

非線形回帰分析による世界各国の肺炎死亡率の要因分析

鈴木孝弘^a・田辺和俊^b

Determinant Analysis of Pneumonia Mortality Rates
of Many Countries Using Nonlinear Regression

Takahiro SUZUKI^a・Kazutoshi TANABE^b

Abstract

Mortality of pneumonia attracts much attention worldwide in relation to COVID-19 mortality, but few papers were presented on the statistical analysis of determinants of pneumonia in various countries. In this study, a large-scale experiment has been done to analyze determinants of pneumonia mortalities of 159 countries using 44 explanatory variables of seven fields and a non-linear regression support vector machine (SVM) technique. It has been found that just seven factors represent the pneumonia mortalities of 159 countries with a high accuracy of determination coefficient of 0.900, and that four factors among them, literacy, birth rate, temperature, and poverty rate significantly contribute to the pneumonia mortality. For the four factors, only literacy shows negative effect on the pneumonia mortality and other three factors have positive effects.

Keywords : pneumonia; mortality; determinant analysis; nonlinear regression

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、現時点において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、肺炎などを発症して命を落とすこともあるとされる。中国のデー

^a 東洋大学自然科学研究室：〒112-8606 東京都文京区白山 5-28-20

Natural Science Laboratory, Toyo University, 5-28-20 Hakusan, Bunkyo-ku, Tokyo 112-8606, Japan

^b 東洋大学現代社会総合研究所：〒112-8606 東京都文京区白山 5-28-20

Institute of Social Sciences, Toyo University, 5-28-20 Hakusan, Bunkyo-ku, Tokyo 112-8606, Japan

タに基づく分析 (Wu & McGoogan, 2020) 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%となっている。軽症で済むかどうかの分かれ道は、体内の防御システムが機能するかどうかにかかっており、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、肺炎などが重症化しやすいとされる。

一般に肺炎は細菌やウイルスなどの病原体が肺の組織に感染して炎症が生じるものである。わが国をはじめとする先進諸国では、心臓疾患、がん、脳疾患が三大死因であるが、発展途上国などの医療が進んでいない地域では肺炎が死因のトップとなっており、世界全体では肺炎を中心とする呼吸器疾患で死亡する人は年間770万人に上る。

ところで、20世紀後半以降、人の健康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“Closing the Gap in a Generation: Health Equity through Action on the Social Determinants of Health” (WHO, 2008) は健康格差の社会的要因として、社会格差、ストレス、幼少期、社会的排除、労働、失業、社会的支援、薬物依存、食品、交通の10項目を挙げている。

このように人の健康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健康に重大な影響を与える要因の解明が重要になる。そのための疫学的手法には様々な手法があるが、重大要因を解明し、それらの影響度を推定する手法としては、多くの地域集団について、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。

筆者らは国内の肺炎の要因解明を目的として、都道府県別の肺炎死亡率について多数の説明変数を用いて重回帰分析を行い、喫煙、短大卒、社会福祉士、ボランティア活動、保健師等が影響要因であることを見出した (鈴木・田辺、2020)。

これまで統計解析手法を用いて各国の肺炎死亡率について要因分析を行った先行研究は多いが、それらの大多数は一国内の患者について少数の臨床データを説明変数として回帰分析を行っている。複数国の肺炎死亡率について複数の説明変数を用いて重回帰分析を行った研究はあるが、ごく少数であり、しかも、それらは少数の国を対象とし、少数の説明変数を用いて解析している。広範かつ多様な国々の肺炎死亡率の要因を多数の変数の中から探索した研究はない。

そこで本研究では、世界中の多数国の肺炎死亡率を目的変数とし、健康・医療、経済・社会、教育・文化等の多分野の多数の説明変数を用いて重回帰分析により一括解析し、それらの中から決定要因を探索する実証分析を行った。

2. 方法

2.1 目的変数

重回帰分析の目的変数にはWHO公表の世界159カ国の肺炎死亡率 (2018年) を用いた。

死亡率上位の10カ国とUnited States、Japan、Germanyなど主要先進国を含む159位までのFinlandなどの死亡率をTable 1に示す。死亡率最高のNigeriaと最低のFinlandでは100倍近い差がある。わが国は159カ国中では84位の中程度であるが、先進38カ国の中では4位と高い。159カ国中の上位10カ国のほとんどはアフリカの途上国である一方、下位には先進国が多い。しかし、途上国は下位にも散見され、157位のMontenegroのように死亡率が非常に低い国もある。

Table 1. Top ten and other major countries in terms of pneumonia mortalities

No	Country	Area	Mortality*	A/D**
1	Nigeria	Central Africa	174.0	D
2	Chad	Central Africa	171.6	D
3	Niger	Central Africa	162.3	D
4	Guinea	West Africa	143.8	D
5	Burkina Faso	West Africa	140.6	D
6	Sierra Leone	West Africa	138.3	D
7	Guinea-Bissau	West Africa	135.7	D
8	Central African Republic	Central Africa	122.7	D
9	Haiti	Central America	122.2	D
10	Mozambique	South Africa	121.2	D
34	Philippines	Southeast Asia	90.4	D
48	India	South Asia	63.1	D
58	Argentina	South America	48.7	D
62	United Arab Emirates	Middle East	45.8	A
64	Brazil	South America	42.7	D
84	Japan	East Asia	24.3	A
97	South Korea	East Asia	21.0	A
100	United Kingdom	Western Europe	19.2	A
114	Russia	Russia	16.2	D
129	China	East Asia	11.7	D
131	United States	North America	10.9	A
140	Germany	Western Europe	9.0	A
147	France	Western Europe	7.9	A
157	Montenegro	Eastern Europe	5.3	D
159	Finland	Northern Europe	2.1	A

* Death rate per 100,000 population, published by WHO (2008).

** A: advanced country, D: developing country.

159カ国の先進国・途上国別の死亡率の分布 (Fig. 1) をみると、先進国は途上国と比較して死亡率が低い国が多い。最高はUnited Arab Emiratesの45.8であり、死亡率0～10の国が最多である。そのため、先進38カ国の平均死亡率は13.8と低い。これに対し、途上国はNigeriaの174.0からMontenegroの5.3まで広範囲に分布している。最多の死亡率は10～20であるが、死亡率90～120の国も比較的多い。そのため、途上121カ国の平均死亡率は57.2であり、先進国より高い。

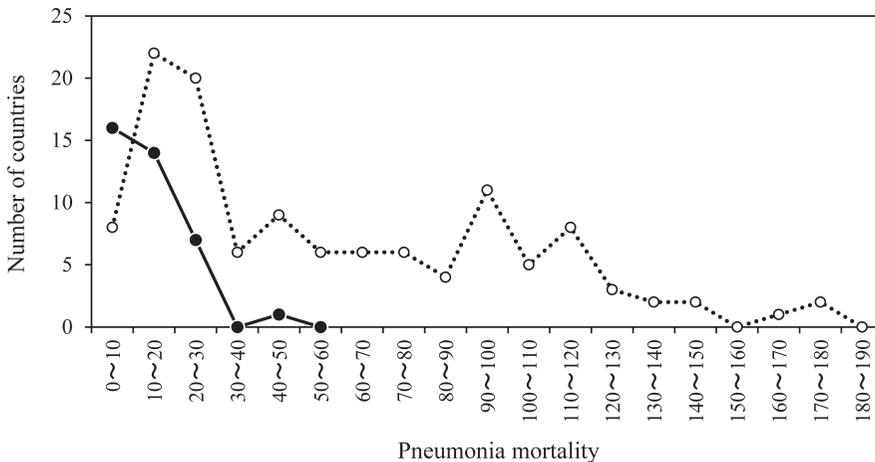


Fig. 1. Distribution of pneumonia mortalities of advanced (●) and developing (○) countries

以上のように、この159カ国は世界中の多様な、かつ、全地域の国々を網羅している。本研究では、肺炎死亡率の要因に関して、国家の発展段階によらない一般性の高い結果を得るために、全159カ国を一括して解析し、決定要因を探索した。

2.2 説明変数

説明変数については、本研究では広範囲の説明変数の中から決定要因を探索するために、先行研究で用いられている指標をできるだけ採用することにした。そこで、WHO等の報告や多くの論文で検証されていること、世界中の多数の国についてデータが入手可能であること等の選択基準から、Table 2に示す医療、健康、経済、文化、人口、地理、環境の7分野の44種の説明変数を採用した。これらの指標の内、人口当たりの数値が算出できるものはその数値を用い、すべての指標は0と1の間に正規化して解析に用いた。

Table 2. Objective and 44 explanatory variables for pneumonia mortalities

Field	Variable	Definition	Source
Objective	Pneumonia mortality	Pneumonia death rate per 100,000 population	WHO
Medical	Health expenditure	Public and private health expenditures per capita	WHO
	Physicians	Number of physicians per 1,000 population	WHO
	Hospitals	Number of hospitals per 1,000 population	WHO
	Hospital beds	Number of hospital beds per 1,000 population	WHO
	Social welfare	Sen social welfare function	WHO
Health	Obesity	Rate of adults of body mass index over 30	WHO
	Calorie intake	Daily calorie intake per capita in 10 ³ kcal	WHO
	Undernourishment	Prevalence of undernourishment (% of population)	WHO
	Physical inactivity	Prevalence of insufficient physical activity among adults	WHO
	Smoking prevalence	Percentage of adults who smoke daily	WHO
	Alcohol consumption	Alcohol consumption among adults	WHO
Economy	Gross domestic product	Per capita gross domestic product	IMF
	Gini coefficient	Measure of inequality of income	WB
	Poverty rate	Poverty ratio at national poverty line per population	WB
	Labor force	Labor force participation rate	WB
	Unemployment rate	Rate of the number of unemployed individuals	IMF
	Aged economic activity	Economic activity in men aged 65 plus	ILO
	Grain self-sufficiency rate	Grain self-sufficiency rate	FAO
	Agriculture	Agricultural land (% of land area)	WB
Culture	Education expenditure	Public expenditure on education per capita	WB
	Literacy	Literacy rate of people ages 15 and above	WB
	Intelligence level	Ratio of tested mental age to chronological age	Eutimes
	School life expectancy	School life expectancy, primary to tertiary education	UN
	University	Percentage of enrollment into university	WB
	Mobile phone	Mobile phone subscribers per 1,000 people	WB
	Internet	Internet users per 100 population	WB
Population	Population density	Number of people per sq. km of land area	WB
	Urban population	Rate of people living in urban areas	WB
	Population growth rate	Population growth rate	WB
	Birth rate	Average annual number of births per 1,000 persons	WB

	Population aging	Rate of population ages 65 and over	WB
	Household member	Members per private household	WB
Geography	Latitude	Average latitude of country	WB
	Temperature	Average yearly temperature	WB
	Sunshine hours	Yearly sunshine hours	WB
	Precipitation	Average precipitation in depth	WB
	Climate Risk	Climate Risk Index	CRI
	Environment	Terrestrial protected area	Rate of terrestrial protected areas to total land area
Forest area		Rate of forest area to land area	WB
Air pollution		Country level of PM10 in $\mu\text{g}/\text{m}^3$	WHO
Indoor air pollution		Rate of access to clean fuels and technologies for cooking	WHO
Water and sewer services		Rate of improved water source and sewer services	WHO
Roads		Roads, paved (% of total roads)	WB
Electric power		Electric power consumption (kWh per capita)	WB

For more detailed definition of variables, see homepages of data sources.

2.3 解析手法

先行研究においては解析に線形重回帰分析（OLS）が適用されてきた。しかし、各種の指標と肺炎死亡率との関係は一般に線形であるものはほとんどなく、複雑な相関関係を示す指標が多い。このような複雑な事象に対する有効な対処策として、非線形回帰分析手法の適用が考えられる。しかし、非線形回帰分析を用いて肺炎死亡率の決定要因を探索した研究は見当たらない。

非線形解析手法の中で近年注目されているサポートベクターマシン（SVM）（大北, 2005；小野田, 2007；阿部, 2011；竹内・鳥山, 2015）は、説明変数の数値に対してカーネルと呼ぶ非線形関数（本稿ではガウス関数）を用いて学習パターンを別の空間（超平面）に写像し、そこで線形回帰を行う。それにより、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。また、高速処理が可能、最適解が一義的に求まる等の利点があるため、データ解析手法として現時点では最も有効な方法とされている。

そこで、本研究ではSVMを用いて多数の説明変数の中から決定要因の探索を試みた。SVMのソフトウェアはLIBSVM ver. 2.89（Chang & Lin）の回帰機能（ ϵ SVR）、カーネル関数はRBFを用いた。説明変数の中から決定要因を探索するためには、SVMモデルの最適化と説明変数の最適化が必要である。そのために、交差検証法（Cross Validation Test）と感度分析法の2つを組み合わせた方法を採用した。後者の感度分析法とは、各変数の感度を計算し、全変数の中で感度が最も低い変数を順次削除しながらSVM解析を行って予測値と実測値の平均二乗誤差（RMSE）が最小となる点を探索する方法である（Tanabe et al, 2013）。筆者らはその有効性を各種の問題で検証している（鈴木・田辺

2019, 2020, 田辺・鈴木2016, 2018, 2019, 2020)。

本研究では、以下の手順で決定要因の探索を行った。

- ①交差検証法によりSVMモデルを最適化するためにLIBSVMのパラメータ ν を159とし、全159カ国のデータをSVMに入力し、モデルの最適化に必要なパラメータ g と c をグリッドサーチしてRMSEが最小になる組み合わせを探す。
- ②各変数の感度を求めるために、当該変数は実際の数値に設定し、その他の変数は159カ国の平均に設定したデータを最適モデルに入力し、出力値を求める。
- ③当該変数の設定値を説明変数、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方法を用いて、各変数の感度を算出する。
- ④全変数の中で感度の絶対値の最も小さい変数を取り除き、以上の操作を繰り返し、159カ国のRMSEが最小になる変数の組み合わせを決定要因とする。

3. 結果

以上の方法により、全159カ国の肺炎死亡率について44種の説明変数から出発して順次減らした結果、7種の変数を用いた時にRMSEが最小となった。この7変数はリテラシー、出生率、気温、貧困率、上下水道サービス、飲酒量、喫煙率である。このSVMモデルの死亡率の実測値と予測値の散布図をFig. 2に示す。

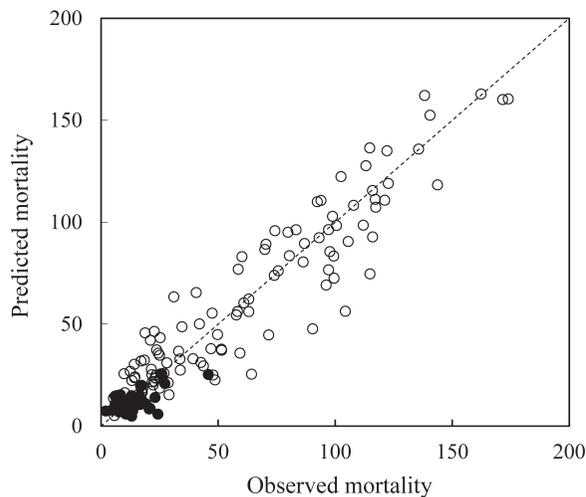


Fig. 2. Scatter plot of observed vs predicted mortalities for advanced (●) and developing (○) countries

この図から世界の多数国の死亡率の実測値がよく再現されていることが分かり、RMSEが14.3と小さい。また、回帰決定係数 (R^2) が0.900と高いことから、統計的に危険率1%で有意の回帰ありと判定される。したがって、この7種の変数は世界159カ国の肺炎死亡

率の決定要因とみなされる。

これらの決定要因7種の内訳とそれらの感度、および式

$$C_i(\%) = \frac{S_i^2}{\sum_{i=1}^7 S_i^2} \times 100 \quad (1)$$

(ここで、 S_i は決定要因*i*の感度)により算出した死亡率への寄与率をTable 3に示す。

Table 3. Determinants of pneumonia mortality

No	Determinant	Field	S*		C** (%)
			Risk factor	Decline factor	
1	Literacy	Culture		-0.115	27.4
2	Birth rate	Population	0.102		21.8
3	Temperature	Geography	0.091		17.3
4	Poverty rate	Economy	0.082		14.1
5	Water and sewer services	Environment		-0.061	7.7
6	Alcohol consumption	Health	0.060		7.6
7	Smoking prevalence	Health	0.044		4.0

S*: sensitivity, C**: contribution rate to mortality

4. 考察

以上の結果を、肺炎死亡率の決定要因を探索した先行研究の結果と比較してみよう。上記のように、これまで統計解析手法を用いて各国の肺炎死亡率について要因分析を行った先行研究は多いが、それらの大多数は一国内の患者について少数の臨床データを説明変数として回帰分析を行った研究である。複数国の肺炎死亡率について複数の説明変数を用いて統計解析を行った研究はあるが、ごく少数であり、しかも、それらはいくつかの問題点がある。

まず、第一の問題点は線形回帰の適用である。本研究では目的変数(死亡率)に対して非線形関係を示す説明変数が多いことから、非線形回帰分析の一手法であるサポートベクターマシン(SVM)を適用し、Table 3に示した良好な回帰結果を得た。これに対して、この7種の説明変数を用いて線形重回帰分析(OLS)を行うと、RMSEは37.2、 R^2 は0.677となり、SVMの結果より再現性が悪くなる。この原因は上記のように死亡率に対して非線形関係にある指標が多いためであり、その例をFig. 3に示す。

この結果は、各種指標と死亡率との関係は非線形性が高いため、先行研究で用いられて

いる線形重回帰分析では精度の良い結果を得ることが難しく、死亡率の決定要因に関して信頼性の高い結果を得るためにはSVM等の非線形解析手法の適用の必要性を示している。

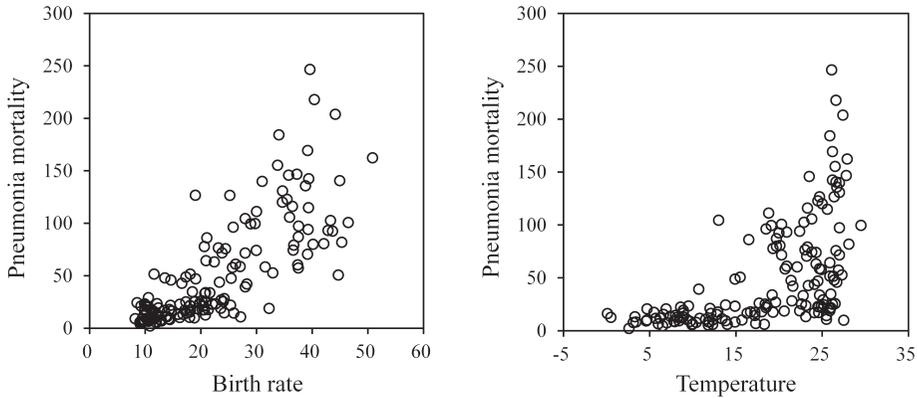


Fig. 3. Correlation plot of birth rate and temperature vs pneumonia mortality

先行研究の第二の問題点は、少数の国について少数の説明変数を用いて解析した研究がほとんどであり、結果の汎用性に欠けることである。Macinko, et al. (2003) はOECD 18カ国の肺炎死亡率について、医師、GpC（国民一人当たりのGDP値）、飲酒、喫煙等の8種の説明変数を用いて解析した。Puro, et al. (2005) は欧州37カ国について、年齢、性別、地域等の5種の説明変数を用いて解析した。Nikolopoulos, et al. (2011) は欧州30カ国の死亡率について、緯度、医療費、GpC、高齢率等の12種の説明変数を用いて解析した。Noordam, et al. (2015) はアフリカ6カ国の子供の肺炎死亡率について、年齢、性別、教育、貧富度等の5種の説明変数を用いて解析した。このように、いずれの研究も少数の国について限定された少数の説明変数を用いて解析している。

これに対し、本研究ではTable 2に示したように、医療、健康、経済、文化、人口、地理、環境の7分野の44種の説明変数の中から要因を探索した。さらに、Table 3に示したように、7種の決定要因は医療を除く文化、人口、地理、経済、環境、健康の6分野にまたがっている。この結果は、死亡率の決定要因に関して汎用性と信頼性の高い結果を得るためには、多分野の多種の説明変数を用いて解析することの必要性を示している。

先行研究の第三の問題点は、各種指標の中で死亡率に有意な影響を示す指標を探索するために相関分析の手法を用いた研究があるが、これらの研究結果の信頼性にも疑問がある。すなわち、本研究で検証した44種の説明変数について、死亡率との単相関係数と感度分析で得られた感度との散布図をFig. 4に示す。44種の説明変数全体では右上がりの相関傾向がみられるが、単相関係数と感度が異符号の変数が多く存在する。このことは、相関分析法で導かれた結果が本研究で採用した感度分析の結果と異なる変数が多く存在することを示し、死亡率に有意な影響を与える変数を探索するためには、感度分析法を適用することの必要性を示している。

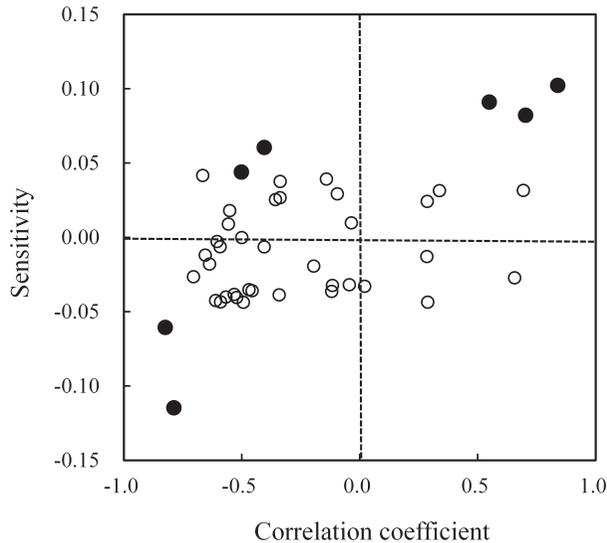


Fig. 4. Scatter plot of correlation coefficients vs sensitivities for determinants (●) and non-determinants (○)

以上を総括すると、先行研究では線形重回帰分析や相関分析の適用、および解析対象（国数）と説明変数の十分さの点に問題があり、それらの研究結果の信頼性と汎用性には疑問が残る。したがって、肺炎死亡率の決定要因に関して信頼性の高い結果を得るためには、本研究のように、多分野の多種の説明変数を用い、SVM等の非線形解析手法を適用することが必要性不可欠であると結論される。

Table 3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とき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、リテラシー（%）と気温（℃）と飲酒量（L）のように単位の異なる要因について、それらの感度の大きさにより、死亡率への影響度の比較が可能になる。また、感度が正の出生率、気温、貧困率、飲酒量、喫煙率の5要因は死亡率の危険要因であり、負のリテラシー、上下水道サービスの2要因は抑制要因であると解釈できる。

また、Table 3に示した決定要因の死亡率に対する寄与率をみると、リテラシー、出生率、気温、貧困率の4要因のみで81%の高率であり、これら4要因が肺炎死亡率の重大要因であるといえる。冒頭に記したように、先進諸国では心臓疾患、がん、脳疾患が三大死因であるが、発展途上国などの医療が進んでいない地域では肺炎が死因のトップとなっている。そこで、途上国の政府が肺炎死亡率抑制策を考える際には、教育制度の充実、出生率の抑制、貧困層の低減などが最も有効であると推論できる。

5. 結論

WHO公表（2018年）の世界159カ国の肺炎死亡率を目的変数とし、医療、健康、経済、文化、人口、地理、環境の7分野の44種の指標を説明変数として用い、非線形重回帰手法SVMで解析し、感度分析法により決定要因を探索した。その結果、わずか7種の指標で世界159カ国の肺炎死亡率が回帰決定係数（ R^2 ）0.900という先行研究よりはるかに高い精度で再現できることを見出した。さらに、7種の指標の中ではリテラシー、出生率、気温、貧困率の4要因のみで81%もの高い影響度を占めることから、肺炎死亡率が高い開発途上国の政府が死亡率抑制策を考える際には、教育制度の充実、出生率の抑制、貧困層の低減などが最も有効であると推論した。

しかし、本研究の結果にも幾つかの課題がある。第1は本研究の結論の一般性の検証である。すなわち、WHOの報告書や先行研究によれば、肺炎死亡率を決定している要因には本研究で取り上げた指標以外に、住居、家族・友人、余暇・趣味・ボランティア活動等、様々な要因が挙げられている。しかし、これらの要因は少数の国しか指標の数値がないため、多数の国を統一的に扱う本研究では採用できなかった。したがって、これらの指標のデータを何らかの方法で入手して解析を行い、その結果を検証することは今後の課題である。

本研究の結果では、肺炎死亡率に対する経済的指標の直接的効果は低いが、間接効果は大きいことが明らかになった。健康の社会的決定要因については階層構造を形成すると考えられている。このような階層構造を考慮した肺炎死亡率の決定要因の分析は本研究の結果の展開として重要なテーマであるので、今後の課題として検討していきたいと考えている。

参考文献

- Chang, C. C. and Lin, C. J. "LIBSVM: A Library for Support Vector Machines," <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
- CRI, "Global Climate Risk Index," <https://germanwatch.org/en/17307>.
- Eutimes, "The European Union Times," <http://www.eutimes.net/2009/11/iq-by-country/>.
- FAO, "United Nations Food and Agriculture Organization Statistics Division," <http://faostat.fao.org/site/567/default.aspx#ancor>.
- ILO, "International Labour Union LABORSTA," <http://laborsta.ilo.org/>.
- IMF, "International Monetary Fund World Economic Outlook Database," <http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/index.aspx>.
- Macinko, J., Starfield, B. and Shi, L. (2003) "The Contribution of Primary Care Systems to Health Outcomes within Organization for Economic Cooperation and Development (OECD) Countries, 1970-1998," *Health Services Research*, Vol. 38, No. 3, pp. 831-865.

- Nikolopoulos, G., Bagos, P., Lytras, T., and Bonovas, S. (2011) "An Ecological Study of the Determinants of Differences in 2009 Pandemic Influenza Mortality Rates between Countries in Europe," *PLOS ONE*, Vol. 6, No. 5, e19432, doi:10.1371/journal.pone.0019432.
- Noordam, A. C., Carvajal-Velez, L., Sharkey, A. B., Young, M. and Cals, J. W. L. (2015) "Care Seeking Behaviour for Children with Suspected Pneumonia in Countries in Sub-Saharan Africa with High Pneumonia Mortality," *PLOS ONE*, Vol. 10, No. 2, e0117919, doi:10.1371/journal.pone.0117919.
- Puro, V., Serraino, D., Piselli, P., Boumis, E., Petrosillo, N., Angeletti, C. and Ippolito, G. (2005) "The Epidemiology of Recurrent Bacterial Pneumonia in People with AIDS in Europe," *Epidemiology and Infection*, Vol. 133, pp. 237-243.
- Tanabe, K., Kurita, T., Nishida, K., Lučić, B., Amić, 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, pp. 565-580.
- UN, "United Nations Development Program Human Development Report," <http://hdr.undp.org/en/>.
- WHO "Closing the Gap in a Generation: Health Equity through Action on the Social Determinants of Health," http://whqlibdoc.who.int/hq/2008/WHO_IER_CSDH_08.1_eng.pdf.
- WHO "World Health Organization World Health Statistics," http://www.who.int/gho/publications/world_health_statistics/en/index.html.
- 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.
- 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.
- 大北剛 (訳) (2005) 『サポートベクターマシン入門』 共立出版.
- 小野田崇 (2007) 『サポートベクターマシン』 オーム社.
- 鈴木孝弘、田辺和俊 (2019) 『サポートベクター回帰による都道府県別乳がん死亡率の要因分析』 東洋大学紀要自然科学篇, Vo. 63, pp. 85-99.
- 鈴木孝弘、田辺和俊 (2020) 『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—』 東洋大学経済論集, 第46巻1号, pp. 15-28.
- 竹内一郎、鳥山昌幸 (2015) 『サポートベクトルマシン』 講談社.
- 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 (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, Vol. 68, No. 1, pp. 175-192.