



# SUMRAY: R and Python Codes for Calculating Cancer Risk Due to Radiation Exposure of a Population

Michiya Sasaki<sup>1</sup>, Kyoji Furukawa<sup>2</sup>, Daiki Satoh<sup>3</sup>, Kazumasa Shimada<sup>3</sup>, Shin'ichi Kudo<sup>4</sup>, Shunji Takagi<sup>5</sup>, Shogo Takahara<sup>3</sup>, Michiaki Kai<sup>6</sup>

<sup>1</sup>Central Research Institute of Electric Power Industry, Abiko, Japan; <sup>2</sup>Biostatistics Center, Kurume University, Kurume, Japan; <sup>3</sup>Japan Atomic Energy Agency, Tokaimura, Japan; <sup>4</sup>Radiation Effects Association, Chiyoda, Japan; <sup>5</sup>Mitsubishi Research Institute, Chiyoda, Japan; <sup>6</sup>Department of Health Science, Nippon Bunri University, Oita, Japan

## ABSTRACT

**Background:** Quantitative risk assessments should be accompanied by uncertainty analyses of the risk models employed in the calculations. In this study, we aim to develop a computational code named SUMRAY for use in cancer risk projections from radiation exposure taking into account uncertainties. We also aim to make SUMRAY publicly available as a resource for further improvement of risk projection.

**Materials and Methods:** SUMRAY has two versions of code written in R and Python. The risk models used in SUMRAY for all-solid-cancer mortality and incidence were those published in the Life Span Study of a cohort of the atomic bomb survivors in Hiroshima and Nagasaki. The confidence intervals associated with the evaluated risks were derived by propagating the statistical uncertainties in the risk model parameter estimates by the Monte Carlo method.

**Results and Discussion:** SUMRAY was used to calculate the lifetime or time-integrated attributable risks of cancer under an exposure scenario (baseline rates, dose[s], age[s] at exposure, age at the end of follow-up, sex) specified by the user. The results were compared with those calculated using another well-known web-based tool, Radiation Risk Assessment Tool (RadRAT; National Institutes of Health), and showed a reasonable agreement within the estimated confidential interval. Compared with RadRAT, SUMRAY can be used for a wide range of applications, as it allows the risk projection with arbitrarily specified risk models and/or population reference data.

**Conclusion:** The reliabilities of SUMRAY with the present risk-model parameters and their variance-covariance matrices were verified by comparing them with those of the other codes. The SUMRAY code is distributed to the public as an open-source code under the Massachusetts Institute of Technology license.

**Keywords:** R, Python, Radiation Exposure, Lifetime Attributable Risk, Risk Model, Variance-Covariance Matrix

## Original Research

**Received** December 20, 2022

**Revision** February 22, 2023

**Accepted** March 5, 2023

**Corresponding author:** Michiya Sasaki

Central Research Institute of Electric Power Industry, 1646 Abiko, Abiko-shi, Chiba 270-1194, Japan

E-mail: [michiya@criepi.denken.or.jp](mailto:michiya@criepi.denken.or.jp)

<https://orcid.org/0000-0002-0444-5459>

This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2023 The Korean Association for Radiation Protection

## Introduction

Estimated lifetime risks of cancer mortality or incidence attributable to exposure to ionizing radiation have been repeatedly included in major reports of international committees including United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR) [1] and International Commission on Radiological Protection

(ICRP) [2, 3]. These risk calculations mainly adopt risk models with parameters obtained by regression analyses of data from the Life Span Study (LSS) of the atomic bomb survivors in Hiroshima and Nagasaki [4] and the baseline reference data (hereinafter the baseline risk) for a population on cancer incidence or mortality and all-cause mortality. The risk model with estimated parameters from the LSS may be applied to the projection of cancer risks at low doses and low-dose rates, with an additional assumption about low-dose or low-dose-rate extrapolation.

Risk projections have been conducted for various purposes: risk projection for an actual population under specific scenarios of medical, occupational, and environmental exposures [5–7]; sensitivity analysis of risk estimates [8]; and development of methodologies of risk assessment [9]. On the other hand, calculational tools nowadays are publicly available; thus, one can assess outcomes calculated using given risk models, assumptions, and statistical data. For example, Radiation Risk Assessment Tool (RadRAT; National Institutes of Health) [10], developed by the National Cancer Institute (NCI) at the National Institutes of Health of the United States, is a web-based risk calculator for estimating the lifetime risk of cancer incidence for populations of selected countries including the United States and Japan as a result of exposure to ionizing radiation at doses below 1 Gy and is based on the risk models mainly developed in Biologic Effects of Ionizing Radiation (BEIR)-VII [11]. LARisk (<https://cran.r-project.org>) [12] is an R package (R Foundation for Statistical Computing) [13] for computing the risk using methods and models similar to those in RadRAT. Nevertheless, we emphasize that there is no open-source code applicable to risk calculations that can be directly traced with the calculation algorithm, model parameters, and numerical database.

Quantitative risk evaluations should accompany comprehensive assessments of sources of uncertainty involved in the evaluated risks (UNSCEAR 2019 report, Annex A [14]). The latest National Aeronautics and Space Administration (NASA) studies in space flight provided the overall uncertainty of risk projection in terms of fold factors and a range of projection risk values at 95% confidence interval (CI) levels [15]. The main sources of uncertainties in evaluated risks include statistical uncertainties of the parameter estimates in the risk models, the method of risk transfer applied to populations with different baselines, and the assumption of applying a dose and dose rate effectiveness factor (DDREF) when projecting an LSS risk model to low-dose or low-dose-

rate exposure scenarios. The epidemiological data from the atomic bomb survivors is the primary data on which most of the current risk projections are based. The cohort follow-up study that began in the 1950s is still ongoing, and the current risk projection covered by an international organization is the incidence rate data reported in 2007 [4]. Either the excess absolute risk (EAR) or excess relative risk (ERR) model is commonly used with effect modifications by sex, age at exposure, and attained age [1, 2]. The calculated risk measures are expressed by lifetime cancer probability (death, incidence) and life expectancy loss [1]. The ICRP uses detriment as a risk indicator for defining tissue weighting factors [2, 3].

With this background in mind, the present study aims to provide a computational code for risk projection and make the code publicly available as a resource for further improvement of risk projection. The Japan Health Physics Society (JHPS) approved to set up a research group tasked with the development of an estimation code for cancer risk associated with radiation exposure. The research group has developed a computational code named SUMRAY to encourage quantitative discussion on cancer risk and its uncertainty following radiation exposure at low doses and low dose rates. To disseminate the code widely, SUMRAY is released in two popular computer languages, that is, R and Python. The R-version SUMRAY is given as an R package called in the user's own R script to calculate the risk probability, and the Python version is dedicated to calculations without any code written by users using American Standard Code for Information Interchange (ASCII) input and output files.

The distribution of the code was also discussed, taking into account copyright, maintenance, and ease of dissemination and updating. It was decided that the code is to be distributed under the Massachusetts Institute of Technology license [16] through the GitHub website (<https://github.com/JapanHealthPhysicsSociety/SUMRAY>) upon approval by the JHPS Board.

## Materials and Methods

### 1. Overview

Fig. 1 illustrates an overview of the SUMRAY code system. The system calculates, as a primary risk measure, the cumulative excess risk (CER) of dying from or developing a specific cancer (specified by the risk model) as a result of a radiation exposure scenario (specified by the configuration data) in a population (specified by the baseline risk). The output from

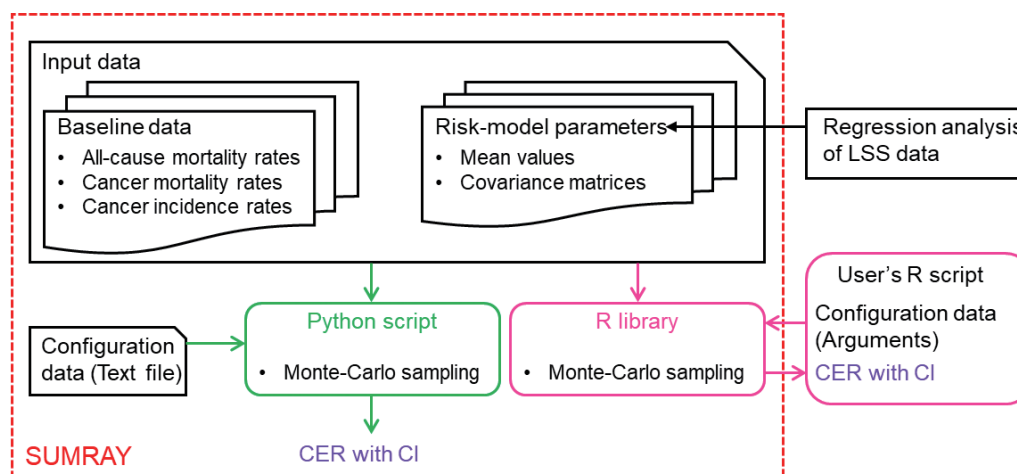


Fig. 1. Overview of the SUMRAY code system. LSS, Life Span Study; CER, cumulative excess risk; CI, confidence interval.

the system includes point estimates as well as interval estimates in terms of 95% CI. The details of input data, configuration data of an exposure scenario, and computational scheme are given in the following subsections. As we discuss the risk models included in the package in this article, the code to estimate a risk model is not included in the SUMRAY package.

In SUMRAY, a calculation part of CER with CI was developed using an R library and a Python script independently to meet the various demands of users in Supplementary Materials and Methods (Supplementary Table S1, Supplementary Figs. S1, S2) ('1. Usage of the R library provided in SUMRAY' and '2. Usage of the Python script provided in SUMRAY'). Both of them employed the same algorithm, but their styles are different. The R library in the SUMRAY package is called from the user's own R script to calculate CER for an exposure scenario configured in that script. The Python script was designed for a stand-alone code for calculating CER in SUMRAY without any programming by users. The users can configure the exposure scenario via a text file, and obtain results as comma-separated values (.csv) files after running the Python script. The usage of the R library and the Python script is provided in the electronic supplements of this article.

Currently, mortality and incidence rates for all solid cancers as a group are targeted for risk projection without individual cancers. Our code can be extended through risk model development and its model fitting to the LSS data since it is desirable to address individual cancers. The risk prediction using the code can be made in cases of acute and chronic exposures without DDREF. The code can reflect the variances of the evaluated risk-model parameters on the risk predictions by a Monte Carlo sampling. The development of the

computational code aims to clarify the process required for risk prediction of the cumulative probabilities of death and morbidity by sex and age. It should be noted; however, that the code may give a false sense of accuracy unless the quantitative uncertainties associated with the overall risk prediction are taken into account.

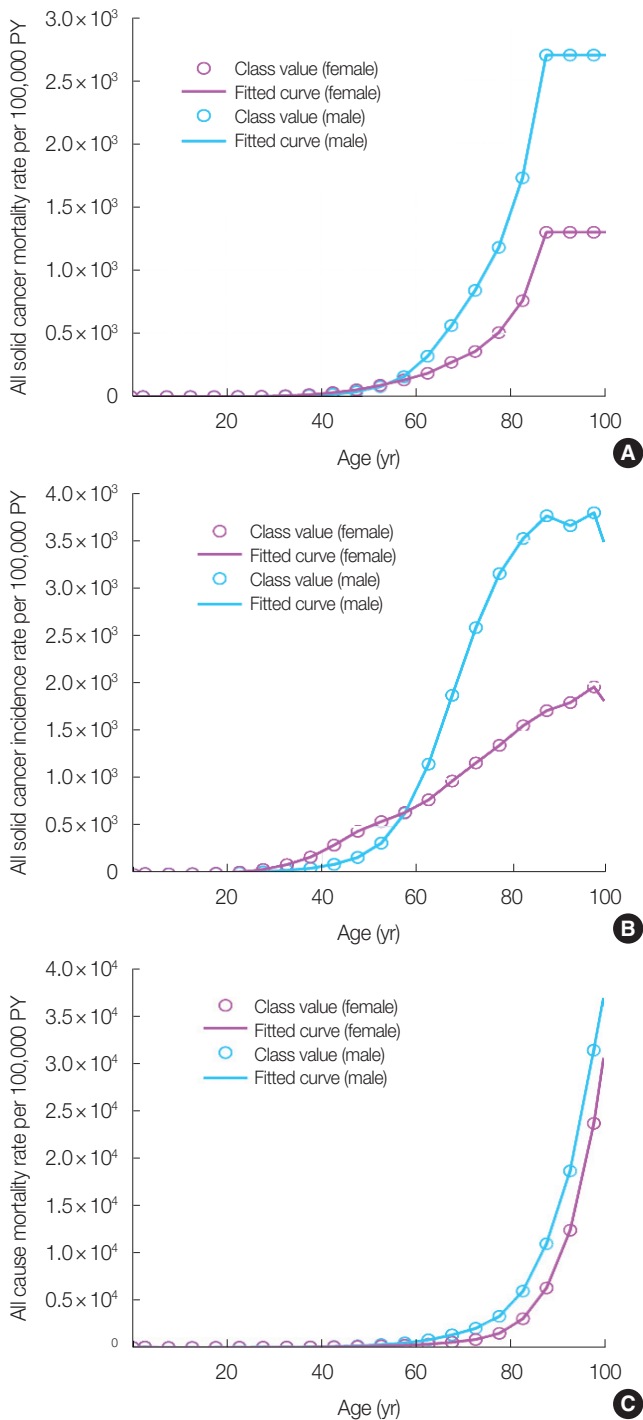
## 2. Baseline Risk

In SUMRAY, risk prediction is made based on the baseline risk of all-solid-cancer mortality and incidence rates  $\lambda_0$  for a target population. The baseline risk defines age-specific cancer mortality and incidence rates and age-specific all-cause mortality.

Default values for the baseline risks in SUMRAY were set to the risks of mortality and incidence from all solid cancers (International Statistical Classification of Diseases and Related Health Problems; International Classification of Diseases 10th Revision, category of cancer from C00 to C80) in the Japanese population in 2018 obtained from the National Cancer Center [17]. These baseline risks were converted to 1-year intervals from 5-year intervals by linear interpolation in SUMRAY. Fig. 2 shows the baseline risks of all-solid-cancer mortality and incidence rates  $\lambda_0$  for the Japanese population in 2018 with fitted curves. The default value of the survival function  $S$  for the target population is obtained from the Japanese statistical data for 2018 [18].

## 3. Equation

The radiation risk models in SUMRAY are the ERR and EAR models, which were created by fitting cancer mortality and incidence risk data from the LSS of atomic bomb survi-



**Fig. 2.** Baseline risks of all-solid-cancer mortality and incidence rates  $\lambda_0$  for the Japanese population in 2018: (A) all-solid-cancer mortality rate, (B) all-solid-cancer incidence rate, and (C) all-cause mortality rate. PY, person-year.

vors in Hiroshima and Nagasaki [4, 19]. The ERR and EAR models for all solid cancers derived from the LSS are used in the same equations in SUMRAY.

$$ERR(d, e, a, s) \text{ or } EAR(d, e, a, s) = (\alpha \cdot d) \cdot \exp\left\{\gamma \cdot \left(\frac{e-30}{10}\right) + \eta \cdot \log\left(\frac{a}{70}\right)\right\} (1 + \nu \cdot s) \quad (1)$$

The variables of the ERR and EAR models are the absorbed dose  $d$  (Gy), age at exposure  $e$ , attained age  $a$ , and sex, where  $s=1$  for males, and  $-1$  for females. The parameters of each model are either the ERR estimate  $\alpha$  (ERR/Gy) or the EAR estimate  $\alpha$  (EAR per 10,000 person-years per Gy) and the effect modifiers of these risk estimates by sex  $\nu$ , age at exposure  $\gamma$ , and attained age  $\eta$ . The variance-covariance matrices were set up for the parameter values ( $\alpha, \gamma, \eta, \nu$ ) of the ERR and EAR models estimated from the published data of LSS using the generalized nonlinear model of the R library. The reason for estimating the variance-covariance matrices of the ERR/EAR models using the R library is that the previous research papers on the LSS [4, 19] only provide the CIs of the ERR/EAR models. In addition, in these previous studies on LSS, Epicure (Risk Science International, Inc.) [20], a commercial statistical software to fit the LSS data and obtain the numerical values of the variance-covariance matrix, was used. Therefore, the variance-covariance matrices of the ERR/EAR models in SUMRAY were estimated using the R library, which is free software. As the default values for the ERR/EAR models in SUMRAY, these parameters of cancer mortality were estimated from the latest LSS data [20] and cancer incidence data (1958 to 1998) [4].

The excess incidence and mortality risks of cancer associated with radiation exposure  $M$  were obtained using either the ERR or EAR model taking into account the baseline cancer incidence and mortality risks  $\lambda_0$  and the latent period of cancer  $F_L$  which is the S-shape function [10]. Here,  $M^{(R)}$  is calculated using the ERR model and  $M^{(A)}$  is calculated using the EAR model.

$$M^{(R)}(d, e, a, s) = \lambda_0(a) \times ERR(d, e, a, s) \times F_L(e, a) \quad (2)$$

or

$$M^{(A)}(d, e, a, s) = EAR(d, e, a, s) \times F_L(e, a) \quad (3)$$

$$F_L(e, a) = [1 + \exp(-(a - e - 7.5))]^{-1} \quad (4)$$

SUMRAY calculates the cumulative excess cancer incidence or mortality risk, CER. For cancer incidence or mortality risk from age at exposure,  $e$  to attained age  $a_{\max}$ , CER is calculated using the following equation.

$$CER(d, e, s) = \int_e^{a_{\max}} M(d, e, a, s) \times S(a|e, s) da \quad (5)$$

Here, the conditional survival function  $S(a|e,s)$  is the probability of surviving to age  $a$ , adjusted for survival with the condition that the probability equals one at the age at exposure ( $e$ ) for the target population [19].

$$S(a|e,s) = \frac{S(a,s)}{S(e,s)} \quad (6)$$

SUMRAY calculates CERs separately based on the ERR and EAR models and then combines these results using a weighted arithmetic mean from the following equation. Here,  $CER^{(R)}$  is calculated from the ERR model and  $CER^{(A)}$  is calculated from the EAR model.

$$CER(d, e, s) = w \times CER^{(R)}(d, e, s) + (1-w) \times CER^{(A)}(d, e, s) \quad (7)$$

where,  $w$  is the weight between the ERR and EAR transfer of risk. Expert choices for the possible weights are available in the code. The weights for estimating risk transfer to a population have been used in radiological protection [2]. Note that the values of DDREF are important in radiation risk assessment. However, SUMRAY does not consider DDREF explicitly. The determination of DDREF is a future issue.

SUMRAY can calculate the radiation risk from chronic and acute exposures. In the case of chronic exposure, we assume that the risk from the annual dose can be integrated over the duration of exposure. The excess cancer incidence and mortality rates after radiation exposure are calculated using the following equation,

$$M(d_e, e, a, s) = \sum_{k=e_1}^{k=e_{\max}} M(d_k, k, a, s) \quad (8)$$

where,  $d_k$  is the absorbed dose at the  $k^{\text{th}}$  exposure accumu-

lated over 1 year from age ( $e_{k-1}$ ) to age  $e_k$ .  $e: [e_1 \dots e_k \dots e_{\max}]$  is a set of ages at exposure.  $d_e$  is a set of the absorbed doses at the set of ages at exposure  $e$ . With the above equation,  $M$  for chronic exposure with dose set  $d_e$  in ages  $e$  is calculated as a function of attained age  $a$ .

CER for chronic exposure is calculated by integrating  $M$  from the first exposure age  $e_1$  to the maximum attained age  $a_{\max}$  with the conditional survival function  $S(a|e_1,s)$  using the following equation.

$$CER(d_e, e, s) = \sum_{e_1}^{a_{\max}} M(d_e, e, a, s) \times S(a|e_1, s) \quad (9)$$

Note that the conditional survival function used to calculate CER is normalized by the age at first exposure ( $e_1$ ).

#### 4. Monte Carlo Sampling

To calculate CER reflecting uncertainties of the estimated risk-model parameters, EAR and ERR given in Equations (2) and (3) were computed repeatedly using parameter values of  $(\alpha_i, \gamma_i, \eta_i, \nu_i)$  at the  $i^{\text{th}}$  Monte Carlo sampling. Fig. 3 shows examples of EAR and ERR curves calculated by the present Monte Carlo method. The values of  $(\alpha_i, \gamma_i, \eta_i, \nu_i)$  were sampled from a multivariate normal distribution around mean values of  $(\alpha, \gamma, \eta, \nu)$  considering their variances and covariances matrix of the maximum likelihood estimates of the parameters of each ERR and EAR model. The size of Monte Carlo sampling is set to  $10^4$  as a default value. Using the simulated EARs and ERRs, CERs were calculated and their CIs were deduced statistically.

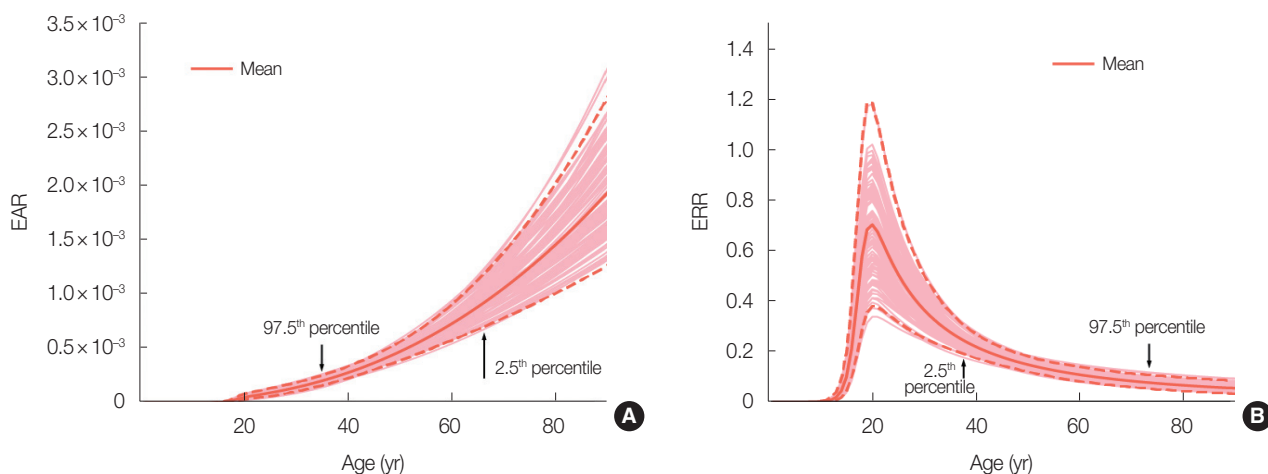


Fig. 3. (A) Excess absolute risk (EAR) and (B) excess relative risk (ERR) curves calculated by Monte Carlo method for female all-solid-cancer incidence from 0.1 Gy acute exposure at 10 years of age.

**Table 1.** Parameter Values and CIs for ERR and EAR Models Estimated Using R Codes and Epicure

Type of cancer (all solid cancers)	Model	Tool	Risk per Gy			Age at exposure (%)			Attained age			Sex		
			$\alpha$	CI <sup>b)</sup>		$\gamma$	CI <sup>b)</sup>		$\eta$	CI <sup>b)</sup>		$\nu$	CI <sup>b)</sup>	
				Min	Max		Min	Max		Min	Max		Min	Max
Incidence: Preston et al. (2007) [4]	ERR	R	0.40	0.38	0.52	-18.9	-30.7	-7.7	-1.74	-2.26	-1.22	1.67	1.26	2.25
		Epicure	0.47	0.40	0.54	-17.0	-25.0	-7.0	-1.65	-2.10	-1.20	1.60	1.31	2.09
	EAR <sup>a)</sup>	R	50.50	42.27	59.30	-28.7	-42.3	-15.8	2.37	1.84	2.93	1.41	1.06	1.94
		Epicure	52.00	43.00	60.00	-24.0	-32.0	-16.0	2.38	1.90	2.80	1.40	1.10	1.79
Mortality: Ozasa et al. (2012) [19]	ERR	R	0.41	0.32	0.51	-33.8	-49.6	-17.9	-0.93	-1.75	-0.12	2.03	1.40	3.11
		Epicure	0.42	0.32	0.53	-29.2	-40.9	-17.4	-0.86	-1.60	-0.06	2.05	1.42	3.13
	EAR <sup>a)</sup>	R	26.40	20.21	32.53	-21.6	-35.7	-7.5	3.40	2.68	4.12	1.14	0.78	1.69
		Epicure	26.40	20.30	32.77	-19.2	-30.8	-7.4	3.38	2.75	4.08	1.10	0.80	1.74

CI, confidence interval; ERR, excess relative risk; EAR, excess absolute risk.

<sup>a)</sup>The EAR per Gy shows the excess cases per 10,000 person-years per Gy.

<sup>b)</sup>The percentiles of CIs are 90% for incidence and 95% for mortality.

## Results and Discussion

All results in this section were calculated with the default reference data provided in the SUMRAY package for the risk model (LSS R14 all-solid-cancer model), baseline risk (all-solid-cancer mortality rates in Japan 2018 [17]), and mortality data (all-cause mortality rates in Japan 2018 [18]). In addition, we employed the value of 7.5 years for the parameter of a latent-period in Equation (4).

### 1. Parameters of ERR/EAR Models

Table 1 shows the ERR and EAR parameter values and CIs used for these default values. Table 1 also shows the parameter values estimated using the Epicure software in previous studies [4, 19] to demonstrate the validity of the ERR/EAR parameter values estimated using the R code. Here, the CIs for mortality and incidence are shown as 95% and 90%, respectively, for comparison with those obtained in previous studies [4, 19] estimated using the Epicure software. The CIs for the ERR and EAR parameters are the uncertainties associated with nonlinear regression analysis. Table 1 shows that the parameter values and CIs estimated using the R code were not significantly different from those estimated using the Epicure software. Hence, the validity of the ERR and EAR parameter values estimated by the R code was confirmed by comparing them with those obtained in previous studies [4, 19].

### 2. Acute Exposure

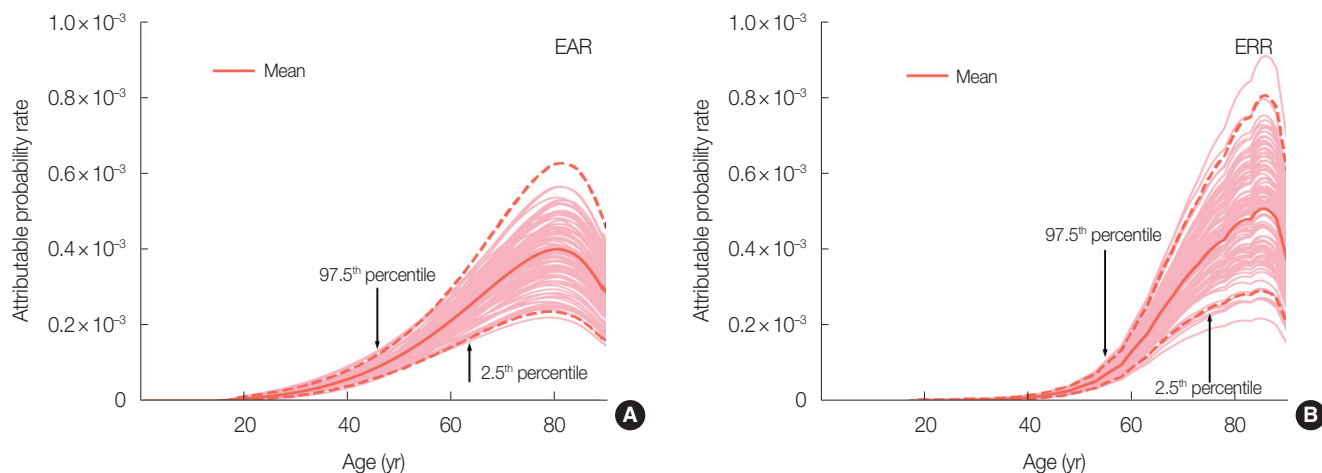
The Supplementary Materials and Methods (3. Acute exposure) shows how to compute the cumulated excess risk for males exposed to 0.1 Gy at age 10 followed up until age 90 with ERR transfer using CER of the R library.

The mean, median, and CI of cumulated excess risk are shown after executing CER. The CI means a 95% CI between the upper and lower percentiles. The default of the weight is 1:0 for ERR and EAR transfers, and means that CER uses the result of only the ERR model. “maxage” indicates the final age at which a cumulative risk calculation is performed. This example shows a cumulative risk until the age of 90. In this example, the lifetime risk until the age of 90 indicates 1.2% (95% CI, 0.75% to 1.8%). The output using plotCER is shown in Fig. 4 to provide the attributable probability rate at an attained age. From the sum of these values, the cumulative risk can be calculated from the age at acute exposure to the age of 90. In the results obtained using the ERR model, there was an irregular variation of attributable probability rates between these ages following the linear interpolation. The absence of variation in the results obtained using the EAR model is due to the fact that the estimated excess risk is independent of the baseline risk.

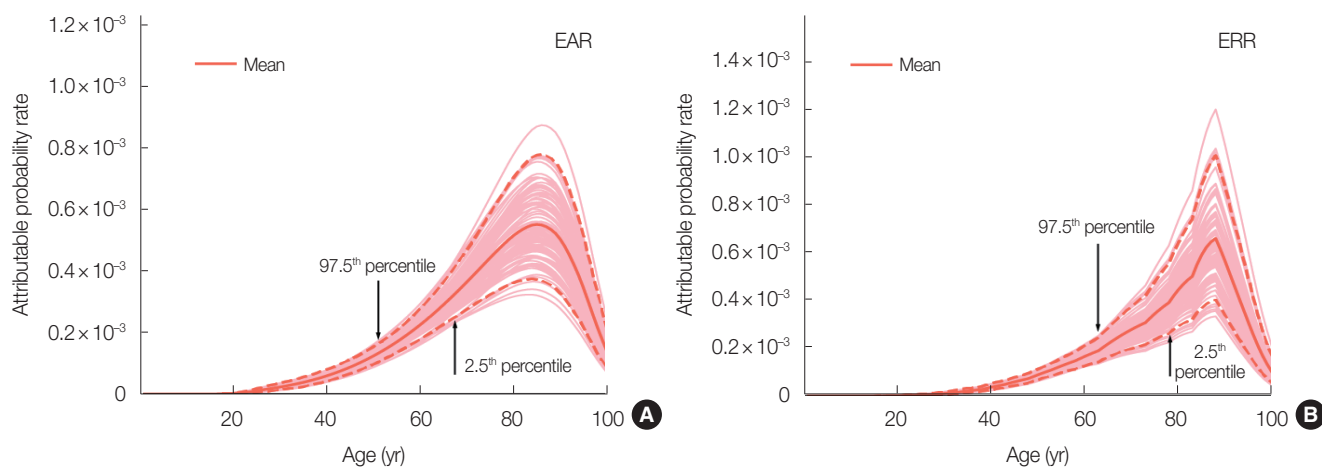
### 3. Chronic Exposure

The Supplementary Materials and Methods (4. Chronic exposure) shows how to compute the cumulated excess risk for females exposed to 0.01 Gy/yr at ages 10 to 19 followed up until age 100 with 7:3 weights for ERR and EAR transfers. The results are similar to those for acute exposure. “agex” means the period of chronic exposure between the ages 10 to 19. “doseGy” means a total dose of 0.1 Gy with an annual dose of 0.01 Gy/yr for 10 years.

Fig. 5 shows the attributable probability rates of all-solid-cancer incidence at attained ages in the case of chronic exposure. There are marked differences between the EAR and ERR models.



**Fig. 4.** Output of attributable probability rates in the case of acute exposure, which are obtained using the plotCER. The solid line and dotted lines show the mean and 95% confidence intervals, respectively. (A) Excess absolute risk (EAR) and (B) excess relative risk (ERR) model.



**Fig. 5.** Output of attributable probability rates in the case of chronic exposure. The solid line and dotted lines show the mean and 95% confidence intervals, respectively. (A) Excess absolute risk (EAR) and (B) excess relative risk (ERR) model.

#### 4. Comparison of CER with Other Risk Calculators

The CERs of cancer incidence calculated with our R and Python codes were compared with those obtained using RadRAT version 4.2.1, which is available on the NCI website and allows web-based calculations. The CER of cancer mortality was also compared with the BEIR-VII estimates [11]. The results are shown in Tables 2, 3. CER was calculated as the cumulated probability until the age of 90, for which CER in our codes was used.

Exposure doses were set at 0.1 Gy for acute exposure at the age of 10 and at 0.01 Gy for chronic exposure for each age from 10 to 19. DDREF was not taken into account in either calculation. The CERs of cancer incidence calculated with our codes were comparable to those obtained using RadRAT, as shown in Table 2. In the comparison, however, one should note that

there are many different methods and assumptions between our codes and RadRAT, and summation over all organs (oral cavity and pharynx, esophagus, stomach, colon, rectum, liver, gallbladder, pancreas, lung, prostate, bladder, kidney, nervous system, thyroid, and leukemia) at risk. The use of organ-specific ERR models in RadRAT would make a slight difference, although both codes used 2010 Japan population data for the cancer incidence baseline. In addition, RadRAT uses the DDREF for all chronic exposure and for acute exposure below 100 mGy using a log-normal distribution with a geometric mean of 1.5. These differences between both codes resulted in slight differences in the risk calculation as shown in the Table 2.

Table 3 provides the CERs of solid cancer mortalities between our codes and BEIR-VII. The CERs were obtained using BEIR-VII as the sum of site-specific CER estimates for the United

**Table 2.** CER of Cancer Incidence for Acute Exposure to 0.1 Gy at Age 10 and Chronic Exposure to 0.01 Gy at Ages 10 to 19

Variable	CER of cancer incidence (%)					
	Male			Female		
	Lower bound	Mean	Upper bound	Lower bound	Mean	Upper bound
Acute						
R <sup>a)</sup>	2.2	2.8	3.4	3.4	4.1	4.9
Python <sup>a)</sup>	2.2	2.8	3.4	3.4	4.1	4.9
RadRAT <sup>b)</sup>	2.0	2.8	3.6	3.8	4.8	6.2
Chronic						
R <sup>a)</sup>	2.0	2.5	3.0	3.1	3.6	4.2
Python <sup>a)</sup>	2.0	2.5	3.0	3.1	3.6	4.2
RadRAT <sup>b)</sup>	0.96	1.7	2.6	1.6	2.8	4.5

The weights for excess relative risk and excess absolute risk transfer are both 0.5:0.5.

CER, cumulative excess risk; RadRAT, Radiation Risk Assessment Tool.

<sup>a)</sup>Lower and upper bounds indicate 95% confidence intervals (2.5% and 97.5% percentiles).

<sup>b)</sup>Lower and upper bounds indicate a 90% uncertainty range.

States population. An assumption was made that the weights for ERR and EAR transfers were site-specific. CERs for the two methods showed close values in females, although there were some differences in males. It seemed that the difference of the cancer mortality baseline caused the difference of CER.

The CERs are calculated as a function of the excess risks and baseline risks of cancer mortality and incidence, as shown in Equation (5). The risk model was obtained by analyzing the LSS data in each site. The risk transfer model to other populations can depend on the baseline cancer mortality/incidence. The weights for ERR and EAR transfers are used for risk assessment of radiation protection issues. These assumptions should be noted when comparing CERs.

## 5. Other Considerations

The development of the codes provided an overall process of cancer risk calculation based on epidemiological data so that they output cumulative probabilities of excess mortality and incidence following acute or chronic exposure. However, estimating the risk of site-specific cancer will be the future issue for risk projection of the codes, although the codes will be able to be utilized for estimating site-specific cancer risk. Lifestyle modification of radiation-related risks, such as smoking history, would be future issues [21]. The overall uncertainty in the risk estimates includes DDREF and risk transfer from the atomic bomb survivors to other populations. Our codes treated risk transfer by considering the weights between ERR and EAR models. The values of

**Table 3.** CER of Cancer Mortality for Acute Exposure to 0.1 Gy at Age 10 and Chronic Exposure to 0.01 Gy at Ages 18 to 65

Variable	CER of cancer mortality (%)					
	Male			Female		
	Lower bound	Mean	Upper bound	Lower bound	Mean	Upper bound
Acute						
R <sup>a)</sup>	0.93	1.3	1.7	1.3	1.7	2.1
Python <sup>a)</sup>	0.92	1.3	1.7	1.3	1.7	2.1
BEIR-VII <sup>a)</sup>	0.45	0.96	2.1	0.71	1.6	3.5
Chronic <sup>b)</sup>						
R <sup>a)</sup>	1.9	2.5	3.3	2.8	3.4	4.1
Python <sup>a)</sup>	1.9	2.5	3.3	2.8	3.4	4.1
BEIR-VII <sup>a)</sup>	1.1	2.1	4.3	1.7	3.3	6.3

The weights for excess relative risk and excess absolute risk transfer are both 0.5:0.5.

CER, cumulative excess risk; BEIR, Biologic Effects of Ionizing Radiation.

<sup>a)</sup>Adjusted considering a dose and dose rate effectiveness factor of 1.5.

<sup>b)</sup>CER in population exposed to 10 mGy per year from ages 18 to 65.

DDREF can be treated with user input data. In addition, although exposure to neutrons is also considered in the dose assessment for atomic bomb survivors, gamma-ray is the main component of the individual radiation dose [22]. Therefore, it should be noted that SUMRAY can provide the risk projection for the whole-body external exposure to gamma-ray, but it does not fit for the risk estimates in the case of both external exposure to high linear energy transfer radiation like neutrons and internal exposure.

The risk calculation may give a false sense of accuracy unless the quantitative uncertainties associated with the overall risk prediction are taken into account. It is necessary to discuss which risk indicator is better than cumulative probability in terms of risk communication.

## Conclusion

The JHPS research group has developed an open-source risk calculation code. The code evaluates the uncertainty of risk predictions taking into account statistical errors when the risk model is fitted to the atomic bomb survivor data and then provides estimates and 95% CIs for attained age and cumulative risk. A future work will extend the present code to allow for site-specific risk assessment.

## Supplementary Materials

Supplementary materials can be found via <https://doi.org/10.14407/jrpr.2022.00213>.



## Conflict of Interest

No potential conflict of interest relevant to this article was reported.

## Acknowledgements

The authors are grateful to Dr. Kawaguchi, Mr. Ohtoshi, and the Planning Committee of Japan Health Physics Society for valuable discussion and advice.

## Ethical Statement

This article does not contain any studies with human participants or animals performed by any of the authors.

## Data Availability

The supporting data and the codes are openly available at <https://github.com/JapanHealthPhysicsSociety/SUMRAY>.

## Author Contribution

Conceptualization: Sasaki M, Furukawa K, Kai M. Data curation: Furukawa K, Satoh D. Project administration: Sasaki M, Takagi S, Kai M. Validation: Shimada K, Kudo S. Writing - original draft: all authors. Writing - review and editing: Sasaki M, Kai M. Approval of final manuscript: all authors.

## References

1. United Nations Scientific Committee on the Effects of Atomic Radiation. Sources, effects and risks of ionizing radiation: UNSCEAR 2012 report. Report to the general assembly. Annex A: Attributing health effects to ionizing radiation exposure and inferring risks. UNSCEAR; 2012.
2. The 2007 Recommendations of the International Commission on Radiological Protection. ICRP publication 103. *Ann ICRP*. 2007;37(2-4):1-332.
3. Ban N, Clero E, Vaillant L, Zhang W, Hamada N, Preston D, et al. Radiation detriment calculation methodology: summary of ICRP Publication 152. *J Radiol Prot*. 2022;42(2):023001.
4. Preston DL, Ron E, Tokuoka S, Funamoto S, Nishi N, Soda M, et al. Solid cancer incidence in atomic bomb survivors: 1958-1998. *Radiat Res*. 2007;168(1):1-64.
5. World Health Organization. Health risk assessment from the nuclear accident after the 2011 Great East Japan earthquake and tsunami, based on a preliminary dose estimation. WHO; 2013.
6. International Atomic Energy Agency. Assessment of prospective cancer risks from occupational exposure to ionizing radiation: IAEA-TECDOC-1985. IAEA; 2021.
7. Kadowaki Y, Hamada N, Kai M, Furukawa K. Evaluation of the lifetime brain/central nervous system cancer risk associated with childhood head CT scanning in Japan. *Int J Cancer*. 2021; 148(10):2429-2439.
8. Sasaki M, Hattori T. Recalculation of age-specific cancer risk due to radiation exposure using the recent mortality data of the Japanese population. *Jpn J Health Phys*. 2018;53(3):154-162.
9. Walsh L, Hafner L, Straube U, Ulanowski A, Fogtman A, Durante M, et al. A bespoke health risk assessment methodology for the radiation protection of astronauts. *Radiat Environ Biophys*. 2021; 60(2):213-231.
10. Berrington de Gonzalez A, Iulian Apostoaiei A, Veiga LH, Rajaraman P, Thomas BA, Owen Hoffman F, et al. RadRAT: a radiation risk assessment tool for lifetime cancer risk projection. *J Radiol Prot*. 2012;32(3):205-222.
11. National Research Council. Health risks from exposure to low levels of ionizing radiation: BEIR VII, phase 2. National Academy Press; 2006.
12. LARisk: an R package for lifetime attributable risk calculation [Internet]. cran.r-project; 2023 [cited 2023 Apr 20]. Available from: <https://cran.r-project.org/web/packages/LARisk/vignettes/LARisk-vignette.html>
13. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing; 2020.
14. United Nations Scientific Committee on the Effects of Atomic Radiation. Sources, effects and risks of ionizing radiation: UNSCEAR 2019 report. Report to the general assembly with scientific annexes. Annex A: Evaluation of selected health effects and inference of risk due to radiation exposure. UNSCEAR; 2019.
15. Simonsen LC, Slaba TC. Ensemble methodologies for astronaut cancer risk assessment in the face of large uncertainties: NASA/TP-2020-5008710 [Internet]. NASA Technical Reports Server (NTRS); 2020 [cited 2023 Apr 20]. Available from: <https://ntrs.nasa.gov/citations/20205008710>
16. The MIT License [Internet]. Open Source Initiative; 2023 [cited 2023 Apr 21]. Available from: <https://opensource.org/licenses/mit-license.php>
17. National Cancer Center Japan. Cancer mortality data in Japan [Internet]. National Cancer Center Japan; 2023 [cited 2023 Apr 21]. Available from: [https://ganjoho.jp/reg\\_stat/statistics/data/dl/index.html#a14](https://ganjoho.jp/reg_stat/statistics/data/dl/index.html#a14)
18. Ministry of Health, Labor and Welfare of Japan. Outline of vital statistics in Japan [Internet]. Ministry of Health Labor and Welfare of Japan; 2023 [cited 2023 Apr 21]. Available from: <https://www.mhlw.go.jp/english/database/db-hw/outline/index.html>
19. Ozasa K, Shimizu Y, Suyama A, Kasagi F, Soda M, Grant EJ, et al.

- Studies of the mortality of atomic bomb survivors, Report 14, 1950–2003: an overview of cancer and noncancer diseases. *Radiat Res.* 2012;177(3):229–243.
20. Preston DL, Lubin JH, Pierce A, McConney ME. *Epicure users guide*. Hirosoft International Corporation; 1993.
  21. Shimada K, Kai M. Lifetime risk assessment of lung cancer incidence for nonsmokers in Japan considering the joint effect of radiation and smoking based on the life span study of atomic bomb survivors. *J Radiat Prot Res.* 2021;46(3):83–97.
  22. Cullings HM, Grant EJ, Egbert SD, Watanabe T, Oda T, Nakamura F, et al. DS02R1: Improvements to atomic bomb survivors' input data and implementation of dosimetry system 2002 (DS02) and resulting changes in estimated doses. *Health Phys.* 2017;112(1): 56–97.