ProZES

ProZES v. 2.3c „Summer 2016“ (02.06.2016)

HELP

INTRODUCTION

The program ProZES (Program for calculation of probability of association between disease and radiation exposure) is developed to serve as a support tool for an expert judgement in legal considerations of compensation claims for diseases following preceding occupational exposure(s). The program is a tool to estimate individualized assigned share of radiation in probability to develop cancer after occupational exposure.

The probability that the occurrence of cancer is caused by radiation is given by the assigned share, or probability of causation Z:

$$Z = \frac{\text{excess rate}}{\text{baseline rate} + \text{excess rate}} .$$

The excess and baseline rates of a disease are estimated using cancer incidence data specific to the target population and radiation risk models, originated from epidemiological studies, which estimates of radiation risk are transferred to the target population. The risk models are mostly based on the cohort of A-bomb survivors in Hiroshima and Nagasaki (the LSS cohort). A calculation of personalised estimate of Z requires also to take into account age of cancer occurrence, exposure history, calendar year and other potential individual factors such as e.g. smoking behaviour. More details on the scientific background and the underlying risk models can be found in a separate technical report.

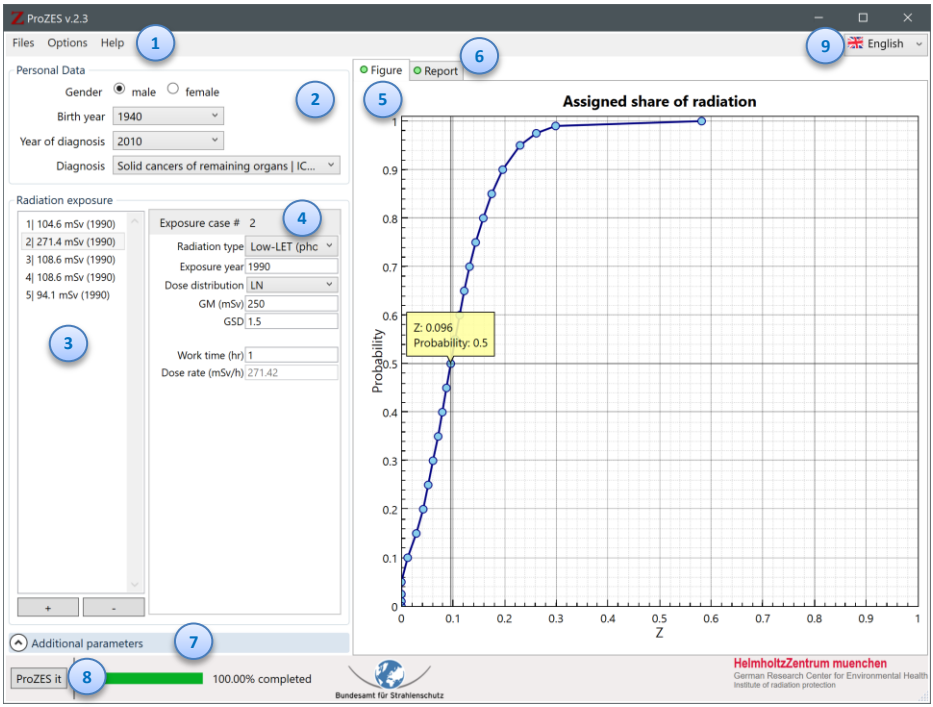
In the present version, the target population is the German population in the period after 1999. Correspondingly, the German-specific disease incidence rates and demographic data are used.

Due to the immanent stochastic nature of cancer and inherent uncertainties of the risk modelling, the estimates of assigned share are probabilistic and are expressed themselves as probability distributions (see example plot in the following screenshot). Therefore, distributions of the assigned share Z are generated in ProZES by Monte Carlo method taking into account uncertainties and correlations of model parameters and of various stochastic correction factors attributed to effects of minimal latency, transfer of risk, dose rates, and others.

To perform calculations, the user should specify diagnosis as well as some individual-specific data: sex, birth year, year of diagnosis. Additionally, history of occupational exposure should be specified accounting for time, duration, radiation type, dose and related uncertainties.

On the below shown screenshot, the following interface elements are indicated:
1 – Main menu
2 – Panel for personal data selection
3 – List of radiation exposures (showing number, mean dose, year)

4 – Input form for the selected exposure case
5 – Plot tab (shown)
6 – Report tab (hidden)
7 – Foldable panel with additional parameters
8 – Button to start computations
9 – Language selection



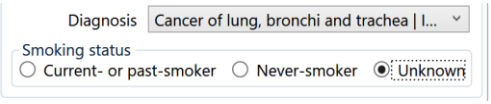
INPUT

To perform risk calculations the program requires corresponding input data. These include data about the individual, the disease, and the radiation exposure(s).

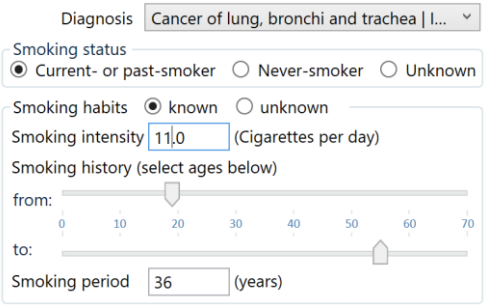
Individual data

Necessary data are: gender, birth year, year of diagnosis. Gender is selected using radiobuttons, the birth year and the year of diagnosis are selected using the dropdown lists. The range of allowed birth years spans from 1900 to the current year, the year of diagnosis spans from 1999 to the current year. The latter is due to constraints of the used database of cancer incidence in Germany, which is based on data of ZfKD (German Centre for Cancer Registry Data at RKI in Berlin-Tempelhof). The most recent incidence data available there are for 2012 (as at the time of release of the program version “Spring 2016”); therefore, if diagnosis year is 2013 or later, then the program does calculations using the most recent data available. If this is the case, then the warning message appears in the report. The same procedure is applied to population data, which are currently available up to 2014.

Additional data need to be specified if lung cancer is selected as a diagnosis. In this case an additional panel appears (see below), which suggest defining individual-specific smoking habits. If option ‘unknown’ is selected, then the smoking habits are randomly sampled during simulations using country-wide information on smoking habits accounting for individual’s gender and age.



If the individual smoking behavior is known and the individual is a life-long non-smoker (‘Never-smoker’ option is selected), then no additional data are required for the program to run. In case the individual is current- or past-smoker, the program shows the panel with parameters characterizing the individual’s smoking habits (see below). If these are unknown (the radio button ‘unknown’ is checked), then the program uses average smoking parameters based on country-wide statistics of smoking habits. Otherwise (the radio button ‘known’ is checked), the user can input the following data: smoking intensity in cigarettes per day, age of starting and age of quitting smoking. Smoking intensity is specified using designated numerical input field, ages are selected using dual slider. The smoking period is calculated and displayed automatically.



Diagnosis

To specify diagnosis, the user selects appropriate individual cancer or group of cancers in the drop-down list ‘Diagnosis’ in the ‘Personal data’ section.

The program implements models of radiation risk for almost all types of primary cancers. Some most frequent types of malignancies are modelled using dedicated models, derived for these specific diagnoses, e.g. cancers of colon, stomach, skin, female breast, cervix and thyroid.

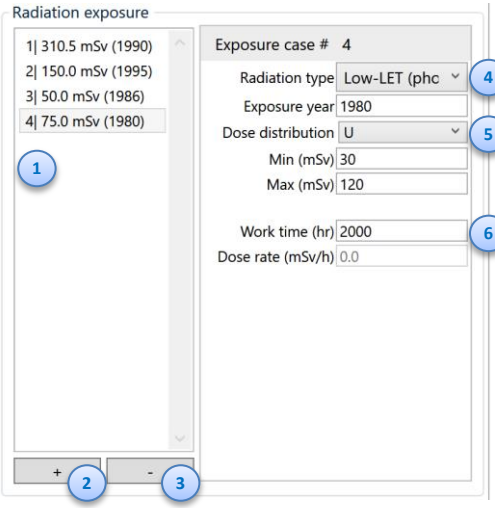
The other types of malignancies are represented by group models. There are seven groups for solid cancers and four groups for cancers of hematopoietic system (‘liquid’ cancers). The cancer groups, individual organs, group codes and disease codes according to ICD10 classifications are shown in the table below.

Type	Model	Organ or Organ Group	Group code	ICD10-code(s)
Solid cancers	**STOMACH**	Stomach cancer	16	C16
	COLON	Colon cancer	18	C18
	LUNG	Lung cancer (incl.trachea)	34	C33,C34
	BREAST	Breast cancer	50	C50
	THYROID	Thyroid cancer	73	C73
	DIG	Cancer of the oral cavity	101	C00-C14
		Esophageal cancer		C15
		Small intestine cancer		C17
		Rectum cancer		C19-C21
		Liver cancer		C22
		Cancer of gallbladder, etc		C23,C24
		Pancreatic cancer		C25
		Other cancers of the digestive system		C26,C48
	REM	Cancer of nasal cavity, etc	102	C30,C31
		Larynx cancer		C32
		Thymus cancer		C37
		Cancer of heart or other intrathoracic organs		C38,C39
		Bone cancer		C40,C41
		Malignant melanoma		C43
		Connective tissue cancer		C47,C49
		Testis cancer		C62
		Adrenal gland cancer		C74
		Cancer of other or unspecified endocrine glands		C75,C76
	GNF1	Uterine cancer/cervix	1031	C53
	GNF2	Uterine cancer/corpus	1032	C54
		Uterine cancer/NOS		C55
		Ovarian cancer, etc		C56
		Other female genital cancer		C51,C52,C57,C58
	GNM	Prostate cancer	104	C61
		Other male genital cancer		C60,C63
	URI	Kidney cancer	105	C64
		Renal pelvis & ureter cancer		C65,C66
		Urinary bladder cancer		C67
		Other urinary system cancer		C68
	BCNS	Cancer of eyes	106	C69
		Cancer of Central nervous system		C70-C72
	SKIN	Non-melanoma skin cancer	107	C44
Hematopoietic cancers	**L1**	Acute lymphoblastic leukemia (ALL)	201	C91.0
		Prolymphocytic leukemia of B-cell type		C91.3
		Lymphoid leukemia/unspecified		C91.9
	L2	Hodgkin disease	202	C81
		Non-Hodgkin disease		C82,C83,C85
		Lymphoma of peripheral and cutaneous T-cell		C84
		Malignant immunoproliferative disease		C88
		Chronic lymphoblastic leukemia (CLL)		C91.1
		Hairy cell leukemia		C91.4
	L3	Acute myeloid leukemia (AML)	203	C92.0
		Sub-acute myeloid leukemia		C92.2
		Myeloid sarcoma		C92.3
		Acute promyelocytic leukemia		C92.4
		Acute myelomonocytic leukemia		C92.5
		Monocytic leukemia		C93
		Other leukemia of specified cell type		C94
		Leukemia of unspecified cell type		C95
		Other or non-specified		C96
	L4	Chronic myeloid leukemia (CML)	204	C92.1

The following diagnoses are not represented in the above table: secondary cancers, multiple localizations, and multiple myeloma. The latter is excluded because of no suitable model of radiation risk could be found from the contemporary A-bomb survival data.

Information on radiation exposures

The interface block for inputting data on radiation exposures consists of two panels. The left panel (see 1 in the below figure) shows a list of input cases of exposure, the right panel shows details of the exposure case, which is currently selected in the left one. The program starts from one default exposure case; additional exposures cases can be added by pressing '+' button (2) in the bottom of the exposure list. Correspondingly, removing items from the exposures list can be done by pressing '−' button (3). Items in the exposure list are generated automatically and show ordinal number of the exposure, arithmetic mean dose and year of the exposure (in parenthesis).



For each exposure case, the following can be specified:

- Radiation type (4)
- Exposure year
- Type of dose distribution (5)
- Parameters of the dose distribution
- Duration of exposure (work time), where relevant (6)

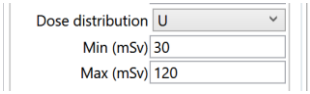
For all diagnoses, low-LET radiation is assumed and the user is expected to input the dose as equivalent dose (mSv). Exposures to high-LET radiation are only considered for lung cancer after exposure to radon, in mines or indoor. For the latter cases, special exposure units are used: working level month (WLM) for occupational exposure in mines and radon concentration in ambient air (Bq m^{-3}) for indoor exposures.

Warning! Z-values for lung cancer after (occupational) exposure to radon in the mines of AG Wismut are only valid for those lung cancer cases who were first employed by AG Wismut from 1960 or later and whose lung cancer was diagnosed within 15 years since the last exposure.

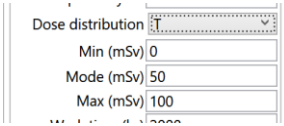
The dose rate window shows dose rate and automatically updates if data are changed by the user. No explicit input of dose rate is necessary.

For each exposure, the exposure dose and related uncertainties must be specified. If no dose uncertainty is known, the normal distribution with standard deviation set to zero (SD=0) can be chosen (see below). The following types of the dose distributions (5) are supported: uniform ('U'), triangle ('T'), normal ('N'), and log-normal ('LN').

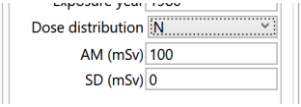
- Uniform distribution 'U': Parameters of the uniform distribution, minimum and maximum values, and their selection are shown in the following screenshot. In this example, doses below 30 mSv or above 120 mSv have zero probability. Between these values each dose is equally likely. The mean dose of this distribution is 75 mSv.



- Triangle distribution 'T': If the triangle distribution is selected, then three parameters have to be defined: minimum, mode and maximum values (see below). Doses below 'Min' or above 'Max' have zero probability. The probability increases from zero at 'Min' dose linearly to a maximum at 'Mode' dose, then linearly decreases back to zero at 'Max' dose.

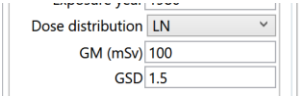


- Normal distribution 'N': Selection of normal (Gaussian) distribution requires specifying two parameters: arithmetic mean (AM) and standard deviation (SD), as shown in the following screenshot.



  Also, the normal distribution can be used to specify a fixed value of dose without uncertainty. For this, zero value of the standard deviation should be selected (SD = 0 mSv, as shown above).

- Log-normal distribution 'LN': The Log-normal distribution is specified via geometric mean (GM) and geometric standard deviation (GSD) as shown in the following screenshot.



  Notice that GSD is dimensionless and cannot be less than or equal to one.

The dose values cannot be negative; therefore, the program automatically converts negative input values to positive ones.

The working time (6) parameter affects risk estimates for low-LET exposure and for exposure to indoor radon. However, the effect of this parameter is different:

- For low-LET radiation, the working time is used to calculate an average dose rate, which is used to quantify a factor affecting risk estimates at low dose rates (dose rate effectiveness factor). For dose rates less than 6 mSv h^{-1}, additional uncertainty of cancer risk estimates is taken into account, resulting in a wider distribution of Z. For larger dose rates, however, there is no influence on the risk estimates, and the distribution of Z remains unaffected.
- For high-LET exposure to radon in indoor ambient air, the whole exposure dose is defined as a product of radon concentration in air (Bq m^{-3}) and the duration of exposure expressed in working hours. Correspondingly, the total exposure to indoor radon is proportional to the duration of exposure. Therefore, this parameter becomes highly influential to the risk estimate and needs to be sufficiently justified in order to obtain plausible estimates of risk. For example, an average annual working time in Germany in the period 2000–2014 accounted for approximately 1408 hours (OECD.stat).

Saving and loading the input data

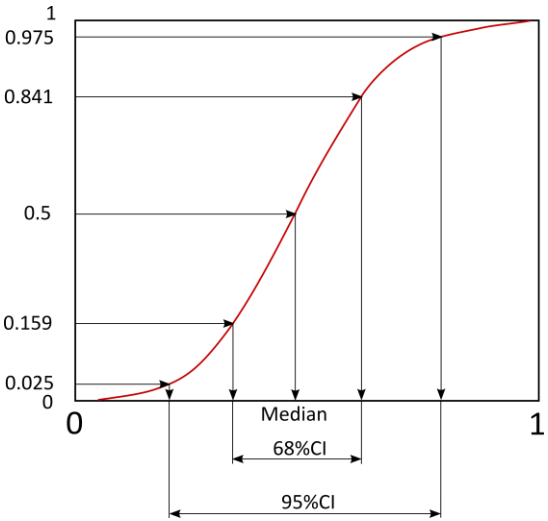
The input data created by the user can be saved to external files for later use. The files are saved as Excel spreadsheets and appear as follows:

Parameter	Value	Comment		Additional data for lung cancer (ICD10:C34)			
Gender	M	(M - male, F - female)		Parameter	Value	Comment	
Birth year	1940			Smoking status		(unknown: -1; never-smoker: 0; ever-smoker: 1)	
Year of diagnosis	2010			Smoking habits		(unknown: 0; known: 1)	
Cancer code	16			Smoking intensity		[cigaretts per day]	
ICD10 code	ICD10:C16	optional		Age smoking began		[age]	
Diagnosis	Stomach cancer	optional		Age smoking quit		[age]	
Exposure #	**Radiation type**	**Exposure year**	**Duration (work-hr)**	**Dose distribution**	**1st parameter**	**2nd parameter**	**3rd parameter**
1	0	1990	1	N	100	0	0
[Stop]							

The input files can be also prepared directly using software, which supports export to Excel 2007 and above format (file extension .xlsx). When preparing input files externally, the user should specify cancer code using values “Group code” from the diagnoses table shown above. The parameters “ICD10 code” and “Diagnosis” are not obligatory for input and do not need to be specified by the user manually. Also, types of radiation exposure are coded as “0” for low-LET radiation, “1” for exposure to radon in mines, and “2” for indoor exposure to radon in air. Types of the dose distributions are coded by their abbreviated names: “U” (Uniform), “T” (Triangle), “N” (Normal), and “LN” (Log-Normal). Meaning of the dose distribution parameters (1^{st}, 2^{nd}, 3^{rd}) is defined by the type of the respective distribution and corresponds to those shown above in the sub-section “Information on radiation exposures”.

OUTPUT

Pressing the button ‘ProZES it’ starts generation of assigned share distribution. The progress of computations is displayed by an indicator right from the button. The result of calculations is the distribution of assigned share (probability of causation) and is reported as a cumulative probability distribution of Z (see illustration below).

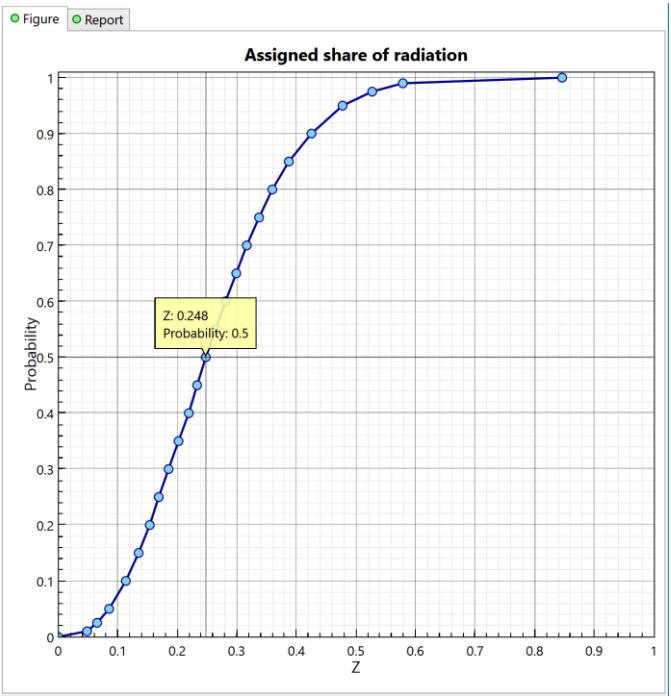


The generated distribution of Z is presented as (a) a plot of cumulative probability distribution and (b) a report with detailed information on the processed case and resulting statistics: median value and confidence intervals. Both tabs for the report and the plot have indicators to show whether they are up-to-date (green, Figure Report) or the user made changes in the input data and the program has to be re-run to update both (red, Figure Report).

Plot of assigned share

An example of the plot of assigned share Z appears below. This is a cumulative probability distribution of assigned share drawn using the percentiles of the generated Z-distribution. The points represent the single percentile values. Clicking on the point in the plot will result in appearance of a tooltip window showing exact numerical values for a given percentile. In the following screenshot, the point corresponding to median of the distribution (50%) is clicked and the tooltip shows the median value of Z equal to 0.248. This means that with 50%-probability Z is lower than or equal to 0.248, and with 50%-probability Z is higher than 0.248. In this sense, Z=0.248 could be considered as a 'central value', with a probability of 24.8% that this cancer was caused by radiation. Correspondingly, clicking other points in the plot will display a tooltip with Z-value for the respective percentile, e.g. the 97.5%-percentile, which is an upper bound of the two-sided 95% confidence interval. That is, the related Z-value, 0.53, means that 2.5% of the generated assigned share values are larger than 53%.

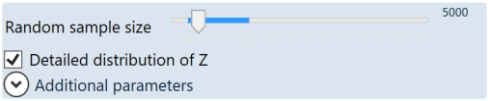
The plot can be saved as bitmap image in PNG-format. This is achieved via menu choice ‘Option-Save plot’. If the user changed input parameters and the plot is not up-to-date, then the menu item has red color and shows a warning in a tooltip.



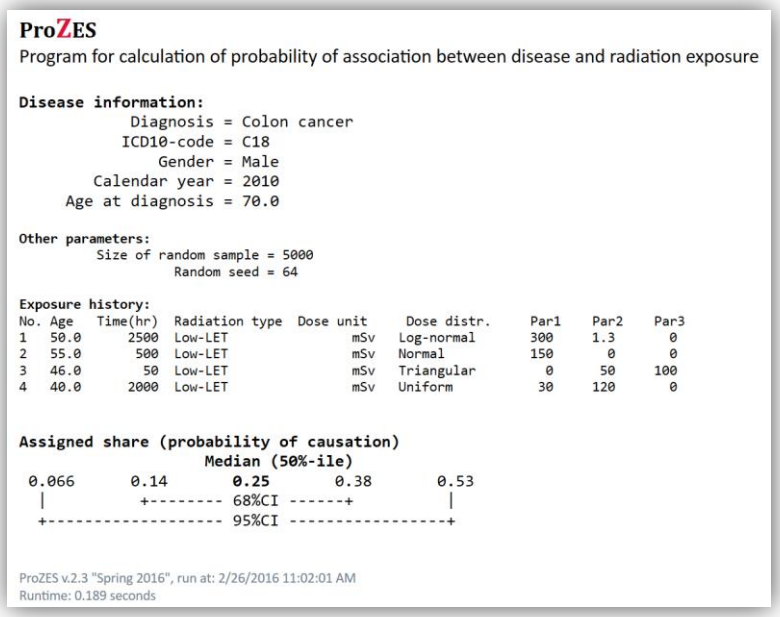
Report

The report with the computed results provides detailed information on all parameters used: individual data, exposure history, and parameters. An example of the output report appears below. The computed assigned share is presented by its median value and two confidence intervals: 68% (“1-sigma” CI) and 95% (“2-sigma” CI).

Optionally, the user can request output of detailed Z-distribution with numerical values of percentiles as shown in the plot window. This can be achieved by checking the box ‘Detailed distribution of Z’ in the foldable panel ‘Additional parameters’ (see below).



The report can be saved as an electronic document in PDF format. This can be achieved via menu choice ‘Options-Save report’. Similarly to saving the plot, if the report is not up-to-date, then the menu item is red and a warning appears in the tooltip for this menu element.



ProZES

Program for calculation of probability of association between disease and radiation exposure

Disease information:

```
            Diagnosis = Colon cancer
           ICD10-code = C18
               Gender = Male
        Calendar year = 2010
     Age at diagnosis = 70.0
```

Other parameters:

```
         Size of random sample = 5000
                   Random seed = 64
```

Exposure history:

No.	Age	Time(hr)	Radiation type	Dose unit	Dose distr.	Par1	Par2	Par3
1	50.0	2500	Low-LET	mSv	Log-normal	300	1.3	0
2	55.0	500	Low-LET	mSv	Normal	150	0	0
3	46.0	50	Low-LET	mSv	Triangular	0	50	100
4	40.0	2000	Low-LET	mSv	Uniform	30	120	0

Assigned share (probability of causation)

```
                    Median (50%-ile)
 0.066       0.14        0.25        0.38        0.53
   |          +--------- 68%CI ------+           |
   +-------------------- 95%CI ------------------+
```

ProZES v.2.3 "Spring 2016", run at: 2/26/2016 11:02:01 AM
Runtime: 0.189 seconds

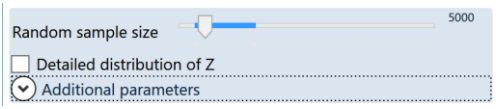
ADDITIONAL PARAMETERS

Random seed

The program initiates a random number generator with the same default seed number at every start. This assures the same output results given the same input parameters. However, in some occasions the user might wish to initiate random number generator with another seed value in order to generate independent realization of probability distribution using the different sequence of pseudo-random numbers. This can be achieved via dialog window, which can be invoked by selecting menu item ‘Option-Random seed’.

Sample size

The default size of the generated sample equals to 5000. The user can change it by moving slider in the foldable panel ‘Additional parameters’ (see below). The sample size can be varied from 10 to 50000 with the recommended sample size larger from 3000. For the sample size larger than 15000 significant changes in the results are unlikely.



Detailed distribution of Z

In the panel ‘Additional parameters’, there is a checkbox ‘Detailed distribution of Z’. Checking it will result in adding a table of detailed percentiles of Z to the report.

GENERAL INFORMATION

Copyright © 2009–2016 Helmholtz Zentrum München and German Federal Office for Radiation Protection. All Rights Reserved.

LICENSE AGREEMENT

Redistribution and use of the Software in source forms, with or without modification, are NOT permitted. Redistribution and use of Software in binary forms without modification are permitted provided that the following conditions are met:

1. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

2. The names of the German Federal Office for Radiation Protection and of Helmholtz Zentrum München may not be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE LICENSOR "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE LICENSOR OR THE AUTHOR BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

THIRD-PARTY components used in the program:

OxyPlot library (oxyplot.org) covered by MIT license (see: http://opensource.org/licenses/MIT)

FlexCel Studio for .NET from TMS Software (http://www.tmssoftware.com/site/flexcelnet.asp) provided under single developer license (see: http://www.tmssoftware.com/site/lic_single2y.asp)

ADDRESS for sending comments and/or bug reports:

German Federal Office for Radiation Protection, Department of Radiation Protection and Health, Ingolstädter Landstraße 1, D-85764 Oberschleißheim, Germany
ProZES@bfs.de