

2023 年度 統計データ分析コンペティション

## 優秀賞 [高校生の部]

### 大腸がん罹患要因の探究と罹患しにくい生活の提案

鈴木 実由（慶應義塾湘南藤沢高等部）

#### 論文の概要

男女の総罹患数が最も多い大腸がんに着目し、生活や環境に関するデータを用いて相関分析や重回帰分析を行うことで、どのような要素が大腸がん罹患に起因しているかを明らかにし、その結果を基に罹患しにくい生活を提案した。

#### 論文審査会コメント

要因の選定や各変数の吟味も自然・丁寧であり、仮説の検討も多角的に行われており、新たな知見のヒントも多く得られている。はずれ値にも有用な情報が含まれている可能性もあるので、機械的に除外するよりは何らかの考察に繋げると良い。

# 大腸がん罹患要因の探究と罹患しにくい生活の提案

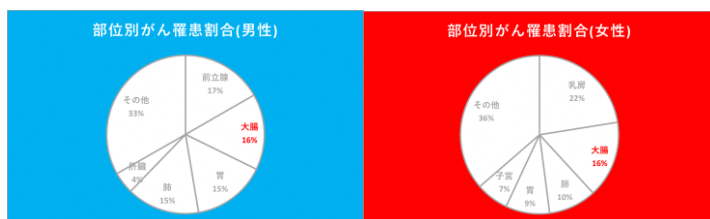
鈴木実由

慶應義塾湘南藤沢高等部

## 1. 研究テーマと目的

科学技術の発達によりがん治療はめまぐるしい発展を遂げているにもかかわらず、未だに、「がん」は日本人の死因の1位となっている。なぜこのような現状となっているかについて、罹患前の対策が十分になされていないからではないかと考えた。対策の不十分さにより、罹患後の治療技術の向上をもってしても、それでも対処しきれないほど、「患者数が多くなっている」ということだ。

部位別のがんの罹患者数を見てみると、男性の1位は「前立腺がん」、女性の1位は「乳がん」であり、その次に、男女ともに「大腸がん」がランクインしている（【図1】）。それぞれ1位は各性別に限られたがんであることもあり、総数では「大腸がん」が1位となっている。大腸がん罹患者数の推移をみると、年々右肩上がりで、増加傾向にあることが分かる（【図2】）。がんの罹患者数を減らすことができれば、必然的に死者数も減る。よって、「がんでも死なない」世界ではなく、「がんにならない」世界を作ること为目标としたいと考えた。本論文では、性別による差がない「大腸がん」に着目し、どのような要素が大腸がんの発症に起因しているのかを調査する。分析結果から「がんにならない」世界を目指して最適な生活を提案する。



【図1】男女別 部位別がん罹患割合(2019年)



【図2】大腸がん罹患者数の推移

【出典】図1:日本対がん協会 「がんの部位別統計」を基に excel で作成

図2: 国立がん研究センター がん統計 「大腸 年次統計」より引用

## 2. 研究の方法と手順

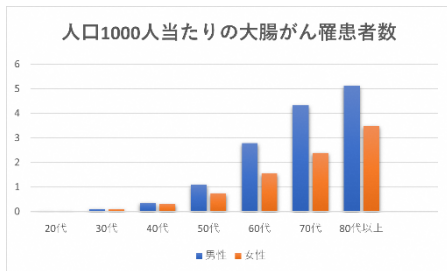
### 2-1 研究の方法

大腸がん罹患の原因となる要素について、都道府県別のデータを用い、「食生活」「生活習慣」「自然環境」「社会・経済環境」の4つのジャンルに分けて考えた。まず、相関を予測したうえで4つをさらに細かく分類し、すべての項目で外れ値の除外を行った。その後、統計ソフトRを用いて分析を行った。考察では、その中から特に関連性の高かったものを取り出し、異なる多様な視点からさらに分析を深め、理由や詳しい因果関係を考察したうえで、大腸がんになりにくい生活の提案をすることとした。

### 2-2 分析する際の患者の区分について

大腸がんは、世代によって罹患者数が大きく異なるうえ、罹患の男女差も確認できる(次ページ【図3】)。これから、大腸がんの罹患要因は、年齢や性別によって異なるのではないかと考えた。よって、「大腸がん患者」として彼らをまとめて分析すると、適切な結果が得られないと考え、次のように分類することにした。

【30代(男/女) 40代(男/女) 50代(男/女) 60代(男/女) 70代(男/女) 80代以上(男/女)】



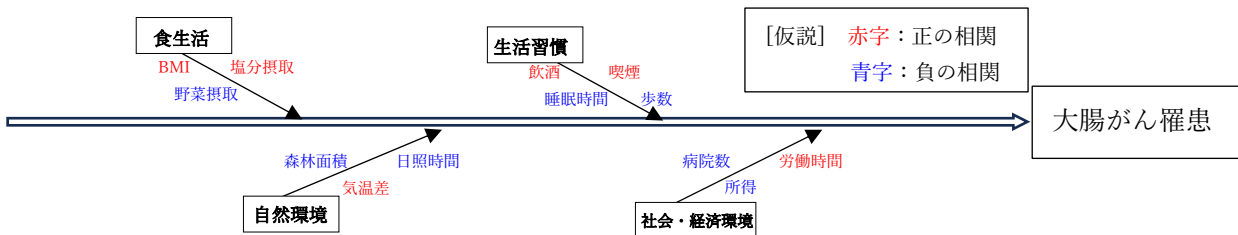
なお、10代以下・20代については患者数が極めて少ないため、分析結果に信頼性がないと考え、分析の対象外とした。

【出典】人口：総務省「人口推計」（2019年）  
大腸がん罹患患者数：国立がん研究センターがん情報サービス「がん統計」（全国がん登録）「都道府県別がん罹患データ」（2016～2019年）を基に excel で作成

【図3】人口 1000 人当たりの大腸がん罹患患者数(2019年)

### 2-3 ジャンルの細分化

前述した4つのジャンルについて、以下の特性要因図のように細分化し、相関があると仮説を立てた。



【図4】特性要因図(仮説)

## 3. データの加工

### 3-1 使用データと作成した指標

使用した変数	出典
最高気温（日最高気温の月平均の最高値）[°C] 最低気温（日最低気温の月平均の最低値）[°C]	SSDSE-B-2023
酒類[円]、食料合計[円]	SSDSE-C-2023
可住地面積[ha](2021年)、1人当たり県民所得.平成23年 基準[千円](2018年)、一般病院数[施設]、一般診療所数 [施設]（2020年）	SSDSE-E-2023
大腸がん罹患患者数[人](2019年)	国立がん研究センターがん情報サービス「がん統計」（全国がん登録）「都道府県別がん罹患データ」（2016～2019年）
BMIの平均値、野菜摂取量の平均値[g/日]、食塩摂取量の平均値[g/日]、歩数の平均値[歩]（2016年）	厚生労働省「国民健康・栄養調査」（平成28年） （熊本県のデータなし）
喫煙率[%]（2019年）	国立がん研究センターがん情報サービス「がん登録・統計」「国民生活基礎調査による都道府県別喫煙率データ」（2001～2019年）
人口 [千人]（2019年）	総務省「人口推計」（2019年）
総数[千人]、7時間以上8時間未満[千人]、8時間以上9時間未満[千人]、9時間以上[千人]（2019年）	厚生労働省「国民生活基礎調査」（2019年）
森林面積割合[%]、日照時間(年間)[時間]（2019年）	総務省「統計でみる都道府県のすがた」（2021年）
1人平均月間実労働時間数 [時間]（事業所規模30人以上）	厚生労働省「毎日勤労統計調査」（平成21年）

指標	計算方法
人口 1000 人当たりの大腸がん罹患患者数※	大腸がん罹患患者数[人] / 人口[千人]
食糧支出のうち酒類の支出が占める割合※	酒類[円] / 食料合計[円]
睡眠時間が 7 時間以上の人の割合※	(7 時間以上 8 時間未満[千人]+8 時間以上 9 時間未満[千人]+9 時間以上[千人]) / 総数[千人]
気温差[°C]	最高気温[°C]-最低気温[°C]
可住地面積に対する病院数[施設 / ha]※	(一般病院数[施設] + 一般診療所数[施設]) / 可住地面積[ha]

※これらについて、以下では 大腸がん罹患患者数、飲酒、睡眠時間、病院数 と表記する。

### 3-2 外れ値の除外

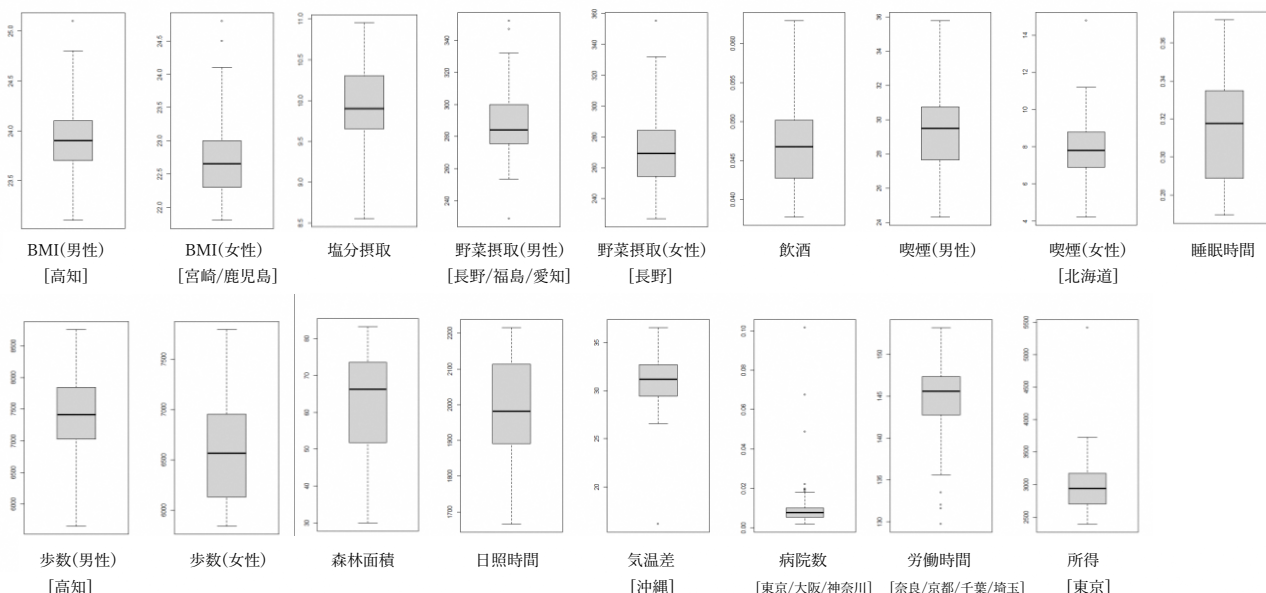
以下のように、それぞれの項目で箱ひげ図を作成し、外れ値は除外した。外れ値の定義は、第一四分位数または第三四分位数から [(第三四分位数-第一四分位数)×1.75] 離れたものとした。【図 5】の○は、この式を満たした外れ値を示している。

なお、食生活・生活習慣を表す項目については、男女の値の相関係数を調べ、**値が 0.80 未満の項目のみ**男女別のデータを用いた。飲酒については、男女別のデータが得られなかったため男女で同じデータを用いた。

以下、算出した男女の相関係数と、それぞれの項目の箱ひげ図を示す。

【表 1】男女のデータの相関係数

	BMI	塩分摂取	野菜摂取	喫煙	睡眠時間	歩数
相関係数	<b>0.467</b>	<b>0.819</b>	<b>0.756</b>	<b>0.437</b>	<b>0.893</b>	<b>0.643</b>



【図 5】それぞれの項目の箱ひげ図

※ [ ]内の都道府県は、外れ値として除外した都道府県である。「病院数」については、外れ値が多数見受けられたが、特に外れた値を示した 3 都府県を除外することとした。

## 4. 分析結果

### 4-1 相関係数の算出

大腸がん罹患患者数と、すべての項目の相関係数を調べ、以下の[凡例 1]に従って色分けをした。(80↑は80代以上を表す。)

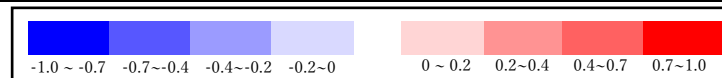
【表 2】大腸がん罹患患者数とすべての項目の相関係数[男性]

	食生活			生活習慣				自然環境			社会・経済環境		
	BMI	塩分摂取	野菜摂取	飲酒	喫煙	睡眠時間	歩数	森林面積	日照時間	気温差	病院数	労働時間	所得
30代	0.095	0.062	0.052	0.034	0.076	0.062	0.032	0.041	-0.105	-0.120	-0.103	0.108	-0.092
40代	0.017	-0.214	0.042	0.325	0.305	0.092	-0.127	-0.224	-0.006	-0.297	-0.192	-0.189	-0.248
50代	0.040	0.024	-0.011	0.229	0.237	0.102	-0.144	-0.025	-0.212	0.032	-0.054	0.186	-0.132
60代	0.162	0.028	-0.063	0.344	0.398	0.281	-0.099	-0.173	-0.394	0.094	-0.092	-0.041	-0.194
70代	0.121	0.096	0.114	0.480	0.309	0.205	-0.040	-0.106	-0.133	0.223	-0.027	-0.000	-0.107
80↑	-0.234	0.099	0.091	0.228	0.234	0.185	-0.050	-0.137	-0.195	0.151	0.061	-0.071	0.170

【表 3】大腸がん罹患患者数とすべての項目の相関係数[女性]

	食生活			生活習慣				自然環境			社会・経済環境		
	BMI	塩分摂取	野菜摂取	飲酒	喫煙	睡眠時間	歩数	森林面積	日照時間	気温差	病院数	労働時間	所得
30代	-0.010	0.272	0.352	0.304	0.147	0.128	0.120	0.126	-0.171	-0.015	-0.056	0.297	-0.264
40代	0.274	-0.078	0.062	0.068	0.042	0.156	-0.131	0.120	-0.282	0.089	-0.211	0.233	-0.311
50代	0.350	0.152	0.340	0.228	0.233	0.106	-0.089	-0.007	-0.142	0.161	0.024	0.149	-0.042
60代	-0.133	0.099	0.191	0.031	0.030	-0.048	0.116	0.162	0.221	0.345	0.039	-0.059	0.108
70代	-0.167	0.307	0.032	0.290	0.209	0.189	-0.060	-0.067	-0.036	0.130	-0.013	0.144	-0.004
80↑	-0.190	0.417	0.279	0.487	0.135	0.263	0.034	0.027	-0.285	0.269	-0.090	0.032	0.031

[凡例 1]相関係数の凡例



その後、すべての項目について無相関の検定を行い、p 値によって、

[凡例 2]に従い\*印をつけた。(色分けは、【表 2】、【表 3】に対応している。)

[凡例 2]p 値の凡例

***: p<0.001	\ : p<0.1
** : p<0.01	× : 0.1 ≤ p
* : p<0.05	

【表 4】p 値の分類

【男性】

【女性】

	【男性】				【女性】			
	食生活	生活習慣	自然環境	社会・労働環境	食生活	生活習慣	自然環境	社会・経済環境
30	×	×	×	×	×	*	*	×
40	×	*	*	*	×	*	*	*
50	×	*	*	*	*	*	*	×
60	×	*	**	**	×	*	*	*
70	×	***	*	*	*	*	*	*
80	×	*	*	*	**	*	***	*

これ以下、相関係数 r の絶対値が  $0 \leq |r| < 0.2$  で「相関がない」、 $0.2 \leq |r| < 0.4$  で「弱い正/負の相関がある」、 $0.4 \leq |r|$  で「正/負の相関がある」と記述することとする。

『食生活』のジャンルは、男性よりも女性に大きな相関を示した。【表 4】で、男性に関しては『食生活』のジャンルすべてで全世代 p 値が 0.1 以上となっており、無相関である可能性を否定できない。一方で女性は、いくつかの項目で弱い正の相関を示しており、【表 4】で p 値を確認しても有意だと言える。特に「塩分摂取」の項目は、高齢者において有意であると判断できる正の相関があり、信頼性が高い。「BMI」「野菜摂

取」に関しても、50代では弱い正の相関がみられた。『生活習慣』のジャンルは、男性・女性共にある程度の相関を示し、特に男性は多くの項目で相関がみられた。男女ともに「飲酒」の項目では、それぞれ70代、80代で有意水準0.1%で有意と判断できるきわめて信頼性の高い正の相関がみられた。女性は「飲酒」以外の項目では相関はみられなかったが、男性は「喫煙」「睡眠時間」でも弱い正の相関がみられた。『自然環境』のジャンルでは、男女ともにあまり相関はみられなかったが、男性は「日照時間」と「気温差」で弱い負の相関、女性は「気温差」で弱い正の相関がみられた。『社会・経済環境』のジャンルでも、男女ともにあまり相関がみられなかった。特に男性はすべての項目でp値が0.1以上で、無相関とする帰無仮説を棄却することができなかった。女性は「所得」の項目で弱い負の相関がみられた。また、世代別に考えると、『食生活』・『生活習慣』のジャンルでは、若い世代より高齢世代の方が大きな相関がみられるうえ、p値が小さく、有意であると判断される傾向にあることが読み取れた。一方で『社会・経済環境』のジャンルでは、若い世代のほうが大きな相関がみられ、有意性も大きい傾向にあった。

#### 4-2 相関係数の捉え方と重回帰分析

相関分析は、各項目と大腸がん罹患数との直線的関係を測ることができる。しかし、相関係数では「因果関係」を表せない。例えば、飲酒者が多いから大腸がん罹患数が多いのか、大腸がん罹患数が多いから飲酒者が多いのか、相関係数だけでは判断できないということである。一方、今回の分析の場合、すべての項目で後者、つまり「大腸がん罹患数が多いから●●が多い(少ない)」とは考えにくい。そのため相関係数だけでその項目が大腸がん罹患数の増加原因となっていると判断してもよいと考える。しかしながら、大腸がん罹患の「原因」を探究することを目的とした本研究の信頼性をより高めるためには、因果関係を調べる「重回帰分析」の必要性も感じる。そこで、「大腸がん罹患数」を目的変数、相関分析より相関が認められた項目を説明変数として、追加で年代ごとに重回帰分析を行った。以下、各項目の回帰係数の推定値を示す。重回帰分析において有意水準5%で有意と判断できたものには下線を引いた。

【表5】重回帰分析における各項目の回帰係数[男性]

	食生活			生活習慣				自然環境			社会・経済環境		
	BMI	塩分摂取	野菜摂取	飲酒	喫煙	睡眠時間	歩数	森林面積	日照時間	気温差	病院数	労働時間	所得
30代													
40代		-4.08e-2		5.18e-1	8.024e-4			-7.35e-5		-7.70e-4			2.53e-6
50代				2.90e-1	1.47e-3				-2.35e-5				
60代				2.30e-1	3.97e-3	5.66e-4			-7.63e-5				
70代				3.68	3.29e-3	-1.72e-3				6.12e-4			
80↑	-4.16e-2			2.03	5.05e-3								

【表6】重回帰分析における各項目の回帰係数[女性]

	食生活			生活習慣				自然環境			社会・経済環境		
	BMI	塩分摂取	野菜摂取	飲酒	喫煙	睡眠時間	歩数	森林面積	日照時間	気温差	病院数	労働時間	所得
30代		9.32e-4	3.28e-5	9.32e-2								2.43e-5	-8.96e-7
40代	-4.87e-4								-9.08e-6		7.20e-1	3.79e-4	-4.17e-6
50代	1.59e-3		1.03e-4	5.53e-2	9.40e-4								
60代									7.72e-5	4.93e-3			
70代		1.47e-2		7.69e-1	2.71e-3								
80↑		2.37e-2	-4.27e-4	3.08		-4.44e-4			-4.35e-5	3.20e-3			

※e-n(nは自然数)は、×10の-n乗を表す。例えば6.25e-2であれば、0.0625となる。

予想通り多くの項目で、相関係数が正/負の場合は、回帰係数も正/負で、「●●が多いから、大腸がん罹患  
者数が多い(少ない)」という因果関係を確認できた。一方、「40代男性所得」「70代男性睡眠時間」「40代女  
性BMI」「40代女性病院数」「80代以上女性野菜摂取」「80代女性睡眠時間」の項目については、相関係数と  
回帰係数の符号が異なるという、係数矛盾現象が発生してしまった。

## 5. 結果の解釈

### 5-1 性別の大腸がん罹患に関する考察

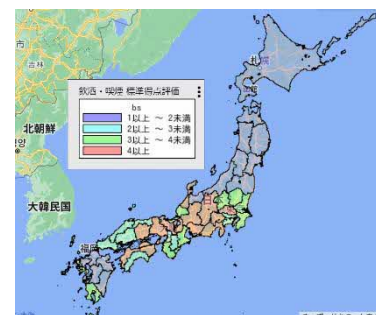
#### 5-1-1 「男性/女性の大腸がん罹患の原因」と判断する項目について

世代を問わず男性/女性全体としてみたとき、2つ以上の世代で同じ符号の回帰係数が算出できた項目を、  
「男性/女性の大腸がん罹患の原因」と考えられる項目として判断することとする(これらの項目は【表4】  
よりp値も小さく結果の信頼性も高い)。つまり、男性は「飲酒」「喫煙」「日照時間」、女性は「塩分摂取」  
「野菜摂取」「飲酒」「喫煙」「日照時間」「気温差」「労働時間」「所得」である。ここで、同じジャンル(食生  
活/生活習慣 など)の中から2つ、原因として判断された場合は、同ジャンル内の項目は項目同士で相関を  
持っている可能性が考えられるため、大腸がん罹患患者数の項目に対する擬似相関の検出を行うこととする。

#### 5-1-2 [男性]「飲酒」「喫煙」について

男性における「飲酒」と「喫煙」の項目について、擬似相関の検出を行ったが、お互いの項目の影響を除  
外してもなおある程度相関がみられたので、擬似相関と判断せず、どちらも直接的な原因と判断することと  
した。そして、両者ともに年代が上がるごとに数値が大きくなる傾向にあることから、「長年の飲酒・喫煙の  
習慣」が大腸がん罹患の原因となっていると考えた。

ここで、各項目で標準得点を求め、それを基に都道府県を1-4で評価、  
日本地図に投影してみた(図6)。飲酒も喫煙も全国平均より多いとき1  
(図中紫色)、飲酒は多いが喫煙は少ないとき2(図中水色)、飲酒は少ないが  
喫煙は多いとき3(図中黄緑色)、飲酒も喫煙も少ないとき4(図中赤色)と  
して色付けした。また、飲酒・喫煙は大腸がん罹患と正の相関があることが  
分かっているので、一般的に1の地域では大腸がん罹患者が多く、4の地  
域では少なくなる傾向にあると捉えてよい。

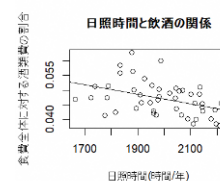


【図6】 飲酒・喫煙 標準得点評価

【図6】より、北海道・東北地方・九州地方に1が偏っており、中部地方から近畿地方にかけて4が偏っ  
ていることが読み取れる。ここから、飲酒・喫煙の文化には「地域性」があると考えられ、その傾向は一時的  
で偶然(2019年のみ)のものとは考えにくく、地域に根付いている文化だと捉えるのが自然である。例え  
ば、1の地域では、多くの人々が長年にわたって飲酒や喫煙を行っていることが予想でき、その行為が「習  
慣化」しているといえる。このように、年代による相関・回帰係数の変化や、地域性という観点から、男性  
は飲酒や喫煙の「習慣化」が大腸がん罹患を招いていると考える。

#### 5-1-3 [男性]「日照時間」について

50代60代において、日照時間と大腸がん罹患には、負の相関がみられた。特に60  
代はp値も小さく、信頼性のある結果となっている。日光が健康に良い影響を与える、  
と考えることも可能だが、「大腸がん」という特定の病気との関連は明言できない。ま  
た、この場合、「日照時間」の指標は何かほかの指標の影響を受けていると考える方  
が自然である。調べてみると、【図7】のように「日照時間」と「飲酒」との間には負



【図7】 日照時間と飲酒の関係

の相関がみられた(相関係数-0.414)。そして「飲酒」の影響を取り除いて偏相関分析を行うと、「日照時間」  
と大腸がん罹患患者数の間の偏相関係数は50代で-0.132、60代で-0.294となった。擬似相関の存在が認められ

たが、60代についてはそれでもなお弱い相関がみられたことから、直接的な相関が全く存在しないわけではないと分かった。

#### 5-1-4 [女性] 「塩分摂取」「野菜摂取」について

「塩分摂取」と「野菜摂取」の項目でも【表3】で相関がみられた30,50,70,80代で擬似相関の検出を行った。お互いの影響を除いた上での大腸がん罹患率とこの間の偏相関係数を表す【表7】より、若い世代での「塩分摂取」、高齢世代での「野菜摂取」の項目では相関がみられなくなり、擬似相関が見破れた。その結果、若い世代では「野菜摂取」、高齢世代では「塩分摂取」が、大腸がん罹患率と正の相関を示すことが分かった。若い世代に関しては、野菜のみではなく、野菜を含む、「食べすぎ」が、高齢世代は、男性の「飲酒・喫煙」同様、塩分過剰摂取の「習慣化」が罹患に影響を及ぼしていると考えられる。

【表7】他方の影響を除いたときの偏相関係数

[女性]	塩分摂取	野菜摂取
30代	0.090	0.268
50代	-0.025	0.271
70代	0.338	-0.158
80↑	0.369	-0.048

#### 5-1-5 [女性] 「飲酒」「喫煙」について

女性の「飲酒」「喫煙」の項目についても、擬似相関の検出を行った。お互いの影響を除いたうえでの偏相関係数が0.2を超えたのは、30,70,80代の「飲酒」のみだった(それぞれ0.330, 0.200, 0.431)。よって、「喫煙」と大腸がん罹患は、擬似相関だとわかった。女性においても、70代80代で相関がみられている、特に80代ではより大きい相関がみられていることから、「習慣化」に目を付けた。そこで今回は「飲酒」のみについて標準得点を求め、男性と同様、地図に投影した(【図8】)。飲酒量が全国平均を大幅に上回っているとき赤色、少し上回っているとき黄色に色づけされている。男性の「飲酒・喫煙」ほどの地域性はみられないが、それでも北海道、東北地方、九州地方は多い傾向にある。男性ほど強い傾向はないものの、女性に関しても「飲酒」の「習慣化」が大腸がん罹患を招く一因となっているといえるだろう。



【図8】飲酒 標準得点

#### 5-1-6 [女性] 「日照時間」「気温差」について

「日照時間」と「気温差」の項目についても、【表3】で相関がみられた60,80代で擬似相関の検出を行った。しかし【表8】のようにお互いの影響を除いても相関がみられたので、擬似相関とは判断しなかった。日光が健康に良い影響を与える可能性が考えられること、気温差が大きいと体調を崩しやすいことから、このような正負の相関がみられたと考えたが、あくまで一般的な「健康」との関連の話であり、「大腸がん」という特定の病気との関連は不明であり明言できない。

【表8】他方の影響を除いたときの偏相関係数(2)

[女性]	日照時間	気温差
60代	0.193	0.282
80↑	-0.366	0.447

#### 5-1-7 [女性] 「労働時間」「所得」について

「日照時間」と「気温差」の項目についても、【表3】で相関がみられた30,40代で擬似相関の検出を行った。しかし【表9】のようにお互いの影響を除いても相関がみられたので、擬似相関とは判断しなかった。それぞれの正負の相関について、労働時間が多いとストレスや生活の乱れから体調を崩しやすい、所得が少ないと食生活が乱れたり病院へ行きにくく病状が悪化したりする、などの理由が考えられるが、特に「労働時間」は一般的な「健康」との関連の話で、「大腸がん」という特定の病気との関連は明言できない。

【表9】他方の影響を除いたときの偏相関係数(3)

[女性]	労働時間	所得
30代	0.196	-0.287
40代	0.265	-0.232

### 5-2 年代別の大腸がん罹患に関する考察

分析結果より、『食生活』『生活習慣』のジャンルでは高齢世代に、『社会・経済環境』のジャンルでは若い世代に、より大きな相関が現れる傾向にあり、『自然環境』のジャンルでは年代による相関の違いが少ないことが読み取れる。『食生活』『生活習慣』では長年の習慣による原因因子の蓄積が罹患を招くこと、『社会・経済環境』では若い世代が労働者世代であり、高齢世代に比べ影響を受けやすいことがこのような傾向を生ん



でいると考える。一方で『自然環境』については年齢に関わらず共通であるため、違いが生まれにくいのだと推測できる。また、多くの項目で30代は相関がみられていないが、30代は他年代に比べ罹患数数がかなり少ないため、適切な結果が得られていない可能性が考えられる。年齢による違いが最も顕著な項目は「BMI」で、若い世代では数値が大きい方が、高齢世代では数値が小さい方が罹患しやすいという興味深い結果となった。年齢に合った体型維持の重要性がうかがえる。

### 5-3 係数矛盾現象についての考察

相関係数と回帰係数の符号が逆転する状態が発生してしまったと前述した。これが起きた項目はすべて相関係数が0.2をほんのわずかに超えるもので、原因は「相関がほとんどなかったため」で、ほとんど意味のない値を算出してしまったと推測する。これらの項目は相関係数を求めた時点で除外するべきであった、つまり、相関を認める基準が低すぎたといえる。さらに、別の原因として「多重共線性」も考えられる。説明変数の中に極めて相関が大きい変数が含まれていたため、算出した回帰式が不安定となった可能性がある。

### 5-4 「遺伝」という原因についての考察

私の母が一度、がん検査で異常が認められた時、医師は「家族や親戚に、がんの人はいますか」と尋ねたそうだ。この医師の言葉から、少なからず「遺伝」が原因となってがん罹患する人はいるのだろうと推測できる。しかし、「遺伝」は数値化することが困難であり、どれくらい罹患に寄与しているのかを考えるのは難しい。そこで、今回分析した項目、つまり生活スタイルに関わる要素だけでどれだけ大腸がん罹患を説明できるかを調べ、逆算してみようと考えた。【表10】に、各年代、すべての項目を説明変数としたときの重回帰分析の決定係数を示す。かなり多くの世代で、今回の項目のみで3割以上説明できた。特に60代男性、80代以上女性では5割を超えている。全体的にみると、高齢世代の方がやや数字が大きい傾向にあり、遺伝だけでなく、生活習慣・生活環境が罹患を招きやすいことがわかる。**たかが3割、されど3割**。生活の工夫によって罹患の可能性を下げられるのなら、大いに意味のある結果だろう。

【表10】重回帰分析の決定係数

	男性	女性
30代	0.061	0.354
40代	0.449	0.268
50代	0.151	0.228
60代	<b>0.533</b>	0.224
70代	0.365	0.460
80代	0.344	<b>0.555</b>

### 5-5 結論

男性は、飲酒や喫煙の習慣化が大腸がん罹患の大きな原因となる可能性があることが分かった。女性は、飲酒・塩分摂取の習慣化や食べすぎ、所得の少なさが原因として考えられる。がんには「遺伝」という自分ではどうすることもできない原因因子があるが、高齢世代は特に、日々の生活の工夫によって罹患を避けられる可能性があると考え。禁酒や禁煙、減塩・適量な食事こそが、罹患しにくい生活といえるだろう。

## 6. 参考文献

- ・藤田紘一郎『病気の9割は免疫力で防げるー大腸がん・乳がん・子宮頸がんを防ぐ！ー』 p8~12,21~23,84~87 (枳出版社 2019年)
- ・国立がん研究センターがん情報サービス「がん統計」(全国がん登録)  
[https://ganjoho.jp/reg\\_stat/statistics/stat/cancer/67\\_colorectal.html](https://ganjoho.jp/reg_stat/statistics/stat/cancer/67_colorectal.html)
- ・厚生労働省「国民健康・栄養調査」(平成28年) <https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00450171&tstat=000001041744&cycle=7&tclass1=000001111535&tclass2val=0>
- ・総務省「人口推計」(2019年)<https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200524&tstat=00000090001&cycle=7&year=20190&month=0&tclass1=000001011679>
- ・厚生労働省「国民生活基礎調査」(2019年)[https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00450061&tstat=000001141126&cycle=7&tclass1=000001141142&tclass2=000001142126&tstat\\_infid=000031964353&tclass3val=0](https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00450061&tstat=000001141126&cycle=7&tclass1=000001141142&tclass2=000001142126&tstat_infid=000031964353&tclass3val=0)
- ・厚生労働省「毎日勤労統計調査」(平成21年) <https://www.mhlw.go.jp/toukei/itiran/roudou/monthly/21/21r/mk21r.html>
- ・総務省「統計でみる都道府県のすがた」(2021年)<https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200502&tstat=00000114949&cycle=0&year=20210&month=0&tclass1=000001149950>