

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－

田 辺 和 俊
鈴 木 孝 弘

I はじめに

日本人の平均寿命は戦後の食料、医療、衛生水準等の改善により、1947年の男性50.1歳、女性54.0歳から、1960年の男性65.4歳、女性70.3歳に伸びた。2012年には男性79.9歳、女性86.4歳に達したが、1947年から1960年までの13年間の伸びを獲得するために1960年から2012年までの52年を要している。WHOのデータによれば、2012年の平均寿命は世界194カ国中で男性は第4位、女性は第1位であり、男女平均では世界の最長寿国であるが、伸びが頭打ちになった昨今の我が国では、寿命の国内格差の解消に関心が集っている。最長寿県と最短寿県の差は戦前には5歳以上あったが、最近では男性の3.6歳の差に対し、女性では1.8歳と減少している。

都道府県の平均寿命の順位は近年かなりの変動があり、特に沖縄、長野、青森の3県が注目されている。男性は沖縄県が1980年と1985年には全国1位であったが、現在は30位と急落し、代わって長野県が1990年以降、1位を維持している。一方、青森県は1960年から現在まで50年以上、最下位のままである。女性は沖縄県が1975年以降、1位を保っていたが、2010年に長野県に抜かれて3位に落ちた。青森県は2000年以降、男性同様、最下位である。最長寿の長野県と最短寿の青森県は経済力(県民所得)や医療水準(病院や医師の数)ではそれほど差はないのに、寿命でかなりの差が付いている原因に多くの関心が集まっている。

長野県の長寿化の躍進には、保健師・保健補導員・食生活改善推進員による塩分・野菜等の摂取、喫煙・飲酒、肥満、健診受診等に関する活動、在宅医療・保育所・図書館・博物館・公民館の充実、高齢者や女性の就業・社会活動・趣味活動の高さ等の要因が指摘されている。一方、青森県の低迷には喫煙率と飲酒率の高さ、塩分摂取量の多さ、野菜摂取量の少なさ、健診受診率の低さ等、また沖縄県の急降下には高カロリーの食事や自動車利用による肥満率の上昇等の要因が挙げられている。これらの多種多様な要因が3県の寿命の変動と相関があることが指摘されている(青森県2002、嘉手川・米盛2003、竹森2003、水島ら2004、竹森ら2005、三上ら2005、竹森2007、綿引・畑2009、白澤2013)。

また、近年では単なる生存年齢である平均寿命でなく、個人の生活の質を重視する健康寿命が注目されている(矢野2000、伏見2003、加藤ら2007、村田2010)。健康寿命とは一般に、ある健康状態で生活することが期待される平均期間またはその指標の総称を指す。健康寿命の定義と指標の算出方法は幾つかあるが、本稿では、21世紀における第2次国民健康づくり運動(健康日本21(第2次))の「健康上の問題で日常生活が制限されることなく生活できる期間」との定義を採用する。日本人の健康寿命はWHOの報告書(WHO 2004)では男性72.3歳、女性77.7歳であり、世界の最長寿国であるが、健康寿命は平均寿命より9歳以上短いため、社会保障負担軽減の面から健康寿命を平均寿命に近づけることに関心が集まっている。健康寿命は

愛知県、静岡県等が上位に並んでいるが、その要因についてはあまり議論されていない。

寿命を決める要因については最近注目されているサーチュイン長寿遺伝子等、遺伝子の影響は2割程度であり、残りは生活習慣や生活環境の影響で決まる。さらに、寿命には医療、健康、栄養等の直接的要因だけでなく、経済、社会、自然・生活環境、文化等の多数の社会経済的要因が関係している（坂井1986, 大蔵2004, 近藤2004, 荒記ら2005, 近藤2005, 川上ら2006, 福田・今井2007, Kagamimoriら2009, 近藤2010, 堀内2010, 杉澤2012, 橋本2012）。そのため、個々の要因と寿命との相関解析では、寿命に重要な影響を与える決定要因を明らかにすることはできない。また、健康寿命は平均寿命と国内順位が異なっているため、平均寿命と健康寿命の決定要因にどのような違いがあるかは興味深い。

そこで、平均寿命や健康寿命を目的変数、幾つかの指標を説明変数とし、線形重回帰分析（OLS）により決定要因を探索する実証的研究が行われている（渡辺・加納1983, 人見ら2002, 鈴木2003, 高・梯2006, 京都大学2007, 栗盛2008, 多門ら2011, 近藤ら2012）。しかし、いずれも少数の説明変数に限定して解析しているため、回帰決定係数が低く、統計的に有意な結果が得られていない。また、決定要因の種類が論文によって全く異なり、結果の信頼性に疑問がある。また、健康・医療、経済・社会等の個別分野に限定して解析しているため、個々の決定要因が平均寿命に与えている相対的影響度が明らかになっていない。さらに、都道府県の平均寿命と健康寿命について解析し、両寿命の決定要因の違いを検討した研究は見当たらない。

先行研究において回帰決定係数の低い一因はOLSの適用にもある。なぜなら、相対所得仮説（Wilkinson 1992）が示すように、各種の指標と寿命との関係は一般に線形ではなく、複雑な相関関係を示す指標が多いからである。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。サポートベクターマシン（SVM）（大北2005, 小野田2007, 阿部

2011）は近年注目されている非線形解析手法であり、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。しかし、SVMを平均寿命や健康寿命の決定要因解析に適用した研究は見当たらない。

そこで、①これまで見出されていなかった新たな決定要因の発見、②得られた決定要因の相対的影響度の解明、および③平均寿命と健康寿命の決定要因の違いの解明を目的として、都道府県別、男女別の平均寿命と健康寿命のデータを目的変数とし、健康・医療、経済・社会、生活・文化等、多分野の多種多様な指標を説明変数として用いてSVMにより一括解析し、決定要因を探索する大規模実証分析を試みた。

II 方法

1) 寿命および各種指標のデータ

国内47都道府県の平均寿命は厚生労働省都道府県別生命表¹⁾（2010年）の数値を、健康寿命は厚生労働科学研究費補助金「健康寿命における将来予測と生活習慣病対策の費用対効果に関する研究」の算定値²⁾（2010年）を用いた。それらの男女別の上位および下位各5都道府県の寿命を表1に示す。最長寿命と最短寿命の差は3.60（歳）（男性の平均寿命）～1.84（歳）（女性の平均寿命）と小さく、かつ、データ数が47しかないため、予測精度の高い回帰モデル作成の困難さが予想される。

そこで、説明変数についてはこれまで見出されていなかった新たな決定要因を発見するためにできるだけ多数の指標を取り上げ、変数選択法によりそれらの中から有効な決定要因を探索した。説明変数の候補になりうる指標の総数は数百種に上るため、先行論文等において解析・検討されてい

表1 平均寿命・健康寿命の上位および下位5都道府県の寿命（歳）（2010年）

No.	平均寿命男		平均寿命女		健康寿命男		健康寿命女	
1	長野県	80.88	長野県	87.18	愛知県	71.74	静岡県	75.32
2	滋賀県	80.58	島根県	87.07	静岡県	71.68	群馬県	75.27
3	福井県	80.47	沖縄県	87.02	千葉県	71.62	愛知県	74.93
4	熊本県	80.29	熊本県	86.98	茨城県	71.32	栃木県	74.86
5	神奈川県	80.25	新潟県	86.96	山梨県	71.20	沖縄県	74.86
	全国平均	79.59	全国平均	86.35	全国平均	70.42	全国平均	73.62
43	長崎県	78.88	岩手県	85.86	岩手県	69.43	徳島県	72.73
44	福島県	78.84	茨城県	85.83	大阪府	69.39	福岡県	72.72
45	岩手県	78.53	和歌山県	85.69	長崎県	69.14	大阪府	72.55
46	秋田県	78.22	栃木県	85.66	高知県	69.12	広島県	72.49
47	青森県	77.28	青森県	85.34	青森県	68.95	滋賀県	72.37

ること、各寿命との相関が比較的高いこと³⁾、長野県、沖縄県、青森県等の寿命順位変動に関連あるとされていること等の基準で、表2に示す60種の指標を採用した⁴⁾。

以上の指標の最新値について人口当たりの数値が算出できるものはその数値を用い、すべての指標は最小値と最大値が0と1になるよう正規化して解析に用いた。説明変数が60種もあるため、それらの記述統計量や相互相関係数は紙面の関係から割愛せざるをえないが、SVMによる解析では、説明変数間に強い相関がある場合でも解析可能であり、多重共線性問題を回避することができる⁴⁾。

2) SVMによる解析

平均寿命男、同女、健康寿命男、同女を目的変数とする4つの場合について、それぞれ決定要因を探索した。SVMのソフトウェアはLIBSVM ver. 3.11 (Chang and Lin, 2013) の回帰機能 (SVR) を、カーネル関数はRBF (ガウス関数) を用いた。

多数の説明変数の中から決定要因を探索するためにはSVMモデルと説明変数の最適化が必要である。前者についてはLIBSVMのSVRでは g (RBFカーネルの γ)、 c (cost)、および p (loss functionの ϵ) の3種のパラメータの最適化が必要であり⁵⁾、交差検証法を用いて行った。

説明変数の最適化には感度分析法を採用した。この方法は、目的変数に対する各説明変数の感度⁶⁾を計算し、感度の絶対値の小さい変数を順次削除しながらSVMモデルを最適化し、目的変数の

平均二乗予測誤差 (RMSE) が最小となる組み合わせを探索する方法であり、筆者らはこの方法の有効性を様々な問題で検証している (田辺ら 2013a, 2013b, 2014a, 2014b, 2014c)。

そこで、交差検証法と感度分析法を組み合わせた以下の手順により決定要因の探索を行った。

- ①47都道府県をランダムに10群に分割し、第1群を予測セット、その他の9群をまとめて学習セットとする。
- ②学習セットについてSVRパラメータ g 、 c 、 p をグリッドサーチして最適条件を探し、この最適モデルに予測セットのデータを入力して寿命の予測値を求める。
- ③第2群以下の各群を予測セットとして以上の操作を繰り返し、全都道府県について寿命のRMSEを求める。
- ④各説明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、その他の変数は全データの平均値に設定したデータセットを作成し、そのデータを最適モデルに入力し、寿命の出力値を求める。
- ⑤当該変数の実測値を説明変数、寿命の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。
- ⑥全説明変数の中で感度の符号が不適当なもの⁷⁾、または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す。
- ⑦説明変数とパラメータ g 、 c 、 p の組み合わせの中で、全データについてのRMSEが最小になる説

表2 用いた説明変数の定義とデータ源

説明変数	定義	出典	説明変数	定義	出典
転入率	人口当たりの転入人口数	1)	生活習慣病	生活習慣病患者の比率	1)
転出率	人口当たりの転出人口数	1)	スポーツ	スポーツをしている人の比率	1)
出生率	合計特殊出生率	1)	ストレス	悩みやストレスを抱えている人の比率	4)
乳児死亡率	出生数当たりの乳児死亡率	1)	宗教	人口当たりの宗教信者数	1)
世帯人数	一般世帯平均人数	1)	医療費	人口当たりの国民医療費(万円)	1)
婚姻率	人口当たりの婚姻件数	1)	健診受診率	健康診断受診率	1)
離婚率	人口当たりの離婚件数	1)	がん受診率	胃・肺・大腸がん受診率の平均値	1)
未婚率	生涯未婚率	1)	健康保険	人口当たりの健康保険被保険者数	1)
独居率	単独世帯率	1)	病院	人口当たりの病院数	1)
高齢化率	65歳以上人口比率	1)	病床	人口当たりの病床数	1)
気温	年平均気温(℃)	1)	医師	人口当たりの医師数	1)
日照	年間日照時間(時間)	1)	保健師	人口当たりの保健師数	1)
降水	年間降水量(mm)	1)	保育所	人口当たりの保育所数	1)
積雪	最深積雪(cm)	1)	生活保護	生活保護世帯の比率	1)
持家率	世帯当たりの持ち家の比率	1)	老人ホーム	人口当たりの老人ホーム数	1)
住宅面積	1住宅当たりの延べ面積(m ²)	1)	高齢者学級	人口当たりの高齢者学級・講座数	1)
災害罹災率	人口当たりの自然災害の罹災者数(2007-2010年の合計)	1)	高齢者ボランティア	ボランティア活動を行っている高齢者の比率	1)
火災死者	人口当たりの火災死者数	1)	県民所得	人口当たりの県民所得(千円)	1)
交通事故死	人口当たりの交通事故死者数	1)	ジニ係数	所得格差の指数	1)
公害	人口当たりの公害苦情件数	1)	仕事	仕事の平均時間	1)
食塩	1日の食塩摂取量(g)	2)	女性労働	女性労働力人口比率	1)
牛乳乳製品	1日の牛乳・乳製品摂取量(g)	2)	高齢有業率	高齢者の有業率	1)
肉類	1日の肉類摂取量(g)	2)	農民率	農林漁業者の人口比率	1)
卵類	1日の卵類摂取量(g)	2)	教育費	世帯の教育費(万円)	1)
緑黄野菜	1日の緑黄色野菜類摂取量(g)	2)	公民館	人口当たりの公民館数	1)
緑茶	二人以上の世帯の緑茶購入量(g)	3)	図書館	人口当たりの図書館数	1)
カロリー	1日の平均摂取カロリー(kcal)	2)	インターネット	インターネットの世帯普及率	1)
肥満	Body Mass Indexの平均値	2)	携帯電話	携帯電話の世帯普及率	1)
喫煙	喫煙習慣の成人の比率	4)	自動車	人口当たりの自家用車保有台数	1)
飲酒	飲酒習慣の成人の比率	4)			
高血圧	高血圧患者の比率	1)			

出典：1) 地域別統計DB, 2) 国民健康栄養調査, 3) 家計調査, 4) 国民生活基礎調査, 各DBのURLは参考文献欄を, 指標の詳しい定義については各DBを参照されたい。

明変数の組み合わせを寿命の決定要因とする。

Ⅲ 結果

以上の方法により, 60種の説明変数の中から決定要因を探索した結果, 13~20種の指標において各寿命のRMSEが最小となった。寿命4種の場合の回帰結果を表3に, 各寿命の実測値vs予測値の散布図を図1に, また, 各寿命の決定要因とその感度を表4に示す。寿命のRMSEは47都道府県の寿命の最大と最小の差と比べて十分小さく, また,

回帰決定係数(R^2)はかなり高い。したがって, 本稿のモデルで平均寿命と健康寿命の都道府県格差が全体的により精度で再現されているといえる⁸⁾

この結果を先行研究と比較すると, 鈴木(2003)は14種の説明変数を用いて解析したが, 回帰決定係数 R^2 は男性0.583, 女性0.479と低い。高・柿(2006)は27種の変数を用いたが, 得られた R^2 は男性0.358, 女性0.472と低い。多門ら(2011)は10変数を用いたが, R^2 は男性0.476, 女性0.324と低い。このように, 先行研究では説明変数の分野が限定

表3 SVMによる回帰結果

	平均寿命男	平均寿命女	健康寿命男	健康寿命女
決定要因の数	17	20	13	15
平均二乗誤差 (RMSE)	0.179	0.139	0.358	0.425
回帰決定係数 (R^2)	0.919	0.897	0.729	0.692
寿命の最大と最小の差	3.60	1.84	2.79	2.95

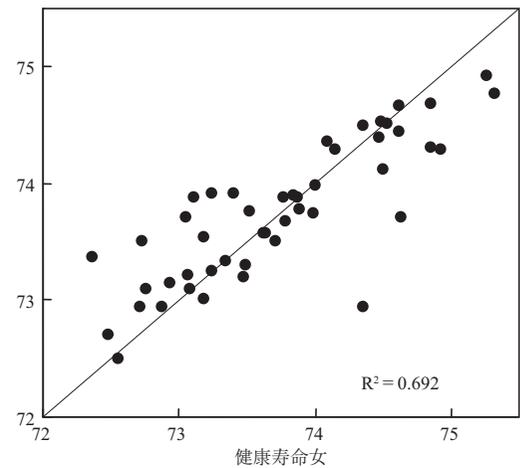
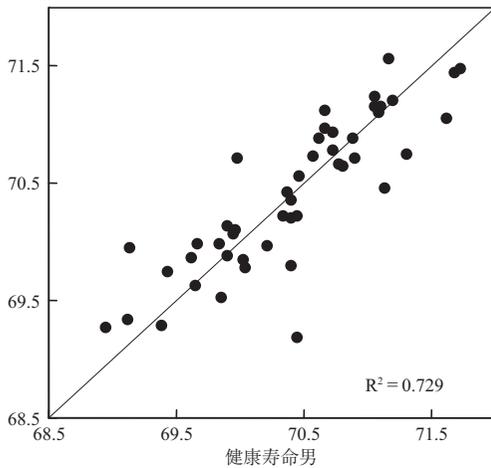
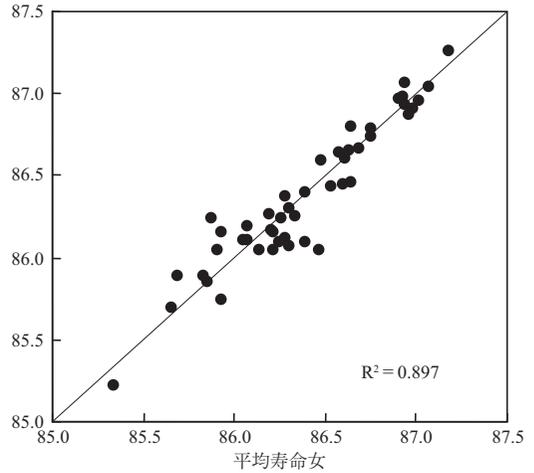
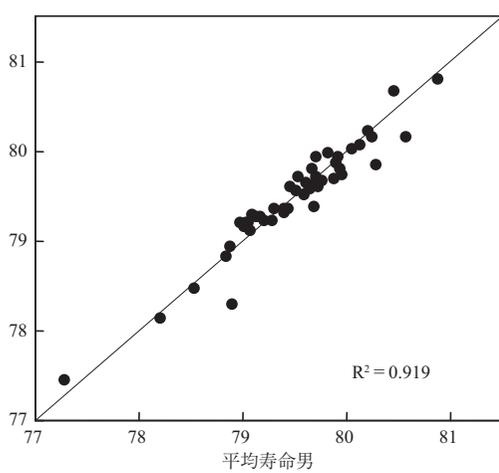


図1 寿命の実測値 (横軸) vs 予測値 (縦軸) の散布図

されているため、回帰決定係数はいずれも低く、寿命の都道府県格差を十分再現できていない。また、これらの論文ではデータ全体でOLSモデルを学習した際の誤差から決定係数を求めているが、本稿では交差検証法で予測性能を厳しく評価しているため、同じデータを用いても決定係数は先行

研究の方法より低くなる。この点を考慮すると、本稿の方が先行研究よりはるかに予測精度の高いモデルが作成できたといえる。

表4に示した決定要因の中には、寿命との相関係数がかなり低いにもかかわらず、SVMで得られた感度が高いものがあることが分かる⁶⁾。このこ

表4 平均寿命・健康寿命の決定要因, その感度, および各寿命との相関係数

No.	平均寿命男	感度	相関係数	No.	平均寿命女	感度	相関係数
1	生活習慣病	-0.386	-0.506	1	健康保険	0.393	-0.411
2	農民率	0.272	-0.179	2	公民館	0.378	0.200
3	健康保険	0.271	-0.436	3	災害罹災率*	-0.288	-0.070
4	図書館	0.245	0.496	4	保育所	0.284	0.405
5	災害罹災率*	-0.202	-0.030	5	高齢者学級	0.283	0.273
6	宗教	0.201	0.334	6	保健師	0.275	0.348
7	生活保護	-0.198	-0.348	7	持家率	0.263	-0.089
8	肉類	0.185	0.331	8	生活習慣病	-0.261	-0.183
9	喫煙	-0.148	-0.532	9	スポーツ	0.259	0.245
10	火災死者	-0.137	-0.592	10	図書館	0.258	0.225
11	降水	0.136	0.023	11	生活保護	-0.240	-0.155
12	公民館	0.132	0.231	12	高齢有業率	0.224	0.128
13	医師	0.129	0.031	13	自動車	-0.220	0.050
14	携帯電話	0.122	0.618	14	県民所得	0.211	0.042
15	インターネット	0.113	0.532	15	牛乳乳製品	-0.206	-0.148
16	公害	0.108	0.386	16	高血圧	-0.206	-0.295
17	保健師	0.106	-0.036	17	カロリー	-0.203	-0.318
				18	食塩	-0.196	-0.331
				19	農民率	0.193	-0.190
				20	降水	0.160	0.217

No.	健康寿命男	感度	相関係数	No.	健康寿命女	感度	相関係数
1	生活習慣病	-0.416	-0.383	1	教育費	0.260	-0.262
2	公害	0.333	0.378	2	ストレス	-0.219	-0.447
3	火災死者	-0.319	-0.444	3	スポーツ	0.205	0.216
4	ストレス	-0.308	-0.160	4	牛乳乳製品	-0.186	-0.281
5	牛乳乳製品	0.278	0.129	5	緑黄野菜	0.185	0.181
6	生活保護	-0.272	-0.618	6	がん受診率	0.169	0.361
7	農民率	0.215	0.098	7	健康保険	0.168	0.238
8	高齢有業率	0.172	0.454	8	自動車	0.164	0.500
9	保健師	0.151	-0.038	9	転出率	-0.161	-0.277
10	降水	0.143	0.234	10	緑茶	0.159	0.203
11	スポーツ	0.137	0.535	11	降水	0.151	0.412
12	乳児死亡率	-0.100	0.133	12	女性労働	0.142	0.507
13	肉類	0.084	0.067	13	仕事	-0.132	0.115
				14	生活保護	-0.127	-0.485
				15	図書館	0.115	-0.245

*災害罹災率は表2のように2007-2010年の自然災害の罹災者数の合計（人口当たり）であり，東日本大震災の被害は含まれない。

とは，先行研究において，寿命との相関係数に基づいて説明変数を選択している論文が多いが，この方法では寿命の決定要因を見逃す可能性があること，また，寿命との相関の高さに基づいて要因の影響度を分析している論文が多いが，このような議論の結果には疑問があることを示唆する。それに対して，本稿では多分野の多数の説明変数を一括して解析し，その中から感度分析法により決

定要因を探索したことで，寿命の決定要因に関して信頼性の高い結果を得ることができたと結論できる。

IV 考察

1) 4種の寿命に対する決定要因の相違

表4に示した決定要因は37種になるが，まず，寿

命4種における決定要因の種類の違いを概観する。37要因の中で、4種の寿命すべてに寄与するものは2要因のみ、寿命3種に寄与するものは7要因、寿命2種に寄与するものは8要因あるが、寿命1種のみ寄与する要因が20要因もある。さらに、寿命2種に寄与する8要因の内では、男性寿命にのみ寄与し、女性寿命に寄与しないものが3要因、逆に女性寿命にのみ寄与し、男性寿命に寄与しないものが1要因、さらに男女の平均寿命のみに寄与し、健康寿命に寄与しないものが2要因、逆に男女の健康寿命にのみ寄与し、平均寿命に寄与しないものが1要因ある。

このことから、本稿で得られた決定要因は4種の寿命に対して共通性が低く、特異的な要因が多いと言える。さらに、各決定要因の寿命への寄与の符号に注目すると、ほとんどの要因は寿命・性別による符号の違いは見られない。しかし、牛乳製品のみは、男性の健康寿命の符号は正であるが、女性の平均寿命と健康寿命の符号は負で短命に寄与するという、きわめて特異的な要因であると言える。この点については次で考察する。

2) 重要決定要因

決定要因37種のすべてについて言及する紙面はないので、その中で寿命4種への寄与が高い重要要因について、その寄与の高さを考察する。

①生活習慣病

生活習慣病患者率は、60種の説明変数の中で最も影響が大きい要因であり、特に男性の平均寿命と健康寿命に対する感度の高さは全要因の中で突出している。この結果は、生活習慣病は現在の日本人の3大死因（がん、心疾患、脳血管疾患）を惹き起し、また、長寿県では生活習慣病の患者が低い（仲都留・大西2008、祖父江2009、綿引・畑2009、竹内・關2013）こととよく対応している。しかし、男性の平均寿命と健康寿命に対しては感度1位であるが、女性の平均寿命に対しては感度8位で、男性より低く、女性の健康寿命に対しては21位で決定要因にならず、男女で大きな違いがあることは興味深い。この点については、生活習慣

病の疾患別・都道府県別・性別の患者数のデータを用いた解析を行うことにより、男女の感度の違いについて解明の可能性がある。

②健康保険・スポーツ・保健師・健康診断・自動車

長野県では保健師・保健指導員による健診受診や肥満防止等に関する活動が長寿化に大きく貢献したといわれているが、本稿の結果では、健康保険加入率は男女の平均寿命および女性の健康寿命に対する感度が生活習慣病に次いで高い。また、スポーツが男女の健康寿命と女性の平均寿命の決定要因となり、さらに、保健師数とがん受診率の感度も高いことから、健康意識・保健活動が日本人の長寿化にとって重要な要因であるといえる。また、沖縄県の短命化には自動車利用による肥満率上昇も挙げられているが、自動車も女性の短命化への影響があることを示している。

③生活保護

37種の決定要因の中で、男女の平均寿命、健康寿命のすべてに対して決定要因となっているのは、生活保護世帯率と年間降水量の2要因のみであるが、生活保護率は感度ははるかに高く、日本人の寿命の重要な決定要因である。長野県は生活保護率が全国で4番目に低く、逆に、青森県は全国4番目に高く、大阪府は全国で最も高い。この感度の符号が負であることは、生活保護世帯の多い地域は病人が多く、短命の人が多（角南1989）ことを示しており、社会福祉政策に示唆を与えている。

④ストレス

ストレスは男女の健康寿命への感度が高く、その符号が負であることから、健康寿命に大きなマイナス要因であるといえる。ストレスに対する男女の抵抗力については、本来、生物的に男性は女性より免疫力が低く、がん等に罹りやすいため、寿命が短いといわれている。しかし、ストレスが男女とも短命化に寄与するという本稿の結果は、女性の社会進出が進み、職場で様々な精神的スト

レスにさらされ、寿命の性差が縮まっている現代社会を象徴している。平均寿命・健康寿命に対するストレスの影響度を分析した先行研究は見当たらない。

⑤牛乳乳製品

2005年に農水省と厚労省が公表した「食事バランスガイド」において、生活習慣病を抑えるために、牛乳を1日1本程度摂取することが推奨されている(津志田2006)。また、戦後の日本人の体位が向上し、寿命が飛躍的に延びた理由の一つに、肉や牛乳など動物性タンパク質の摂取量の増加があると考えられている。そこで、肉類と牛乳乳製品を説明変数に取り上げた。肉類は男性の平均寿命と健康寿命にはそれぞれ正の寄与が認められたが、女性の2種類の寿命についてはほとんど寄与が認められなかった。一方、牛乳乳製品は上記のように男性の健康寿命には正の寄与になったが、女性の健康寿命には逆に負の寄与がある結果になった。男性の場合、牛乳が男性の死亡原因の上位を占める胃がん、脳血管疾患、循環器疾患の抑制に有効であるという医学論文がある(角南1989)。また、男性の高齢者の生存率の向上に牛乳が有効であるという論文もある(Shibata et al. 1992)。その他、牛乳が男性の長寿化に有効であるとする報告は多い。

一方、女性の寿命と牛乳摂取量との関係はまだよく分かっていないが、スウェーデン人の中高年男女について、牛乳の摂取量が多くなるほど死亡率が上がるという論文がごく最近、発表された(Michaëlssonら2014)。この原因として、牛乳に多く含まれるD-ガラクトースが老化を促進させ、寿命を縮める効果が動物実験で認められているとしている。本稿の結果は彼らの論文を支持するようにも思えるが、日本人の牛乳摂取量は欧米人の1/3程度であり、牛乳のタンパク質やD-ガラクトース以外の成分の寄与の検討も含め、牛乳の寿命に及ぼす影響については今後の多くの検証を待たねばならない。

⑥農民率・高齢有業率

長野県が長寿県である理由として農民の比率が多い(農民率：全国2位)ことが挙げられている(祖父江2009, 竹尾ら2012, 白澤2013)。同県は山間地が多いため、傾斜地で作業する農民は足腰が鍛えられ、健康人が多いといわれている。また、長野県のPPK(ぴんぴんころり)運動に象徴されるように、野菜を自給する小規模農家で働く高齢者が非常に多く、高齢者有業率も全国1位である(水野・青山1998, 七田2010)。そのため、男女の平均寿命および男性の健康寿命への感度が高く、決定要因となったのではないかと推測される。

⑦図書館・公民館

図書館・公民館等の充実も長野県の長寿化に寄与があるとされている(白澤2013)。事実、長野県の公民館数は、2位の新潟県の2倍以上の1位であり、公民館での減塩運動や交流活動が長寿に貢献したとされている(信州大学2006, 七田2010)。逆に、寿命最下位の青森県は図書館や公民館を利用する交流活動が少ないことが短寿命の1因とされている(青森県2002, 竹森2003)。したがって、これら文化・交流施設の充実が日本人の長寿に有効であることが本稿の解析で裏付ける結果となった。

⑧公害

公害苦情件数は男性寿命の決定要因となり、その感度の符号が正であることは、苦情件数の多い県ほど男性寿命が長いことを示している。しかし、公害苦情件数が女性の平均寿命と負相関があるという報告(稲葉2009)の方が理解しやすい。確かに、公害苦情件数の総数で上位の大都市圏では平均寿命が短い、本稿で用いた人口当たりの苦情件数では、男性の平均寿命第1位の長野県が第3位、男性の健康寿命第1位の愛知県が第4位であり、本稿の結果を裏付ける。また、公害に対する苦情が多く健康意識の高い都道府県、あるいは長野県のように教育水準が高く、議論好きな都道府県は長命の人が多いという解釈もできる。

3) その他の特記要因

これまで寿命に大きな影響を与えるとされ、多くの先行研究で検証されてきたが、本稿の結果では感度が低く、決定要因にならなかった幾つかの要因について考察する。

①喫煙・飲酒

長野県では保健補導員による喫煙・飲酒等に関する活動が長寿化に大きく貢献したといわれており、都道府県の平均寿命に対する喫煙率や飲酒率の影響を分析した先行研究は多い(角南1989, 旭ら2001a, 2001b, 2003, 鈴木2003, 高・梯2006, 京都大学2007, 多門2011)。しかし、説明変数が限定的なため、得られた結果が著者により異なり、信頼性に疑問がある。

長野県では飲酒率は低くないが、喫煙率は低いとされている(祖父江2009)。しかし、本稿のデータで長野県は、飲酒率が全国で低い方から29位、喫煙率は男性が20位、女性が17位であり、両要因とも長寿化に多少の影響はあるかもしれないが、決定的な影響度ではない。したがって、予想に反して、飲酒率はどの寿命に対しても決定要因とならず、喫煙率も男性の平均寿命についてのみ決定要因となったという本稿の結果が理解できよう。

②食塩・野菜・緑茶

長野県の長寿化には減塩が最も大きく貢献したといわれているが、塩分摂取量は女性の平均寿命に18位で決定要因となったものの、その他の寿命に対する感度はいずれも低く、決定要因に入っていない。その理由としては、寿命が短い東北地方の県は未だ塩分摂取量が多い(男性の摂取量1位は青森県、女性は福島県、2位は男女とも山形県)が、長野県は減塩に成功したといっても、塩分摂取量は未だ男性が4位、女性6位であり、全国的規模では塩分摂取量は長寿化にそれほど大きな影響を与えていないとみられる。

緑黄色野菜の摂取も長野県の長寿化に貢献したとされて、それを実証する論文もある(角南1989, Shimazuら2014)。長野県は野菜王国といわれるほど生産量が多いが、男性の野菜摂取量は全国で

24位、女性は16位であり、それほど多い方ではなく、むしろ短命の秋田県や岩手県の方が摂取量は多い。したがって、野菜摂取量が女性の健康寿命のみ決定要因に入り、その他の寿命に対しては決定要因にならなかったという本稿の結果は納得できよう。

緑茶の消費量については健康寿命第一位の静岡県の高齢化に寄与しているという説から説明変数に取り上げたが、女性の健康寿命にのみ寄与するという結果になった。確かに、女性の健康寿命1位の静岡県は緑茶の消費量が全国で1位であるが、男性の健康寿命1位の愛知県は消費量18位で、それほど多くない。緑茶の消費量2位は奈良県、3位は京都府、最下位は沖縄県であり、全体的には東日本で多く、西日本で少ない傾向があるが、寿命との関連はつかみにくい。

③降水

日本人の平均寿命は一般に南国に比べて北国の都道府県が低い傾向にあるため、平均気温や積雪量等の地理的要因も説明変数に取り上げた。しかし、平均気温、積雪量は決定要因にならず、降水量のみが感度はそれほど高くないものの、生活保護と同じように4種の寿命に対して決定要因となった。この結果は、日本の死亡構造が脳心血管疾患型であり、国土の降水量等が平均寿命の決定要因になるとする解釈(渡辺・加納1983)、および高齢者が屋外で活動する際、降水量等の気候条件の変動が平均寿命に影響する可能性があるという解釈(北島・太田2004)で説明できる。平均気温や積雪量が決定要因にならなかった原因としては、最長寿の長野県と最長寿の青森県でこれらの数値があまり変わらないことが考えられる。

④県民所得・ジニ係数

経済指標と平均寿命に関しては、格差が大きいほど寿命が短いという相対所得仮説(Wilkinson 1992)があり、多くの先行研究で検証されている(鈴木2003, 水島ら2004, 近藤2005, 2010, 川上ら2006, 京都大学2007, 多門ら2011)が、結果が一致していない。その原因はやはり説明変数の範囲

が限定的であるためであり、県民所得とジニ係数のどちらも決定要因とならなかったという本稿の結果の方が信頼できると思われる。したがって、相対所得仮説を支持する結果は得られなかったが、この仮説は世界中の多数の国について検証されている仮説であり、1国内の地域では成立しない可能性があるからであろう。

V 結論

我が国の47都道府県別および男女別の平均寿命および健康寿命のデータを目的変数とし、60種の指標を説明変数として用い、サポートベクターマシン(SVM)により解析した。人口、環境、栄養、健康、経済、文化の多分野の説明変数の中から感度分析法により決定要因を探索し、以下の結果を得た。

- ①13~20種の指標で都道府県別および男女別の平均寿命および健康寿命を平均二乗誤差(RMSE) 0.139~0.425歳、回帰決定係数(R^2) 0.692~0.919という、先行研究より高い精度で再現するモデルを構築できた。
- ②得られた決定要因の種類と影響度は、平均寿命と健康寿命、男性と女性の違いにより異なるが、総合的には生活習慣病等の健康関連要因が寿命に大きな影響を与えることを明らかにした。
- ③生活習慣病等の既発見の要因の他に、ストレス等の新たな要因が寿命に大きな影響を与えることを見出した。
- ④非線形回帰分析手法SVMを用い、多数の説明変数の中から感度分析法により寿命の決定要因を探索する手法の有効性を実証した。

今後の課題としては、第1はデータの拡充による解析結果の信頼性の検証がある。本稿では都道府県ごとの寿命の平均値を利用したため、データ数は少なく⁹⁾、求めた決定要因は都道府県間の寿命の差を説明するものであって、個人の長寿化に有効な要因を説明するものではない¹⁰⁾。日本人の健康長寿化政策に反映させる結果を得るためには、個人単位のミクロな各種データを利用した統

計解析も不可欠であろう。また、数年間の時系列データを利用する方法も有効であり、このような多様なデータを用いた多角的な解析を行い、平均寿命・健康寿命の決定要因に関してより信頼性の高い結果を導くことにより、健康長寿化政策に生かすことが可能になる。筆者らはこれら多面的なデータの解析を今後計画している。

第2は寿命に関する因果構造の解析の必要性がある。健康の社会的決定要因は多数が相互に関連しあっており、ミクロ(生物個体)からメソ(個人・家族)、さらにマクロ(国家・世界)まで複雑な階層構造を形成するとされている(近藤2005, 川上ら2006, 近藤2010)。この問題に対するアプローチとして、共分散構造分析手法を用いた因果構造モデルの解析がある(豊田ら1992, 星2013, Kodamaら2013)。しかし、現時点では共分散構造解析は線形モデルに限られており、SVMのような非線形回帰手法の導入が待たれる。このような因果構造を考慮した寿命の決定要因の分析は本稿の展開として重要なテーマであり、今後の課題として検討することを計画している。

(平成26年7月投稿受理)

(平成27年3月採用決定)

注

- 1) <http://www.mhlw.go.jp/toukei/saikin/hw/life/tdfk10/> (2013年12月5日最終確認)。
- 2) <http://toukei.umin.jp/kenkoujyumyou/> (2013年12月5日最終確認)。
- 3) 具体的には相関係数の絶対値が0.1以上のもの。
- 4) SVMでは説明変数の数がデータ数より多い場合や、説明変数間の相関が高い場合でも解析可能である。これはカーネル回帰という手法の採用によるものであるが、詳細は赤穂(2008)を参照されたい。
- 5) SVMの原理の詳しい説明や用語・記号の意味については大北(2005)、小野田(2007)、阿部(2011)を参照してほしい。
- 6) 感度分析法で算出される感度は、他の変数を固定し、当該変数のみ変動させたときの目的変数の変動から算出されるため、当該変数の純粋な影響度である。これに対し、目的変数との相関係数には他の説明変数の影響が含まれている(間接効果)。
- 7) 例えば、病院は感度の符号が負になるため、不

- 適当と判断して決定要因探索の早い段階で削除した。病院の符号が負になる理由は短命の東京都等の大都市県と比べて長野県等の長寿県では病院数が少ないためである。
- 8) 図1に示すように実測値と予測値のばらつきは平均寿命では非常に小さいが、健康寿命ではやや大きい。この原因としては説明変数の不足の可能性がある。
- 9) 本稿では都道府県別のデータを用いたため、交差検証法の学習セットは42ないし43データ、予測セットは4ないし5データとなり、十分な大きさとはいえない。しかし、個人別の個票データと比較すると、各都道府県単位での平均値の方が信頼性は高いと思われる。
- 10) 個人データと都道府県別データとでは結果が異なるという「生態学的誤謬 (Ecological Fallacy)」の問題があり、より正確な結果を得るためには個人データを用いるべきという指摘がある (森1987)。
- 参考文献**
- 青森県企画振興部統計情報課 (2002) 「青森県の男性の平均寿命について～社会生活統計指標 (都道府県の指標2002) からみて～」青森県統計活用レポート。
- 赤穂昭太郎 (2008) 『カーネル多変量解析 非線形データ解析の新しい展開』岩波書店。
- 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 (2001a) 「都道府県別観察による喫煙率と疾患別死亡率の関連」『厚生学の指標』48 (10) : 11-15。
- 旭 伸一, 多治見守泰, 大木いずみ, 尾島俊之, 中村好一, 岡山 明, 松村康弘, 柳川 洋 (2001b) 「都道府県別にみた飲酒率と疾患別年齢調整死亡率の相関」『厚生学の指標』48 (15) : 10-17。
- 旭 伸一, 渡邊 至, 多治見守泰, 大木いずみ, 尾島俊之, 中村好一, 小栗重統, 岡山 明, 松村康弘, 柳川 洋 (2003) 「都道府県別喫煙率, 飲酒率と疾患別死亡率の関係」『厚生学の指標』50 (1) : 1-6。
- 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。
- 荒記俊一, 北村文彦, 金 会慶 (2005) 「日本人の出生, 死亡, 寿命の疫学: 第2次世界大戦後の社会, 経済および人口要因が及ぼした影響に関する研究」『生体の科学』56 (2) : 150-156。
- 稲葉 裕 (2009) 「健康の地域格差及び性差と社会経済要因, 環境要因, 生活要因の関連」『科学研究費補助金研究成果報告書』課題番号19500613。
- 大北 剛 (訳) (2005) 『サポートベクターマシン入門』共立出版。
- 大蔵 暢 (2004) 「日本はなぜ世界一の長寿国か? 社会疫学の視点から」『総合臨牀』53 (8) : 2228-2231。
- 小野田 崇 (2007) 『サポートベクターマシン』オーム社。
- 家計調査, <http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001117248> (2013年12月5日最終確認)。
- 嘉手川繁三, 米盛徳一 (2003) 「長寿・死亡に関する社会的・経済的要因の検討」『沖縄国際大学総合学術研究紀要』6 (1) : 49-92。
- 加藤昌弘, 川戸美由紀, 橋本修二, 林 正幸, 中村好一 (2007) 「保健医療福祉統計に基づく高齢者の平均自立期間の推移」『厚生学の指標』54 (7) : 41-46。
- 川上憲人, 小林廉毅, 橋本英樹 (2006) 『社会格差と健康—社会疫学からのアプローチ』東京大学出版会。
- 北島晴美, 太田節子 (2004) 「都道府県別平均寿命の分布の変遷と気候の影響」『信州大学山地環境教育研究センター研究報告』3 : 53-75。
- 京都大学 (2007) 『健康と経済社会的属性との関係に関する調査研究報告書』<http://www.esri.go.jp/jp/prj/hou/hou029/hou29.pdf> (2013年12月5日最終確認)。
- 栗盛須雅子, 福田吉治, 八幡裕一郎 (2008) 「介護保険統計を用いた高齢者健康指標の提案と指標の関連要因」『老年社会科学』30 (3) : 383-392。
- 高 俊珂, 梯 正之 (2006) 「都道府県別の平均寿命と社会・経済指標および栄養指標との関連性」『広島大学保健学ジャーナル』5 (2) : 62-69。
- 国民健康栄養調査, http://www.e-stat.go.jp/SG1/estat/GL08020103.do?_toGL08020103_&listID=000001118468&requestSender=estat (2013年12月5日最終確認)。
- 国民生活基礎調査, http://www.e-stat.go.jp/SG1/estat/GL08020103.do?_toGL08020103_&listID=000001045678&requestSender=dsearch (2013年12月5日最終確認)。
- 近藤克則 (2004) 「なぜ社会経済的因子が健康に影響するのか」『公衆衛生』68 (4) : 306-310。
- 近藤克則 (2005) 『健康格差社会—何が心と健康を蝕むのか』医学書院。
- 近藤克則 (2010) 「幸福・健康の社会的決定要因—社会疫学の視点から」『科学』80 (3) : 290-294。
- 近藤克則, 芦田登代, 平井 寛, 三澤仁平, 鈴木佳代 (2012) 「高齢者における所得・教育年数別の死亡・要介護認定率とその性差—AGESプロジェクト縦断研究」『医療と社会』22 (1) : 19-30。
- 坂井博通 (1986) 「日本人の60歳時平均余命と社会経済的要因の関連に関する一考察」『人口問題研究』(180) : 46-51。
- 七田恵子 (2010) 「長野県高齢者の健康に関する指標の検討」『佐久大学看護研究雑誌』2 (1) : 51-58。

- 白澤卓二 (2013) 『長寿県長野の秘密』しなのき書房。
- 信州大学人文学部社会学研究室 (2006) 『長野県の郷土と文化—第3回調査報告』八十二文化財団。
- 杉澤秀博 (2012) 「健康の社会的決定要因としての社会関係：概念と研究の到達点の整理」『季刊・社会保障研究』48 (3) : 252-265。
- 鈴木健二 (2003) 「各種社会指標と都道府県別生命表の関係」『厚生指標』50 (5) : 30-35。
- 角南重夫 (1989) 「最近における我が国の平均寿命と医療および保健指標、食料等との関係」『民族衛生』55 (3) : 144-149。
- 祖父江逸郎 (2009) 『長寿を科学する』岩波書店。
- 竹内光、關雅夫 (2013) 「平成22年都道府県別生命表における平均寿命の地域差分析」『厚生指標』60 (16) : 32-39。
- 竹尾恵子、七田恵子、桶田真吾 (2012) 「佐久市における平均寿命について」『佐久大学看護研究雑誌』4 (1) : 3-14。
- 竹森幸一 (2003) 「生命表による青森県の死亡構造の解析」『青森保健大雑誌』5 (1) : 7-15。
- 竹森幸一、三上聖治、工藤奈織美 (2005) 「市町村別平均寿命の全国順位の変化からみた長野県と沖縄県の平均寿命の解析」『厚生指標』52 (10) : 36-45。
- 竹森幸一 (2007) 「青森県および長野県の市町村別たばこ喫渡本数と主要死因別標準化死亡比との関連」『厚生指標』54 (3) : 21-28。
- 田辺和俊、栗田多喜夫、西田健次、鈴木孝弘 (2013a) 「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』23 (1) : 70-91。
- 田辺和俊、鈴木孝弘 (2013b) 「サポートベクターマシンを用いたエコロジカル・フットプリント値の決定要因の分析」『日本エネルギー学会誌』92 (12) : 1205-1211。
- 田辺和俊、鈴木孝弘 (2014a) 「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』(188) : 44-67。
- 田辺和俊、鈴木孝弘 (2014b) 「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」『厚生指標』61 (13) : 23-30。
- 田辺和俊、鈴木孝弘 (2014c) 「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』(189) : 57-66。
- 多門隆子、黄裕美、吉田幸恵、小川由紀子、今木雅英 (2011) 「地域行政基礎データを用いた健康格差に関する研究」*Journal of Life Science Research*, 9 : 13-21。
- 地域別統計データベース, <http://www.e-stat.go.jp/SG1/chiiki/CommunityProfileTopDispatchAction.do?code=2> (2013年12月5日最終確認)。
- 津志田藤二郎 (2006) 「フードバランス」『日本食品科学工学会誌』53 (8) : 448-449。
- 豊田秀樹、前田忠彦、柳井晴夫 (1992) 『原因を探る統計学 共分散構造分析入門』講談社。
- 仲都留隆、大西雄基 (2008) 「都道府県別生命表による平均寿命の地域差分析」『厚生指標』55 (5) : 44-53。
- 橋本英樹 (2012) 「健康格差の実証研究—方法論的課題と展望—」『医療と社会』22 (1) : 5-17。
- 福田吉治、今井博久 (2007) 「日本における「健康格差」研究の現状」*Journal of National Institute of Public Health*, 56 (2) : 56-62。
- 伏見恵文 (2003) 「方法論からみた地域別生命表と健康寿命」『厚生指標』50 (5) : 47-49。
- 星且二 (2013) 「共分散構造分析を活用する因果構造分析事例」『日本健康教育学会誌』21: S82-83。
- 堀内四郎 (2010) 「日本人の寿命伸長：要因と展望」『人口問題研究』66 (3) : 40-49。
- 三上聖治、竹森幸一、工藤奈織美 (2005) 「市区町村別平均寿命の変化からみた長野県と沖縄県の平均寿命の解析」『厚生指標』52 (10) : 36-45。
- 水島豊、水野宏、浦田幸朋、鈴木克彦 (2004) 「青森県における社会生活指標と平均寿命の現状」*Gerontology*, 16 (4) : 73-77。
- 水野肇、青山英康編 (1998) 『PPKのすすめ：元気に生き抜き、病まずに死ぬ』紀伊國屋書店。
- 村田久 (2010) 「健康寿命と平均寿命—地域住民の健康寿命延伸を目指して—」*ESTRELA*, (194) : 38-44。
- 森幸雄 (1987) 「生態学的データ利用における誤謬の問題—ロビンソンの生態学的誤謬問題を中心として—」*Sociologica*, 12 (1) : 23-38。
- 森川千鶴子 (2001) 「健康寿命に影響を与える慢性疾患と大豆食品摂取状況の日本と韓国3地域における調査研究」『看護学統合研究』3 (2) : 18-27。
- 矢野栄二 (2000) 「平均寿命から健康寿命—その試みと国際比較」『日本医師会雑誌』123 (6) : 823-827。
- 渡辺毅、加納克己 (1983) 「重回帰分析による平均寿命に関する考察」『医学と生物学』106 (1) : 9-12。
- 綿引信義、畑栄一 (2009) 「男子平均寿命の国内格差について」『厚生指標』56 (1) : 9-14。
- Amundadottir, L., Kraft, P., Stolzenberg-Solomon, R., Fuchs, S. H., Petersen, G., Hassan, M. M., Holly, E. A., Hunter, D. J., Hutchinson, A., Jackson, A., Jacobs, V., Jenab, M., Kaaks, R., Klein, A. P., Kooperberg, C., Kurtz, C., Li, D., and Hoover, R. N. (2009) “Genome-wide Association Study Identifies ABO Blood Group Susceptibility Variants for Pancreatic Cancer,” *Nature Genetics*, 41 (9) : 986-90.
- Chang, C. C. and Lin, C. J. “LIBSVM-A Library for Support Vector Machines,” <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

- edu.tw/~cjin/ libsvm/ (2013年12月5日最終確認).
- Kagamimori, S., Gaina, A. and Nasermoaddeli, A. (2009) "Socioeconomic Status and Health in the Japanese Population," *Social Science & Medicine*, 68 : 2152-2160.
- Kodama, S. and Furuhashi, T. (2013) "Associations among Diet Quality, Socioeconomic Status and Healthy Life Expectancy – Structural Equation Modeling with Related Indicators of 47 Prefectures in Japan –," *Medical and Biology*, 157 (6) : 917-925.
- Michaëlsson, K., Wolk, A., Langenskiöld, S., Basu, S., Lemming, E. W., Melhus, H., and Byberg, L. (2014) "Milk Intake and Risk of Mortality and Fractures in Women and Men: Cohort Studies," *British Medical Journal*, 349, doi: <http://dx.doi.org/10.1136/bmj.g6015> (Published 28 October 2014).
- Shibata, H., Nagai, H., Haga, H., Yasumura, S., Suzuki, T., and Suyama, Y. (1992) "Nutrition for the Japanese elderly," *Nutrition and Health*, 8 (2/3) : 165-175.
- Shimazu, T., Wakai, K., Tamakoshi, A., Tsuji, I., Tanaka, K., Matsuo, K., Nagata, C., Mizoue, T., Inoue, M., Tsugane, S., and Sasazuki, S. (2014) "Association of Vegetable and Fruit Intake with Gastric Cancer Risk among Japanese: A Pooled Analysis of Four Cohort Studies," *Annals of Oncology*, doi: 10.1093/annonc/mdu115, (First published online: March 11, 2014).
- WHO (2004) "The World Health Report 2004," http://www.who.int/whr/2004/en/report04_en.pdf?ua=1 (2015年2月17日最終確認).
- Wilkinson, R. G. (1992) "Income Distribution and Life Expectancy," *British Medical Journal*, 304 (6829) : 165-168.

(たなべ・かずとし 東洋大学客員研究員)
(すずき・たかひろ 東洋大学教授)