

LARisk: An R package for lifetime attributable risk from radiation exposure

Juhee Lee¹, Young Min Kim^{1*}, Yeongwoo Park¹, Eunjin Jang¹, Jinkyung Yoo¹, Songwon Seo², Eun Shil Cha³, Won Jin Lee³

1 Department of Statistics, Kyungpook National University, Daegu, Republic of Korea

2 Laboratory of Low Dose Risk Assessment, National Radiation Emergency Medical Center, Korea Institute of Radiological and Medical Sciences, Seoul, Republic of Korea

3 Department of Preventive Medicine, College of Medicine, Korea University, Seoul, Republic of Korea

* Corresponding author

Email: kymmyself@knu.ac.kr

Abstract

The lifetime radiation associated risk of cancer incidence has been of interest at least seventy years after atomic bombings of Hiroshima and Nagasaki in 1945. Since most of radiation epidemiological studies are retrospective, it is important to assess the potential magnitude of radiation-related cancer risk. The R-package so-call **LARisk** implements to easily calculate lifetime attributable risks (LAR) of radiation associated cancer incidence presenting bootstrap-based confidence intervals of the estimated risks. The main characteristics of the **LARisk** package are to flexibly compute the LAR-associated risks, effectively summarize large data, and promptly update domain-knowledge such as new regression coefficients for excess relative risk or excess absolute risk models, population information, etc. The manuscript provides a complete guide for usage of the **LARisk** package, which will enable to contribute the prospective studies in radiation epidemiology.

Keywords: Lifetime attributable risk, Prospective study, Radiation epidemiology, Radiation-associated cancer

Introduction

The lifetime radiation related risk of cancer incidence remains elevated for at least seventy years after atomic bombings of Hiroshima and Nagasaki, Japan in 1945 [1]. Estimating radiation related cancer risk has made steady progress [1, 2], but most studies have been retrospective. Forward projection of radiation related cancer risk is essential to allow predictive studies of irradiated groups to ultimately control excess radiation exposure occurring, e.g. from medical radiography (computed tomography, CT) scans [3, 4].

There are several reports to discuss lifetime radiation related cancer risk projection models. The U.S. Nuclear Regulatory Commission developed the NUREG model to predict health effects from radiation exposure at nuclear power plants, considering the dose and dose-rate effectiveness factor (DDREF) and age-specific risk coefficients [5]. The U.S. National Research Council proposed the BEIR VII model [6] for radiation exposed groups, focusing on cancer risk projection from low linear energy transfer

(LET) radiation exposure with weighted average estimates from excess relative risk (ERR) and excess absolute risk (EAR) models [6]. The BEIR VII risk models were based primarily on the Life Span Study (LSS) study from the Radiation Effects Research Foundation (RERF), which has advantages of its large population, including of both sexes and all ages, long follow-up, and high-quality mortality and cancer incidence data [6]. However, LSS is based on single acute exposure and still limited in its ability to precisely estimate site-specific cancer risks. Therefore, we acknowledge that the BEIR VII models need to be updated with other studies of chronic low-level radiation exposure with the longer duration follow-up.

The lifetime attributable risk (LAR) means the number of excess event incidences for an irradiated group compared with a non-radiation exposed group. The LAR risk projection model is defined as

$$\int_{a_1+L}^{a_{\max}} M(D, e, a, s) \frac{S(a, s)}{S(e, s)} da, \quad (1)$$

where L is the latency period for each cancer, $M(D, e, a, s)$ is a risk model, and the ratio $S(a, s)/S(e, s)$ is the conditional probability of a person alive and cancer-free at age-at exposure (e) to approach an attained age (t). We used cancer-free survival function after removing cancer incidence in the risk projection models [7]. In radiation epidemiology, LAR is the probability of premature cancer incidence attributable to radiation exposure in a representative member of the population [8–10] over their lifetime. World Health Organization (WHO) used LAR for Fukushima health risk assessment [11], and many other studies also have forecast radiation related cancer risk using LAR [3, 12–22].

The US National Cancer Institute ‘RadRAT’ program estimates lifetime risk of cancers induced by radiation exposure [23] based on BEIR VII [6], which estimated radiation related cancer risk transferred from Japanese (atomic bomb survivors) to US populations, and considered uncertainties related to risk model regression coefficients, e.g. ERR and EAR models (see [23] and [6] for more details), minimum latency, DDREF and radiation dose exposure distributions and population transfer weights. The ‘RadRAT’ program can be accessed for free at <https://radiationcalculators.cancer.gov/radtrat/> and is widely used to assess lifetime cancer risks attributable to radiation [20, 24, 25]. Although ‘RadRAT’ was designed to be easy and simple to use, it is difficult to flexibly use batch files for large data analyses or multiple cases, or ensure timely knowledge updates, e.g. baseline incidence rate in a specific population, life tables, etc. Thus, we developed an **LARisk** R package to calculate lifetime attributable risk for radiation related cancer built over the online RadRAT program. **LARisk** provides significantly enhanced flexibility by incorporating functions to project LAR for the batch file considering modified baseline incidence rate and/or population transfer weights, and DDREF application and a summary function based on average LAR for each case. In addition, **LARisk** uses the parametric bootstrap method to compute risk estimates and confidence intervals because the population distributions for uncertainties are known. The parametric bootstrap method is more powerful and provides small variances compared with the nonparametric bootstrap if the distribution assumption is known and the sample sizes are relatively small.

The remainder of this paper is organized as follows. Section **LAR models** describes the LAR risk models in detail, including the underlying ERR and EAR functions and discusses statistical uncertainties. Section **Illustration of LARisk package** discusses input information, output values, arguments, and result interpretation in the **LARisk** package; briefly describes **LARisk** options; Section **An example with chronic exposure rate** uses data examples to describe how to use **LARisk**. Finally, Section **Concluding remarks** summarizes the paper.

LAR models

Risk model

Risk models for the **LARisk** package were based on those developed for the RadRAT program to estimate lifetime risk following low-dose radiation exposure. ERR (Excess Relative Risk) and EAR (Excess Absolute Risk) models were used for exposed individuals compared with unexposed individuals, based on a dose response function, $\rho(D)$ with radiation dose D in Gray (Gy), and effect modification, $\varepsilon(\mathbf{X}) = \exp(\boldsymbol{\alpha}^T \mathbf{X})$ where \mathbf{X} is a vector of several covariates including age at exposure, attained age, sex, etc. and $\boldsymbol{\alpha}$ is a vector of regression coefficients. $\rho(D)$ can be linear, linear-quadratic, quadratic, linear-threshold, nonparametric, etc. For example,

$$\rho(D) = \begin{cases} \beta_1 D & \text{Linear} \\ \beta_1 D + \beta_2 D^2 & \text{Linear-Quadratic} \\ \beta_2 D^2 & \text{Quadratic.} \end{cases}$$

See [6], [1], [26], [27] and [28] for more details.

LARisk incorporates both ERR and EAR for a risk model (1) to compute LAR estimates of solid cancer and leukemia. The standard form of ERR or EAR models for solid cancer, except thyroid and breast, is defined as

$$\beta_s D \exp \left[\alpha_1 e^* + \alpha_2 \log \left(\frac{a}{60} \right) \right], \quad (2)$$

where the risk is assumed to be linearly related to the radiation dose, a is attained age, and e is age at exposure, which are considered to be effect modifications where

$$e^* = \begin{cases} (e - 30)/10 & \text{for } e < 30 \\ 0 & \text{for } e \geq 30. \end{cases}$$

The form for thyroid and breast cancer is slightly different from the standard model. Thyroid only depends on age at exposure effect modification. The breast cancer model is defined as

$$\beta_s D \times \exp \left[\alpha_1 e^{**} + \alpha_2 \log \left(\frac{a}{50} \right) \right], \quad (3)$$

where $e^{**} = (e - 25)/10$, α_2 is equal to 3.5 for $a \leq 50$ and 1.0 elsewhere. Gallbladder, brain/CNS, and thyroid cancers only consider a risk model based on ERR to calculate LAR, and breast cancer utilizes a risk model based on only EAR. A standard form of a risk model for leukemia takes the model as

$$\beta_s D(1 + \theta D) \exp \left[\alpha_1 e^* + \alpha_2 \log \left(\frac{t}{25} \right) + \alpha_3 e^* \log \left(\frac{t}{25} \right) \right], \quad (4)$$

which is linear-quadratic in dose and age at exposure, and time since exposure (t) is considered as the effect modification. For chronic exposure, the risk model is linear with dose, i.e., $\theta = 0$. A generalized nonlinear model, in particular a Poisson nonlinear model, is required to directly estimate the parameters [6, 26, 28, 29].

It is possible that LAR can be obtained from the multiplicative or additive model controlling weights. The multiplicative model assumes that radiation induced excess cancer risk proportionally increases with baseline risk as

$$\begin{aligned} \text{Multiplicative model : } M(D, e, a, s) &= \text{ERR}(D, e, a, s) \times \lambda_0(a, s) \\ &= \text{ERR}(D, e, a, s) \times m(a) \times \frac{S(a, s)}{S(e, s)}, \end{aligned}$$

where $\lambda_0(a, s)$ is the excess risk for non-exposed group. On the other hand, the additive model assumes that radiation induced excess cancer risk increases independently of baseline risk as

$$\text{Additive model : } M(D, e, a, s) = \text{EAR}(D, e, a, s) \times \frac{S(a, s)}{S(e, s)}.$$

Both LAR_{ERR} and LAR_{EAR} , are LAR components based on ERR and EAR models, respectively, defined as

$$\begin{aligned} \text{LAR}_{\text{ERR}} &= \int_{a_1+L}^{a_{max}} \text{ERR}(D, e, a, s) \times m(a) \times \frac{S(a, s)}{S(e, s)} da \\ \text{LAR}_{\text{EAR}} &= \int_{a_1+L}^{a_{max}} \text{EAR}(D, e, a, s) \times \frac{S(a, s)}{S(e, s)} da, \end{aligned} \quad (5)$$

where $m(a)$ is population cancer incidence rate, and $S(a, s)/S(e, s)$ is the probability of surviving to age a , conditional on survival to age e , given sex s . Thus, the LAR model in the **LARisk** package is computed as

$$\text{LAR} = \frac{w \times \text{LAR}_{\text{ERR}} + (1 - w) \times \text{LAR}_{\text{EAR}}}{DDREF} \quad \text{for all solid cancer,}$$

and

$$\text{LAR} = (\text{LAR}_{\text{ERR}})^w \times (\text{LAR}_{\text{EAR}})^{1-w} \quad \text{for leukemia,}$$

where w is a population transfer weight.

Uncertainty assessment

Uncertainties including statistical uncertainties must be considered when calculating LAR, since we use probability distribution functions based on radiation epidemiology and Monte Carlo (MC) simulation. The uncertainties are as follows.

1. risk model coefficients,
2. dose and dose-rate effectiveness factor,
3. minimum latency,
4. risk population transfer between populations (e.g. Japan to US),
5. radiation doses, and
6. baseline incidence rate.

Uncertainties 1–4 including statistical uncertainties have the most critical impact to assess lifetime cancer risk attributable to radiation, whereas uncertainties 5 and 6 cause LAR variability regardless of the model. The baseline incidence rate in particular (uncertainty 6) differs with population and year, generating different incidence rate distributions by age.

To calculate radiation induced future cancer risk, we substitute the formula (2) into (5) for solid cancer as

$$\begin{aligned} \text{LAR}_{\text{ERR}} &= \int_{e+L}^{a_{max}} \beta_{\text{ERR},s} \times D \times \exp \left\{ \alpha_{\text{ERR},1} e^* + \alpha_{\text{ERR},2} \log \left(\frac{a}{60} \right) \right\} \times m(a) \times \frac{S(a, s)}{S(e, s)} da \\ \text{LAR}_{\text{EAR}} &= \int_{e+L}^{a_{max}} \beta_{\text{EAR},s} \times D \times \exp \left\{ \alpha_{\text{EAR},1} e^* + \alpha_{\text{EAR},2} \log \left(\frac{a}{60} \right) \right\} \times \frac{S(a, s)}{S(e, s)} da. \end{aligned} \quad (6)$$

The regression coefficients, $\beta_{\text{ERR},s}$, $\beta_{\text{EAR},s}$, $\alpha_{\text{ERR},1}$, $\alpha_{\text{EAR},1}$, $\alpha_{\text{ERR},2}$, and $\alpha_{\text{EAR},1}$, were determined from radiation epidemiologic research with atomic bomb (A-bomb) survivors

Table 1. The regression coefficients of all radiation related solid cancer risk models [6, 23]

ERR regression coefficients for solid cancers				
Cancer site	β_{Male}	β_{Female}	α_1	α_2
Oral	0.23 (<0.0–0.66)	0.53 (0.13–1.24)	–0.30	–1.40
Esophagus	0.51 (<0.0–1.13)	0.82 (<0.0–3.10)	–0.30	–1.40
Stomach	0.21 (0.11–0.40)	0.48 (0.31–0.73)	–0.30	–1.40
Colon	0.63 (0.37–1.10)	0.43 (0.19–0.96)	–0.30	–1.40
Rectum	0.12 (<0.0–0.38)	0.12 (<0.0–0.38)	–0.30	–1.40
Gallbladder	–0.018 (<0.0–0.29)	–0.018 (<0.0–0.29)	–0.30	–1.40
Pancreas	0.36 (<0.0–0.88)	0.36 (<0.0–0.88)	–0.30	–1.40
Liver	0.32 (0.16–0.64)	0.32 (0.10–1.00)	–0.30	–1.40
Lung	0.32 (0.15–0.70)	1.4 (0.94–2.10)	–0.30	–1.40
Breast	-	-	-	-
Ovary	-	0.38 (0.10–1.40)	–0.30	–1.40
Uterus	-	0.05 (<0.0–0.22)	–0.30	–1.40
Prostate	0.12 (<0.0–0.69)	-	–0.30	–1.40
Bladder	0.50 (0.18–1.40)	1.65 (0.69–4.00)	–0.30	–1.40
Kidney	0.34 (<0.0–1.00)	0.34 (<0.0–1.00)	–0.30	–1.40
Brain/CNS	0.71 (0.26–1.34)	0.24 (0.09–0.47)	–0.30	–1.40
Thyroid	0.53 (0.14–2.00)	1.05 (0.28–3.90)	–0.83	0.00
Remainder	0.84 (0.45–1.69)	0.80 (0.33–1.93)	–0.30	–1.40

EAR regression coefficients for solid cancers				
Cancer site	β_{Male}	β_{Female}	α_1	α_2
Oral	0.44 (0.08–1.10)	0.29 (0.06–0.66)	–0.41	0.50
Oesophagus	0.88 (0.11–2.10)	0.14 (<0.0–0.63)	–0.41	2.80
Stomach	4.90 (2.70–8.90)	4.90 (3.20–7.30)	–0.41	2.80
Colon	3.20 (1.80–5.60)	1.60 (0.80–3.20)	–0.41	2.80
Rectum	0.34 (0.09–1.10)	0.34 (0.09–1.10)	–0.41	2.80
Gallbladder	-	-	-	-
Pancreas	0.49 (0.09–1.10)	0.49 (0.09–1.10)	–0.41	2.80
Liver	2.20 (0.90–5.30)	1.00 (0.40–2.50)	–0.41	4.10
Lung	2.30 (1.10–5.00)	3.40 (2.30–4.90)	–0.41	5.20
Breast	-	10.00 (7.00–14.20)	–0.50	3.50, 1.00
Ovary	-	0.70 (0.20–2.10)	–0.41	2.80
Uterus	-	1.20 (<0.0–2.60)	–0.41	2.80
Prostate	0.11 (<0.0–1.0)	-	–0.41	2.80
Bladder	1.20 (0.40–3.70)	0.75 (0.30–1.70)	–0.41	6.00
Kidney	0.31 (0.08–0.68)	0.31 (0.08–0.68)	–0.41	2.80
Brain/CNS	-	-	-	-
Thyroid	-	-	-	-
Remainder	2.73 (1.55–4.82)	1.06 (0.49–2.28)	–0.41	2.80

Note that the values in the parentheses represent 95% confidence intervals for each regression coefficient.

at RERF, which has conducted health related research among A-bomb survivors in Hiroshima and Nagasaki, Japan, for more than 70 years. Tables 1 and 2 show the regression coefficients for solid cancers and leukemia, respectively. For leukemia, we replace the formula (2) by (4) to derive a similar expression to the formula (6).

The uncertainties of risk model coefficients for solid cancers focus on sex-specific main effect parameters, whereas other parameters are specific constants derived from

Table 2. Regression coefficients for radiation related leukemia risk models [6, 23]

ERR and EAR models for leukemia			
	β_{Male}	β_{Female}	α_1
ERR	1.10 (0.10–2.60)	1.20 (0.10–2.90)	−0.40 (−0.78 to 0.00)
EAR	1.62 (0.10–3.60)	0.93 (0.10–2.00)	0.29 (0.00–0.62)
	α_2	α_3	θ
ERR	−0.48 (−1.10 to 0.20)	0.42 (0.00–0.96)	0.87 (0.16–15.00)
EAR	0.00	0.56 (0.31–0.85)	0.88 (0.16–15.00)

Note that the values in the parentheses represent 95% confidence intervals for each regression coefficient.

A-bomb survivor epidemiology studies. Table 1 shows uncertainties for ERR and EAR model parameters. Dose coefficient, β , for stomach, colon, liver, lung, breast, ovary, bladder, and thyroid cancer is assumed to follow log-normal distribution, hence regression coefficient distributions in LAR models for prostate and uterus cancer are also considered to be normally distributed. Uncertainty in β is described by cumulative distribution functions (CDFs) for oral, oesophagus, gallbladder, pancreas, rectum, kidney, and brain/CNS cancers. α_2 of breast cancer has two values. See the formula 3. In contrast to solid cancer, uncertainty for leukemia is associated with five parameters. Since the covariance matrix for regression coefficients in the leukemia LAR model should be considered, we calculate an approximate estimator for lifetime risk variance using the delta method following the BEIR VII committee [6].

The dose and dose-rate effectiveness factor (DDREF) combines two concepts of a low-dose effectiveness (LDEF) and a dose-rate effectiveness factor (DREF) [6, 30, 31]. LDEF means a ratio of the slope of a linear-quadratic dose-response function for acute exposure and the slope of the linear term of the same model, whereas DREF is computed when an acute exposure is divided and the number of fractions is large enough. ICRP suggested to utilize a DDREF of 2 to reduce solid cancer rates obtained from moderate-to-high acute dose studies for low-dose or low-dose rate exposures [31, 32]. The DDREF distribution is considered as a lognormal with geometric mean= 1.5 and standard deviation= 1.35 [23]. DDREF is always applied for chronic exposure without dose-dependence. In acute exposures, DDREF is applied by comparing dose limit, D_L , and exposure dose. If dose < D_L (generated 30 – 200 mGy using log-uniform distribution), DDREF is considered as used. DDREF is not applied for leukemia, because the uncertainty associated with low-dose risks is modeled using a linear term.

Latency uncertainty for radiation-induced cancers should be considered, because each person has cancer different development time after radiation exposure. Latency distribution is commonly represented by an S-shaped curve with the parameter in the center of the S-shaped function generated from the triangular distribution for solid cancer, thyroid and leukemia as $T(5, 7.5, 10)$, $T(3, 5, 7)$ and $T(2, 2.25, 2.5)$, respectively. An uncertain latency adjustment is between 4 and 11 years after radiation exposure for solid cancer, 2.5 and 7.6 years for thyroid cancer, and 0.4 and 4.1 years for leukemia [33].

We must also consider effect of transferring the risk from Japanese population to population of interest. Because parameters of risk model coefficients are commonly obtained using the Japanese A-bomb survivors cohort, the effect is explained by the weighted mean of LAR values based on ERR and EAR model estimates. Initially we use the RadRAT weight to transfer the risk from Japanese to US populations. Thus, the weight $w = 0.7$ is assigned to the ERR model and the remaining ($1 - w = 0.3$) weight to the EAR model for most cancers. Lung cancer weight is reversed, and only the EAR model for breast is used. For thyroid, gallbladder, and brain/CNS, only ERR is used for calculation of LAR.

Radiation dose uncertainty can be applied to a fixed value or various probability distributions, including lognormal, normal, triangular, log-triangular, uniform, and log-uniform.

The baseline incidence rate depends on population information.

Parametric bootstrap

Algorithm 1: Computation of parametric bootstrap confidence intervals for risk estimates

Result: a bootstrap $(1 - \alpha) \times 100\%$ confidence interval for LAR
Compute the LAR estimates such as LAR;
for *Generate bootstrap samples from distributions of uncertainties* $b = 1, \dots, B$
do
| Calculate the LAR estimates;
end
 $100(\alpha/2)\%$ and $100(1 - \alpha/2)\%$ percentiles of LAR are the $(1 - \alpha) \times 100$
bootstrap confidence interval for the LAR;

Since uncertainties' distributions for estimation of LAR are already known, we apply the parametric bootstrap to compute the confidence intervals for all risk estimates. The parametric bootstrap needs distribution assumption for generation and the other procedure is identical to the general nonparametric bootstrap. The parametric bootstrap is more powerful, in that the method provides small variances compared with the nonparametric bootstrap if the distribution is assumed to be right and the sample sizes are relatively small. The parametric bootstrap algorithm is provided in Algorithm 1.

Illustration of LARisk package

Main arguments

The LARisk package has 3 main functions for estimating lifetime attributable risks such as LAR, LAR_batch and LAR_group. The primary argument for the LAR function is as follows.

```
LAR(data, basedata, sim=300, seed=99, current=as.numeric(substr(Sys.Date(),1,4)), ci=0.9, weight=NULL, DDREF=TRUE, basepy=1e+05)
```

The argument 'data' should include some prerequisite information such as sex, and birth year(s), exposure year, distributions of exposure dose, radiation dose, cancer sites, and exposure rate, shown as follows.

```
> data=data.frame(sex = 'male', birth = 1900, exposure = 1980, site = 'stomach', exposure_rate = 'chronic', dosedist = 'fixedvalue', dose1 = 10, dose2=NA, dose3=NA)
```

For 'site', we provide the irradiated organ or cancer site. In 'dosedist', 'dose1', 'dose2', and 'dose3' arguments, we insert the distribution of the exposure dose such as 'fixedvalue', 'lognormal', 'normal', 'triangular', 'logtriangular', 'uniform' or 'loguniform' with appropriate parameters for each distribution. For instance, if the exposure dose has a normal distribution with mean of 2.3 and standard deviation of 0.8, we set 'dose1=2.3', 'dose2=0.8', and 'dose3=NA'. If the dose has fixed value of 3.2, 'dose1=3.2', 'dose2=NA' and 'dose3=NA' should be employed. If new population information should be considered, lifetime and cancer incidence rate tables are employed in a 'basedata' argument as follows.

```
LAR(data, basedata = list(lifeTable, incidenceTable))
# lifeTable      : lifetime table
# incidenceTable: cancer incidence rate table
```

LARisk has default for these tables, which were made in 2010 and 2018, South Korea such as 'life2010', 'incid2010', 'life2018' and 'incid2018'. Lifetime and cancer incidence rate tables should also have the specified format as follows.

```
> head(incid2010) ## cancer incidence rate table of the Korean
in 2010.
Site Age Rate_m Rate_f
1 oral 0 0.2 0.1
2 oral 1 0.2 0.1
```

Also, the cancer incidence rate table consists of 'Site', 'Age', 'Rate_m', and 'Rate_f' where 'Rate_m' and 'Rate_f' are incidence rates of each cancer site according to male and female, respectively.

The argument 'weight' is employed to compute LAR values through the weighted average of LAR values based on ERR and EAR models. For example, if the 'weight' for LAR based on between EAR and ERR models for stomach cancer is set to 0.5, run the below code.

```
LAR(data, basedata, weight=list(stomach = 0.5))
```

LARisk sets the default weights to 0.7 for most cancers, 0.3 for lung cancer, 0.0 for breast and 1.0 for thyroid for LAR based on EER (see Table 3).

Table 3. Weights for each cancer site

Cancer site	ERR	EAR	weight
Most cancer	0.7	0.3	0.7
Lung	0.3	0.7	0.3
Breast	0.0	1.0	0.0
Thyroid	1.0	0.0	1.0
Gallbladder	1.0	0.0	1.0
Brain/CNS	1.0	0.0	1.0

Dose and dose-rate effectiveness factor (DDREF) is the logical option to select whether or not to consider 'DDREF' for computing LAR values. The value 'DDREF' is to modify the effect of exposure to low-dose. The value of 'DDREF' is considered differently as per exposure rate. However, if the site is leukemia, DDREF dose not apply even if DDREF = TRUE.

Estimating LAR values using LAR, LAR_batch and LAR_group

To describe how to use functions installed in the **LARisk** package, we use the toy data 'nuclear', also installed in **LARisk**, assuming scenario of the data in which everyone is exposed to radiation at the same time in 2011. The data includes 20 people for 10 males and 10 females, and the radiation exposed age has ranges from 3 to 81 years with acute exposed rate and fixed dose distribution.

```
> nuclear1 <- nuclear[nuclear$ID=="ID01",]
> print(nuclear1)
ID sex birth exposure site exposure_rate dosedist
      dose1 dose2 dose3 distance
ID01 female 1973 2011 ovary acute fixedvalue
      50.06989 NA NA 1
ID01 female 1973 2011 oesophagus acute fixedvalue
      50.37462 NA NA 1
```


LAR, **LAR_batch**, and **LAR_group** functions provide 3 main estimated values such as lifetime risk, future risk and lifetime baseline risk. Among them, ‘LAR’ and ‘F_LAR’ in Table 4 have mean values and confidence limits (lower and upper) for each cancer site, solid cancer and total (solid cancer+leukemia), respectively. See Table 4 for more details.

Table 4. Values in the LARisk package

Value	Description
LAR	Lifetime attributable risk (Cancer incidence probability per 100,000 persons after exposed year)
F_LAR	Future LAR (LAR after current year)
LBR	Lifetime baseline risk (Cumulative baseline probability per 100,000 persons of having cancer over the maximum lifetime without radiation exposure after exposed year)
BFR	Baseline future risk (Cumulative baseline probability per 100,000 persons of having cancer over the maximum lifetime without radiation exposure after current year)
LFR	Lifetime fractional risk (=LAR/LBR)
TFR	Total future risk (= F_LAR + BFR)

The **LAR** function prints the total LAR, future LAR, baseline future risk, and total future risk. If you want to obtain more detailed result, you can use the ‘summary’ function. The ‘summary’ function provides the person’s gender and year of birth, and risks by cancer type, confidence level, and current year. In the result, the ‘LAR’ tab includes site-specific LAR, LBR, and LFR, whereas the ‘Future LAR’ tab contains site-specific future LAR, BFR, and TFR.

```

> LAR(nuclear1, basedata = list(life2010, incid2010))
LAR:
Lower      Mean      Upper
359.9479   671.3751 1203.8111

Future LAR:
Lower      Mean      Upper
F.LAR      337.3697   597.7655 1032.216
BFR        15992.9219 15992.9219 15992.922
TFR        16330.2917 16590.6874 17025.137

> summary(LAR(nuclear1, basedata = list(life2010, incid2010)))
Information:
sex birth
female 1973

LAR:
Lower      Mean      Upper      LBR      LFR
lung       74.5065   156.5247   241.1717   3630.3464 0.0431
ovary      4.4880    13.3756    27.6491    697.6095 0.0192
bladder    13.0799   30.4405    58.7444    452.0590 0.0673
thyroid    83.9986   368.7603   877.5446   7318.8901 0.0504
remainder  33.5177   89.1794    169.3327   4237.3935 0.0210
oesophagus 0.2000    4.0450     10.4666    116.7920 0.0346
rectum     0.2670    9.0496     21.7567    2157.6294 0.0042
leukemia   0.0000    0.0000     0.0000     0.0000    NaN
solid      359.9479  671.3751   1203.8111  18610.7199 0.0361
total      359.9479  671.3751   1203.8111  18610.7199 0.0361

Future LAR:
Lower      Mean      Upper      BFR      TFR
lung       73.4884   154.5469   238.6027   3574.2001 3728.7470
ovary      4.0861    12.0693    25.2806    597.5191 609.5885
bladder    13.0173   30.3006    58.5236    452.9404 483.2410
thyroid    68.7104   303.3691   745.8708   5117.1159 5420.4850
remainder  32.3341   84.9052    160.5547   4066.5835 4151.4887
oesophagus 0.1978    3.9420     10.2253    115.9270 119.8690
rectum     0.2558    8.6324     20.6083    2068.6359 2077.2682
leukemia   0.0000    0.0000     0.0000     0.0000    0.0000
solid      337.3697  597.7655   1032.2155  15992.9219 16590.6874
total      337.3697  597.7655   1032.2155  15992.9219 16590.6874

```


The average LAR values in the toy data ‘nuclear’ can be computed by ‘distance’ as follows.

```

> ex_group1 <- LAR_group(nuclear, pid = nuclear$ID, group =
  nuclear$distance, basedata = list(life2010, incid2010))
> summary(ex_group1)
summaries of LAR result : Group 1

Group Information:
sex count birth
female 35 1962.600
male 45 1962.222

LAR:
Lower Mean Upper LBR LFR
stomach 50.9850 66.2373 87.8936 2926.1714 0.0226
colon 43.4301 56.1274 72.3353 2160.0832 0.0260
liver 7.0657 10.4933 15.2987 808.9772 0.0130
lung 39.1538 53.4791 70.6998 3097.9930 0.0173
breast 2.8636 4.2821 6.0646 331.2576 0.0129
ovary 0.3206 0.9554 1.9749 49.8292 0.0192
uterus 0.0216 0.1491 0.3288 68.4665 0.0022
prostate -15.4917 6.2743 27.6037 960.8121 0.0065
bladder 10.5122 16.1598 23.8353 430.7553 0.0375
brain/cns 0.7414 1.3918 2.2998 46.3592 0.0300
thyroid 67.7618 185.2396 350.2187 1648.0018 0.1124
remainder 34.5468 56.8793 88.6948 912.3992 0.0623
oral 0.3038 1.3365 2.7663 105.2675 0.0127
oesophagus 2.7255 5.3313 8.1798 208.5186 0.0256
rectum 0.3667 1.3777 2.5946 618.5623 0.0022
gallbladder -6.2708 -0.9336 4.2742 420.9955 -0.0022
pancreas 3.4377 7.4064 11.9408 493.6795 0.0150
kidney 1.0530 3.2226 6.5984 203.1165 0.0159
leukemia 0.0713 0.2012 0.5677 19.9927 0.0101
solid 339.2210 475.4092 659.3305 15491.2456 0.0307
total 339.4012 475.6104 659.5318 15511.2384 0.0307

Future LAR:
Lower Mean Upper BFR TFR
stomach 48.2294 63.3542 84.7154 2490.0371 2553.3913
colon 40.9639 53.4541 69.9636 1902.1381 1955.5921
liver 6.0418 9.0587 13.2437 571.8794 580.9381
lung 38.4091 52.5642 69.9525 2877.2243 2929.7885
breast 2.3678 3.5644 5.0447 157.2215 160.7859
ovary 0.2919 0.8621 1.8058 42.6799 43.5420
uterus 0.0203 0.1258 0.2655 41.5060 41.6318
prostate -15.5887 6.2226 27.5255 881.1613 887.3839
bladder 10.3276 16.0169 23.7049 387.4721 403.4890
brain/cns 0.6220 1.1373 1.8552 35.4464 36.5837
thyroid 60.0224 170.7748 330.0385 1233.1081 1403.8830
remainder 33.7594 54.9135 84.8616 837.8480 892.7615
oral 0.2668 1.2582 2.6651 87.9061 89.1643
oesophagus 2.4342 4.4985 6.7130 162.4755 166.9740
rectum 0.3337 1.2961 2.4649 524.0167 525.3128
gallbladder -6.0139 -0.8791 4.1981 361.3716 360.4925
pancreas 3.2450 7.1431 11.5463 453.7292 460.8723
kidney 0.9721 3.1387 6.5336 183.9500 187.0887
leukemia 0.0827 0.1485 0.2667 13.7139 13.8624
solid 318.3267 448.5040 621.6234 13231.1713 13679.6753
total 318.4752 448.6524 621.7718 13244.8852 13693.5376

Confidence Level: 0.9
Current Year: 2022

summaries of LAR result : Group 2

Group Information:
sex count birth
female 12 1987.417
male 5 1956.000
.
.
.

```

Saving outputs

LARisk includes a function `write_LAR` which writes outputs from `LAR`, `LAR_batch`, and `LAR_group` into a CSV file.

```
> write_LAR(x, filename)
```

The format of the saved file is illustrated in Table 5.

Table 5. Format to save outputs using a write_LAR function

	Lower	Mean	Upper	F.Lower	F.Mean	F.Upper	LBR	BFR	LFR	TFR
site-name
solid
total

An example for chronic exposure rate with a toy data 'organ'

In this section, we consider a toy data 'organ' which consists of the 20 radiation exposed people over various periods with 14 males and 6 females having the job information 'occup'. The data is shown as follows.

```
> head(organ)
  ID sex birth exposure      site exposure_rate dosedist
1 ID01 male 1985      2011 oesophagus      chronic fixedvalue
  0.001954895      NA      NA      1
2 ID01 male 1985      2011 kidney      chronic fixedvalue
  0.003855487      NA      NA      1
```

Suppose that we want to calculate LAR values with the current year=2021 using the Korean population in 2018. First, we compute LAR values with 'ID01' as follows.

```
> organ1 <- organ[organ$ID=='ID01',]
> ex_organ1 <- LAR(organ1, baseda=list(life2018, incid2018),
  current=2021)
> ex_organ1
LAR:
  Lower Mean Upper
1.1149 1.6981 2.5132

Future LAR:
  Lower Mean Upper
F.LAR 1.1132 1.6759 2.4744
BFR 6694.6423 6694.6423 6694.6423
TFR 6695.7555 6696.3182 6697.1167

> summary(ex_organ1)
Information:
sex birth
male 1985

LAR:
  Lower Mean Upper LBR LFR
thyroid 0.4673 0.9709 1.8205 1771.543 5e-04
oesophagus 0.1025 0.1824 0.2729 1048.947 2e-04
rectum 0.0978 0.2385 0.4111 2893.126 1e-04
kidney 0.1416 0.3064 0.5160 1338.149 2e-04
leukemia 0.0000 0.0000 0.0000 0.000 NaN
solid 1.1149 1.6981 2.5132 7051.764 2e-04
total 1.1149 1.6981 2.5132 7051.764 2e-04

Future LAR:
  Lower Mean Upper BFR TFR
thyroid 0.4507 0.9512 1.7587 1453.492 1454.443
```

oesophagus	0.1025	0.1823	0.2728	1055.583	1055.766
rectum	0.0977	0.2379	0.4108	2877.117	2877.355
kidney	0.1406	0.3045	0.5153	1308.450	1308.755
leukemia	0.0000	0.0000	0.0000	0.000	0.000
solid	1.1132	1.6759	2.4744	6694.642	6696.318
total	1.1132	1.6759	2.4744	6694.642	6696.318

Confidence Level: 0.9
Current Year: 2021

The estimated total LAR values of the person 'ID01' is 1.6981 with the 90% confidence interval (1.1149, 2.5132) and the future risk is 1.6759 with the 90% confidence interval (1.1132, 2.4744). The person 'ID01' is a man born in 1985. 'ID01' was exposed to thyroid, oesophagus, 'rectum', and kidney except for leukemia. Thus, The LAR values of leukemia are all 0.

We also compute LAR values by group - male and female using `LAR_group` as follows.

```
> ex_organ2 <- LAR_group(organ, pid=organ$ID, group=organ$sex,
  basedata=list(life2018, incid2018), current=2021)
> summary(ex_organ2)
summaries of LAR result : Group female

Group Information:
sex count birth
female 166 1976.313

LAR:
Lower Mean Upper LBR LFR
colon 0.6832 1.1392 1.7064 1080.1190 0.0011
lung 2.3739 2.8990 3.5941 1353.2750 0.0021
uterus 0.1447 0.2874 0.4696 338.1363 0.0008
bladder 0.9560 1.3026 1.7539 155.9481 0.0084
remainder 2.9905 4.5565 6.7692 1875.6871 0.0024
oral 0.1509 0.2311 0.3413 130.7238 0.0018
oesophagus 0.0203 0.0768 0.1583 18.0691 0.0043
rectum 0.2375 0.4315 0.6480 873.1753 0.0005
gallbladder -0.1736 -0.0274 0.1067 290.7825 -0.0001
pancreas 0.0661 0.1089 0.1608 283.0072 0.0004
leukemia 0.0792 0.1800 0.4092 81.1517 0.0022
solid 9.3445 11.0056 13.3345 6398.9234 0.0017
total 9.5265 11.1856 13.5145 6480.0751 0.0017

Future LAR:
Lower Mean Upper BFR TFR
colon 0.6658 1.0964 1.6314 1051.6541 1052.7505
lung 2.3673 2.8906 3.5837 1352.5231 1355.4137
uterus 0.1423 0.2825 0.4686 296.3270 296.6095
bladder 0.9503 1.2916 1.7359 154.4180 155.7096
remainder 2.8942 4.4153 6.5409 1792.3971 1796.8124
oral 0.1444 0.2181 0.3177 119.1196 119.3377
oesophagus 0.0201 0.0762 0.1568 18.0923 18.1685
rectum 0.2346 0.4186 0.6327 835.6204 836.0390
gallbladder -0.1735 -0.0274 0.1066 292.4045 292.3771
pancreas 0.0653 0.1079 0.1591 281.1836 281.2915
leukemia 0.0944 0.1773 0.3603 77.5662 77.7435
solid 9.1382 10.7697 12.9982 6193.7399 6204.5095
total 9.3172 10.9470 13.1770 6271.3060 6282.2530

Confidence Level: 0.9
Current Year: 2021

summaries of LAR result : Group male
.
.
.
Confidence Level: 0.9
Current Year: 2021
```

The above results illustrate that the estimated average LAR values (90% confidence intervals) for female and male are 11.1856(9.5265, 13.5145) and

27.1674(23.8700, 28.7939), respectively. **LAR_group** can consider two group variables such as gender and occupation in this toy 'organ' data as follows.

```
> ex_organ3 <- LAR_group(organ, pid=organ$ID, group=list(organ$sex, organ$occup), basedata=list(life2018, incid2018), current=2021)

> print(ex_organ3, max.id=2)
LAR result of female.1

LAR:
Lower Mean Upper
4.7547 6.2051 8.0758

Future LAR:
Lower Mean Upper
F.LAR 4.7348 6.1773 8.0388
BFR 3746.8142 3746.8142 3746.8142
TFR 3751.5490 3752.9915 3754.8530
---

LAR result of male.1

LAR:
Lower Mean Upper
37.9314 41.3941 45.0670

Future LAR:
Lower Mean Upper
F.LAR 32.3847 35.4936 38.7843
BFR 12300.1490 12300.1490 12300.1490
TFR 12332.5337 12335.6426 12338.9333

The results for 5 groups are omitted.
```

Concluding remarks

We proposed the **LARisk** package built in R to calculate projected and/or lifetime radiation related cancer risks after exposure. We described and considered the various uncertainties in the lifetime risk calculations to construct this package. Three main flexible and useful functions were developed, and many arguments/options are available to promptly update new knowledge from radiation epidemiology and related fields. Of particular interest is the ability to input population information, replacing cancer incidence rates and survival functions for the population of interest and observed year to be used for lifetime risk calculation. This greatly simplifies to employ new baseline incidence rate and population survival rate data for a given year. Detailed description of the arguments/options and related data formats, etc. are provided in the **LARisk** vignette (<https://cran.r-project.org/web/packages/LARisk/vignettes/LARisk-vignette.html>).

LARisk is an effective and simple method to calculate excess lifetime risk from radiation exposure, and can be applied to any population of interest.

References

1. Preston DL, Ron E, Tokuoka S, Funamoto S, Nishi N, Soda M, et al. Solid cancer incidence in atomic bomb survivors: 1958–1998. *Radiation research*. 2007;168(1):1–64.
2. Carpeggiani C, Paterni M, Caramella D, Vano E, Semelka RC, Picano E. A novel tool for user-friendly estimation of natural, diagnostic and professional radiation risk: Radio-Risk software. *European journal of radiology*. 2012;81(11):3563–3567.

3. Journy NMY, Lee C, Harbron RW, McHugh K, Pearce MS, Berrington de González A. Projected cancer risks potentially related to past, current, and future practices in paediatric CT in the United Kingdom, 1990–2020. *British journal of cancer*. 2017;116(1):109.
4. Berrington de Gonzalez A, Mahesh M, Kim KP, Bhargavan M, Lewis R, Mettler F, et al. Projected cancer risks from computed tomographic scans performed in the United States in 2007. *Archives of Internal Medicine*. 2009;169(22):2071–2077.
5. Evans JS, Abrahamson S, Bender MA, Boecker BB, Gilbert ES, Scott BR. Health effects models for nuclear power plant accident consequence analysis. Rev. 2 PART II, NUREG/CR-4214. U.S. Nuclear Regulatory Commission; 1993.
6. NRC. Health risks from exposure to low levels of ionizing radiation: BEIR VII phase 2 (National Research Council and others). vol. 7. National Academies Press; 2006.
7. Seo S, Lee DN, Jin YW, Lee WJ, Park S. Usefulness of cancer-free survival in estimating the incidence from radiation exposure. *Journal of Radiological Protection*. 2018;38(3):N17–N24.
8. Thomas D, Darby S, Fagnani F, Hubert P, Vaeth M, Weiss K. Definition and estimation of lifetime detriment from radiation exposures: principles and methods. *Health Physics*. 1992;63(3):259–272.
9. Vaeth M, Pierce DA. Calculating excess lifetime risk in relative risk models. *Environmental Health Perspectives*. 1990;87:83–94.
10. Kellerer AM, Nekolla EA, Walsh L. On the conversion of solid cancer excess relative risk into lifetime attributable risk. *Radiation and environmental biophysics*. 2001;40(4):249–257.
11. WHO. Health risk assessment from the nuclear accident after the 2011 Great East Japan earthquake and tsunami, based on a preliminary dose estimation. vol. 7. World Health Organization; 2013.
12. Faletra FF, D’Ángeli I, Klersy C, Averaimo M, Klimusina J, Pasotti E, et al. Estimates of after a single radiation exposure from 64-slice computed tomographic coronary angiography. *Heart*. 2010;96(12):927–932.
13. Koral K, Blackburn T, Bailey AA, Koral KM, Anderson J. Strengthening the argument for rapid brain MR imaging: estimation of reduction in lifetime attributable risk of developing fatal cancer in children with shunted hydrocephalus by instituting a rapid brain MR imaging protocol in lieu of Head C. *American Journal of Neuroradiology*. 2012;33(10):1851–1854.
14. Aw-Zoretic J, Seth D, Katzman G, Sammet S. Estimation of effective dose and lifetime attributable risk from multiple head CT scans in ventriculoperitoneal shunted children. *European Journal of Radiology*. 2014;83(10):1920–1924.
15. Ivanov VK, Kashcheev VV, Chekin SY, Menyaylo AN, Pryakhin EA, Tsyb AF, et al. Estimating the lifetime risk of cancer associated with multiple CT scans. *Journal of Radiological Protection*. 2014;34(4):825–841.
16. Journy N, Ancelet S, Rehel JL, Mezzarobba M, Aubert B, Laurier D, et al. Predicted cancer risks induced by computed tomography examinations during childhood, by a quantitative risk assessment approach. *Radiation and Environmental Biophysics*. 2014;53(1):39–54.

17. Salibi PN, Agarwal V, Panczykowski DM, Puccio AM, Sheetz MA, Okonkwo DO. from CT among patients surviving severe traumatic brain injury. *American Journal of Roentgenology*. 2014;202(2):397–400.
18. Hoang JK, Reiman RE, Nguyen GB, Januzis N, Chin BB, Lowry C, et al. Lifetime Attributable Risk of Cancer From Radiation Exposure During Parathyroid Imaging: Comparison of 4D CT and Parathyroid Scintigraphy. *American Journal of Roentgenology*. 2015;204(5):W579–W585.
19. Andersson M, Eckerman K, Mattsson S. Lifetime attributable risk as an alternative to effective dose to describe the risk of cancer for patients in diagnostic and therapeutic nuclear medicine. *Physics in Medicine and Biology*. 2017;62(24):9177–9188.
20. Journy NM, Lee C, Harbon RW, McHugh MS K Pearce, Berrington de Gonzalez A. Projected cancer risks potentially related to past, current, and future practices in paediatric CT in the United Kingdom, 1990-2020. *British Journal of Cancers*. 2017;116(1):109–116.
21. Tamura M, Sakurai H, Mizumoto M, Kamizawa S, Murayama S, Yamashita H, et al. Lifetime attributable risk of radiation-induced secondary cancer from proton beam therapy compared with that of intensity-modulated X-ray therapy in randomly sampled pediatric cancer patients. *Journal of Radiation Research*. 2017;58(3):363–371.
22. Branchini M, Del Vecchio A, Gigliotti CR, Loria A, Zerbi A, Calandrino R. Organ doses and lifetime attributable risk evaluations for scoliosis examinations of adolescent patients with the EOS imaging system. *European Journal of Trauma and Emergency Surgery*. 2018;123(4):305–313.
23. Berrington De Gonzalez A, Apostoaei AI, Veiga LHS, Rajaraman P, Thomas BA, Hoffman FO, et al. RadRAT: a radiation risk assessment tool for lifetime cancer risk projection. *Journal of Radiological Protection*. 2012;32(3):205.
24. Lee WJ, Choi Y, Ko S, Cha ES, Kim J, Kim YM, et al. Projected lifetime cancer risks from occupational radiation exposure among diagnostic medical radiation workers in South Korea. *BMC cancer*. 2018;18(1):1206.
25. Fresco R, Spera G, Meyer C, Cabral P, Mackey JR. Imaging radiation doses and associated risks and benefits in subjects participating in breast cancer clinical trials. *The oncologist*. 2015;20(7):702–712.
26. Furukawa K, Cologne JB, Shimizu Y, Cullings HM. Predicting future excess events in risk assessment. *Risk Analysis*. 2009;29(6):885–899.
27. Furukawa K, Misumi M, Cologne JB, Cullings HM. A Bayesian semiparametric model for radiation dose-response estimation. *Risk Analysis*. 2016;36(6):1211–1223.
28. Richardson DB, Cardis E, Daniels RB, Gilies M, Haylock R, Leurand K, et al. Site-specific solid cancer mortality after exposure to ionizing radiation: A cohort study of workers (INWORKS). *Epidemiology*. 2018;29(1):31–40.
29. Turner H, Firth D. Generalized nonlinear models in R: An overview of the gnm package. ESRC National Centre for Research Methods; 2007.

30. Rühm W, Woloschak GE, Shore RE, Azizova TV, Grosche B, Niwa O, et al. Dose and dose-rate effects of ionizing radiation: a discussion in the light of radiological protection. *Radiation and Environmental Biophysics*. 2015;54(4):379–401.
31. Little MP, Pawel D, Misumi M, Hamada N, Cullings HM, Wakeford R, et al. Lifetime Mortality Risk from Cancer and Circulatory Disease Predicted from the Japanese Atomic Bomb Survivor Life Span Study Data Taking Account of Dose Measurement Error. *Radiation Research*. 2020;194(3):259–276.
32. ICRP. The 2007 Recommendations of the International Commission on Radiological Protection. ICRP Publication 103; 2007.
33. Land C, Gilbert E, Smith JM, Hoffman FO, Apostoaei AI, Thomas B, et al. Report of the NCI-CDC working group to revise the 1985 NIH radioepidemiological tables. US Department of Health and Human Services, National Institutes of Health ; 2003.