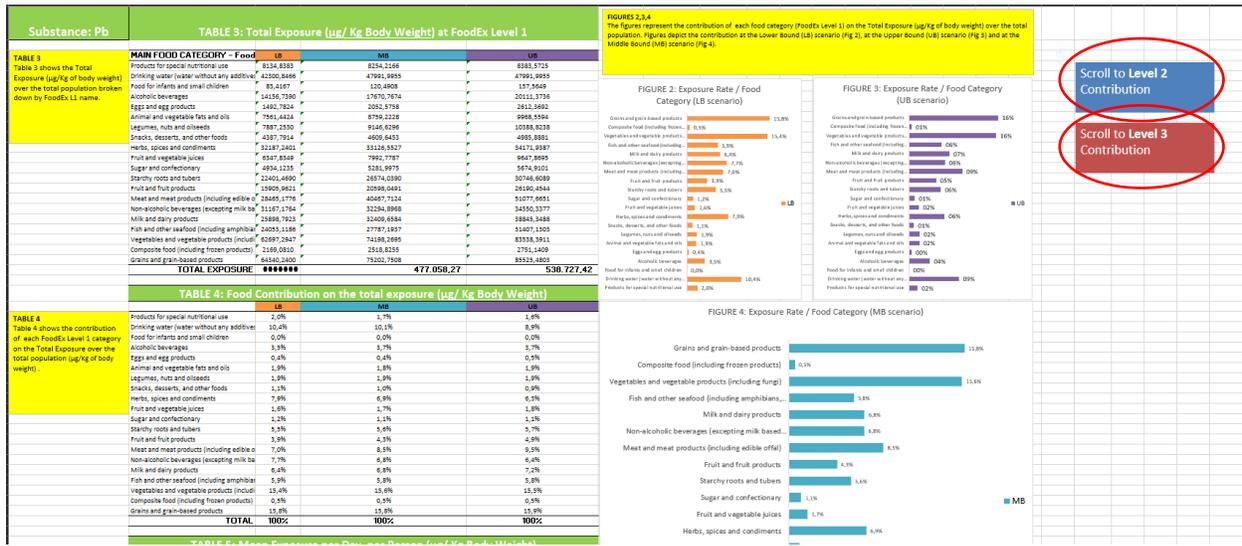
**Table 5:** Mean exposure per Day per Person by each food category (Level 1).

This is the values of Table 1 divided by the Total Consumption Days(i.e. the total number of days across all consumers). The Total Consumption Days is projected to the total population using the weighting factors. [See Table 5 Note in ImproRisk on how this calculation is performed].

The formula used is  $\text{TotalExposure}/\text{TotalConsumptionDays}$

*One can use the scroll buttons at the right (see Figure 13) to scroll to the other sections of the worksheet for level 2 and level 3.*

Figure 13 Contribution at FoodEx1 level 1



Tables can then be **sorted** in the following manner (Figure 14): e.g. for Table 3 of ImproRisk:

1. Select the whole table including the titles (FoodEx1 Level 1, LB, MB, UB)
2. On the excel menu bar ribbon (horizontal menu) go to **Data** and then click on the **Sort** buttons
3. Make sure the "My data has headers" is enabled (clicked)
4. Select the **Sort by** title [Food name or just the LB, MB or UB values]
5. Press OK
6. Press **Shift+F9** to recalculate the worksheet.

FIGURE 4 in ImproRisk shows the contribution of each food FoodEx1 L2 category to the Total Exposure (µg/Kg of body weight) over the total population, at the MB scenario. Users of the Excel 2013 or later versions can filter the visualisation by LB and UB scenarios as well (Figure 15).

Figure 14: Contribution at FoodEx1 level 1 - Sort the contribution by Food name or Scenario value

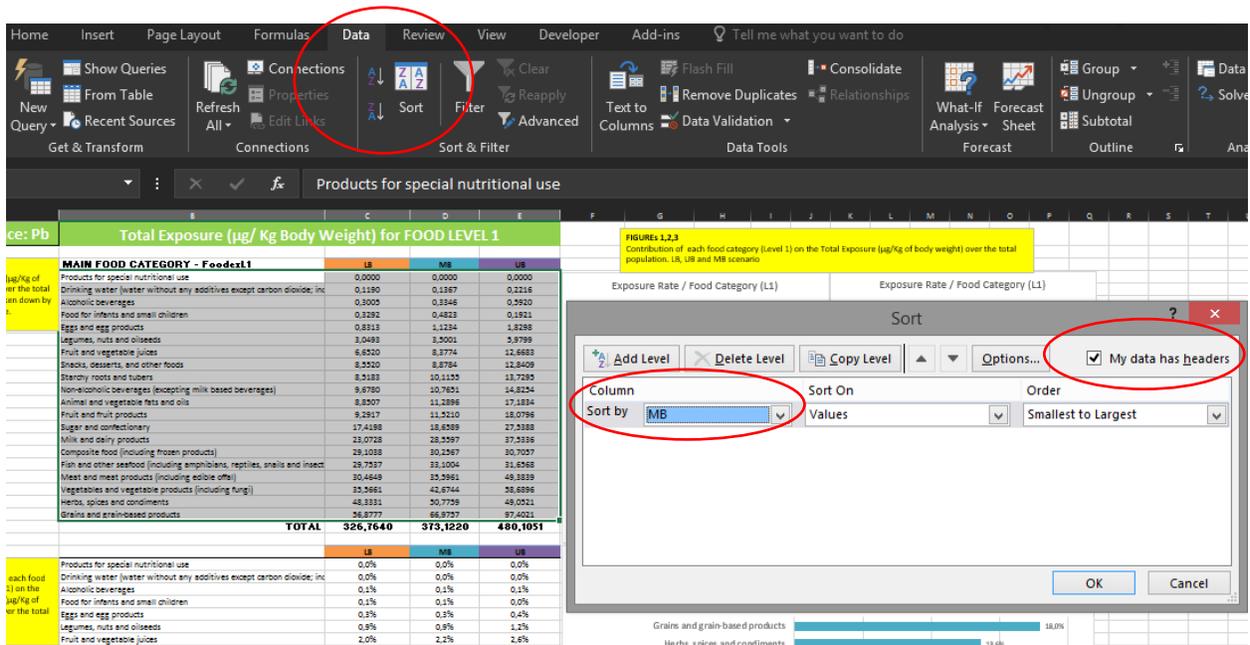
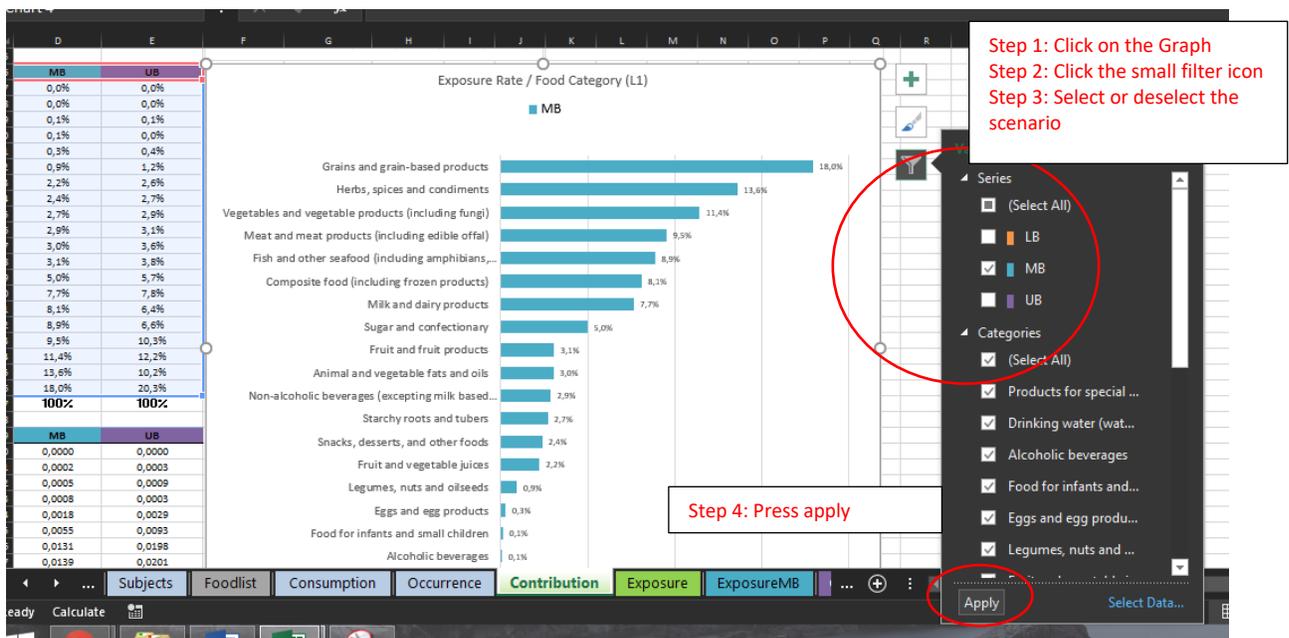


Figure 15: How to view the food contribution in LB, MB and UB scenario





## Level 3 contribution

A pivot table where the user can query the exposure across level1, level2 and get estimates of the contribution at level 3 to:

### a) Total Exposure

The contribution of the level 3 food item to the total exposure of the selected level 1 food groups. If no level 1 food group is selected, then this is the total exposure over all level 1 food groups combined.

### b) Within Level 2

Contribution within the parent level 2 exposure

Figure 17 Contribution at FoodEx1 level 3- A pivot table

**Substance: Pb**      **TABLE 6a: Total Exposure and Contribution to the Total Exposure (µg/ Kg Body Weight) at FoodEx Level 3**      Scroll to TOP

**Table 6a.**  
Table 6a shows the contribution of Level 3 food items at the MB scenario.  
• Column "Intake" depicts the total exposure for each food item.  
• Column "Contribution to the Total Exposure" depicts the contribution of each food item to the total exposure  
• Column "Contribution at Level2" depicts the contribution within the parent Level 2 exposure

Consumed food at Level 1 (Multiple Items)

Consumed food at Level 2	LEVEL_3	Intake	Contribution to the Total Exposure	Contribution within Level 2
Cocoa beverage		614,1148388	1,52%	100,00%
	Hot chocolate	614,1148388	1,52%	100,00%
Coffee (Beverage)		8093,83925	20,09%	100,00%
	-Other-Coffee (Beverage)	8093,83925	20,09%	100,00%
Fruit juice		5849,946788	14,52%	100,00%
	-Other-Fruit juice	5849,946788	14,52%	100,00%
Soft drinks		15788,08621	39,19%	100,00%
	-Other-Soft drinks	15788,08621	39,19%	100,00%
Tea (Infusion)		7798,856534	19,36%	100,00%
	-Other-Tea (Infusion)	7798,856534	19,36%	100,00%
Mixed fruit juice		206,3966359	0,51%	100,00%
	-Other-Mixed fruit juice	206,3966359	0,51%	100,00%
Vegetable juice		1017,394072	2,53%	100,00%
	-Other-Vegetable juice	1017,394072	2,53%	100,00%
Concentrated fruit juice		494,3631202	1,23%	100,00%
	-Other-Concentrated fruit juice	494,3631202	1,23%	100,00%
Fruit nectar		206,2331081	0,51%	100,00%
	-Other-Fruit nectar	206,2331081	0,51%	100,00%
Fruit and vegetable juices		218,4449459	0,54%	100,00%
	-Other-Fruit and vegetable juices	218,4449459	0,54%	100,00%
<b>Grand Total</b>		<b>40287,6755</b>	<b>100,00%</b>	

Consumed food at Level 1

- Food for infants and small children
- Fruit and fruit products
- Fruit and vegetable juices**
- Grains and grain-based products
- Herbs, spices and condiments
- Legumes, nuts and oilseeds
- Meat and meat products (including edi...
- Milk and dairy products
- Non-alcoholic beverages (excepting mil...**
- Products for special nutritional use
- Snacks, desserts, and other foods
- Starchy roots and tubers

The PivotTable is autorefreshed after each model run. However, should you need to refresh the graphs Go to Review->UnprotectSheet and then click on the table and go to Analyze (or PivotTable Analyze)->Refresh

Table 6a in the worksheet "Contribution" calculates the contribution of each level-3 food item to the total intake of the population (Column D) and the contribution within Level 2 (Column E). The contribution is currently calculated only at the MB occurrence scenario.

In the example shown in (Figure 18), we can see that in the level-2 category CHEESE, there is a total intake of 9.77 µg/Kg, which corresponds to the 5.78% of the total exposure. Edam cheese has a total of 6.22 µg/Kg that corresponds to

3.68% of the total population intake, and 63.2% of the total intake due to Cheese. Should you want to filter the total intake to a Level 1 category, you can simply use the excel slicer to the right of the table. When a food category at Level 1 is selected through the slicer, then all the contribution to the total intake correspond to the intake due to the Level 1 category.

Level 3 items that occurrence values were not provided, are all grouped into an "other" category. These food items use the aggregated occurrence value provided at the Level 2 in the worksheet "OccurrenceL2". For example, at the CHEESE level-2 category, the "-other-CHEESE" item (6.22 µg/Kg) encapsulates all the level-3 cheese items that occurrence values were not provided in the "OccurrenceL3" worksheet, but an aggregated occurrence of Cheese was provided in the worksheet "OccurrenceL2". This highlights the need for still providing (although not necessary for the model to function) aggregated occurrence values at Level 2 for the food items that are in the food survey, but no level-3 occurrence value is available.

Figure 18 Contribution at Level 3. Worksheet "Contribution". Contribution to the total exposure and within Level 2

Consumed food at Level 2	LEVEL_3	Intake	Contribution to the Total Exposure	Contribution within Level 2
Brassica vegetables	-Other-Brassica vegetables	1,115,611,93	0,66%	100,00%
Bulb vegetables	-Other-Bulb vegetables	0,623,536,89	0,37%	100,00%
Cheese	-Other-Cheese	9,779,607,89	5,78%	100,00%
Chocolate (Cocoa) products	-Other-Chocolate (Cocoa) products	4,408,118,45	2,60%	100,00%
Concentrated milk	-Other-Concentrated milk	0,326,078,62	0,19%	100,00%
Confectionery (non-chocolate)	-Other-Confectionery (non-chocolate)	0,971,204,54	0,57%	100,00%
Cream and cream products		0,410,956,44	0,24%	100,00%

## Worksheet Exposure

In this worksheet, the exposure assessment ( $\mu\text{g}/\text{Kg}$  body weight) is performed for each subject in the food survey.

### TABLE 7

The total exposure for each subject that participated in the food study, is calculated for:

a. The total survey duration – TOTAL EXPOSURE

The total exposure for each subject over the full survey period.

b. Daily or Weekly depending on the tolerable reference value type of the chemical.

The DAILY or WEEKLY exposure is calculated using the following formula (see Figure 19):

$$(TotalExposure/ CON\_DAYS) \times Exposurefactor$$

where:

*exposurefactor* is either 1 if DAILY reference value type, or 7 if WEEKLY reference value type. This value is set via the VBA code during updating of the Occurrence data.

*CON\_DAYS* = Number of survey days of each subject (captured by the VBA code when updating the Consumption data).

### TABLE 8

Summary statistics over the whole population adjusted using weight coefficients.

A. The weighted mean exposure

The weighted mean exposure is calculated using the following formula:

$$\text{Weighted mean exposure} = \frac{\sum_{i=1}^n y_i w_i}{\sum_{i=1}^n w_i}$$

where  $n$  is the number of subjects in the sample,  $y_i$  is the exposure for the  $i_{th}$  subject, and  $w_i$  is the respective weight coefficient for the  $i_{th}$  subject.

Essentially, instead of each data point (exposure of each subject in the sample) contributing equally to the final average, each individual contributes differently based on the weight coefficient. If all the weights are equal (e.g. 1), the weighted mean is the same as the unweighted mean.

B. The 95<sup>th</sup> percentile

### C. Percentage above the reference value

Similarly, a percentage of the total population that exceeds the reference value is calculated and the weighting factor is also considered.

Figure 19: Worksheet Exposure - Calculation of DAILY or WEEKLY exposure of each subject

Exposure at Individual Level										Chemical Substance: Pb		Contaminant		TABLE 8: Summary Statistics for Exposure of the overall population (µg/ Kg Body Weight).	TABLE 8: Overall Weighted Exposure over the Population (WCOEFF adjusted)			
Total µg/Kg of Pb for each individual										Reference value (µg/Kg b.w.)		Benchmark Dose Level (BMDL)			LB	MB	UB	
Number of participants: 272		Sum of Weight Coefficients: 551.204								Type		DAILY			Mean	95 <sup>th</sup> Percentile	Percentage ABOVE the Reference value	
TABLE 7: Individual Exposure										TOTAL EXPOSURE			DAILY EXPOSURE			Margin of Exposure (MOE)		
SUBJECTID	GENDE	AI	WEIGI	AREA	POP_CLASS	WCOEF	CON_DA	YS	NDaysSubExp_M	can_LB	can_MB	can_UB	SubExp_L	SubExp_M	SubExp_U	LB	MB	UB
70000	MALE	24.3	76.7	Lemesos	Adults	1993	2	0.7846	0.991815	1.1568	0.3923	0.4959	0.5784					
70001	MALE	32.8	60.1	Lemesos	Adults	1993	2	0.6873	0.808383	0.9328	0.3436	0.4042	0.4664					
70012	MALE	41.3	64.7	Pafos	Adults	1993	2	0.7918	0.929923	1.0395	0.3959	0.4650	0.5197					
70013	FEMALE	44.5	72	Pafos	Adults	2059	2	0.6316	0.738038	0.8476	0.3158	0.3690	0.4238					
70015	FEMALE	21.5	82	Pafos	Adults	2059	2	0.3533	0.413451	0.4715	0.1767	0.2067	0.2358					
70019	MALE	23.9	56	Pafos	Adults	1993	2	1.3200	1.468318	1.6111	0.6600	0.7342	0.8055					
70020	FEMALE	63.4	92.8	Pafos	Adults	2059	2	0.3726	0.461144	0.5415	0.1863	0.2306	0.2708					
70034	FEMALE	49.1	57	Lemesos	Adults	2059	2	0.6239	0.736158	0.8161	0.3120	0.3681	0.4080					
70035	MALE	62.9	89.6	Lemesos	Adults	1993	2	0.9047	1.108694	1.2683	0.4523	0.5543	0.6341					
70036	MALE	31.4	94.6	Lemesos	Adults	1993	2	0.4833	0.610148	0.7051	0.2417	0.3051	0.3526					
70037	FEMALE	44.8	138	Pafos	Adults	2059	2	0.3584	0.462587	0.5508	0.1792	0.2313	0.2754					
70038	MALE	64.6	91	Pafos	Adults	1993	2	0.7209	0.884681	1.0170	0.3605	0.4429	0.5085					
70039	MALE	24.6	99	Pafos	Adults	1993	2	0.6102	0.732389	0.8412	0.3051	0.3662	0.4206					
70054	FEMALE	20.8	49.5	Lemesos	Adults	2059	2	0.8711	1.042081	1.2109	0.4356	0.5210	0.6054					
70056	FEMALE	34.1	67	Pafos	Adults	2059	2	0.6647	0.849582	0.9852	0.3323	0.4248	0.4926					
70057	MALE	60.6	109	Pafos	Adults	1993	2	0.6355	0.806404	0.9345	0.3178	0.4032	0.4672					
70058	MALE	32.5	67	Pafos	Adults	1993	2	1.1537	1.360925	1.5017	0.5769	0.6805	0.7509					
70062	FEMALE	64.8	75.5	Lemesos	Adults	2059	2	0.4758	0.592252	0.7001	0.2379	0.2961	0.3501					
70063	FEMALE	21	57	Lemesos	Adults	2059	2	0.6499	0.780421	0.8822	0.3250	0.3902	0.4411					
70064	FEMALE	47.3	75.6	Lemesos	Adults	2059	2	0.5020	0.595664	0.6894	0.2510	0.2978	0.3447					
70065	MALE	58.1	75	Lemesos	Adults	1993	2	0.9771	1.179749	1.3557	0.4885	0.5899	0.6779					
70071	MALE	53	78.8	Lemesos	Adults	1993	2	0.5757	0.734798	0.8682	0.2879	0.3674	0.4341					
70090	MALE	64.4	90.7	Lemesos	Adults	1993	2	0.5159	0.643258	0.7342	0.2580	0.3216	0.3671					
70093	MALE	63.9	94.7	Lemesos	Adults	1993	2	0.6171	0.767920	0.8652	0.3086	0.3840	0.4326					
70094	FEMALE	51.5	48.9	Lemesos	Adults	2059	2	4.3243	4.442601	4.5593	2.1621	2.2213	2.2797					
70095	FEMALE	64.1	64.9	Lemesos	Adults	2059	2	0.6031	0.734237	0.8267	0.3015	0.3671	0.4134					
70096	FEMALE	63.4	76.7	Lemesos	Adults	2059	2	0.3724	0.444690	0.5074	0.1862	0.2323	0.2537					
70099	FEMALE	24.9	92.2	Lemesos	Adults	2059	2	0.8318	0.957996	1.0671	0.4159	0.4790	0.5335					

**TABLE 7:**  
Table 7 contains the aggregated exposure for each individual in the food survey.

- The TOTAL EXPOSURE is the total (summation) of the exposure over the total number of the consumption days recorded in the food survey for each subject
- The DAILY or WEEKLY exposure is calculated using the following formula:  
(TOTAL EXPOSURE/ CON\_DAYS)\*exposurefactor

where :  
exposurefactor is either 1 or 7 if the type of the reference value is set as DAILY or WEEKLY respectively, through the dropdown list in the Occurrence Template

CON\_DAYS = Number of Consumption Days during the food survey

## Worksheet ExposureMB-LB-UB

In this worksheet, a probability distribution of the exposure (histogram) and a cumulative distribution at the LM or MB or UB scenarios of exposure are created (Figure 20).

### Summary statistics

An overview of the MB scenario is as follows:

The summary statistics in Table 13 of the worksheet are all weighted.

The frequency weights (i.e. "repeat"-type weights - integers counting the number of occurrences for each observation) and the estimators used are considered unbiased<sup>2</sup>. Analogous calculations are performed in the worksheets "ExposureLB" and ExposureUB".

The 25<sup>th</sup> percentile (Q1), the median (Q2) and the 75<sup>th</sup> percentile (Q3) are all weighted.

The weighted standard deviation is calculated using the following procedure:

The weighted variance<sup>3</sup> is

$$s^2 = \frac{\sum_{i=1}^n w_i * (y_i - \mu)^2}{\sum_{i=1}^n w_i}$$

or...

$$s^2 = \frac{\sum_{i=1}^n (w_i * y_i^2) - \mu^2 * \sum_{i=1}^n w_i}{\sum_{i=1}^n w_i - 1}$$

where n is the number of subjects in the sample,  $y_i$  is the exposure for the  $i_{th}$  subject,  $w_i$  is the respective weight coefficient for the  $i_{th}$  subject, and  $\mu$  is the weighted mean exposure.

**therefore, the weighted standard deviation is**

$$s = \sqrt{\frac{\sum_{i=1}^n (w_i * y_i^2) - \mu^2 * \sum_{i=1}^n w_i}{\sum_{i=1}^n w_i - 1}}$$

---

<sup>2</sup> [https://en.wikipedia.org/wiki/Weighted\\_arithmetic\\_mean#Frequency\\_weights](https://en.wikipedia.org/wiki/Weighted_arithmetic_mean#Frequency_weights)

<sup>3</sup> [https://en.wikipedia.org/wiki/Weighted\\_arithmetic\\_mean#Frequency\\_weights](https://en.wikipedia.org/wiki/Weighted_arithmetic_mean#Frequency_weights),  
<https://stats.stackexchange.com/questions/47325/bias-correction-in-weighted-variance>



## Frequency distribution table

TABLE 12 of the worksheet is the frequency Distribution Table at the MB Scenario

Width of the bin is set via the Friedman-Diaconis (F-D)<sup>5</sup> rule. The F-D rule works well in practice (Scott, D. (1992))

The bin width is set to:

$$Factor * IQR * N^{-3}$$

This is a modification of the Friedman-Diaconis rule:

IQR=Interquartile range

N=Sample Size

Factor = increments of 1 starting from 1.

The original F-D rule is with a factor of 2.

Using the grey Spin Button (see Figure 20) within the chart enables the user to modify the Friedman-Diaconis rule by increasing or decreasing the factor (with a step of 1). This will change the height of the bin and get a better visualisation of the histogram.

Then, the number of Bins is set to:

$$(Max-Min)/binswidth$$

The Max and Min are the Maximum and minimum observed value of the exposures across the population, respectively.

---

<sup>5</sup> [https://en.wikipedia.org/wiki/Freedman%E2%80%93Diaconis\\_rule](https://en.wikipedia.org/wiki/Freedman%E2%80%93Diaconis_rule)

Figure 21: Worksheet ExposureMB - Example of a cumulative distribution of exposure

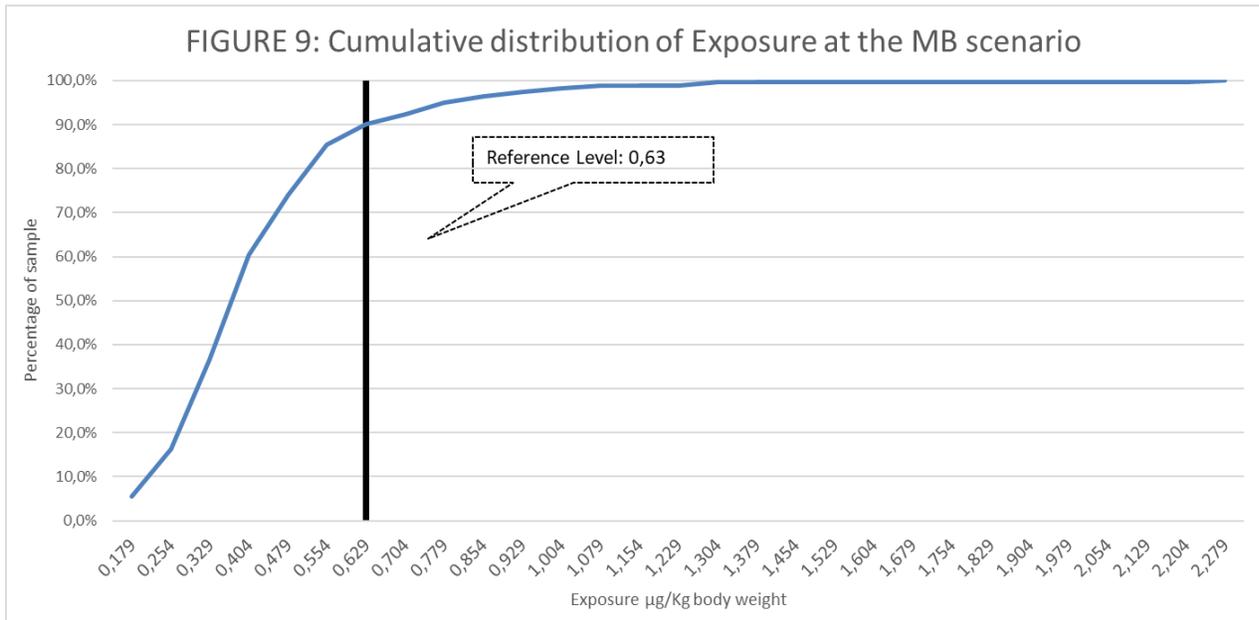
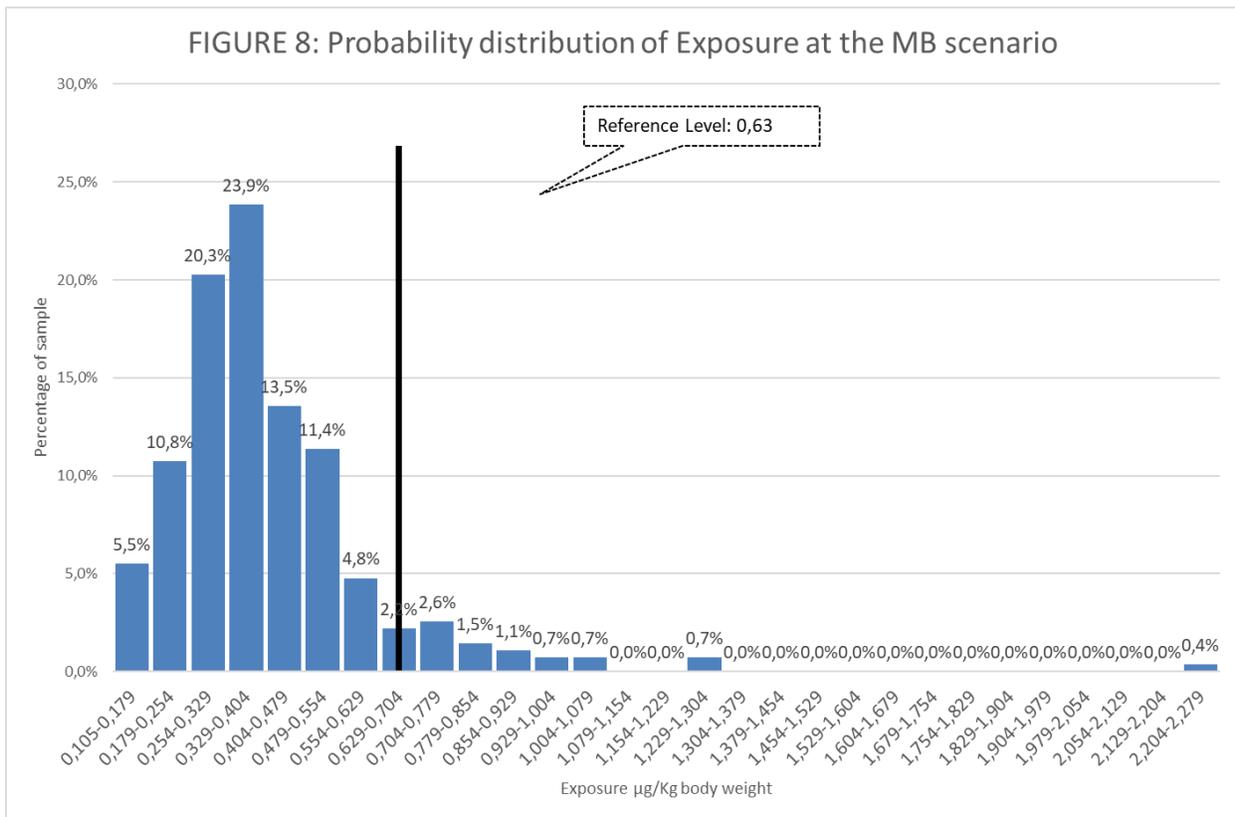


Figure 22: Worksheet ExposureMB - Example of a probability distribution of exposure (histogram)



# Worksheet GENDER

The bin width is the same as the width already set in the Exposure MB (Figure 14).

The statistics needed for the calculation of the height are taken those of the MB scenario i.e. IQR, Sample Size, Max and Min exposure.

The grey Spin Button within the chart modifies the factor of the modified Friedman Diaconis rule by increasing or decreasing it (with a step of 1) as with the ExposureMB worksheet.

Figure 23: Worksheet GENDER

The screenshot displays the 'Worksheet GENDER' interface with the following components:

- Chemical Reference Value:** Pb, Contaminant, Benchmark Dose Level (BMDL), Type: DAILY.
- TABLE 18:** Frequency Distribution Table of MALES and FEMALES at the MB Scenario. Columns include Bin, C.M.F., N, C.M., % C.M., N, C.F., % C.F.
- TABLE 19:** Summary statistics for the Exposure at the MB scenario at the POPULATION level. Columns include MALE and FEMALE for n (population), n (sample), Mean, Mean Exposure, St. Deviation of Exposure, St. Error.
- TABLE 20:** Histogram parameters including Bin width, FreqBin, Factor, and a note about bin width calculation.
- TABLE 21:** Comparison between FEMALES and MALES (t-test) showing Cohen's D and Cohen's D guidelines.
- TABLE 22:** TOTAL EXPOSURE vs REFERENCE VALUE, showing Percentage of Population ABOVE.
- Figure 12:** Cumulative Distribution of Exposure at the MB scenario [By GENDER]. A line graph showing cumulative percentage of population vs exposure (µg/kg b.w.) for MALES and FEMALES. A vertical line indicates the Reference Level at 0.63.
- Figure 13:** Probability Density Distribution of Exposure at the MB scenario [By GENDER]. A histogram showing the percentage of population vs exposure (µg/kg b.w.) for MALES and FEMALES. A vertical line indicates the Reference Level at 0.63.

## Comparison between FEMALES and MALES

Comparison between exposure levels of Male and Female is performed using the Cohen's d statistic. (Cohen, J. (1977)) is a measure of quantification of the difference (effect size).

Cohen's D guidelines:

D ~ 0,20 -> small difference

D ~ 0,50 -> moderate difference

D > 0,80 -> large difference

For example, if D is more than 0.80 then the observed difference is considered to be large.

The model produces comparative graphs for Male and Female such as probability and cumulative distribution of exposure (Figure 24, Figure 25)

Figure 24: Cumulative distribution of the exposure by Gender

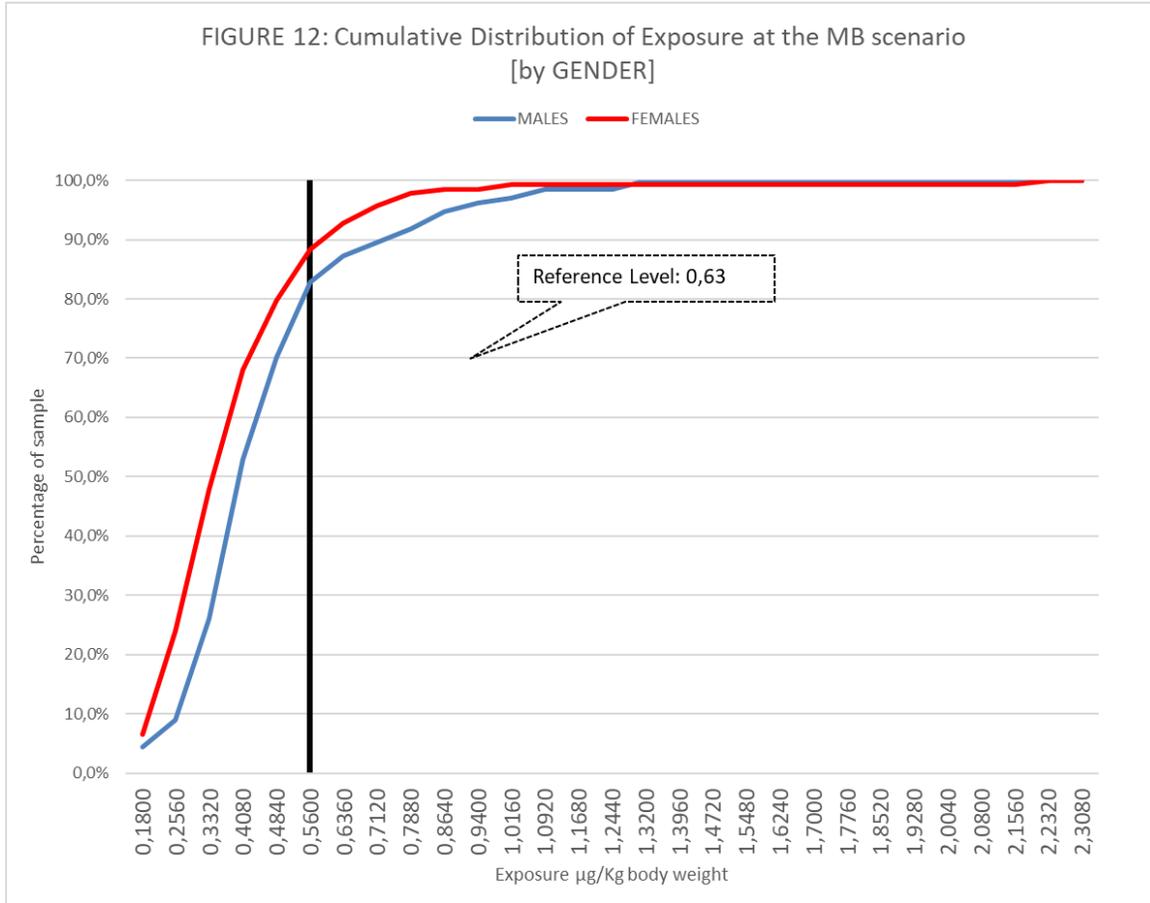
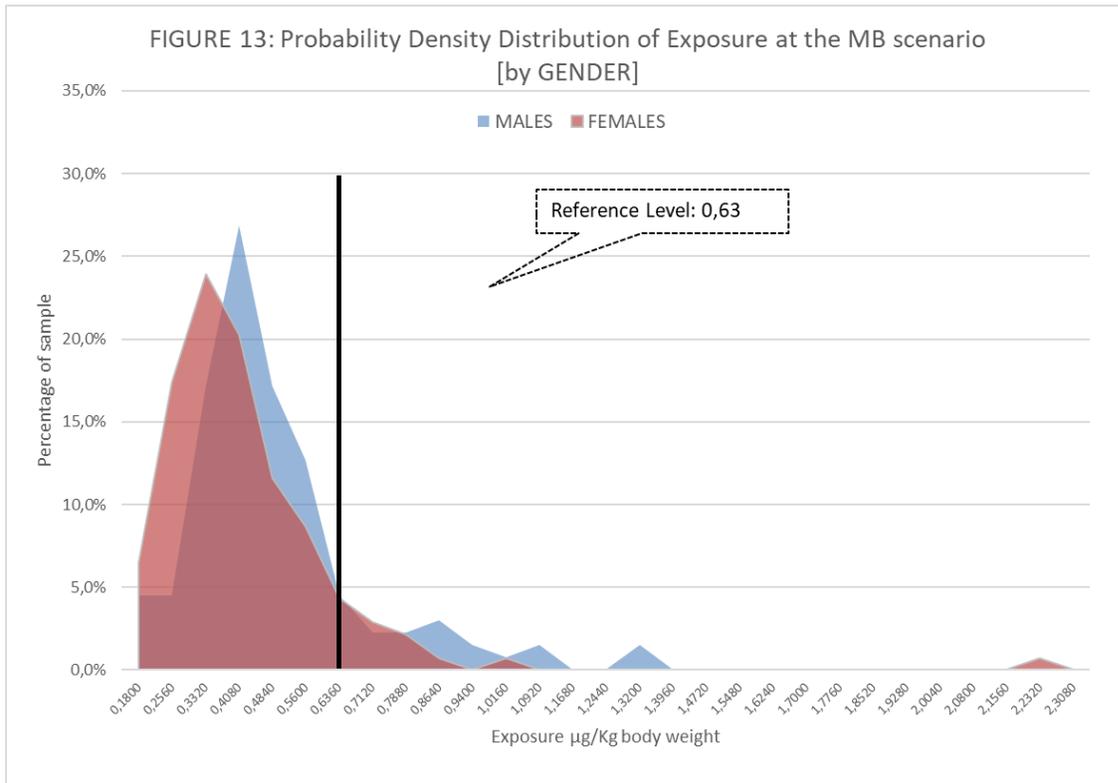


Figure 25 Probability distribution of the exposure by Gender



## Worksheet AREA

ImproRisk is using built in formulas to identify the number and name of areas that are included in the *Subject* worksheet and creates a table of aggregated statistics (Figure 27). ImproRisk can handle up to 20 different Areas.

ImproRisk also produces comparative graphs for the mean exposure across area (Figure 28), and proportion of exceeding the tolerable reference value across the AREA. (Figure 29)

*Figure 26: Statistics of the mean Exposure in the MB scenario by AREA of the individual. 95% error is the error ( $\pm$ ) in the mean exposure. The error is 1.96 times the Standard Error (SE) of the estimated mean exposure*

TABLE 23: Summary statistics for the Exposure at the MB scenario at the population level						
Total N=272	Lemesos	Pafos	Lefkosia	Ammochos	Larnaka	
N	135841	73002	184267	36534	121560	
Min	0,193	0,132	0,105	0,139	0,134	
Max	2,2213	0,8436	1,0312	0,6419	0,6955	
Mean Exposure	0,4691	0,4268	0,4345	0,3257	0,3444	
St. Deviation of Exposure	0,2890	0,1574	0,2112	0,1420	0,1138	
St.Error	0,0353	0,0262	0,0221	0,0335	0,0147	
Lower Limit	0,3999	0,3753	0,3911	0,2601	0,3157	
Upper Limit	0,5383	0,4782	0,4780	0,3913	0,3732	
95% error	0,0692	0,0514	0,0434	0,0656	0,0288	

Figure 27: Mean Exposure in the MB scenario by AREA of the individual. The error lines on the bars represent the lower and upper limits for the 95% confidence interval of the true mean exposure.

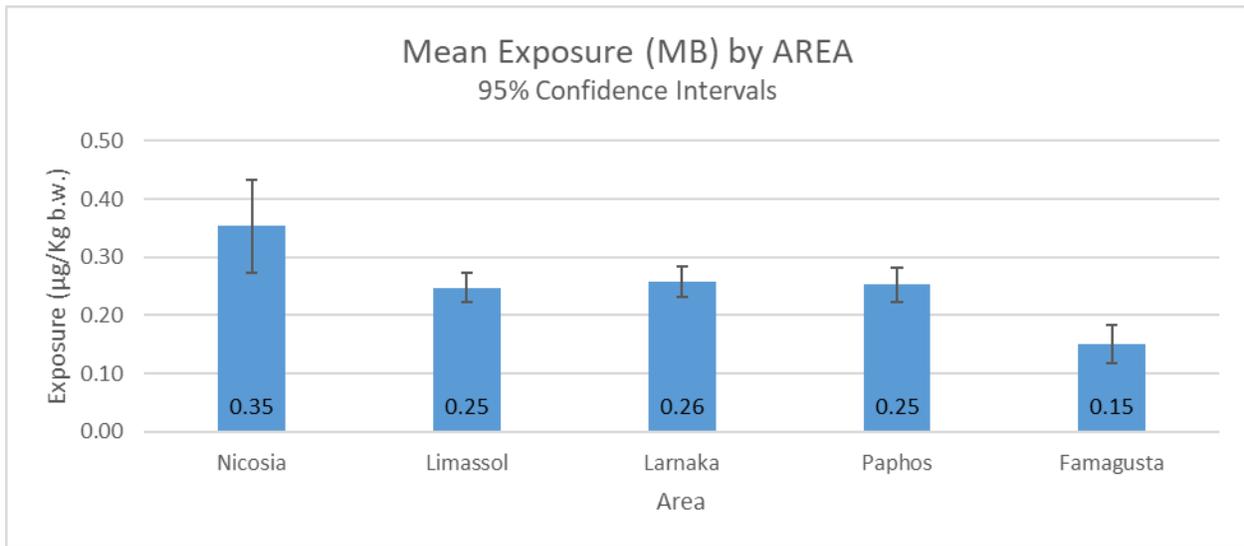
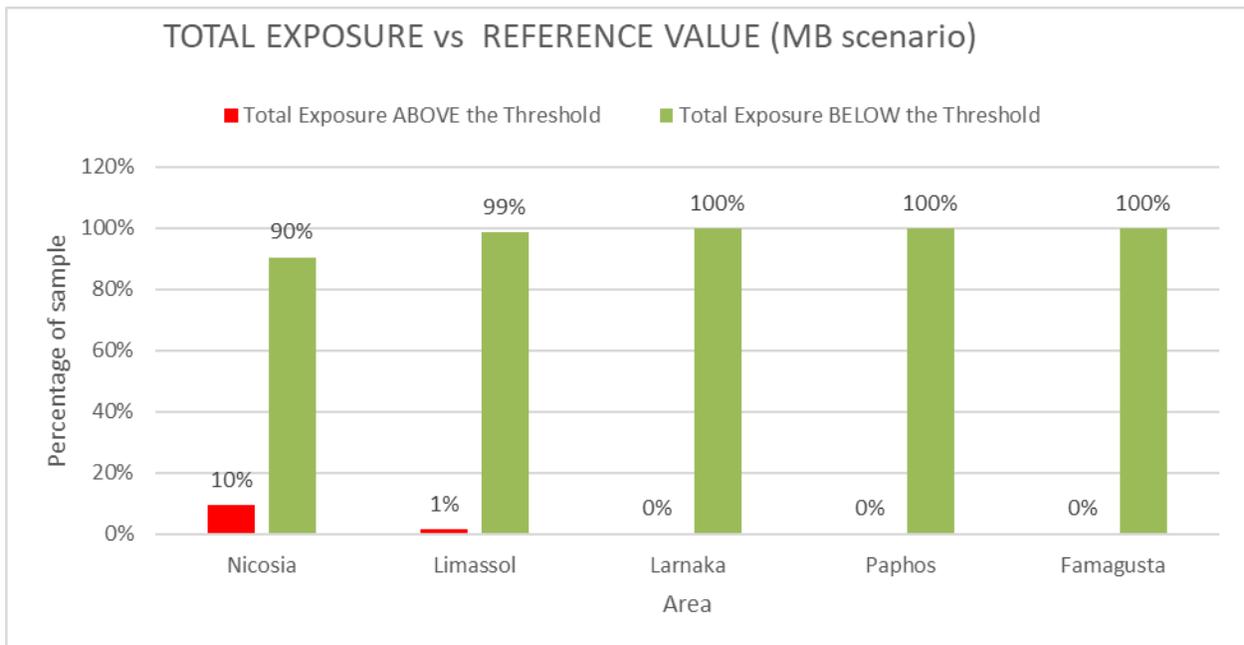


Figure 28: TOTAL EXPOSURE vs REFERENCE VALUE (MB scenario). The green bars represent the proportion of the sample that is below the tolerable reference value of exposure to the chemical. The red bars represent the proportion of the sample which exceeds the tolerable reference value.

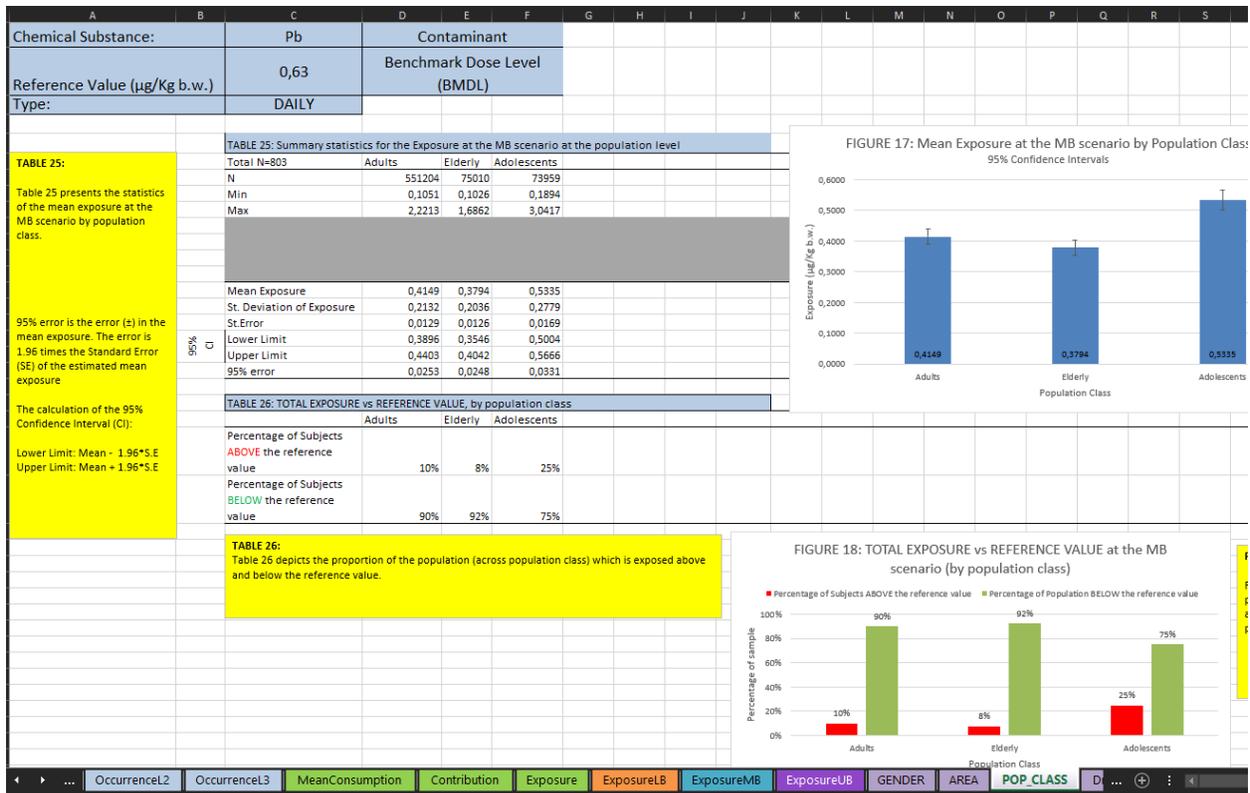


## Worksheet POP\_CLASS

ImproRisk is using built in formulas to identify the number and the name of Population Classes that are included in the **Subject** worksheet and creates a table of aggregated statistics (Figure 30). ImproRisk can handle up to 20 different Population Classes.

ImproRisk also produces comparative graphs for the mean exposure across each Population Class and the proportion of the sample that exceeds the reference value across a Population Class.

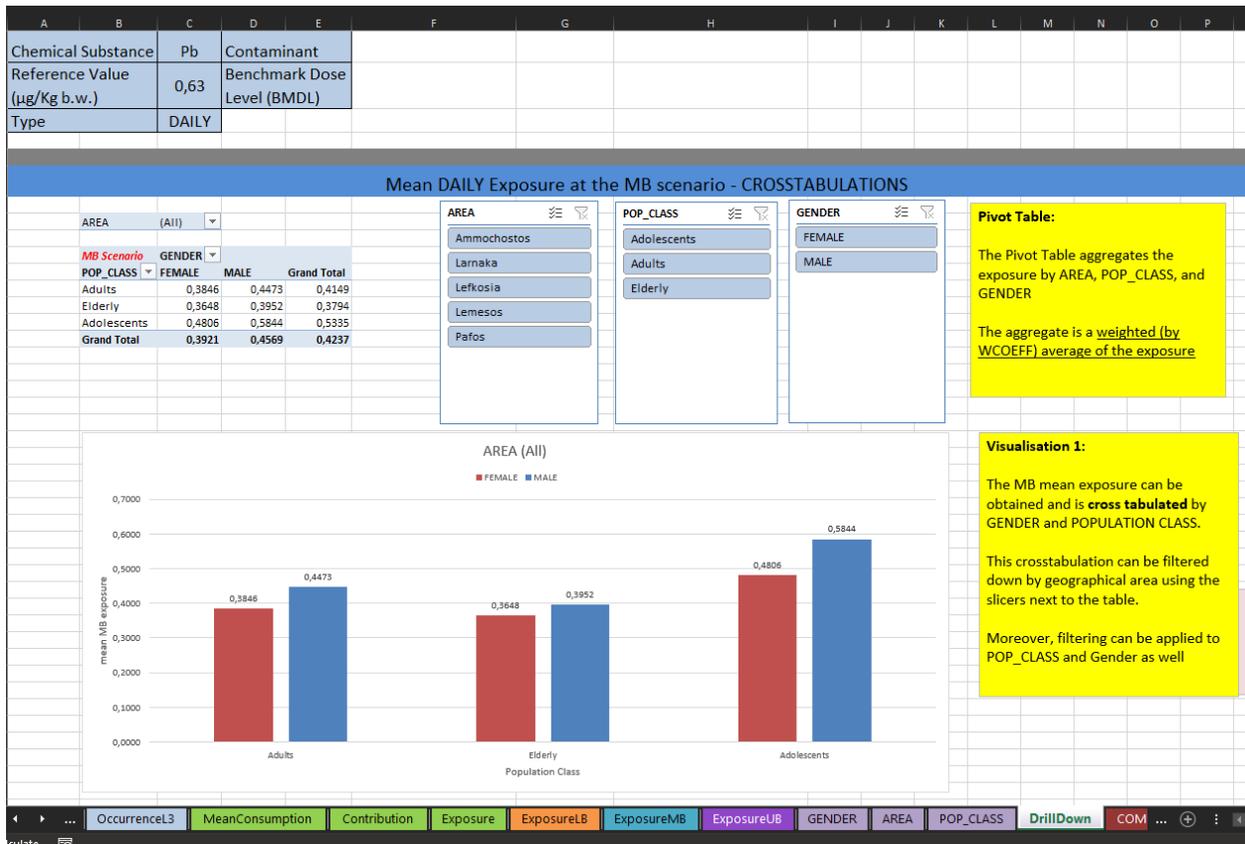
Figure 29: Worksheet POP\_CLASS



## Worksheet DrillDown

A pivot table connected with a pivot chart for subgroup analysis is provided. The user can slice the data and aggregate the exposure across levels of AREA, Population Class and Gender. Use the Slicers to select/deselect options and get the average exposure of subgroups

Figure 30 Worksheet DrillDown



## Worksheet **COMPREHENSIVE**

In this worksheet, one can calculate the mean exposure based on mean consumption data rather than consumption data at the individual level.

Mean consumption data can be retrieved (manually) via the EFSA data warehouse (Figure 32) where lies the EFSA Comprehensive European Food Consumption Database. The Comprehensive Food Consumption Database is a source of information on food consumption across the European Union (EU). It contains detailed data for a number of EU countries.

EFSA published the "Guidance in the use of the EFSA Comprehensive European Food Consumption Database in Exposure Assessment" found in <http://www.efsa.europa.eu/en/efsajournal/pub/2097>.

The user can retrieve the data via the EFSA data warehouse or alternatively can download the full datasets (in excel format .xlsx) from the Comprehensive data webpage<sup>6</sup>. The researcher can select from different datasets (chronic, acute, consumer based, etc) (see Figure 33). The excel files contain several worksheets each containing information down to specific FoodEx1 levels. Note that *ImproRisk* utilises the FoodEx1 level 2 categories.

### **How to use the worksheet**

The food categories down to FoodEx1 Level 2 are pre-installed (Columns A,B) (Figure 31). The user must fill in the Mean Consumption data retrieved from the EFSA comprehensive dataset to columns C,D and E.

Do not copy-paste the full dataset you have retrieved from EFSA! Not all countries have information on all 160 food categories of FoodEx1 level 2! The ImproRisk worksheet allows for the input of data for 160 food categories.

The Mean Occurrence is automatically retrieved from the Occurrence worksheet within ImproRisk.

The Exposure is then calculated via Excel formulas.

Figure 31 COMPREHENSIVE worksheet

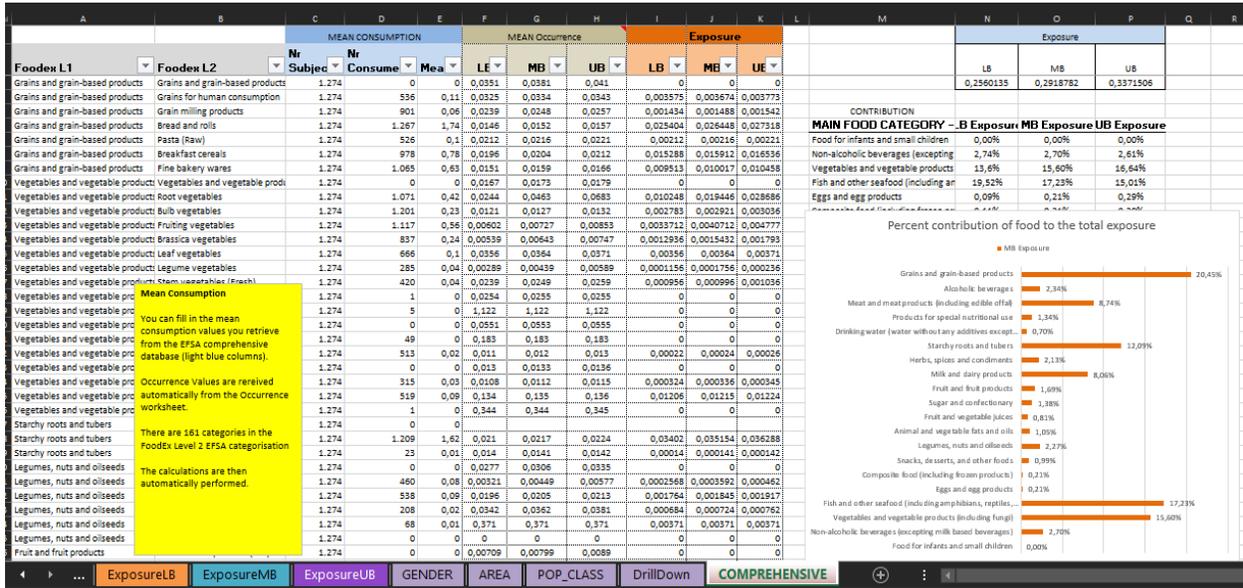


Figure 32 The EFSA Data Warehouse

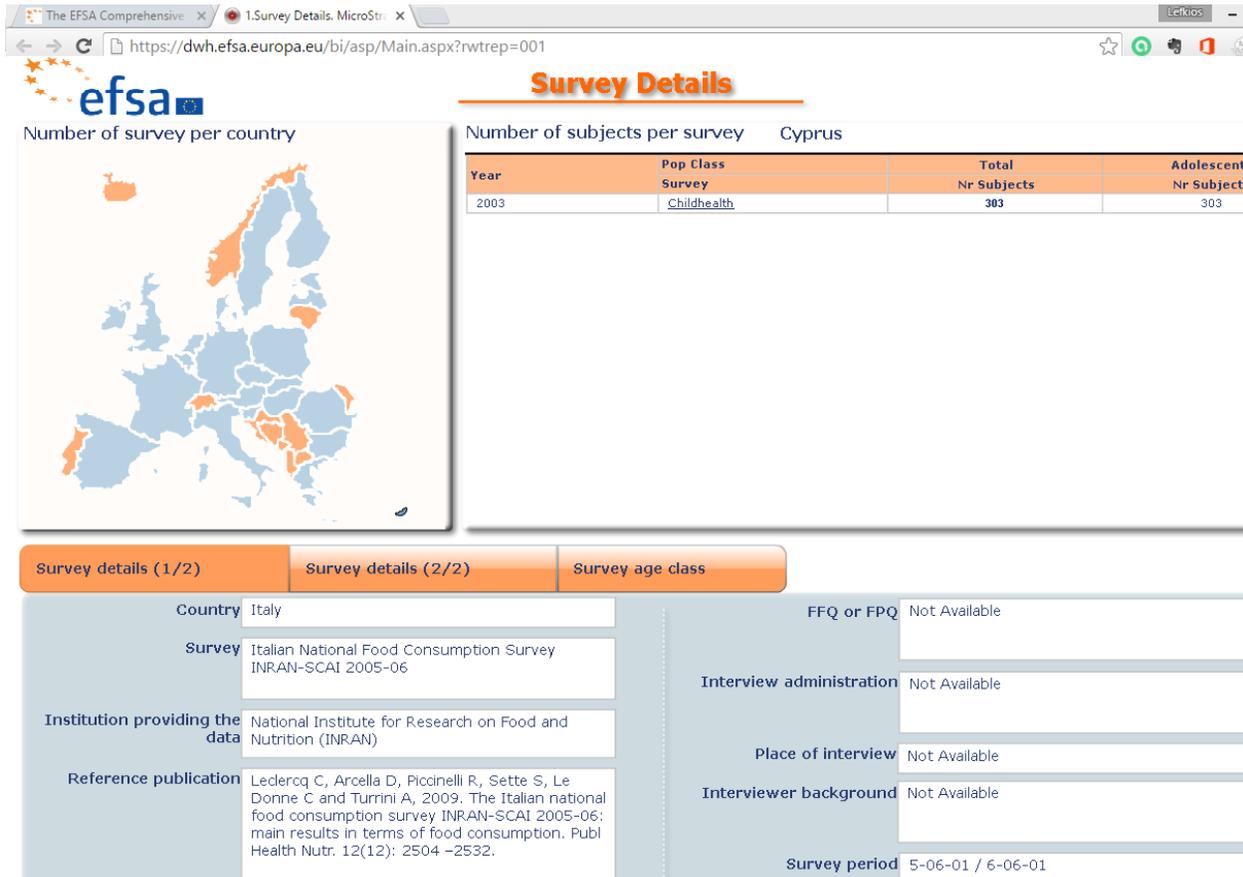


Figure 33 Screenshot for the EFSA Comprehensive European Food Consumption Database (<http://www.efsa.europa.eu/en/food-consumption/comprehensive-database>)

Chronic food consumption statistics		
Intake	All subjects	Consumers only
grams per day* (g/day)	 	 
grams per day per kilogram of body weight* (g/kg bw per day)	 	 

## Worksheet LOG

This worksheet keeps a log of all the Exposure Assessments that are performed with ImproRisk.

Figure 34: LOG worksheet

*****LOG FILE*****										
ImproRisk 2.0.3										
		Reference value		MEAN Exposure µg/ kg body weight			Percentage Above the Reference Value			
		(µg/ Kg b.w.)	Type	LB	MB	UB	LB	MB	UB	
Date	Chemical Substance									NOTES
										DATAFILE NAME
27/06/2019	***									New Consumption Data are
27/06/2019	***									Subjects_Consumption_EUMENU Lot2 (N=803).xlsx
27/06/2019	***									Subjects_Consumption_EUMENU ALL SUBJECTS (exc.Pregnant) (N=1641).xls
27/06/2019	***									Subjects_Consumption_EUMENU ALL SUBJECTS (exc.Pregnant) (N=1641).xls
28/06/2019	***									Subjects_Consumption_EUMENU Elderly (N=260)1.xlsx
28/06/2019	***									Subjects_Consumption_EUMENU Lot1 (N=838).xlsx
28/06/2019	***									Subjects_Consumption_EUMENU Adults (N=272).xlsx
28/06/2019	Type the name of the cher	0	WEEKLY	Mean	2,48	2,9	Percent	100,0%	100,0%	Occurrence Example - Level 3 - EFSA Pb.xlsm
28/06/2019	Pb	0,63	DAILY	Mean	0,35	0,41	Percent	6,2%	9,9%	Occurrence Example - Level 3 - EFSA Pb.xlsm
28/06/2019	Pb	0,63	DAILY	Mean	0,35	0,41	Percent	6,2%	9,9%	Occurrence Example - Level 3 - EFSA Pb.xlsm
28/06/2019	Pb	0,63	DAILY	Mean	0,35	0,41	Percent	6,2%	9,9%	Occurrence Example - Level 3 - EFSA Pb.xlsm
28/06/2019	Aflatoxin B1	0,87	DAILY	Mean	0	0	Percent	0,0%	0,0%	Occurrence Template - Level 2 & Level 3_AFB1_Test.xlsm
28/06/2019	***									New Consumption Data are
28/06/2019	***									Subjects_Consumption_EUMENU Elderly (N=260).xlsx
28/06/2019	Pb	0,63	DAILY	Mean	0,58	0,76	Percent	39,6%	53,8%	Subjects_Consumption_EUMENU Infants (N=266).xlsx
28/06/2019	Aflatoxin B1	0,87	DAILY	Mean	0	0	Percent	0,0%	0,0%	Occurrence Example - Level 3 - EFSA Pb.xlsm
28/06/2019	***									Occurrence Template - Level 2 & Level 3_AFB1_Test.xlsm
28/06/2019	Pb	0,63	DAILY	Mean	0,4	0,48	Percent	12,6%	18,7%	New Consumption Data are
01/07/2019	***									Subjects_Consumption_EUMENU ALL SUBJECTS (exc.Pregnant) (N=1641).xls
31/10/2019	***									Occurrence Example - Level 3 - EFSA Pb.xlsm
										New Consumption Data are
										Subjects_Consumption_EUMENU Adults (N=272).xlsx
										Subjects_Consumption_EUMENU Lot2 (N=803).xlsx

CLEAR THE LOG

**Date:** The date that the Exposure Assessment was performed.

**Compound:** The chemical substance that was investigated.

**Reference value:** The reference value.

**Type:** The type of the reference value (DAILY or WEEKLY).

**Mean Exposure:** LB, MB, UB: The mean exposure for the LB, MB and UB scenarios in micrograms per body Kg body weight (µg/ Kg b.w.)

**Percentage ABOVE the Tolerable Intake Level:** for the LB, MB and UB scenario.

Also, whenever new consumption or occurrence data are installed, then the file name is stored for quick reference.

# Updating the Occurrence and Consumption

## Occurrence

To update the occurrence FoodEx1 level-2 or level-3, an occurrence template is provided - "Occurrence Template - Level 2 & Level 3.xlsm" and is a macro enabled workbook.

The occurrence template has two worksheets; Level2 and Level3, for providing aggregated occurrence values at Level 2 and Level 3 categorisation respectively.

### Level2

The template has the same structure as with the **OccurrenceL2** worksheet in the ImproRisk.

The current version of ImproRisk performs exposure assessments on the average (mean) occurrences in the LB, MB, and UB scenario. The user can add occurrence values to the template only on columns H, I and J as seen in Figure 35.

Then three parameters must be defined:

1. The name of the chemical substance
2. The substance category; Select one of Additive, Pesticide, Veterinary Drug Residue, Contaminant, Genotoxic-Carcinogen
3. The Reference Value in mg/Kg b.w; intake level in µg/Kg of body weight as provided by EFSA
4. Type of Reference value; Select one of Acceptable Intake, Tolerable Intake, Provisional Maximum Tolerable Intake, Benchmark Dose Level (BMDL)
5. The reference value type; i.e. it is the DAILY or WEEKLY intake.

Note that the template name (i.e. file name) is not important. Name the resulting file appropriately for easier reference. **What is of importance, it is the name of the worksheet within the template that needs to be named "Level2".**

Figure 35: Occurrence Template – Worksheet Level2 for updating the occurrence at FoodEx1 level 2

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
	Chemical Substance	Type the name of the chemical		Type in the name of the chemical											
	Substance Category			Click the cell and Select from the drop down list											
	Reference value (µg/Kg b.w.)			Type in the reference value											
	Type of Reference value			Click the cell and Select from the drop down list											
	Type			Click the cell and Select											
				mg/kg											
				min			mean			median			P95		
	FoodEx1_name	FoodEx2_name	No of Samples	LB	MB	UB	Occur_Mean_LB	Occur_Mean_MB	Occur_Mean_UB	LB	MB	UB	LB	MB	UB
	Grains and grain-based products	Grains and grain-based products													
		Grains for human consumption													
		Grain milling products													
		Bread and rolls													
		Pasta (Raw)													
		Breakfast cereals													
	Vegetables and vegetable products (including fungi)	Fine bakery wares													
		Vegetables and vegetable products (including fungi)													
		Root vegetables													
		Bulb vegetables													
		Fruiting vegetables													
		Brassica vegetables													
		Leaf vegetables													
		Legume vegetables													
		Stem vegetables (Fresh)													
		Sugar plants													
		Sea weeds													
		Tea and herbs for infusions (Solid)													
		Cocoa beans and cocoa products													
		Coffee beans and coffee products (Solid)													
		Coffee imitates (Solid)													
		Vegetable products													
	Fungi, cultivated														
	Fungi, wild, edible														
	Starchy roots and tubers	Other starchy roots and tubers													
		Potatoes and potatoes products													
	Legumes, nuts and oilseeds	Starchy roots and tubers													
		Legumes, beans, dried													
		Legumes, beans, green, without pods													
		Legumes, nuts and oilseeds													
		Oilseeds													
		Other seeds													
		Tree nuts													

### Level 3

The worksheet "Level3" (Figure 36) is where the user will provide aggregated occurrence values for up to a 1000 food items at the Level 3. The user can *click on the green banner "CLICK HERE to load the autocomplete SearchBox"*, and a smart autocomplete feature is activated for quickly and correctly accessing the level-3 food items. When a food item is selected (or typed in), the corresponding Level-2 and Level-3 names appear at columns B and C. Then, the user then needs to provide the occurrence values in columns I;J;K for the mean LB, MB and UB scenario respectively. In case of duplicate values, these are indicated with a pink cell-background. If duplicate values are present when this template is loaded into ImproRisk, the model considers only the first entry (top to bottom).

The same occurrence template can be used to import only occurrences at Level 2. Providing occurrence values at level-3 is optional,



Figure 37: Worksheet UPDATE DATA

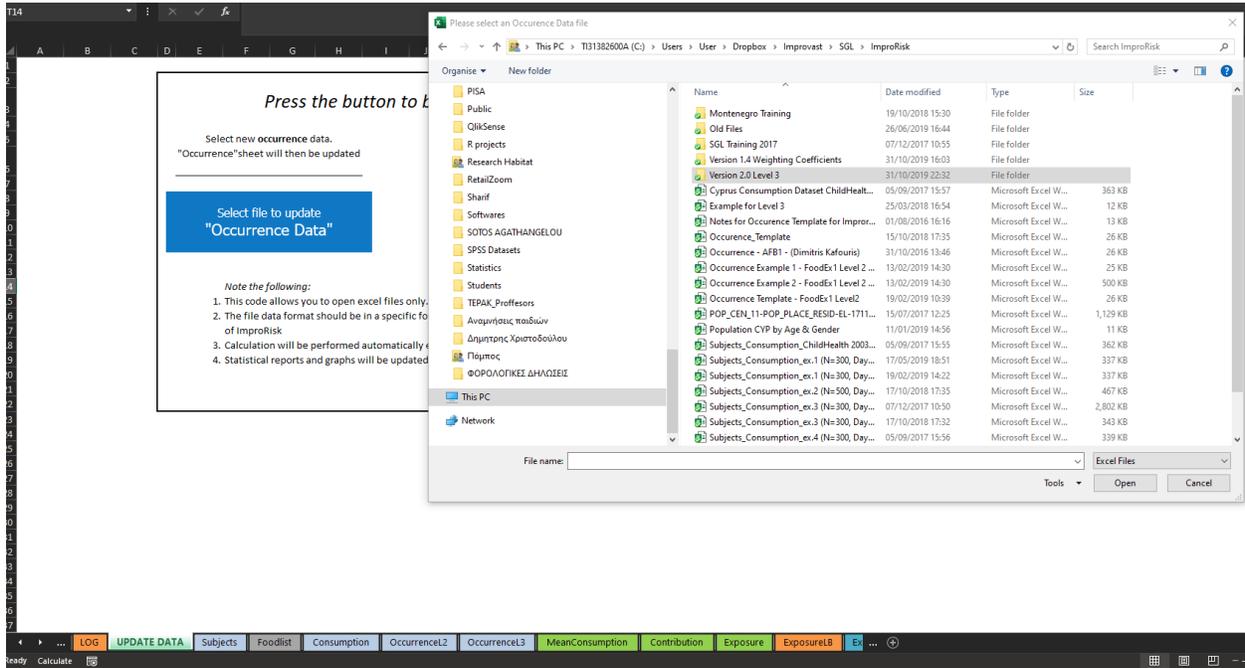
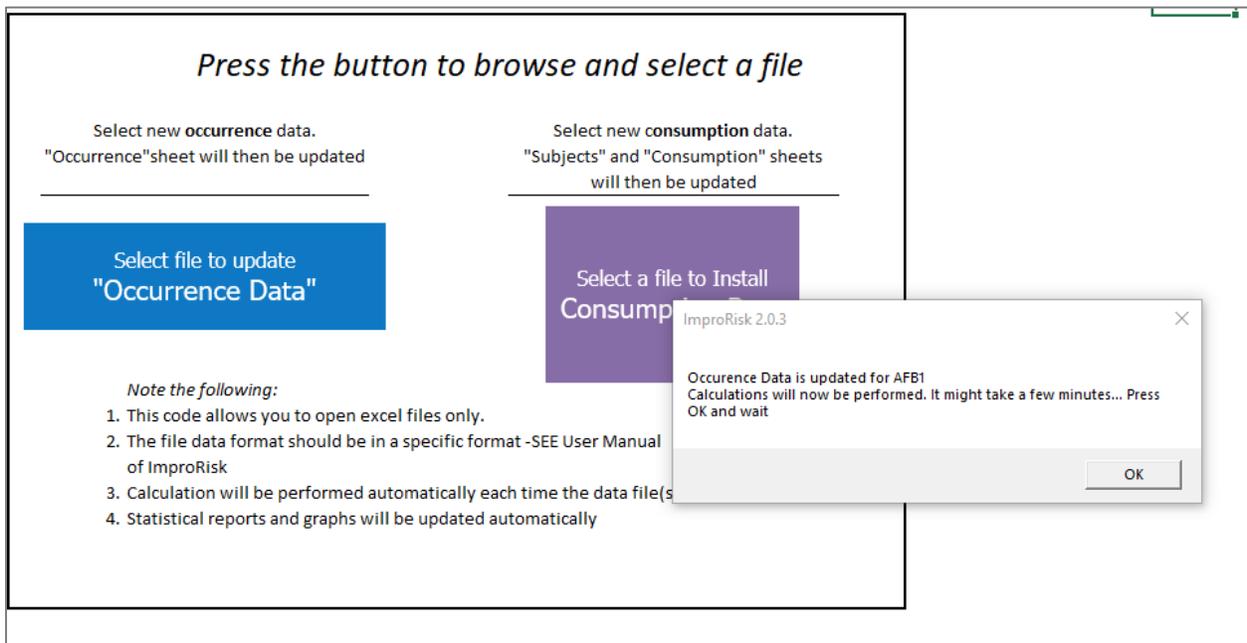


Figure 38: Worksheet UPDATE DATA - Notifications when updating the Occurrence data



## Consumption

The Subject or the Consumption data, when in need to be updated or changed, a template is provided (Subjects\_Consumption\_Template.WeightingCoefficients.xlsx).

There are two worksheets in the template:

### Worksheet "Subjects"

This worksheet will contain the participants' demographic characteristics

When filling in the demographics, use:

- MALE, FEMALE for the Gender
- Numerical values for AGE and WEIGHT
- Text values for AREA and POP\_CLASS.
- SUBJECTID must be a unique identifier; either numeric or a combination of text and number
- POP\_CLASS; It is recommended to use the EFSA population classes approach (Figure 39)
- WCOEFF; The weight coefficient for each subject

**!!DO NOT change this worksheet's name**

Figure 39: Population Class by EFSA

POP_CLASS - AGE	
Infants*	3-11 months
Toddlers *	1-2 years
Other children*	3-9 years
Adolescents*	10-17 years
Adults*	18-64 years
Elderly*	65-74 years
Very elderly*	≥75

*\*\*Guidance on the EU Menu methodology, EFSA Journal 2014;12(12):3944*

## Worksheet "Consumption"

This worksheet will contain the participants' food consumption occasions

Please note the following:

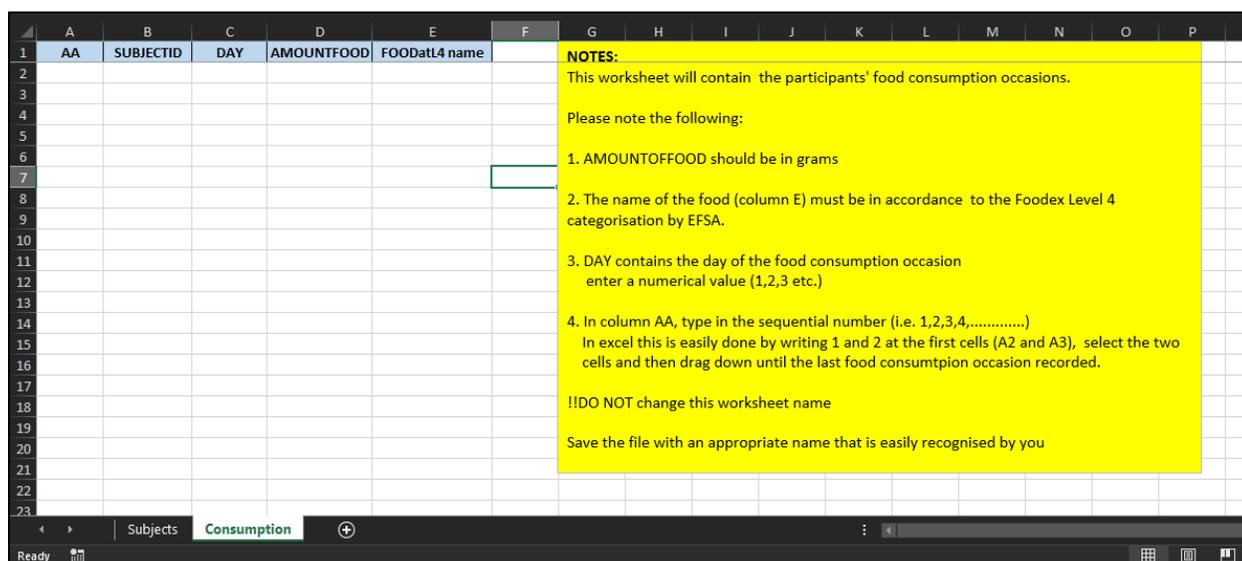
- AMOUNTOFFOOD should be in grams
- The name of the food (column E) must be in accordance to the FoodEx1 Level 4 classification by EFSA.
- DAY contains the day of the food consumption occasion. **Enter a numerical value (1,2,3 etc.)**
- In column SERIAL, type in the sequential number (i.e. 1,2,3,4,...) *In excel this is easily done by writing 1 and 2 at the first cells (A2 and A3), select the two cells and then drag down until the last food consumption occasion recorded.*

**!!DO NOT change this worksheet's name**

Figure 40: Template for updating the Subjects and Consumption data – Subjects worksheet

B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
GENDER	AGE	WEIGHT	AREA	POP_CLASS	WCOEFF			NOTES:						POP_CLASS - AGE			
								In this worksheet, the participants' demographic characteristics will be added.						Infants*	3-11 months		
								When filling in the demographics:						Toddlers *	1-2 years		
								1. In column A, "SUBJECTID" must be a univocal value.						Other children*	3-9 years		
								2. In column B, "GENDER", use MALE or FEMALE.						Adolescents*	10-17 years		
								3. In columns C & D, "AGE" & "WEIGHT", use numerical values.						Adults*	18-64 years		
								4. In columns E & F, "AREA" and "POP_CLASS", use text values.						Elderly*	65-74 years		
								5. In column G, insert the weighting coefficient "WCOEFF" for each subject to adjust for non-representativeness of your sample. <b>If the sample is representative of your population, add the value "1" for every subject.</b>						Very elderly*	≥75		
								Please Note:						**Guidance on the EU Menu methodology, EFSA Journal 2014;12(12):3944			
								A. It is recommended to use the EFSA population classes approach (see next note).									
								B. Use the import button "Select a file to Install Consumption Data" on the "UPDATE DATA" worksheet within the ImproRisk model, to install consumption data.									
								C. The sample size is automatically calculated when this template is installed in the ImproRisk model.									
								!!DO NOT change this worksheet's name									
								Save the file with an appropriate name that is easily recognised by you									

Figure 41: Template for updating the Subjects and Consumption data – Consumption worksheet



Save the file with an appropriate name that is easily recognised by the user

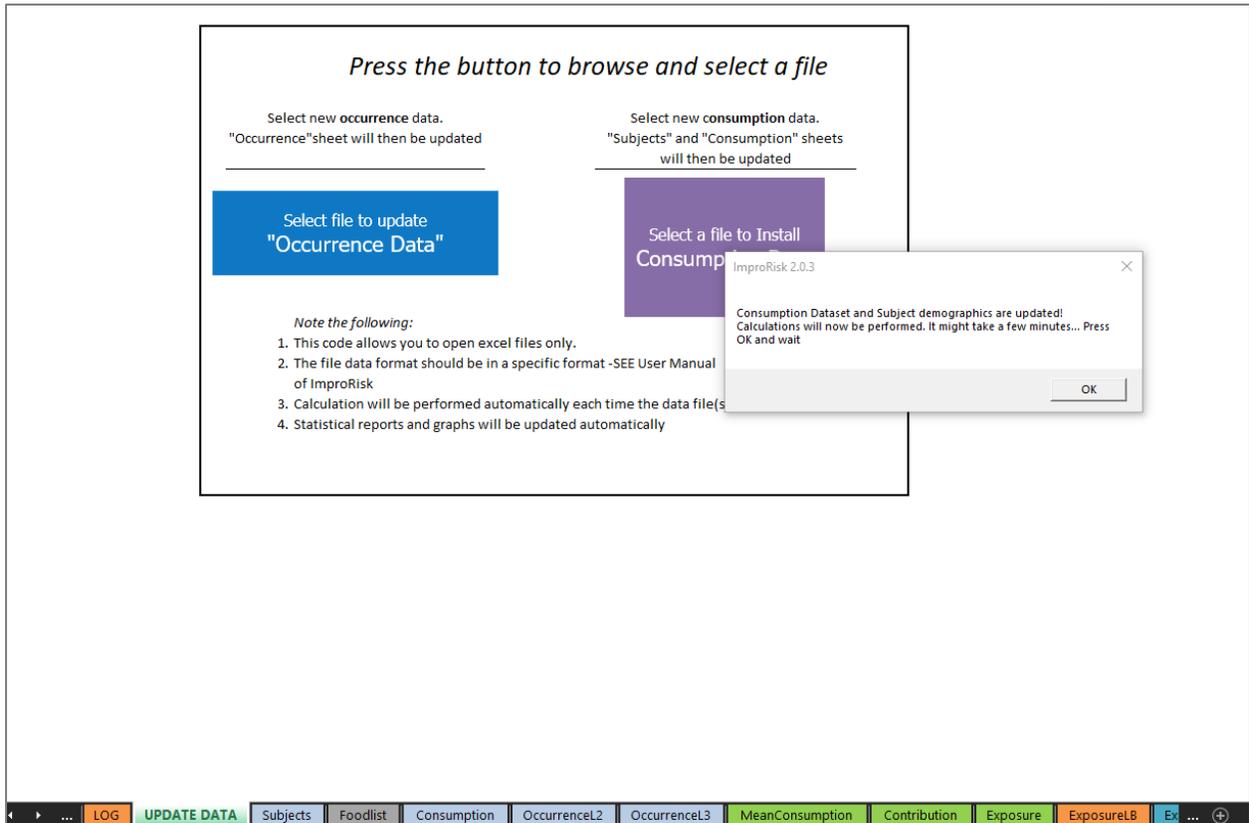
When the template is saved, then use the *UPDATE DATA worksheet* to install the consumption data

### Steps:

1. Go to worksheet UPDATE DATA in the ImproRisk model
2. Click on the light purple button for selecting a consumption file
3. The user will be prompted to locate and select the file in the local disk
4. Navigate to the file that holds the consumption file and click "Open"
5. The model will notify the user that Consumption and Subject demographics are updated and that calculations will be executed.
6. Wait for the model to until a message appears that informs that "All Done!"

**ALL the calculations and reports are automatically performed! Navigate to the other worksheets for the reporting.**

Figure 42: Notification the new Subject and Consumption data are updated



## Weighting coefficients

ImproRisk since version 1.4 accommodates weighting coefficients to adjust the sample for non-representativeness within the population. Exposure statistics e.g. mean exposure and standard error will now be adjusted to reflect the population from which the sample was taken from.

ImproRisk uses frequency weights<sup>6</sup> and not reliability weights. The formulae for adjusted estimates derive unbiased estimates for frequency weights and not for reliability weights.

All subjects in the food survey sample will carry a “weight coefficient”. This coefficient will be a number (>1) that indicates the number of persons in the population that this individual represents. If for example an individual’s weight coefficient is 500, then that individual represents 500 persons in the population of the same population class.

Weight coefficients, or simply “weights” need to be calculated beforehand by the risk assessor. The weight coefficients will then be installed in ImproRisk within the Subjects\_Consumption\_Template.WeightingCoefficients.xlsx template provided (see

---

<sup>6</sup> [https://en.wikipedia.org/wiki/Weighted\\_arithmetic\\_mean#Frequency\\_weights](https://en.wikipedia.org/wiki/Weighted_arithmetic_mean#Frequency_weights)

Subjects and Consumption template section). **In case there are no weight coefficients assigned, the weight for each individual is set to "1"** (i.e. subject represents only him/herself)

Typically, weight coefficients are calculated using (recent) census data of the population.

## Calculating weights

A typical way of calculating weights is as follows:

Let's assume the user wants to adjust the sample by **Gender** and **Age Group**.

First, a (cross tabulation) table of real **population counts** in each Gender by Age combination needs to be calculated. This is where the census data will be used.

Assuming that the age group has 3 levels (Children, Adults, Elderly) and Gender has 2 levels (Males, Females), the cross tabulation gives a table of the total number of persons in each Gender by Age class, a total of 6 Classes (2X3=6).

The corresponding **sample counts** will be calculated from the Sample data again by tabulation.

The weight to be applied, **is the ratio of the Population counts to the Sample counts for the Class to which the person belongs**. This means that subjects that belong to the same "population" class will carry the same weighting coefficient, i.e. Adult men will all have the same "weight", adult women the same, etc.

In case the user wants to adjust for more than three population attributes (e.g. Gender, Age and Area), it is essential to note of the possibility (depending in the sample size and sampling strategy) that the sample crosstabulation may give many cells with no data or very small sample numbers. This would give very high weights to some individuals and may lead to inaccurate results.

A methodology for calculating weights for a non-representative sample is described on page 64 of the guide to the 2007 Australian National Children's Nutrition and Physical Activity Survey:

[https://www.health.gov.au/internet/main/publishing.nsf/Content/589EFDBF5E7B916FCA257BF000211E08/\\$File/user-guide-v2.pdf](https://www.health.gov.au/internet/main/publishing.nsf/Content/589EFDBF5E7B916FCA257BF000211E08/$File/user-guide-v2.pdf)

## Example - Cyprus

An example is provided to calculate the weight coefficients for the Cyprus sample (n=300) to adjust for the population (N=706652) (Figure 43).

The crosstabulation for the sample counts (blue table) shows the distribution of gender and age group in the sample. The crosstabulation for the population counts (green table) shows the respective counts as extracted by tabulating the latest census data (CYSTAT, 2011)<sup>7</sup>

The pink table (Weights rounded), shows the resulting weight coefficients for the 6 classes

e.g. The weight coefficient for Male Adolescents will be  $52862/19=2782$

e.g. The weight coefficient for Male Adults will be  $259221/77=3367$

etc.

It can be derived that the female proportion in the sample is 62.7% (188/300) while the female proportion in the population is 51.4% (362869/706652). This means that females are overrepresented in the sample. This is now adjusted when the weight coefficient for females for all age groups, is lower than that of males.

Note that the weights do not need to be integers. In this example, we round the weights to the nearest integer for simplicity reasons.

Figure 43: Weight coefficients calculated for the Cyprus population for a sample of 300

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	SUBJECTID	GENDER	AGE	WEIGHT	AREA	POP_CLASS	WCOEFF						
2	1001	FEMALE	12	81	Nicosia	Adolescents	2094		<b>Sample Counts</b>				
3	1002	FEMALE	15	43	Nicosia	Adolescents	2094						
4	1003	FEMALE	62	69	Nicosia	Adults	2302		<b>Count of subjects</b>				
5	1004	FEMALE	63	65	Nicosia	Adults	2302		<b>GENDER</b>	<b>POP_CLASS</b>			
6	1005	FEMALE	64	72	Nicosia	Adults	2302		MALE	Adolescents	Adults	Elderly	Grand Total
7	1006	FEMALE	65	71	Nicosia	Adults	2302			19	77	16	112
8	1007	FEMALE	29	57	Nicosia	Adults	2302		FEMALE	24	121	43	188
9	1008	FEMALE	65	64	Nicosia	Adults	2302		<b>Grand Total</b>	<b>43</b>	<b>198</b>	<b>59</b>	<b>300</b>
10	1009	MALE	60	62	Nicosia	Adults	3367						
11	1010	MALE	50	104	Nicosia	Adults	3367		<b>Population Real Counts</b>				
12	1011	MALE	58	69	Nicosia	Adults	3367			Adolescents	Adults	Elderly	Total
13	1012	MALE	45	59	Nicosia	Adults	3367		MALE	52.862	259.221	31.700	343.783
14	1013	MALE	70	58	Nicosia	Elderly	1981		FEMALE	50.254	278.554	34.061	362.869
15	1014	MALE	44	61	Nicosia	Adults	3367		Total	103.116	537.775	65.761	706.652
16	1015	MALE	64	91	Nicosia	Adults	3367						
17	1016	FEMALE	73	89	Nicosia	Elderly	792						
18	1017	FEMALE	74	70	Nicosia	Elderly	792		<b>WEIGHTS (rounded)</b>				
19	1018	FEMALE	75	88	Nicosia	Elderly	792			Adolescents	Adults	Elderly	
20	1019	FEMALE	45	85	Nicosia	Adults	2302		MALE	2.782	3.367	1.981	
21	1020	MALE	33	77	Nicosia	Adults	3367		FEMALE	2.094	2.302	792	
22	1021	FEMALE	50	65	Nicosia	Adults	2302						
23	1022	MALE	71	85	Nicosia	Elderly	1981						
24	1023	MALE	50	88	Nicosia	Adults	3367						
25	1024	MALE	50	96	Nicosia	Adults	3367						

7

[http://www.cystat.gov.cy/mof/cystat/statistics.nsf/populationcondition\\_22main\\_en/populationcondition\\_22main\\_en?OpenForm&sub=2&sel=2](http://www.cystat.gov.cy/mof/cystat/statistics.nsf/populationcondition_22main_en/populationcondition_22main_en?OpenForm&sub=2&sel=2)

## Subjects and Consumption template

There is an updated version of the Subjects&Consumption template called *Subjects\_Consumption\_Template.WeightingCoefficients.xlsx* and it can be downloaded at [www.improrisk.com](http://www.improrisk.com)

An additional column (column G – “WCOEFF”) contains the weighting coefficient of each subject (Figure 44). Note that the values should be frequency weights<sup>8</sup> and not reliability weights.

An example is shown in Figure 45

**IMPORTANT! When there are no weight coefficients assigned, the “WCOEFF” column is filled in with a factor of 1 for every subject (see Figure 46).**

Figure 44: New template for installing new consumption data. Worksheet “Subjects”

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	SUBJECTID	GENDER	AGE	WEIGHT	AREA	POP_CLASS	WCOEFF			NOTES:						POP_CLASS - AGE				
2										In this worksheet, the participants' demographic characteristics will be added.						Infants*	3-11 months			
3																Toddlers *	1-2 years			
4																Other children*	3-9 years			
5										When filling in the demographics:						Adolescents*	10-17 years			
6										1. In column A, "SUBJECTID" must be a univocal value.						Adults*	18-64 years			
7										2. In column B, "GENDER", use MALE or FEMALE.						Elderly*	65-74 years			
8										3. In columns C & D, "AGE" & "WEIGHT", use numerical values.						Very elderly*	≥75			
9										4. In columns E & F, "AREA" and "POP_CLASS", use text values.										
10										5. In column G, insert the weighting coefficient "WCOEFF" for each subject to adjust for non-representativeness of your sample. If the sample is representative of your population, add the value "1" for every subject.										
11																				
12																				
13																				
14																				
15																				
16																				
17																				
18										Please Note:										
19																				
20										A. It is recommended to use the EFSA population classes approach (see next note).										
21																				
22																				

<sup>8</sup> [https://en.wikipedia.org/wiki/Weighted\\_arithmetic\\_mean#Frequency\\_weights](https://en.wikipedia.org/wiki/Weighted_arithmetic_mean#Frequency_weights)

Figure 45: Example of weighting coefficients (see WCOEFF) in the Subjects & Consumption template- Worksheet "Subjects"

	A	B	C	D	E	F	G	H	I	J	K
1	SUBJECT	GENDI	AGE	WEIGH	AREA	POP_CLAS	WCOEFF			POP_CLASS	AGE
2	1001	FEMALE	12	81	Nicosia	Adolescents	2094			Infants	<1
3	1002	FEMALE	15	43	Nicosia	Adolescents	2094			Toddlers	1≤3
4	1003	FEMALE	62	69	Nicosia	Adults	2302			Other children	3≤10
5	1004	FEMALE	63	65	Nicosia	Adults	2302			Adolescents	10≤18
6	1005	FEMALE	64	72	Nicosia	Adults	2302			Adults	18≤65
7	1006	FEMALE	65	71	Nicosia	Adults	2302			Elderly	65≤75
8	1007	FEMALE	29	57	Nicosia	Adults	2302			Very elderly	≥75
9	1008	FEMALE	65	64	Nicosia	Adults	2302			*EFSA, 2011b	
10	1009	MALE	60	62	Nicosia	Adults	3367				
11	1010	MALE	50	104	Nicosia	Adults	3367				
12	1011	MALE	58	69	Nicosia	Adults	3367				
13	1012	MALE	45	59	Nicosia	Adults	3367				
14	1013	MALE	70	58	Nicosia	Elderly	1981				
15	1014	MALE	44	61	Nicosia	Adults	3367				
16	1015	MALE	64	91	Nicosia	Adults	3367				
17	1016	FEMALE	73	89	Nicosia	Elderly	792				
18	1017	FEMALE	74	70	Nicosia	Elderly	792				
19	1018	FEMALE	75	88	Nicosia	Elderly	792				
20	1019	FEMALE	45	85	Nicosia	Adults	2302				
21	1020	MALE	33	77	Nicosia	Adults	3367				
22	1021	FEMALE	50	65	Nicosia	Adults	2302				

Figure 46: Subjects & Consumption template. If no weigh coefficients are assigned, the "WCOEFF" column is filled in with 1

	A	B	C	D	E	F	G	H	I	J	K
1	SUBJECT	GENDI	AGE	WEIGH	AREA	POP_CLAS	WCOEFF			POP_CLASS	AGE
2	1001	FEMALE	12	81	Nicosia	Adolescents	1			Infants	<1
3	1002	FEMALE	15	43	Nicosia	Adolescents	1			Toddlers	1≤3
4	1003	FEMALE	62	69	Nicosia	Adults	1			Other children	3≤10
5	1004	FEMALE	63	65	Nicosia	Adults	1			Adolescents	10≤18
6	1005	FEMALE	64	72	Nicosia	Adults	1			Adults	18≤65
7	1006	FEMALE	65	71	Nicosia	Adults	1			Elderly	65≤75
8	1007	FEMALE	29	57	Nicosia	Adults	1			Very elderly	≥75
9	1008	FEMALE	65	64	Nicosia	Adults	1			*EFSA, 2011b	
10	1009	MALE	60	62	Nicosia	Adults	1				
11	1010	MALE	50	104	Nicosia	Adults	1				
12	1011	MALE	58	69	Nicosia	Adults	1				
13	1012	MALE	45	59	Nicosia	Adults	1				
14	1013	MALE	70	58	Nicosia	Elderly	1				
15	1014	MALE	44	61	Nicosia	Adults	1				
16	1015	MALE	64	91	Nicosia	Adults	1				
17	1016	FEMALE	73	89	Nicosia	Elderly	1				
18	1017	FEMALE	74	70	Nicosia	Elderly	1				

# General Information

## Automatic recalculation

When new Occurrence or Consumption data are installed, internal VBA code recalculates all the formulas within ImproRisk. Automatic recalculation is disabled just after the calculations are performed.

However, the model recognizes when changes are made to certain worksheets (worksheets that contain information that have a direct impact on the exposure assessment), as soon as the user leaves the worksheet a message appears requesting permission to recalculate. Changes in Graphs do not count as changes, but changes in Pivot tables do count as changes.

In any case, any user who wishes to recalculate *the active worksheet*, this can be done when pressing *Shift-F9*. For recalculating *the entire model*, it can be done with *F9*.

If the user wants to activate Automatic recalculation at any time (though not suggested since calculations may take significant amount of time depending on the subjects and food consumption occasions), the user should select Formulas in the horizontal menu bar(ribbon) and then executes the Calculation Options (Figure 44).

Figure 47: Message when changes are detected in the ImproRisk model

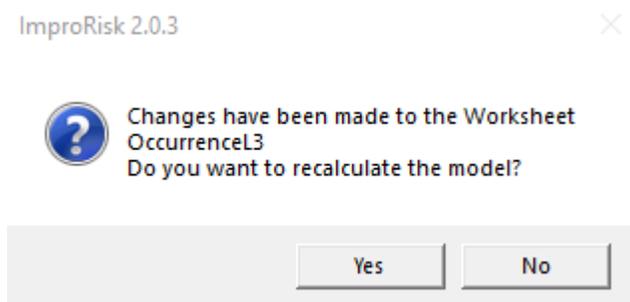
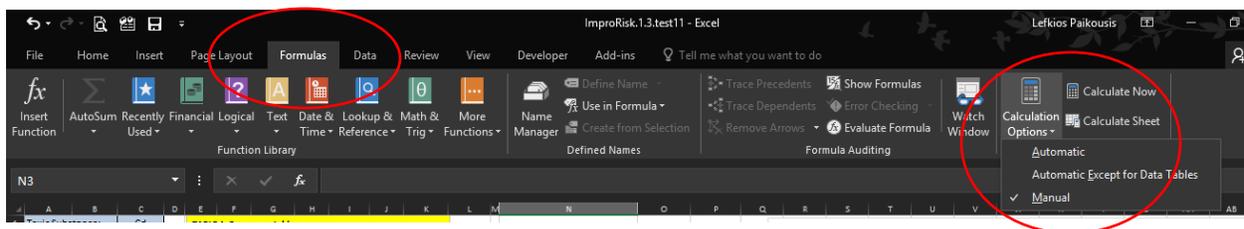


Figure 48: Automatic and Manual recalculation



## Locked workbook and worksheet

ImproRisk and the worksheets are “locked for editing” meaning that the user cannot perform certain operations within the worksheets.

### The user cannot:

1. Delete, move or change the name of the Worksheets
2. Insert, delete, move or change the columns within the worksheets
3. Type in data in the cells. All the cells are locked

This means that if the user wants to update the Occurrence or Consumption database, the templates must be used and installed via the UPDATE DATA worksheet.

### The user is allowed to:

1. Resize the worksheet’s Columns and Rows
2. Format all the cells or columns
3. Edit the graphs
4. Use the PivotTables
5. Use AutoFilter
6. Edit all the objects in the workbook (i.e. note, graphs etc.)

## Notes within the worksheet

Within the worksheets there are several yellow notes that can be moved around. To move them, just click inside the note, and then “drag” them by touching on one of the edges.

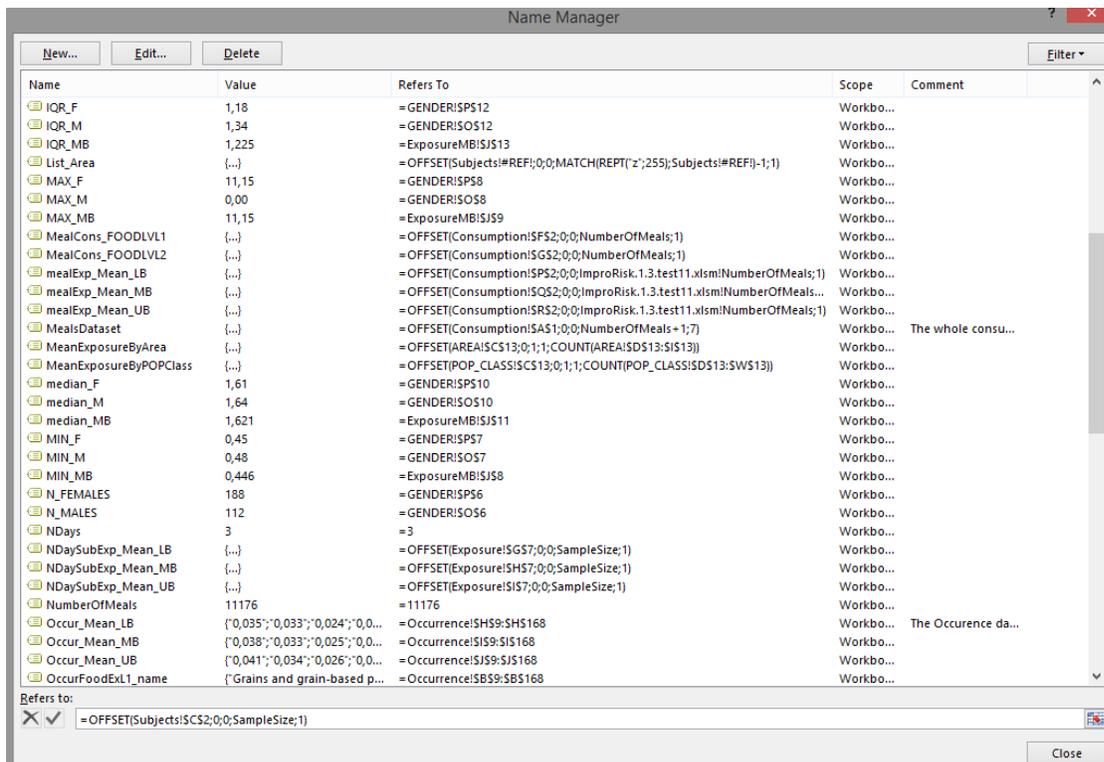
## Named Ranges in Excel

Named Ranges are Excel names that refer to cells, a range of cells, a constant value, or a formula. ImproRisk uses those names in formulas, to replace values or cell references.

By using names, the formulas are much easier to read, understand and maintain.

i.e. named range *SUBJECTID* refers to all subjectID’s in the sample, *NDays* holds the number of days of the food survey, *SubExp\_MB* holds the total exposure (either Daily or Weekly) of each individual in the sample, etc.

Figure 49: Sample of Named Ranges in ImproRisk



## Statistical Terms

### Cumulative Frequency

Cumulative means "how much so far"<sup>9</sup>. The total of a frequency and all frequencies so far in a frequency distribution. It is the 'running total' of frequencies.

To have cumulative totals, just add up the values as you go.

Scores: 1,1,2,2,2,2,2,3,3,3,3,4,4,5

Score	Frequency	Cumulative Frequency
1	2	2
2	5	7
3	4	11
4	2	13
5	1	14

Cumulative Frequency for Score 3  
is  $2+5+4 = 11$

Figure 50: Cumulative frequency example. Thank you [www.mathisfun.com](http://www.mathisfun.com)

### Cumulative percentage

The total of a percentage and all percentages so far in a frequency distribution. It is the 'running total' of percentages.

In order to have cumulative totals, the values can be added up as the user moves on.

Otherwise, it is the Cumulative frequency divided by the total number of subjects. A nice description of the cumulative percentage use and benefits is presented by the state statistical service of Canada.<sup>10</sup>

### Quartiles

When the distribution of exposure is described in the **ExposureMB** or **GENDER** worksheets, summary statistics named Q1, Q3 and IQR are presented.

Q1 = 1<sup>st</sup> quartile or 25<sup>th</sup> percentile.

Median = 2<sup>nd</sup> quartile or 50<sup>th</sup> percentile

<sup>9</sup> <https://www.mathisfun.com/definitions/cumulative-frequency.html>

<sup>10</sup> <http://www.statcan.gc.ca/edu/power-pouvoir/ch10/5214864-eng.htm>

Q1 = 3<sup>rd</sup> quartile or 75<sup>th</sup> percentile

IQR = Q3-Q1

**e.g. if the Q1 of exposure** is 1.19 µg/Kg b.w., this means 25% of the sample are exposed UP TO 1.19. In other words, 25% of the subjects have exposure below 1.19 µg/Kg b.w.

**e.g. if the Q3 of exposure** is 2.53, this means 75% of the sample are exposed UP TO 2.53 µg/Kg. In other words, 25% of the sample are exposed more than 2.53 µg/Kg b.w.

**e.g. If the Median exposure** is 1.64 µg/Kg b.w. then half of the sample is exposed less than 1.64 µg/Kg b.w. and half above 1.64 µg/Kg b.w.

## Probability Distribution (histogram)

It is better known as the histogram<sup>11</sup>.

A histogram is a graphical representation of the distribution of **exposure values** of all the subjects in the food survey grouped into ranges. It is similar to a Bar Graph, but in a Histogram each bar is for a range of data.

The entire range of exposure values is divided into a series of intervals—and then a count is obtained of how many exposure values (i.e. subjects) fall within each interval. The bins are specified as consecutive, non-overlapping intervals of the exposure values. The bins (intervals) must be adjacent, and they are of equal size. In ImproRisk the size of the bin is called “height”. It is calculated via the Friedman-Diaconis rule<sup>12</sup> and it is very robust in practice (Freedman D. 1981, Scott D. 1992) when compared to Sturges’ rule that is used by a variety of statistical software<sup>13</sup>

For example, in Figure 47, 21.7% of the sample has exposure between 1.19 and 1.56 µg/Kg b.w.

## Cumulative Distribution

It is a graphical representation of the Cumulative Percentages (see *Cumulative percentage*). In other words, it is the running total of the bars in a histogram. The end of the running total will be 100%.

---

<sup>11</sup> <https://en.wikipedia.org/wiki/Histogram>

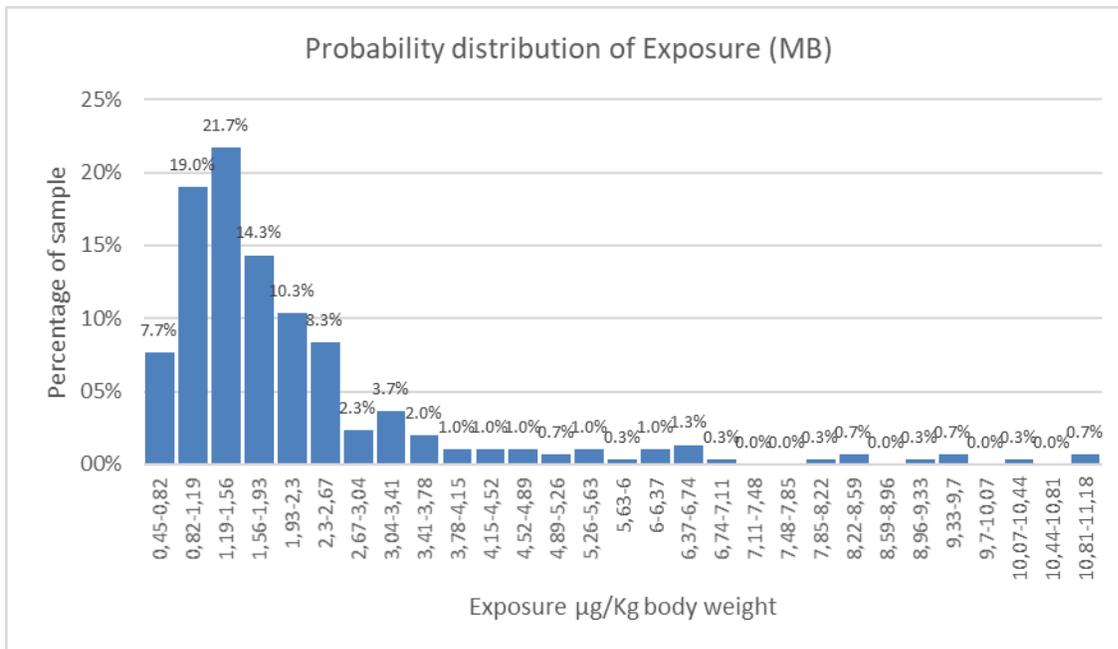
<sup>12</sup> [https://en.wikipedia.org/wiki/Freedman%E2%80%93Diaconis\\_rule](https://en.wikipedia.org/wiki/Freedman%E2%80%93Diaconis_rule)

<sup>13</sup> <http://robjhyndman.com/papers/sturges.pdf>

It is a better visualisation when two groups are compared as to their exposure i.e. Male and Female. In Figure 48 the red line for Female is higher than Male. This means that more females are below<sup>14</sup> any exposure value than men are.

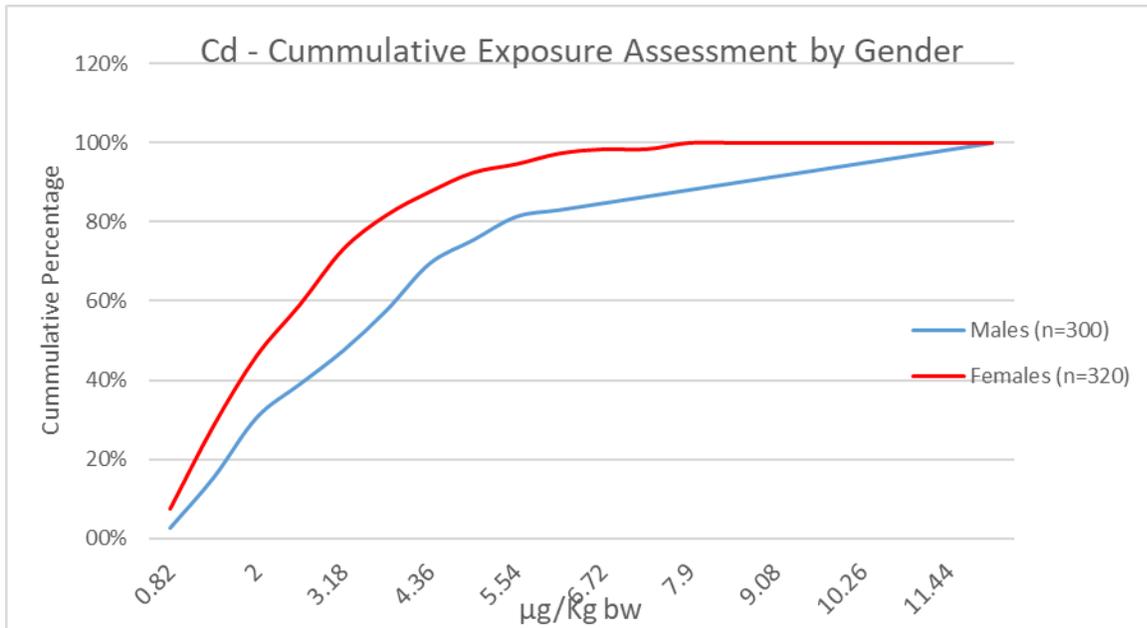
In general, the higher the curve, the better off.

Figure 51: Example of a histogram



<sup>14</sup> Remember that the cumulative is a running total – it is the percentage UP to and including a particular exposure

Figure 52: Cumulative probability across Gender



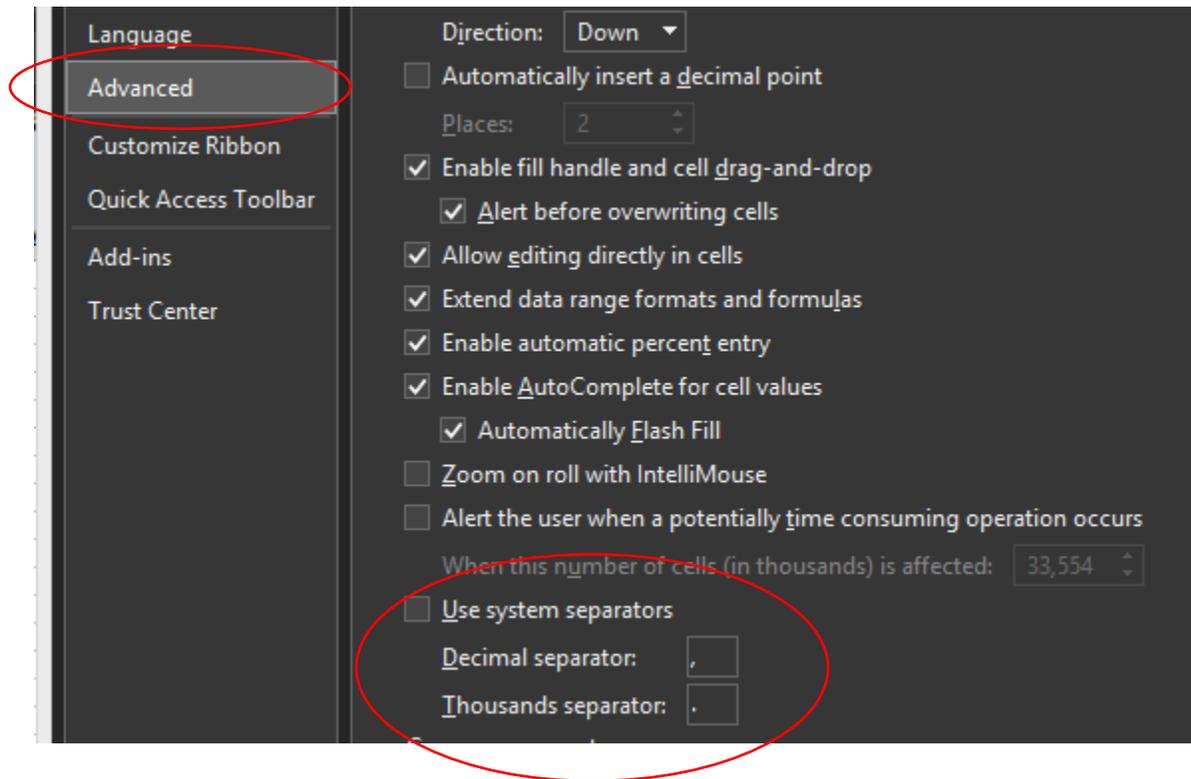
## Decimal and Thousand separator

Whatever the decimal or thousand separators were during building the ImproRisk, when opening the model values will adjust to your local settings.

If the user wishes to change the decimal and thousand separators in Excel, the user should do the following:

Go to **File** -> move down and click on **Options**-> move down and click on **Advanced**-> On the right panel look for the "Use system separators section" (Figure 49).

Figure 53: Changing the decimal and thousand separators in Excel



## Pop Up Menu for quick navigation across worksheets

Using VBA code, a Pop-Up Menu is created that enables the user to quickly navigate through the worksheets.

The Pop-Up Menu is activated by either of the following ways:

### 1. Press **CTRL-m**

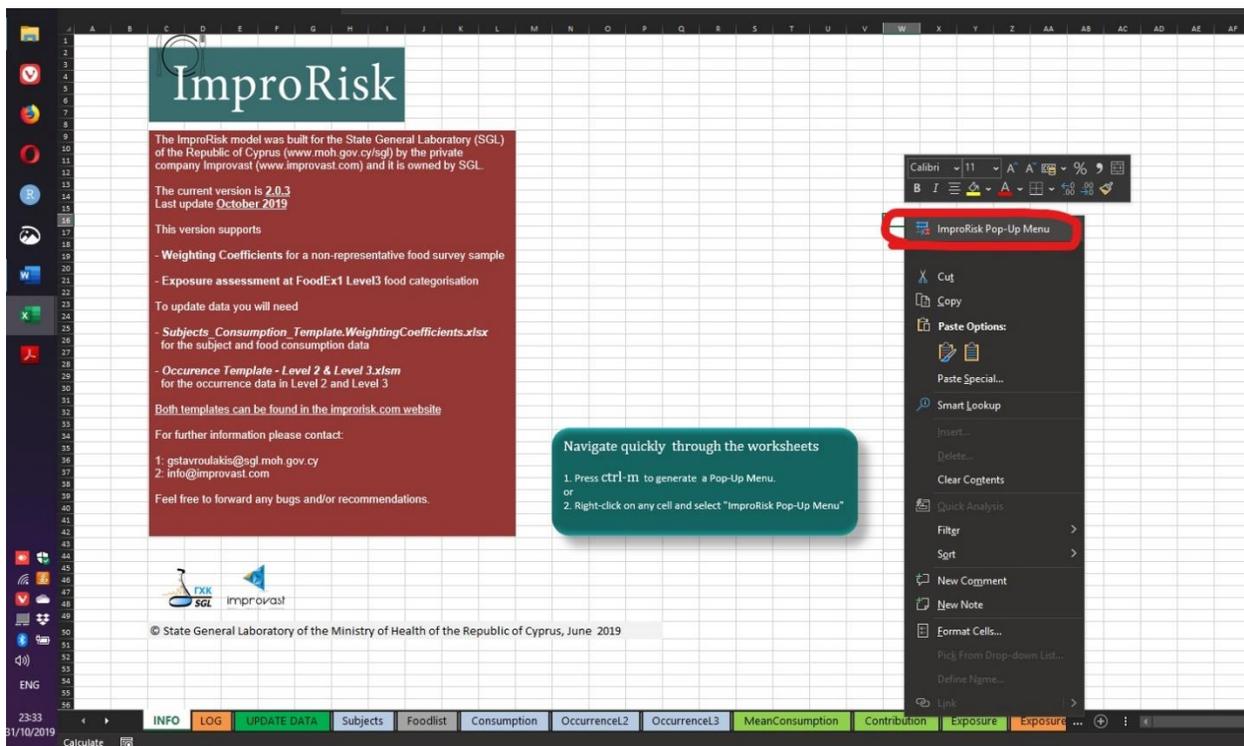
The Pop-Up menu will appear. The user can select any worksheet and will be redirected to it.

### 2. **Right-click** on any cell and select "ImprovRisk Pop-Up Menu"

Then the Pop-Up menu will appear

Both ways can be applied at any time within any worksheet.

Figure 54 Pop Up Menu for navigating the ImproRisk worksheets (Right-click on any cell within any worksheet)



## References

1. Freedman, David; Diaconis, Persi (December 1981). "On the histogram as a density estimator: L2 theory" *Probability Theory and Related Fields* (Heidelberg: Springer Berlin) 57 (4): 453–476.
2. Scott, D. (1992). *Multivariate density estimation*. New York: Wiley.
3. STATISTICS CANADA. Cumulative percentage. [online] Statcan.gc.ca. Available at: <http://www.statcan.gc.ca/edu/power-pouvoir/ch10/5214864-eng.htm> [Accessed 10 May 2016].
4. Cohen, J. (1969) *Statistical Power Analysis for the Behavioral Sciences*. NY: Academic Press
5. Coe, R. (2002). It's the Effect Size, Stupid. What effect size is and why it is important. Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England
6. University, Penn State. Comparing Two Independent Means. [online] STAT 200 Elementary Statistics. Available at: <https://onlinecourses.science.psu.edu/stat200/node/60> [Accessed 5 May 2016].
7. European Food Safety Authority; Evaluation of the FoodEx, the food classification system applied to the development of the EFSA Comprehensive European Food Consumption Database. *EFSA Journal* 2011; 9(3):1970. [http://www.efsa.europa.eu/sites/default/files/scientific\\_output/files/main\\_documents/1970.pdf](http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/1970.pdf)
8. European Food Safety Authority; Use of the EFSA Comprehensive European Food Consumption Database in Exposure Assessment. *EFSA Journal* 2011;9(3):2097. [http://www.efsa.europa.eu/sites/default/files/scientific\\_output/files/main\\_documents/2097.pdf](http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/2097.pdf)
9. European Food Safety Authority; Cadmium dietary exposure in the European population. *EFSA Journal* 2012;10(1):2551 [http://www.efsa.europa.eu/sites/default/files/scientific\\_output/files/main\\_documents/2551.pdf](http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/2551.pdf)
10. European Food Safety Authority; Lead dietary exposure in the European population. *EFSA Journal* 2012;10(7):2831. [59 pp.] [http://www.efsa.europa.eu/sites/default/files/scientific\\_output/files/main\\_documents/2831.pdf](http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/2831.pdf)