

# 肺炎死亡率の社会経済的要因の分析 — 都道府県別の肺炎対策の視点から —

鈴木 孝 弘

田 辺 和 俊<sup>1)</sup>

## 目次

- 1 はじめに
- 2 データと方法
- 3 結果と考察
- 4 地域の肺炎対策への試論
- 5 おわりに

## 参考文献

## 1 はじめに

2019年12月、中国湖北省武漢市で発生した新型コロナウイルス肺炎は全世界に蔓延し、収束の気配が見えない。このウイルスに感染すると、風邪のような症状だけで済むこともあれば、肺炎を発症して命を落とすこともある。中国のデータに基づく分析によれば、新型コロナウイルス感染患者の81%は軽症で、重症は14%、重篤は5%、死亡は2.3%となっている（Wu and McGoogan [2020]）。死亡率についてWHOは2%程度と発表しているが、現時点において日本の致死率は0.1%（高齢者0.3%、若年者0.01%）と諸外国より低くなっている。

一般に肺炎の多くは感染症で、気道から侵入した細菌やウイルスなどの病原体が肺の中で炎症を引き起こす。軽症で済むかどうかの分かれ道は、体の防御システムが機能するかどうかにかかっている。高齢者や基礎疾患のある人は免疫力が低下しているため、防御システムがうまく機能せず、重症化しやすいとされる。わが国の主な死因別の死亡率の推移（図1）をみると、1980年頃から肺炎による死亡者数が急増し、2011年から2016年まで死因の第3位であった。1980年代以降の急増の原因は、急速な高齢化の進行により肺炎に罹患して死亡する高齢者が増加したためとされている。

---

1) 現代社会総合研究所

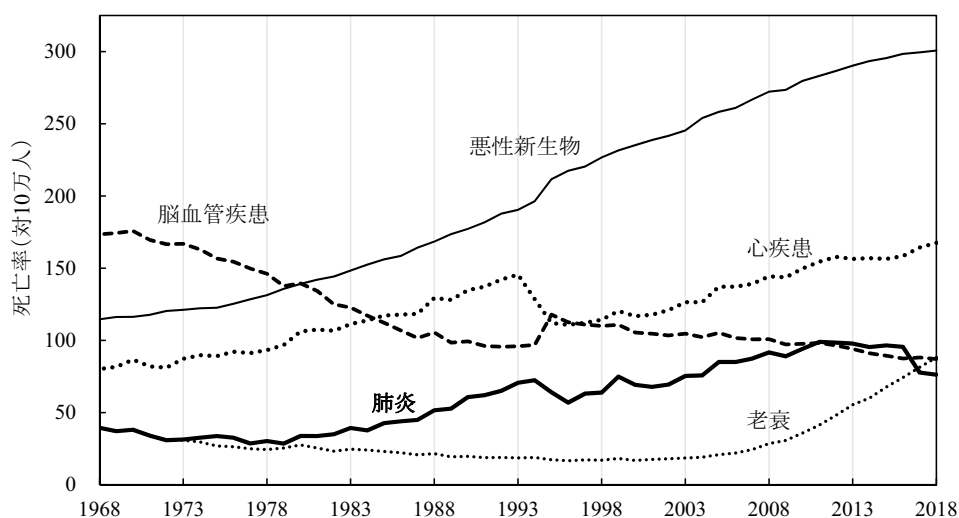


図1 日本人の主要死因別死亡率の推移（「平成30年人口動態調査」による）

ところで、20世紀後半以降、個人や集団の健康や病気が遺伝等の先天的要因だけでなく、経済、社会、文化、環境等の後天的要因によっても影響されることが広く認められるようになった。WHOの報告書“The Determinants of Health”は健康格差の社会的影響要因として、社会格差、ストレス、社会的排除、労働、失業、社会的支援等の10項目を挙げている。わが国の健康格差については、所得や失業等の経済・労働要因、学歴等の教育要因、病院や医師等の医療資源要因の影響が大きいとされている。このように人の死には非常に多くの要因が影響すると考えられることから、国や地域自治体が対策を行うためには、多数の要因の中から死亡に重大な影響を与える要因の解明とその相対的影響度の推定が重要になる。

このように複雑な疾患原因を解明するために、罹患率や死亡率のデータを解析して要因を探索する疫学研究が行われてきている。その疫学的手法を大別すると、第一に、検証対象の要因を持つ集団と持たない集団について長期間追跡し、両群間での死亡率等を統計的に比較することで要因を検索するコホート研究がある。第二は、着目する疾患の症例群と対照群について、過去に遡って調査した要因の影響の違いを分析する症例対照研究（ケースコントロール研究）である。第三は、過去に発表された関連論文を総合評価し、症例に対する要因の影響を解析するメタアナリシス研究である。しかし、これら3種の方法は、調査や解析が大規模になり、膨大な時間や経費がかかることや、多数の要因について解析し、それらの相対的影響度を評価することは容易でない、といった難点がある。

疾患に対する各種要因の影響度の推定法としては、観察集団内の個人ごとの疾病発生の有無と要

因との相関を解析する横断的研究（クロスセクショナル研究）と、国や地域などの異なる集団間の疾病の頻度（罹患率、死亡率等）と要因の相関を分析する生態学的研究（地域相関研究）がある。しかし、個々の要因と死亡率との相関係数には他の要因の影響が含まれるため、これらの手法で各要因の相対的影響度を推定することは難しい。

多数の要因について影響度をより正確に推定する方法として、死亡率を目的変数、複数の要因を説明変数として重回帰分析を行う方法がある。この方法で肺炎の死亡率について解析を行った先行研究はあるが、限定的範囲の少数の説明変数を用いて解析した研究のみである。肺炎の死亡には多数の要因が関与し、特に喫煙や受動喫煙は大きな影響を与えるとされている。そのため、それら重要な説明変数を用いずに重回帰分析を行った解析の結果には疑問がある。

以上の背景から本研究では肺炎の死亡率に焦点を当て、その地域格差の要因分析を目的とする実証研究を試みた。2で政府統計の集計データを用いて都道府県別の肺炎死亡率を目的変数とし、多数の社会経済的指標を説明変数として非線形性を考慮した重回帰分析を行い、死亡要因を探索した。3で本研究の結果を先行研究と比較して考察し、4で肺炎死亡率が国内最高の青森県を対象として、地域の肺炎死亡率低減策を試論した。

## 2 データと方法

### 2-1 目的変数

目的変数には、「平成29年人口動態特殊報告」で公表されている肺炎の都道府県別年齢調整死亡率を用いた。男性、女性および男女平均の死亡率（図2）を見ると、男女間および都道府県間の格差がかなり大きいことが明らかである。男性1位の青森県の死亡率は最下位の長野県の2倍近い。男性での順位は1位青森、2位山口、3位徳島、45位宮城、46位香川、47位長野であるが、女性では1位鹿児島、2位山口、3位青森、45位島根、46位鳥取、47位長野であり、男女とも青森県の高さと長野県の低さが際立っている。男女の死亡率の都道府県傾向はよく類似し、男女間の相関係数は0.850と高い。そこで、地域の肺炎対策を検討する観点からは、男女別に解析するよりは一括して取り扱った方が死亡要因を明らかにする上で好ましいと考えられることから、重回帰分析の目的変数には男女の平均を用いた。

### 2-2 説明変数

説明変数については、肺炎死亡に関する多変量解析を行っている先行研究で検証された指標をすべて採用し、さらに、肺炎への影響が議論されているが、これまで統計的に検証されていない指標を追加し、総計57種の説明変数（表1）を採用した。分野別の変数の数を見ると、生活習慣分野が9種、医療・福祉分野が11種、経済分野が11種、労働分野が7種、人口・世帯分野が7種、教育分

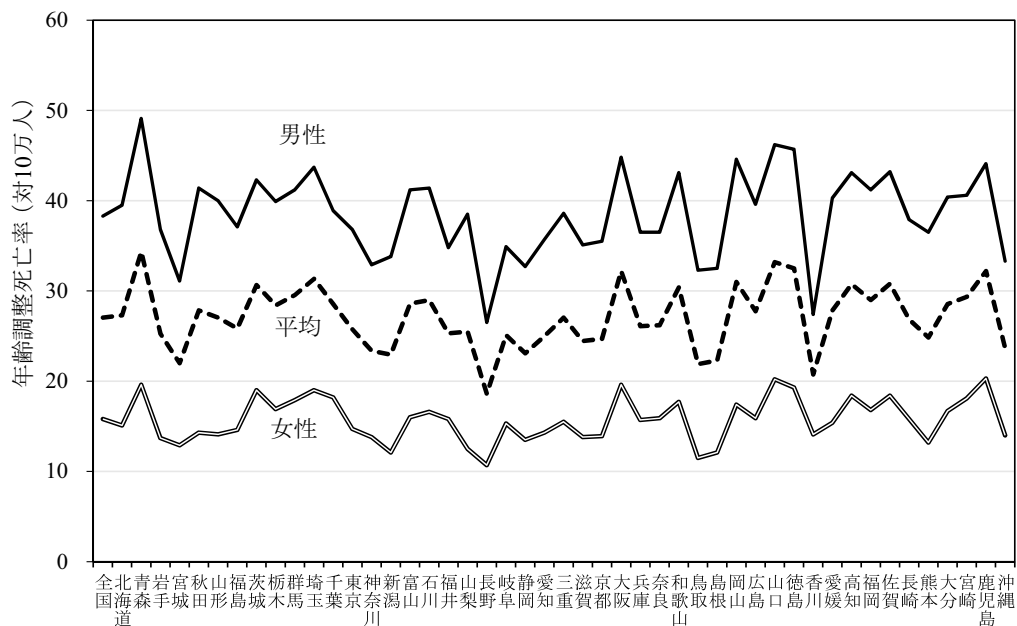


図2 都道府県別の肺炎の年齢調整死亡率（「平成29年人口動態特殊報告」による）

野が4種、地理・環境分野が8種である。その内、先行研究で検証された変数は22種であり、これまで未検証の変数の方が多い。各都道府県のデータは各種の政府統計調査から入手し、最小0と最大1となるよう規格化して解析に用いた。

### 2-3 分析方法

死亡率要因分析の先行研究では、最も一般的な線形重回帰分析（OLS）が多用されてきた。しかし、本研究の説明変数の中には、死亡率に対して非線形関係にあるものが多いため、OLSでは統計的に有意な結果を得ることは難しい。先行研究では、このような非線形関係に対処するために、一部の説明変数の2乗の項の追加や、対数変換を行った論文がある。また、説明変数間の交絡効果に対処するために2変数の積の項を追加して解析した論文がある。さらに、説明変数間に相関の高い組がある場合、OLSでは多重共線性効果により回帰分析が不安定になるため、高相関の組の一方を削除して解析した論文がある。さらに、本研究のように説明変数が目的変数のデータ数より多い場合、OLSではそのままの解析は不可能なため、主成分分析等の手法を用いて説明変数を減らして解析した論文がある。しかし、これらの対処はad hoc的なものであり、完全な解決策とはいえない。

本研究では、これらの問題を解決するために非線形重回帰分析の手法としてサポートベクターマシン（SVM）を適用し、ソフトウェアはLIBSVM（Chang and Lin [2016]）の回帰機能を用いた（SVM

の原理や特徴、用語についてはCristianini and Shaw-Taylor Jr. [2000]、小野田 [2007]、阿部 [2011] 参照)。SVMは、説明変数の数値に対してカーネルと呼ぶ非線形関数を用いて学習パターンを別の空間（超平面）に写像し、その空間で線形回帰を行う。この操作により、説明変数の元の数値での非線形回帰が可能になり、目的変数と説明変数の間の任意の関係に対して高精度の回帰結果が得られる。さらに、SVMでは、①説明変数間の交絡効果が予想される場合でもこの効果が自動的に考慮されるため、変数の積の項の追加は不要、②変数間に強い相関がある場合でも解析可能であり、多重共線性問題は生じない、③目的変数のデータ数以上の説明変数を用いた解析が可能、などの利点がある。以上のSVMの多くの利点はカーネル回帰という手法の採用によるものである。

多数の説明変数の中から要因を探索するためには、有効な変数を抽出する変数選択が必要である。一般に重回帰分析では、説明変数の中に有効でないものがあると過学習状態に陥り、既存データに対する学習誤差は減少するが、未知データに対する予測誤差は増大する。そのため、必要最小限の説明変数を抽出する必要がある。本研究では、迅速な変数選択法として感度分析法を採用した。これは、目的変数に対する各説明変数の感度を計算し、感度の低い変数を順次削除しながらSVMモデルを学習最適化し、死亡率の予測誤差が最小となる説明変数の組み合わせを探索する方法である。筆者らはこの感度分析法による変数選択の有効性を様々な問題で実証している (Tanabe, Kurita, Suzuki, et al. [2013]、田辺・鈴木 他 [2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020])。

また、先行研究では、全データでOLSモデルを学習した際の結果に対して、平均二乗誤差 (RMSE) や回帰決定係数 ( $R^2$ ) 等の指標を計算して要因を探索していることが多いが、この方法では回帰モデルの性能を厳密に評価していない。本研究では、回帰性能をより厳密に評価する方法として、1個抜き交差検証法 (LOOCVT) を採用した。そこで、LOOCVTと感度分析を組み合わせた以下の手順により要因の探索を行った。

- ①LOOCVTによりSVMモデルを最適化するために、学習データを用いてSVMのモデルパラメータ  $g$  と  $c$  をそれぞれグリッドサーチし、全都道府県について死亡率の予測値と実測値との平均誤差 (RMSE) が最小になる最適条件を探索する。
- ②当該変数は実際の数値のまま、その他の変数は全データの平均値に設定したデータセットを最適モデルに入力し、出力値を求め、当該変数の実測値を説明変数、出力値を目的変数とする単回帰分析を行い、回帰直線の傾きをその変数の感度とする。
- ③全説明変数の中で感度の絶対値の最も小さい変数を取り除き、①～③の操作を繰り返し、RMSE が最小になる説明変数の組み合わせを死亡率の要因とする。

表1 解析に用いた説明変数の定義とデータ出典

分野	説明変数	定義 (単位)	出典	
生活習慣	喫煙*	毎日または時々喫煙している人の割合 (%)	国民生活基礎調査	
	飲酒	1日の平均飲酒量 (合)	国民生活基礎調査	
	高齢者学習	人口100万人当たりの高齢者学級・講座数 (箇所)	社会教育調査	
	老人クラブ	高齢者人口10万人当たりの老人クラブ数 (箇所)	厚生省報告例 (社会福祉関係)	
	スポーツ	スポーツの年間行動者率 (%)	社会生活基本調査	
	高齢者スポーツ	高齢者のスポーツの年間行動者率 (%)	社会生活基本調査	
	社会体育施設	人口100万人当たりの社会体育施設数 (箇所)	社会教育調査	
	ボランティア活動	ボランティア活動の年間行動者率 (%)	社会生活基本調査	
	高齢者ボランティア活動	高齢者のボランティア活動の年間行動者率 (%)	社会生活基本調査	
	医療・福祉	病院数*	人口10万人当たりの一般病院数 (箇所)	医療施設調査
		病床数*	人口10万人当たりの一般病院病床数 (床)	医療施設調査
医師*		人口10万人当たりの医療施設に従事する医師数 (人)	医師・歯科医師・薬剤師調査	
看護師*		人口10万人当たりの医療施設に従事する看護師数 (人)	衛生行政報告例	
保健師*		人口10万人当たりの保健師数 (人)	衛生行政報告例	
老人ホーム*		高齢者人口10万人当たりの老人ホーム数 (箇所)	社会福祉施設等調査	
介護療養型医療施設		高齢者人口10万人当たりの介護療養型医療施設数 (箇所)	介護サービス施設・事業所調査	
訪問介護員		人口10万人当たりの訪問介護員 (ホームヘルパー) 数	介護サービス施設・事業所調査	
介護老人福祉施設		高齢者人口10万人当たりの介護老人福祉施設数 (箇所)	介護サービス施設・事業所調査	
社会福祉士		人口10万人当たりの社会福祉士数 (人)	社会福祉振興・試験センター	
高齢者受療		高齢者の人口10万人当たりの受療率 (%)	患者調査	
経済	県民所得*	人口1人当たりの県民所得 (千円)	県民経済計算	
	個人所得*	雇用人1人当たりの所得 (千円)	県民経済計算	
	賃金*	一般労働者の基本賃金 (千円)	賃金構造基本統計調査	
	ジニ係数*	所得格差を表す指数	全国消費実態調査	
	衛生費	人口1人当たりの衛生費 (千円)	地方財政統計年報	
	医療費	人口1人当たりの国民医療費 (千円)	国民医療費調査	
	保健医療費	二人以上の世帯での保健医療費 (千円)	家計調査	
	高齢者医療費	被保険者1人当たりの後期高齢者医療費 (千円)	後期高齢者医療事業年報	
	社会福祉費	人口1人当たりの社会福祉費 (千円)	地方財政統計年報	
	老人福祉費	人口1人当たりの老人福祉費 (千円)	地方財政統計年報	

	生活保護費	被保護実人員1人当たりの生活保護費 (千円)	地方財政統計年報	
労働	労働力	労働力人口比率 (%)	国勢調査	
	一次産業*	第一次産業就業者比率 (%)	国勢調査	
	二次産業*	第二次産業就業者比率 (%)	国勢調査	
	三次産業*	第三次産業就業者比率 (%)	国勢調査	
	失業率	完全失業率 (%)	国勢調査	
	共働き	共働き世帯の割合 (%)	国勢調査	
	高齢就業者	高齢就業者の割合 (%)	国勢調査	
	人口・世帯	高齢化率*	国勢調査	
		核家族世帯	核家族世帯の割合 (%)	国勢調査
		単身世帯	単身世帯の割合 (%)	国勢調査
	高齢世帯	高齢者のいる世帯の割合 (%)	国勢調査	
	高齢夫婦世帯	高齢夫婦のみの世帯の割合 (%)	国勢調査	
	高齢单身世帯	高齢单身世帯の割合 (%)	国勢調査	
	持家世帯*	持家世帯の割合 (%)	国勢調査	
教育	学歴	最終学歴までの平均年数 (年)	建築動態統計調査	
	中卒	最終学歴が中学校卒業者の割合 (%)	国勢調査	
	高卒	最終学歴が高校卒業者の割合 (%)	国勢調査	
	短大・大学卒	最終学歴が短大・高専・大学・大学院卒業者の割合 (%)	国勢調査	
		都道府県庁所在地の標高 (m)	気象庁	
地理・環境	標高*	都道府県庁所在地の年平均気温 (°C)	気象庁	
	気温*	都道府県庁所在地の年平均相対湿度 (%)	気象庁	
	湿度*	都道府県庁所在地の年間日照時間 (時間)	気象庁	
	日照時間	都道府県庁所在地の年間降水量 (mm)	気象庁	
	降水量*	都道府県庁所在地の年平均大気汚染物質量濃度 (ppm)	気象庁	
	大気汚染	下水道普及率 (%)	統計年鑑	
	下水道*	下水道普及率 (%)	水道統計	
	下水道	下水道施設等実態調査	下水道施設等実態調査	

\* 先行研究で検証された変数。

### 3 結果と考察

以上の方法により、57種の説明変数の中から低感度の変数を順次削除し、有意な要因を探索した結果、12種の変数を用いた場合に死亡率の予測値と実測値のRMSEが最小となった。このときの回帰決定係数（ $AR^2=0.635$ ）から、この12種の説明変数は47都道府県の死亡率格差を危険率1%で有意に説明する要因であると判定される。また、得られた各要因の死亡率への相対的影響度について考察するために、要因*i*の感度 $S_i$ から式

$$C_i(\%) = \frac{S_i^2}{\sum_{i=1}^{12} S_i^2} \times 100 \quad (1)$$

により死亡率に対する寄与率 $C_i$ を推定した。要因12種の内訳、死亡率に対する感度、および寄与率を表2に示す。

表2に示した要因の感度は、感度分析において、他の説明変数は固定し、当該変数のみ変化させた場合の死亡率の変化から求めたことから、死亡率に対する当該要因の正味の影響度を表わしている。したがって、感度が正の要因は死亡率の上昇要因（すなわち、危険要因）であり、負の要因は下降要因（すなわち、防御要因）であると解釈できる。また、喫煙（%）と社会福祉士数（人）と衛生費（円）のように単位の異なる要因についても、それらの感度の大きさにより、死亡率への影響度についての考察が可能になる。さらに、上式で求めた寄与率により、各要因の死亡率への相対的影響度についての議論が可能になる。

要因12種を分野別に集計すると、1位は医療・福祉分野が4要因（社会福祉士、保健師、訪問介

表2 要因の内訳、死亡率に対する感度、および寄与率

	要因	分野	感度		寄与率 (%)
			上昇要因	下降要因	
1	喫煙*	生活習慣	0.194		10.2
2	短大・大学卒	教育		-0.183	9.1
3	社会福祉士	医療・福祉		-0.183	9.1
4	ボランティア活動	生活習慣		-0.182	9.1
5	保健師*	医療・福祉		-0.178	8.6
6	訪問介護員	医療・福祉		-0.177	8.5
7	社会体育施設	生活習慣		-0.176	8.5
8	衛生費	経済		-0.172	8.1
9	高齢夫婦世帯	人口・世帯	0.170		7.9
10	生活保護費	経済	0.166		7.5
11	高齢者受療	医療・福祉	0.166		7.5
12	高齢単身世帯	人口・世帯	0.148		6.0

\*先行研究で検証された要因。



護員、高齢者受療)で寄与率33.7%、2位は生活習慣分野が3要因(喫煙、ボランティア活動、社会体育施設)で27.7%、3位は経済分野が2要因(衛生費、生活保護費)で15.6%、4位は人口・世帯分野が2要因(高齢夫婦世帯、高齢単身世帯)で13.9%、5位は教育分野が1要因(短大・大学卒)で9.1%となる。また、要因12種の内、先行研究で検証された要因は保健師と喫煙の2種のみであり、未検証の要因の数と寄与がはるかに多い。これらの結果は肺炎の死亡要因が広汎な分野にまたがっており、本研究のように多分野の多種多様な説明変数の中から要因を探索する解析手法が不可欠であると結論できる。

これに対し、肺炎死亡率について多変量解析を行った先行研究は狭い分野の少数の説明変数を用いた研究のみである。岩本ら [1981] は肺炎等の死亡率について都道府県の第一次・第二次産業就業率、個人所得等の7種の説明変数を用いて主成分分析を行った。山中・中村 [1997] は東京都等の5地域における肺炎死亡率について平均気温、平均湿度、降水量の3種の説明変数により重回帰分析を行った。旭ら [2001] は肺炎等の死亡率と喫煙率との相関を分析した。長谷川 [2015] は肺炎等の死亡率について県民所得、ジニ係数、老人ホーム数等の5種の説明変数を用いて重回帰分析を行った。小出・森田 [2019] は肺炎死亡率と医師数、高齢化率、喫煙率等の13種の変数との相関を分析した。これらの先行研究はいずれも少数の説明変数を用いており、その他に肺炎死亡率に大きく寄与する未検証の要因が存在している可能性が高いため、これらの結果には疑問が残る。

また、先行研究の一部で採用されている死亡率と説明変数との相関係数を用いた解析法にも問題があると考えられる。本研究の結果、12種の要因を含む57種の全説明変数について、感度分析で得

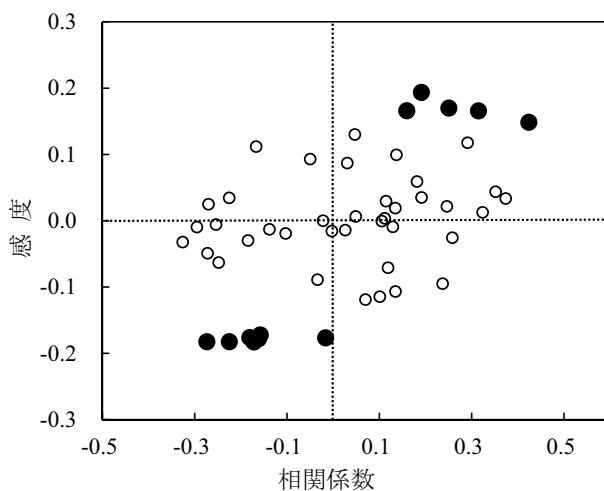


図3 全説明変数についての死亡率との相関係数vs感度のプロット

(黒丸：12種の要因、白丸：その他の変数)

られた感度と、死亡率に対する単相関係数との散布図を図3に示す。要因の中には相関係数の絶対値が低いものがあること、また、要因にならなかった変数の中には相関係数の絶対値が高いものがあること、さらに、感度と相関係数が異符号のものがあることが分かる。上記のように、表2に示した要因の感度は、感度分析において他の変数は固定し、当該要因のみ変化させた場合の死亡率の変化から求めたことから、死亡率に対する当該要因の正味の影響度を表わしているのに対し、相関係数には他の変数の影響が含まれている。これらの結果は、先に述べたように相関係数に基づいた分析の結果には疑問があることを示唆する。

次に、表2の要因の内の主要なものについて考察する。まず、要因1位となった喫煙については、多くの疫学研究からも肺炎に対する喫煙の重大性が指摘されている。日本学術会議[2008]は「脱タバコ社会の実現に向けて」と題した要望書の中で、喫煙は肺炎リスクの80~90%を占めるとした。また、厚生労働省の喫煙の健康影響に関する検討会[2016]は報告書「喫煙と健康」の中で、たばこの健康影響について疫学研究などの科学的知見を系統的にレビューし、たばこ慢性閉塞性肺疾患(COPD)との因果関係をレベル1(科学的証拠は因果関係を推定するのに十分である)と判定した。しかし、肺炎と喫煙の関係について統計分析を行った先行研究では、喫煙の影響の重大性は得られていない。旭ら[2001]は喫煙率と肺炎死亡率との相関を分析し、男性では相関がなく、女性のみ有意な相関があるとした。小出・森田[2019]は肺炎死亡率と医師や病院の数、高齢化率、喫煙率等との相関を分析し、病院や看護師数の数との相関はあるが、喫煙率は有意の相関がないとした。したがって、肺炎死亡に対して喫煙の影響が高いことを実証したのは本研究が初めてである。

さらに、この喫煙と要因2位の短大・大学卒、4位のボランティア活動、および7位の社会体育施設との4要因には共通するキーワードとして、健康意識が想定される。すなわち、高学歴の人ほど、禁煙に留意し、運動やボランティア活動を活発に行っているため肉体的および精神的健康度が高く、肺炎の罹患率や死亡率が低いと考えられるからである。これら4要因の寄与率の合計36.9%は医療・福祉分野の4要因(社会福祉士、保健師、訪問介護員、高齢者受療)の寄与率33.7%を超えることから、肺炎予防には社会インフラの整備以上に、個人の健康意識の向上が有効であると推論される。

本研究では表1のように医療・福祉分野では11種の説明変数を取り上げて重回帰分析を行い、その内、社会福祉士、保健師、訪問介護員、高齢者受療の4種が寄与率(33.7%)で要因となり、肺炎死亡には医療・福祉分野の要因の影響が最大であるという結果が得られた。しかし、先行研究でこの分野の変数の効果を検証したのは小出・森田[2019]のみである。彼らは肺炎死亡率と医師・看護師・助産師・薬剤師・保健師・病院の数等との相関を分析し、看護師や病院の数との間に相関があるが、保健師との相関は有意でないとした。しかし、彼らが用いた相関分析では他の要因の影響が排除できないため、彼らの結果の信憑性には疑問がある。

はじめに記したように、肺炎の死亡率が1980年頃から急増し、2011年から2016年まで死因の第3位になった原因は、急速な高齢化の進行により肺炎に罹患して死亡する高齢者が増加したためとされている。そこで、本研究のSVMによる重回帰分析では表1のように高齢者に関連する12種の説明変数を取り上げて解析したところ、高齢夫婦世帯・高齢者受療・高齢単身世帯の3変数が要因となり、その寄与率合計21.4%で肺炎死亡率に影響するという結果が得られた。これまで高齢者に関する要因を検証した研究としては疫学的見地からの論文は多いが、本研究のように社会経済的変数の影響を統計分析した論文は小出・森田 [2019] のみである。彼らは肺炎死亡率と医師・病院の数、高齢化率、喫煙率等との相関を分析し、看護師や病院数の数との間に相関があるが、高齢化率は有意でないとした。

経済分野では衛生費と生活保護費の2要因が寄与率15.6%で肺炎死亡率に影響するという結果が得られたが、これまで経済変数を検証した先行研究は岩本ら [1981] のみである。彼らは都道府県の主要死因別死亡率について産業就業率、個人所得等の7種の経済社会指標を用いて主成分分析を行い、第二次産業就業率と平均気温が農村地域の肺炎死亡率に大きな寄与をもたらすとしたが、衛生費や生活保護費のような都道府県の所要経費については検証していない。肺炎では死者のほとんどが65歳以上の高齢者であり、団塊の世代が75歳となる2025年には75歳以上が全人口の18%となり、死者数がさらに増加することが予想される。このことは社会保障関係費が急増する「2025年問題」をさらに深刻化するのではないかと危惧されている。

地理・環境分野については8種の説明変数を取り上げたが、どれも要因とはならなかった。これに対し、先行研究では幾つかの変数が要因となっている。岩本ら [1981] は都道府県の主要死因別死亡率について就業率、気温、個人所得等の7種の社会地理環境指標を用いて主成分分析を行い、第二次産業就業率と平均気温が農村地域の肺炎死亡率に大きな寄与をもたらすとした。山中・中村 [1997] は東京都等の5地域における肺炎死亡率と平均気温、平均湿度、降水量との関係を重回帰分析し、冬季の低温とともに夏季の高温が危険因子であるとした。長谷川 [2015] は肺炎死亡率について、平均気温、標高等を説明変数として重回帰分析し、死亡率には標高が有意な影響を及ぼすことを見出した。しかし、これらの先行研究の分析では喫煙等の重大変数が含まれていないため、これらの結果には疑問がある。

#### 4 地域の肺炎対策への試論

次に、肺炎死亡率が国内最高である青森県について、その原因と対策を考察する。そのため、12種の要因の指標値の最小と最大の都道府県、および青森県と、死亡率最下位の長野県について、指標値と危険度順位を表3に示す。ここで、危険度順位は、死亡率に対する要因の危険度を表すために、喫煙等の危険要因については指標値の降順、社会福祉士等の防御要因については昇順での順位

表3 要因の指標値の最小と最大の都道府県および青森県と長野県の指標値と危険度順位

要因	感度符号*	最小		最大		青森県		長野県	
		指標値	該当県	指標値	該当県	指標値	順位**	指標値	順位**
喫煙	+	17.0	奈良	27.6	北海道	25.9	2	20.0	31
短大・大学卒	-	18.5	秋田	39.8	神奈川県	18.8	2	28.6	31
社会福祉士	-	109	茨城	216	新潟	123	3	162	28
ボランティア活動	-	20.6	大阪	33.9	滋賀	22.4	3	32.3	40
保健師	-	23.5	神奈川県	79.3	島根	54.2	25	77.2	46
訪問介護員	-	29.9	岡山	179.3	神奈川県	89.4	37	64.9	22
社会体育施設	-	138	大阪	983	長野	547	29	983	47
衛生費	-	39.5	神奈川県	101.5	熊本	69.0	36	63.7	29
高齢夫婦世帯	+	7.3	沖縄	15.1	山口	11.1	37	13.2	12
生活保護費	+	1,563	岩手	2,159	東京	1,620	44	1,873	12
高齢者受療	+	10.6	長野	16.9	佐賀	12.2	33	10.6	47
高齢単身世帯	+	8.3	滋賀	16.5	高知	12.1	15	10.5	29

\*感度符号の-は下降要因、+は上昇要因。

\*\*危険度順位は上昇要因については指標値の降順、下降要因については昇順の順位。

である。

この危険度順位を見ると、長野県は多くの要因について順位が低く、全体的に危険要因は少ない。これに対し、青森県では全体的に危険度順位の高い要因が多く、特に喫煙、短大・大学卒、社会福祉士、ボランティア活動の4要因の順位の高さが同県の死亡率国内1位に大きな影響を与えている。そこで、これら4要因の結果を基に、青森県の死亡率改善策を考える。

青森県では現在、健康づくりの推進や自殺対策などの短命県返上への模索を実施しているが、表3の結果を基に青森県の死亡率改善策を考えてみると、まず、社会福祉士の増員が考えられる。青森県の社会福祉士数（人口当たり）は全国で3番目に少ない123人であり、これは最多の新潟県の6割弱にとどまる。社会福祉士の充実が青森県のみならず、全国の自治体の死亡率低減に費用対効果の高い対策であると考えられる。

青森県は降雪日数が2位の雪国県であり、喫煙率2位などの健康意識の低さ、無関心につながり、最終的に死亡率の高さを招いていると推測される。そこで、青森県の根本的な死亡率低減対策には健康意識の向上が重要であると考えられ、そのための一つの方策として第一次産業からの産業構造の転換を提案する。同県が雪国県である点を考慮すると、IoT、ビッグデータ、AIなどICT（情報通信技術）産業の推進が有効と思われる。山形県なども同様の施策を進めているが、その見本になるのは自然条件が似ている北欧のフィンランドとエストニアであり、両国の新興IT企業は世界に進出している。青森県内の大学や高専の情報関係学部の誘致と奨学金制度の拡充強化、さらにIT企業の誘致に成功すれば、教育水準の改善が図られ、これらを通じて健康意識の向上、ひいては死亡率の低下が期待できよう。

## 5 おわりに

本研究では、肺炎の死亡要因を明らかにするために、都道府県別の年齢調整死亡率を目的変数とし、57種の社会経済的指標を説明変数とし、サポートベクターマシンによる非線形重回帰分析を行った。その結果、47都道府県の死亡率を有意に再現する12種の要因が得られ、その中では社会福祉士や保健師等の医療・福祉要因と、喫煙等の生活習慣関連要因の寄与がかなり大きいこと、経済分野の要因も有意の寄与を果たしていること等、いくつか興味ある結果が得られた。さらに、12種の要因が5分野にまたがっていることから、多分野の多数多様な説明変数の中から要因を探索する本研究の解析手法の有効性を実証した。

現在、世界中の人間は新型コロナウイルスとの悪戦苦闘を余儀なくさせられている。しかし、ウイルスだけでなく、細菌や微生物などによる感染症は広く生物界に存在しており、100年前のスペイン風邪等、人類は感染症と闘いながら歴史を刻んできた。また、感染症は致死率の高いものから低いものまで常に人間とともにあり、そこから逃れることはできないといわれている。したがって、人間は感染症に対する対策をいつまでも続ける努力が要求される。

我々はこれらの問題に対処すべく、今後、肺炎だけでなくさまざまな感染症の死亡率について、国内・国外の多種多様なデータを用いた解析を行い、要因を検証する研究を行うことを計画している。

### (参考文献)

- Chang, C-C., and Lin, C-J. [2016], "LIBSVM—A Library for Support Vector Machines," <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
- Cristianini, N., and Shaw-Taylor, J. [2000], "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods," Cambridge: Cambridge University Press : 大北剛 訳 [2005], 『サポートベクターマシン入門』 共立出版.
- Tanabe, K., Kurita, T., Nishida, K., Lucic, B., Amic, D., and Suzuki, T. [2013], "Improvement of Carcinogenicity Prediction Performances Based on Sensitivity Analysis in Variable Selection of SVM Models," *SAR and QSAR in Environmental Research*, Vol. 24, No. 7-9, pp. 565-580.
- Wu, Z., and McGoogan, J-M. [2020], "Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China; Summary of a Report of 72,314 Cases From the Chinese Center for Disease Control and Prevention," *The Journal of the American Medical Association*, Vol. 323, No. 13, pp. 1239-1242. doi:10.1001/jama.2020.2648.
- 旭伸一・大木いずみ・谷原真一・尾島俊之・中村好一・岡山明・松村康弘・柳川洋 [2001], 「都道府県別観察による喫煙率と疾患別死亡率の関連」, 『厚生指標』, Vol. 48, No. 10, pp. 11-15.
- 阿部重夫 [2011], 『パターン認識のためのサポートベクトルマシン入門』, 森北出版.
- 岩本美江子・百々栄徳・米田純子・上田洋一 [1981], 「多変量解析による我国の死因別死亡率と社会環境に関する研究」, 『日本衛生学雑誌』, Vol. 36, No. 4, pp. 678-686.

- 小野田崇 [2007], 『サポートベクターマシン』, オーム社.
- 小出淳貴・森田一三 [2019], 「肺炎の都道府県別年齢調整死亡率と医療・社会環境の地域相関研究」, 『日本赤十字豊田看護大学紀要』, Vol. 14, No. 1, pp. 45-54.
- 厚生労働省喫煙の健康影響に関する検討会 [2016], 「喫煙と健康」,  
<https://www.mhlw.go.jp/file/05-Shingikai-10901000-Kenkoukyoku-Soumuka/0000172687.pdf>
- 田辺和俊・栗田多喜夫・西田健次・鈴木孝弘 [2013], 「サポートベクター回帰を用いた158カ国の国債格付けの再現」, 『情報知識学会誌』, Vol. 23, No. 1, pp. 70-91.
- 田辺和俊・鈴木孝弘 [2014], 「サポートベクターマシンを用いた世界各国の幸福度の決定要因の実証分析」, 『経済分析』, No. 188, pp. 44-67.
- 田辺和俊・鈴木孝弘 [2014], 「非線形回帰分析による世界各国の貧困の決定要因の解析」, 『海外社会保障研究』, No. 189, pp. 57-66.
- 田辺和俊・鈴木孝弘 [2014], 「サポートベクターマシンを用いた世界各国の平均寿命の決定要因の実証分析」, 『厚生生の指標』, Vol. 61, No. 13, pp. 23-30.
- 田辺和俊・鈴木孝弘 [2015], 「サポートベクターマシンを用いた所得格差の決定要因の実証分析」, 『情報知識学会誌』, Vol. 25, No. 3, pp. 223-242.
- 田辺和俊・鈴木孝弘 [2015], 「平均寿命および健康寿命の都道府県格差の解析－非線形回帰分析による決定要因の探索－」, 『季刊社会保障研究』, Vol. 51, No. 2, pp. 198-210.
- 田辺和俊・鈴木孝弘・中川晋一 [2016], 「サポートベクター回帰による都道府県別肺がん死亡率の関連要因に関する検討」, 『保健医療科学』, Vol. 65, No. 6, pp. 598-610.
- 田辺和俊・鈴木孝弘 [2016], 「出生率の都道府県格差の分析」, 『厚生生の指標』, Vol. 63, No. 5, pp. 13-21.
- 田辺和俊・鈴木孝弘 [2017], 「母子世帯の都道府県別就業率の推定と要因分析」, 『労働科学』, Vol. 93, No. 6, pp. 165-177.
- 田辺和俊・鈴木孝弘 [2018], 「都道府県の相対的貧困率の計測と要因分析」, 『日本労働研究雑誌』, Vol. 60, No. 2/3, pp. 45-58.
- 田辺和俊・鈴木孝弘 [2018], 「都道府県別全がん死亡率に及ぼす生活習慣要因の影響度分析－自治体のがん対策の視点から－」, 『厚生生の指標』, Vol. 65, No. 11, pp. 15-21.
- 田辺和俊・鈴木孝弘 [2019], 「サポートベクター回帰による都道府県別自殺率の要因分析」, 『情報知識学会誌』, Vol. 29, No. 3, pp. 247-267.
- 田辺和俊・鈴木孝弘 [2020], 「サポートベクター回帰における感度分析による変数選択の有効性の検証－都道府県別全死因死亡率の影響要因の分析－」, 『統計数理』, 印刷中.
- 日本学術会議 [2008], 「脱タバコ社会の実現に向けて」, <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-20-t51-4.pdf>.
- 長谷川卓志 [2015], 「本邦人口動態統計からみた死亡割合とその地域差を決定する要因について」, 『社会医学研究』, Vol. 32, No. 1, pp. 41-45.
- 山中伸一・中村泰人 [1997], 「肺炎・気管支炎死亡率と気温の関係－我が国5地域における比較－」, 『日本気象学会雑誌』, Vol. 34, No. 1, pp. 45-52.