

小特集

第 76 回日本公衆衛生学会総会 シンポジウム 6

疫学研究の意義とその活用を検討する－放射線に関連した労働者の健康を守るために－

2017 年 10 月 31 日 (火) 13:30～15:00 第 8 会場(かごしま県民交流センター3F 大研修室 2)

報告ならびに発表者による論考 (4)

放射線疫学の課題:マーケティングの観点から

濱岡豊 (慶応大学商学部)

hamaoka@fbc.keio.ac.jp

はじめに

福島原発災害以降、放射線被曝の問題が重要化している。筆者は修士課程では原子力工学を学んだが、その後、研究分野を変更し、マーケティング・リサーチなどを大学で教えている。マーケティングでは、例えば広告を増加した場合に、売上が増加するのか?といった問題をデータ分析によって明らかにしている。

福島核災害以降、「100mSv 未満の線量なら発がんリスクなし¹⁾」「低線量域では、被ばく線量が約 150 mGy あたりから統計学的に有意ながん罹患リスク増加がみられている。²⁾」といった説明がされてきた。

データ分析は筆者の専門分野なので、どのような分析が行われるのか、放射線疫学に関する論文をよみ、公開されているデータを用いて再分析を行ってきた。放射線疫学にはいくつかの問題があるが(濱岡 2015a)、本稿では、個人データを集計して分析すること、サンプルを限定して推定することの問題を紹介する。そして、原発従業員の個票データを用いることによって、集計したデータでは検出できなかった被曝の影響を検出できることを示す(濱岡 2015b)。統計学に関する内容も含むが、詳細に説明すると読みにくくなるため、かなりいい加減に記述する。統計学の知識のある方は自分で補足しながら、また知識のない方は統計学の本を参考にして頂きたい。

マーケティングにおけるデータ分析の例

疫学よりもマーケティングの方が身近な話題なので、それからはじめよう。10 店舗の広告費と売上高

の仮想例を図表 1 に示した。このデータについて、広告費を横軸、売上高を縦軸にプロットした(図表 2(a))。二つの変数には右上がりの直線的な関係があることがわかる。回帰分析という手法は、測定された二つの変数 x と y に次式のような直線的な関係があることを想定して、データにもっともよくあてはまるように、二つのパラメータ α と β を推定する手法である。

$$y = \alpha + \beta x$$

ここで α は「切片」とよばれ、広告費 x がゼロの場合の売上高 y を意味する。 β は直線の「傾き」を表し、広告費が 1 単位、(ここでは百万円)増加したときの売上高の変化分を意味する。上式で $\beta = 0$ とすると、 $y = \alpha$ となり、 y は x と無関係に一定の値 α となる。このように、 β がゼロか否かは、 x と y の関係の有無を表す重要な意味をもつ。このため、 $\beta = 0$ を帰無仮説として統計的検定を行う。

図表 2(a-1)には、回帰分析の推定結果を示した。傾きのパラメータの推定値は 0.593 となっている。正の値が得られているが、わずか 10 店舗のデータから推定したので、誤差を考慮して結論付ける必要がある。そのための統計的検定には表にある t 値や P 値が用いられる。サンプルサイズに依存するのだが t 値の絶対値が 2 よりも大きければ、 P 値が 0.05 よりも小さくなり、 $\beta = 0$ という帰無仮説を棄却することが多い。

ここでの t 値は 26.59 となっており、2 よりもずっと大きく P 値も 0 に近い。このため、 $\beta = 0$ という(帰無)仮説は棄却される。つまり推定された 0.593 は 0 ではないといえることになる。つまり、広告費が百万円多い店は、売上高が 0.593 億円高いといえる³。この結果を踏まえて、マーケティング担当者は広告を増加させるという意思決定をすることになる。

データのカテゴリ化による問題

ここまでの分析は店舗レベルのデータをそのまま用いて分析したが、次のようにデータを加工して分析するとどうなるだろうか。つまり、広告費については、「5 百万円未満」「5 百万円以上、1 千万円未満」「1 千万円以上、2 千万円未満」「2 千万円以上」に 4 区分する。同様に売上高については「3 億円未満」「3 億円以上、10 億円未満」「10 億円以上」の 3 区分とする。

図表 1(b)には、このようにカテゴリ化したデータを、図表 2(b)には、これの散布図を示す。生のデータと比べると直線的な関係がボケていることがわかる。図表 2 (b-1)には、このデータを用いた回帰分析の推定結果も示した。傾きのパラメータは 0.448 となっている。 t 値は 4.93 と 2 を越え P 値も 0.001 であり、 $\beta = 0$ という帰無仮説は棄却される。ただし、 t 値は生のデータを用いたときの 26.59 と比べると大幅に小さくなっている。このように生のデータをカテゴリ化すると、データがもっていた情報が失われ、変数間の有意な関係を検出しにくくなるのである。

なお、ここでは広告費については 4 区分、売上高については 3 区分としたが、いくつかの区間に区分するのか、さらに区間の幅をどう設定するのかといった恣意性の問題もある。

データを限定することによる問題

前節ではすべてのデータを用いたが、一部のデータを用いるとどうなるだろうか。直感的にはサンプルサイズが小さくなるので、変数間の影響を検出しにくくなることが予想される。図表 1 のデータについて、広告費の少ない 5 店舗のデータ、つまり図表 2(a)の左隅の 5 点のみを用いて同様に推定した。

図表 2(a-2)に示すように、傾きの係数は 0.413、t 値は 2.53 とさらに小さくなった。この場合 P 値は 0.05 つまり 5%よりも大きいので、 $\beta = 0$ という帰無仮説は 5%水準では棄却できないことになる。なお、カテゴリ化したデータで同様に分析したが、推定自体が不可能であったため、結果は示していない。これは図表 1(b)で、広告費が少ない 4 店舗(店舗 1,3,6,8)のデータは広告費 5 億円未満、売上高 3 億円未満と同じ値となり、直線を引けないためである。このように、サンプルを限定することも変数間の関係を検出しにくくするのである。

図表 1 10 店舗の広告費と売上高の仮想データ

(a) 生のデータ				(b) カテゴリ化したデータ	
生のデータ				カテゴリ化したデータ	
店舗番号	店舗番号	広告費(百万円)	売上高(億円)	広告費(百万円)	売上高(億円)
1	1	1	2	<5	<3
2	2	25	15	20 \geq	<10
3	3	5	3	<5	<3
4	4	16	8	<20	<10
5	5	29	18	20 \geq	10 \geq
6	6	2	1	<5	<3
7	7	32	19	20 \geq	10 \geq
8	8	3	2	<5	<3
9	9	6	4	<10	<10
10	10	10	6	<10	<10

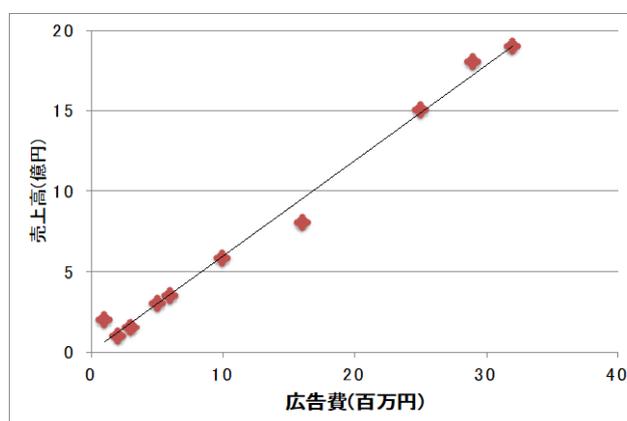
注)「カテゴリ化したデータ」の記号は下記の通り。

広告費 <5:「5 百万円未満」、<10:「5 百万円以上、1 千万円未満」、<20:「1 千万円以上、2 千万円未満」、20 \geq :「2 千万円以上」

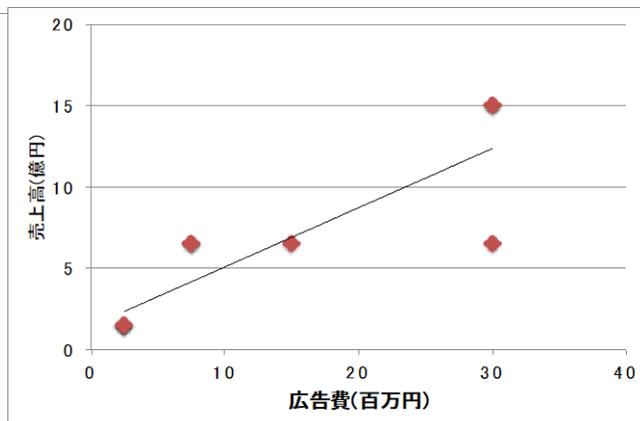
売上高 <3:「3 億円未満」、<10:「3 億円以上、10 億円未満」、10 \geq :「10 億円以上」

図表2 散布図と回帰分析の結果

(a) 生のデータ



(b) カテゴリ化したデータ



(a-1)生のデータを用いた回帰分析の結果

	係数	t 値	P-値
切片	0.036	0.09	0.927
広告費(百万円)	0.593	26.59	0.000***
N=10, R ² =0.989			

(b-1)カテゴリ化したデータを用いた回帰分析の結果

	係数	t 値	P-値
切片	1.053	0.78	0.460
広告費(百万円)	0.448	4.93	0.001***
N=10, R ² =0.752			

(a-2)生のデータ(広告費の少ない5店舗)を用いた回帰分析の結果

	係数	t 値	P-値
切片	0.797	1.263	0.296
広告費(百万円)	0.413	2.534	0.085*
N=5, R ² =0.680			

注1)カテゴリ化したデータでは、各カテゴリの中央値を用いてプロット、分析した。なお、広告費「2千万円以上」については3千万円、売上高「10億円以上」については15億円とした。

注2) ***:1%水準で有意 **:5%水準で有意 *:10%水準で有意 *なし 10%水準で有意でない。

個人レベルのデータ分析のためのモデル

ここまでは、yについて店舗の売上高という集計されたデータを考えてきた。店舗の売上は、個人の購買を集計したものであり、店舗には多くの消費者が訪れることを考えれば、サンプルサイズが多い個人レベルのデータを用いた方が、より詳細に広告や価格などの影響を検出できるはずである。

現在では、買い物に行くとレジで商品のバーコードを読み、価格がレジに表示されるようになっている。メンバーズカードによって個人を識別できれば、「誰が」「いつ」「何を」「いくらで」購入したのかも特定できる。1980年代以降の店頭の情報化によって、マーケティングではこのような個人レベルの情報を簡単に集められるようになった。

このような個人レベルの購買を分析するには、ある商品を購入した(1)、しない(0)を y 、価格や広告への露出回数を x として分析すればよい。ただし、 $y=0,1$ とした回帰分析には問題があるため、購入する確率 $P(y=1)$ を説明するモデルが開発されている。確率について、ロジスティック関数を想定した次式は「二項ロジットモデル」と呼ばれ、回帰分析の一種と考えることができる。

$$P(y=1)=1/(1+\exp(-\alpha -\beta x))$$

店舗には多くの製品があふれている。消費者の意思決定を考えると、二項ロジットモデルのように、個別の商品を買ったか、買わなかったかを分析するのではなく、複数の商品の中から一つを選ぶという選択行動を想定する方が適切だろう。このように複数の商品($j=1\sim J$)の中から j' を選ぶ確率は、次式で与えられる。ここで x_j は商品 j の価格などである。これは多項ロジットモデルと呼ばれる。

$$P(j' \text{ を選択}) = \exp(-\alpha_{j'} - \beta x_{j'}) / \sum \exp(-\alpha_j - \beta x_j)$$

放射線疫学における分析の問題点

ここまでは、マーケティングにおけるデータを例に論じてきたが、放射線疫学では、被曝量 x と例えば「がんによる死亡率」 y との関係性を分析する。そこで用いられるのも前述の回帰分析である。扱う対象は異なるが、用いる手法は同じなのである。

放射線疫学では例えば、原爆被爆者や原子力施設従業者「個人」の被曝量を測定、記録し、健康状態を追跡する。このように個人レベルのデータが収集されているにも係わらず、放射線疫学の主要な研究では、被曝量を「10mSv 未満」「10mSv 以上、50mSv 未満」のようにカテゴリ化し、年齢などの連続変数も同様に加工して層別集計したデータを用いて分析している⁴。前述のように、このように集計すると、被曝量の影響が検出しにくくなるはずである。

原爆被爆者については、被曝量を 22 区分して集計したデータが公開されている⁵。ここまでに指摘したカテゴリ化による問題を確認するために、このデータについて、オリジナルの被曝量区分 22 カテゴリを 11 カテゴリ、6 カテゴリへと集計して推定した(図表 3)。集計にともなって、傾きの推定値が 0.413 から 0.408、0.391 と小さくなる一方、 t 値と同様に解釈できる z 値も小さくなっており、変数間の有意な関係を検出しにくくなっていることが確認できる。

図表3 被曝量区分の集計度を変更した推定結果

	係数	z値	P値	レコード数
22カテゴリ	0.413	8.07	0.000***	53782
11カテゴリ	0.408	7.84	0.000***	36275
6カテゴリ	0.391	7.34	0.000***	29036

注)線量のカテゴリ数と区間は以下の通り(単位 Gy)。

・オリジナルデータ:22 カテゴリ: (<0.005, <0.02, <0.04, <0.06, <0.08, <0.1, <0.125, <0.15, <0.175, <0.2, <0.25, <0.3, <0.5, <0.75, <1.0, <1.25, <1.5, <1.75, <2.0, <2.5, <3.0, 3.0≥)

・11 カテゴリ:(<0.02, <0.06, <0.1, <0.15, <0.2, <0.3, <0.75, <1.25, <1.75, <2.5, 2.5≥)

・6 カテゴリ:(<0.06, <0.15, <0.3, <1.25, <2.5, 2.5≥)

・固形がん死について結腸線量を用いて推定。係数は過剰相対リスク係数/Gy。分析の詳細は濱岡(2015a)を参照のこと。

前節で紹介したように、サンプルを限定するとどうなるだろうか。これについては、寿命調査13報 (Preston et al.2003)の結果を紹介しよう。彼らは固形がん死と被曝量について、被曝線量範囲を限定して分析した(図表4)。例えば0-0.05Svまでのサンプルを用いた場合、係数は0.93だが、t値は1.09、P値は0.150であり有意ではない。0-0.1Svまでを用いても有意ではないが、0-0.125Svつまり125mSvまでを用いると有意になっている。はじめに紹介した「100ミリシーベルト以下の低線量域では、明らかな発がんリスクの増加は確認されていません。」といった記述は、このようにサンプルを限定した分析結果踏まえたものだと考えられる。

前述のように、サンプルを限定すると変数間の関係を検出しにくくなる。全サンプルを用いた分析を行うべきである。そもそも、(個人レベルの)データをカテゴリ化して、集計し、さらにその一部を取り出すという分析は不適切である。

図表4 被曝線量範囲を限定した分析

推定に用いた線量範囲 Sv	係数	t値	P値
0-0.05	0.93	1.09	0.150
0-0.1	0.64	1.16	0.300
0-0.125	0.74	1.95	0.025**
0-0.15	0.56	1.75	0.045**
0-0.2	0.76	2.62	0.003***
0-0.5	0.44	3.67	<0.001***
0-1	0.47	4.70	<0.001***
0-2	0.54	7.71	<0.001***
0-4	0.47	9.40	<0.001***

注)P値は片側検定。***:1%水準で有意 ** :5%水準で有意 * :10%水準で有意

出所) LSS13報(Preston et al.2003)のTable 4より作成。

Hanford 他米国核関連 3 施設従業者調査データの再分析

残念ながら日本では放射線疫学に関する個票データは公開されていない。一方で、米国では多くのデータセットが公開されている⁶。ここでは、比較的サンプルサイズが大きい Gilbert et al.(1993)らのデータセットを用いて、個人レベルのデータを分析した例を紹介する。

これは米国の核関連 3 施設(Hanford、Oak Ridge、Rocky Flats)従業者 4.5 万人を 40 年以上追跡したデータであり、累積被曝量は平均 21-36mSv となっている(図表 5)。個人レベルのデータだが、Gilbert et al.は、被曝量を「10mSv 未満」「10mSv 以上、50mSv 未満」「50mSv 以上、100mSv 未満」「100mSv 以上、200mSv 未満」「200mSv 以上、400mSv 未満」「400mSv 以上」に 6 区分した。区分数や区間の設定には恣意性の問題があるのは前述の通りだが、このデータでは死亡者 7,863 名のうち 5,424 名が「10mSv 未満」に集中している。低線量の被曝の影響を分析するならば、この部分はもっと細かく分割すべきだろう。この他、年齢などの連続量についても同様にカテゴライズして集計したデータに対して、Mantel Haenszel トレンド検定、ポアソン回帰を適用し、被曝量は有意とならなかったとしている(図表 6 左)⁷。

このデータについては、匿名化した個票データが公開されているので、集計せず個人レベルのモデルを適用した。まず、観測終了までに各死因で死亡したか否かについて、「2 項ロジット・モデル」で推定した。図表 6 に示したように、白血病、外因性死を除いて係数は正で有意となった。さらに、固形がん、白血病、その他のがん、非がん、外因性、不明のいずれかで亡くなることを考慮した「多項ロジットモデル」で分析した。二項ロジット同様、白血病、外因性以外は正で有意な係数が得られた。

このように、集計されたデータを用いた分析では有意な関係を検出できなかったが、個人レベルのデータを用いれば有意な影響が検出できるのである。

図表5 Hanford 他3施設データの概要

	全サンプル			分析対象		
	Hanford	Oak Ridge	Rocky Flats	Hanford	Oak Ridge	Rocky Flats
合計	44,156	8,318	7,616	33,973	6,743	6,788
性別						
男性	31,488	8,318	7,616	25,705	6,743	6,788
女性	12,668	0	0	8,268	0	0
フォローアップ						
開始	1944	1943	1952	1944	1944	1952
終了	1989	1984	1987	1989	1984	1987
累積線量 (mSv)						
平均	23.5	17.3	32.2	25.4	21.1	35.6
中央値	3.0	1.4	7.4	3.7	3.5	9.7
最大	1477.0	1144.0	726.0	1477.0	1144.0	726.0
死因						
全体	9771	1433	794	7012	1208	719
ガン	2390	352	214	1732	316	194
固形ガン	2133	302	186	1540	271	171
白血病	87	28	10	62	26	10
その他ガン	170	22	18	130	19	13
非ガン	6145	891	479	4446	741	437
外因性	911	172	100	618	137	87
不明	325	18	1	216	14	1

注)分析対象は一年間に250mSv以上を被曝した者および従業期間6ヶ月未満の者等を除く。

図表6 Hanford 他3施設従業者データを用いた推定結果

	集計データを用いた分析 (Guilbert et. al. 1993)		個人レベルデータを用いた分析 (濱岡 2015b)	
	トレンド検定 検定統計量 (a)	ポアソン回帰 係数 (b)	二項ロジット 係数(d)	多項ロジット 係数 (e)
全体	-0.25		2.55**	
がん	-0.04	-0.0 (<0-0.8)	2.22**	
(白血病を除く)		0.0 (<0-0.8)	2.37**	
固形がん			1.88*	1.70*
白血病		-1.0 (<0-2.2)	-0.38	-0.40
その他がん			2.02*	2.22**
非がん	-0.08		1.78*	2.50**
外因性	-1.85*		-0.14	-0.29
不明	-1.46		2.48**	2.50**

(a) Mantel-Haenszel 方による傾向性の検定 (Gilbert et al.1993 の Table II).

(b) ポアソン回帰による ERR(過剰相対リスク係数) と 90%信頼区間(同 Table VI)

(c)二項もしくは多項ロジットの推定値。

(d)当該の死因で死亡したか否かを二項ロジットで推定した結果。

(e)「生存」を基準として各死因での死亡を多項ロジットで推定した結果。

注)***:1%水準で有意 ** :5%水準で有意 * :10%水準で有意

まとめと低線量被曝問題への市民によるアプローチ

本稿では、マーケティングにおける回帰分析の例を紹介し、データをカテゴリ化したり、一部を取り出すことによって変数間の有意な関係が得られにくくなることを示した。そして、原爆被爆者のデータを用いて、このことが放射線疫学のデータでも成立していることを確認した。さらに、個票データが公開されている米国のデータに個人レベルのモデルを適用し、集計レベルの分析では検出できなかった被曝の有意な影響が検出されることを示した。

ここで紹介したのは原爆被爆者、米国 3 施設のデータだが、被曝量などの連続量をカテゴリ化し、分析するという方法は、Chernobyl 周辺住民(Likhtarov et al. 2006)、Mayak 従業員(Gilbert et al. 2013)、Techa 川流域住民(Schonfeld et al. 2013)、国内の原発従業員(放射線影響協会 2015)、原発従業員の国際比較(Cardis et al.2007; Leuraud et al.2015)などでも用いられている。

分析には放射線影響研究所で長くデータ分析に携わった研究者が開発した、Epicure というソフトウェアの AMFIT というルーチンが用いられていることも共通している。このソフトウェアが開発された 1980 年代は、コンピュータの処理能力が低く、推定には一晩かかったという⁸。しかし、現在ではそのような制約は存在しない。例えば、本稿で紹介した二項ロジット、多項ロジットモデルの推定は、筆者の個人所有パソコンでも数秒で推定できる。

集計する事のメリットは現在となっては、匿名化できることぐらいしか思い当たらない。本稿では、「放射線疫学」の問題点として批判したが、例えば、同分野の Nagataki et al.(1994)、Imaizumi et al.(2015)は原爆被爆者のうち約 2 万人に対して行っている、成人健康調査対象者について、被曝量と甲状腺結節との関係を分析した。個人レベルのデータに二項ロジットモデルを適用し、2668 名というサンプルサイズでも被曝量の有意な係数が得られている。この他、英国での自然被ばくと小児がん(Kendall et al.2012)⁹、台湾のコバルト汚染アパート(Hwang et al. 2008)でも二項ロジットモデル等が用いられ、有意な係数が得られている。

放射線防護基準は、原爆被爆者データをカテゴリ化したデータを用いて推定した結果が用いられており、リスクを過小評価している可能性が高い。「主流」放射線疫学が、個別データを集計して分析している理由はわからないが、個人レベルのデータで再分析すべきである。

このような時代遅れの分析をしないように、当該分野の専門家への統計モデリングについての教育が必要である。なお、「100mSv 以下にサンプルを限定すると、係数 $\beta = 0$ という帰無仮説が棄却されなかった」という結果を、はじめに紹介したように「100mSv 以下では健康への影響がない」と解釈する例もみられる。これは仮説検定の意味やサンプルを限定することの問題が理解されていないためである。専門家の統計的知識の向上とあわせて、このような情報を伝えるメディア関係者、さらには市民の統計リテラシーの向上も重要である。

最後に、本稿で紹介した再分析が可能になったのは、データが公開されているからである。原爆被爆者、

原子力施設従業員、福島県甲状腺調査など、日本でも放射線関連のデータは収集されている。より適切な分析を行うためにも、匿名化された個票データを一定のルールのもとに公開すべきである。

謝辞等

・広島、長崎被曝者データ

用いたデータは広島および長崎の放射線影響研究所（放影研）から入手したものである。放影研は、日本の厚生労働省（厚労省）ならびに米国のエネルギー省（DOE）により資金提供を（後者については、その一部を米国学士院に対する DOE 研究助成金 DE-HS0000031 を通じて）受けている公益財団法人である。この報告書に示した結論は著者のものであり、必ずしも放影研またはその資金提供機関の判断を反映するものではない。

・米国核施設労働者データ

Access to nuclear worker data was granted by the US DOE CEDR project. The protocol and results of this study were not reviewed by the DOE. The results and conclusions do not necessarily reflect those of the US Government or DOE.

・本研究の遂行には慶應大学学事振興資金および科研費(課題番号 15K03674)から支援を受けた。

参考文献

- Cardis E. et al.(2007), "The 15-Country Collaborative Study of Cancer Risk among Radiation Workers in the Nuclear Industry: Estimates of Radiation-Related Cancer Risks," *Radiation Research*, 167 (4), 396-416.
- Gilbert, E. S. et al. (2013), "Lung Cancer Risks from Plutonium: An Updated Analysis of Data from the Mayak Worker Cohort," *Radiat Res*, 179 (3), 332-42.
- Gilbert, Ethel S., Donna L. Cragle, and Laurie D. Wiggs (1993), "Updated Analyses of Combined Mortality Data for Workers at the Hanford Site, Oak Ridge National Laboratory, and Rocky Flats Weapons Plant," *Radiation Research*, 136 (3), 408-21.
- Imaizumi, M. et al. (2015), "Association of Radiation Dose with Prevalence of Thyroid Nodules among Atomic Bomb Survivors Exposed in Childhood (2007-2011)," *JAMA Intern Med*, 175 (2), 228-36.
- Kendall, G. M. et al. (2012), "A Record-Based Case- Control Study of Natural Background Radiation and the Incidence of Childhood Leukaemia and Other Cancers in Great Britain During 1980- 2006," *Leukemia*, 27 (1), 3-9.
- Leuraud et al.(2015), "Ionising Radiation and Risk of Death from Leukaemia and Lymphoma in Radiation-Monitored Workers (Inworks): An International Cohort Study," *The Lancet Haematology*, 2 (7), e276-81.
- Likhtarov, I. et al. (2006), "Post-Chornobyl Thyroid Cancers in Ukraine. Report 2: Risk Analysis," *Radiation Research*, 166 (2), 375-86.
- Nagataki, S.et al.(1994), "Thyroid Diseases among Atomic Bomb Survivors in Nagasaki," *JAMA*, 272 (5), 364-70.

- Ozasa, K. et al. (2012), "Studies of the Mortality of Atomic Bomb Survivors, Report 14, 1950– 2003: An Overview of Cancer and Noncancer Diseases," *Radiation Research*, 177 (3), 229-43.
- Preston, DL et al. (2003), "Studies of Mortality of Atomic Bomb Survivors. Report 13. Solid Cancer and Noncancer Disease Mortality: 1950-1997," *Radiation Research*, 160 (4), 381-407.
- Schonfeld, S. J., L. Y. Krestinina, S. Epifanova, M. O. Degteva, A. V. Akleyev, and D. L. Preston (2013), "Solid Cancer Mortality in the Techa River Cohort (1950-2007)," *Radiation Research*, 179 (2), 183-89.
- 放射線影響協会 (2015), 原子力発電施設等放射線業務従事者等に係る疫学的調査(第 V 期調査 平成 22 年度~平成 26 年度): <http://www.rea.or.jp/ire/pdf/report5.pdf>.
- 濱岡豊 (2015a), "広島・長崎被曝者データの再分析," *科学* (9 月号), 875-88.
- (2015b), "長期低線量被曝研究からの知見・課題と再分析," *科学* (10 月号), 985-1006.

注

- ¹日経メディカル 2011/3/29 <http://medical.nikkeibp.co.jp/leaf/all/hotnews/int/201103/519126.html>
- ² (財)放射線影響研究所主席研究員・児玉和紀「低線量被ばくのリスク管理に関するワーキンググループ第 1 回会合 発表概要 2011 年 11 月 9 日」
<https://www.cas.go.jp/jp/genpatsujiko/info/twg/dai1/gaiyou1.pdf>
- ³ 売上が得られそうな店舗に広告費を多く配分したという逆の可能性もあるが、ここでは簡単のため、広告費→売上という因果関係を想定した解釈とする。
- ⁴ 何人亡くなったかというカウントデータなので、ポアソン分布を想定した回帰分析が用いられる。
- ⁵ 寿命調査(Life Span Survey) 14 報([Ozasa et al.\(2012\)](#))で用いられたデータが下記から公開されている。
<http://www.rerf.jp/library/dl/index.html>
- ⁶ 米国エネルギー省 Comprehensive Epidemiologic Data Resource (CEDR)。下記のページから申請すれば利用可能である。 <https://www3.ornl.gov/CEDR/>
- ⁷ ポアソン回帰については被曝者データと同じく Epicure の AMFIT を用いている。
- ⁸ Epicure を開発した Preston 氏の放射線関連学会国際会議 ICRR2015 京都での報告。
- ⁹ 本稿で紹介した研究はコホート研究だが、この研究ではケース・コントロール法である。