

『放射線被ばくによるがんのリスク推定に利用できるコンピュータプログラムを公開』
～オープンソースにより計算方法を広く専門家間で共有～

【ポイント】

- 日本保健物理学会の専門研究会は、放射線被ばくによる発がんのリスク推定に利用できるコンピュータプログラム(名称：SUMRAY)を開発しました。
- 専門家間で計算方法を広く共有するため、プログラムのソースコード^{*1}とデータベースをインターネット上で無償公開し、プログラムの自由な改変や二次利用を認めています。
- 本プログラムは、1回被ばくあるいは慢性被ばく^{*2}からの放射線リスクを、ある年齢までのがんに対する罹患もしくは死亡の確率(累積リスク)として計算します。
- 計算したリスクの不確かさについても、プログラムで使用したデータのばらつきを加味して算出します。

論文: M Sasaki, K Furukawa, D Satoh, K Shimada, S Kudo, S Takagi, S Takahara, M Kai; SUMRAY: R and Python codes for calculating cancer risk due to radiation exposure of a population. JRPR, 2023.

URL: <https://doi.org/10.14407/jrpr.2022.00213>

プログラム公開先: <https://github.com/JapanHealthPhysicsSociety/SUMRAY>

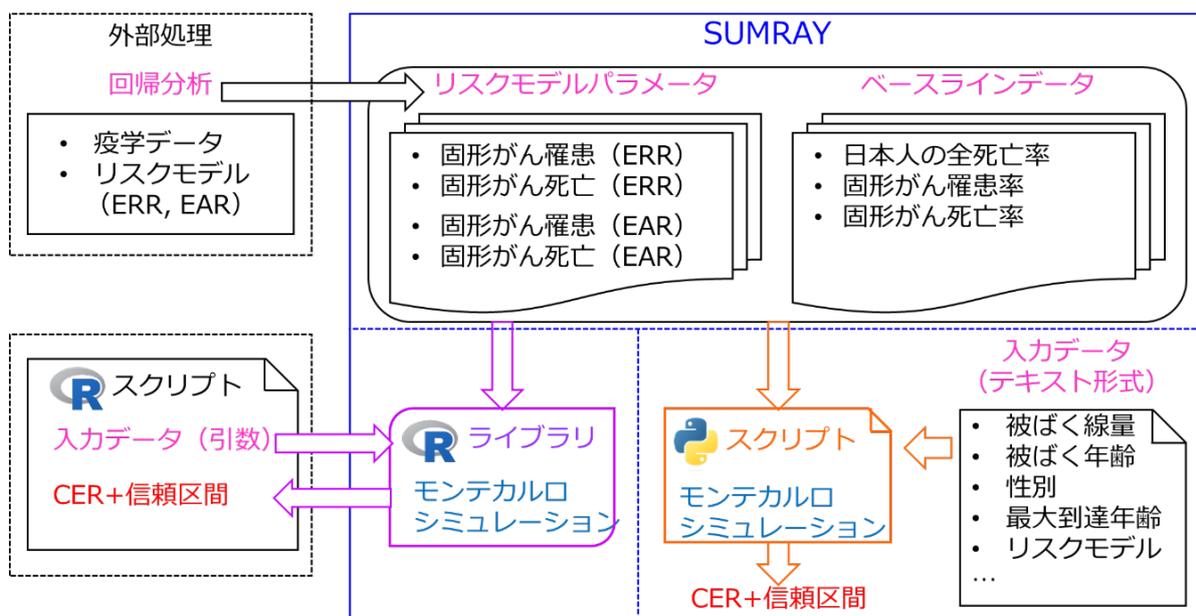


図 1 開発した放射線被ばくによる発がんのリスク推定に利用できるコンピュータプログラム

【概要】

放射線の健康影響は社会の大きな関心事です。福島第一原子力発電所事故が起きたときに、放射線の健康影響が起きる可能性は被ばくの大きさ(mSv:ミリシーベルト)とどのように関係しているのか、ネットやメディアで話題となりました。

これまでも、低線量(100mSv 以下)での放射線被ばくによる発がんの可能性(リスク)は科学的に議論され、リスクの定量的な計算は国際機関でも行われています。しかし、一般に確率を使って表現されることの多いリスクは理解することの難しい数値です。そこで、背後に使われている情報やモデルによる計算方法を広く伝えていくことで、リスクのより深い理解につながることを私たちは目指しました。

そのために、日本保健物理学会の専門研究会では、放射線リスク計算を行う専門家やリスク計算をより深く理解したい専門家と共有するオープンソースプログラム(SUMRAY)を作成し公開しました。将来的には、計算モデルなどを拡張し一般の方でも使いやすい計算ツールに発展させていきたいと考えています。

計算プログラム SUMRAY は、利用者の多いプログラミング言語^{※3}である R 言語^{※4}および Python^{※5} によって作成されたリスク推定プログラムです。本プログラムでは、原爆被爆者の疫学調査^{※6} 研究と同様のリスクモデル^{※7}(図 1 の ERR^{※8}、EAR^{※9})に基づき、ある年齢までにがん死亡するあるいは罹患する確率(図 2 左は到達した年齢でのリスク、図 2 右はそれらを合算した累積のリスク)を計算します。また、その推定リスクは、疫学データのばらつきを考慮してモンテカルロシミュレーション^{※10} を用いて評価した信頼区間とともに計算されます。

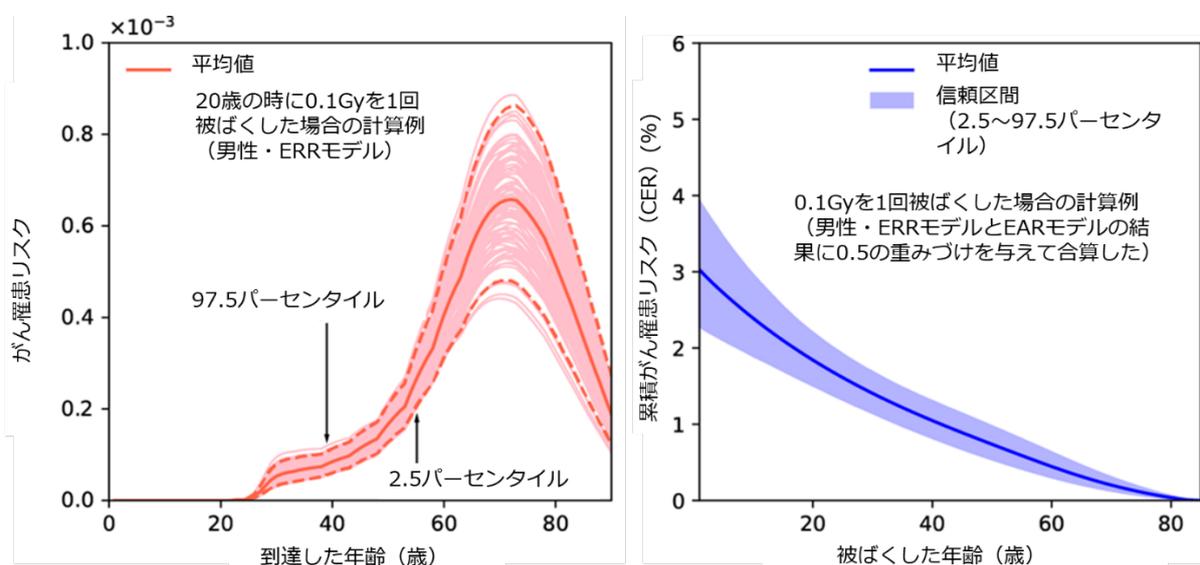


図 2 リスク計算結果の例。左図は 20 歳の時に 0.1Gy^{※11} を 1 回被ばくした場合の到達年齢におけるがん罹患リスクを示し、右図は 0.1Gy をさまざまな被ばく年齢で 1 回被ばくした場合の 90 歳に至るまでの累積がん罹患リスクを 2.5~97.5 パーセンタイル^{※12} の信頼区間^{※13} とともに示す。

(最後に伝えたいこと)

一般の方が放射線リスクを計算するのに利用できるコンピュータプログラムは、すでに、米国国立がん研究所が誰でもアクセスできる環境で Web 上に公開しています (<https://radiationcalculators.cancer.gov/radrat/>)。私たちも、一般の方が使用できるような計算ツールを目指した研究を行う必要があります。一方で、リスク推定値を使用する上で理解しておくべき点をわかりやすく伝えていく努力が必要です。そうでなければ、リスク推定値が誤解や間違った使われ方をしてしまうことでリスク計算そのものの信頼を失っていくでしょう。この観点から私たちは、まず計算に使用する手法とデータのすべてをオープンソースプログラムとして公開しました。今後も人々にとって有用なリスク情報となるように発展させていきたいと考えています。

【用語説明】

※1 ソースコード

人間が理解・記述しやすい言語やデータ形式で書き記されたコンピュータプログラムのこと。

※2 慢性被ばく

ある期間にわたって比較的低い線量を繰り返し被ばくすること。

※3 プログラミング言語

ソースコードを記述してコンピュータプログラムを作成するための言語。プログラミング言語ごとに記述仕様が決められている。

※4 R 言語

統計解析を主目的に設計されたプログラミング言語。オープン・ソース・ソフトウェアであるため、誰でも無料で利用することができる。

※5 Python

さまざまな開発に対応できる汎用プログラミング言語。オープン・ソース・ソフトウェアであるため、誰でも無料で利用することができる。

※6 疫学調査

対象とする集団において、疾病の原因と思われる因子と疾病との関連を調べる調査。放射線被ばくと発がんとの関連については、原爆被爆者を対象とした寿命調査(LSS, Life Span Study)が大規模疫学データとして特に重視される。

※⁷ リスクモデル

疫学調査によるデータを統計的に処理することによって、データの背景事象、例えば「放射線を被ばくするとがん死亡率が直線的に増加する」等を数学的に表したものの。低線量での慢性被ばくのようにデータが十分でない場合には、モデルは推定のためのツールとして利用される。

※⁸ ERR(Excess Relative Risk, 過剰相対リスク)

放射線被ばく集団における疾病の罹患率や死亡率を、放射線に被ばくしなかった集団における疾病の罹患率や死亡率(ベースラインデータ)で割ったもの(相対リスク)から、ベースラインデータでのリスク(すなわち 1)を引いたもの。相対リスクのうち、放射線被ばくに起因するリスク因子による過剰な部分を指す。

※⁹ EAR(Excess Absolute Risk, 過剰絶対リスク)

放射線被ばく集団における疾病の罹患率や死亡率から、放射線に被ばくしなかった集団における疾病の罹患率や死亡率(ベースラインデータ)を引いたもの。

※¹⁰ モンテカルロシミュレーション

乱数を発生させて行うシミュレーション。SUMRAY ではリスク推定の不確かさを計算するために、モデルのパラメータの不確かさを確率分布として扱い、繰り返し計算するシミュレーションを行っている。

※¹¹ Gy(グレイ)

放射線により物質に与えられたエネルギーを物質の質量で割った単位(J/kg)の特別な名称。

※¹² パーセンタイル

データを小さい数字から大きい数字に並べ替えて百分率で表した際に、何パーセント目に位置するかを表す量。50 パーセンタイルはデータの中央値を意味する。

※¹³ 信頼区間

統計学で母集団の真の値が含まれることの確からしさを表現する数値範囲のこと。例えば 95%信頼区間とは、母集団の真の値が 100 回試験して 95 回は当該数値範囲に含まれることを意味する。

【問い合わせ先】

一般社団法人日本保健物理学会 事務局
TEL: 03-6205-4649 / FAX: 03-6205-4659
E-mail: exec.off@jhps.or.jp