

生理的年齡予測の重回帰分析(II)

吉川, 和利
Institute of Health Science Kyushu University

松本, 壽吉
Institute of Health Science Kyushu University

大柿, 哲朗
Institute of Health Science Kyushu University

<https://doi.org/10.15017/482>

出版情報 : 健康科学. 9, pp.147-157, 1987-03-28. Institute of Health Science, Kyushu University
バージョン :
権利関係 :

生理的年齡予測の重回帰分析 (II)

吉川 和利 松本 壽吉 大柿 哲朗

A Multiple Regression Model for Physiological Age (II)

—on the Case of Adult Males—

Kazutoshi KIKKAWA, Jukichi MATSUMOTO
and Tetsuro OGAKI

Summary

In this paper, a predictive equation for chronological age (AGE) from various physical performance test, anthropometric measurements, and internal medical indices. Fifty-five healthy adult males, aged 32 to 59 years old, volunteered as subjects in this survey.

Data for this study were collected during the course of another study of various indices on the degree of health and fitness and life-style. Information collected included the followings: vital capacity (VITAL), grip strength (GRIP), back strength (BACK), standing trunk flexion (FLEX), standing blinded balance (BALANCE), heart rates after stepping test by Margaria's method (HR1, HR2), heart rates at rest (HR), estimated % fat from skinfolds thickness (FAT), and systolic and diastolic blood pressure (SBP, DBP) at rest.

Forward stepwise regression analysis examined the reliability of the physiological, internal medical and anthropometric measurements as the indicator of age.

At the each step, the statistics of degree of fitness, that is, Akaike's AIC, Mallows' Cp, Schwarz criterion and R^* (adjusted for d.f. R) adding to multiple correlation coefficient R and SEE (standard error of estimates) were calculated.

The results obtained were as follows.

1) The specified features were not shown from descriptive statistics of dependent and independent variables. Multi-colinearity was not shown from independent variables correlation coefficient matrix (11×11).

2) In terms of following variables order, all variables were entered into equation; BACK, VITAL, SBP, GRIP, HR, DBP, HR2, HR1, FAT, BALANCE, FLEX. The multiple correlation coefficient R , and standard error of estimates (SEE) were 0.6317, 7.37 (age of years), respectively.

3) The equations obtained at each step, which included independent variables at m step, in general, were evaluated by several statistics such as criteria described above. According to these criteria, following described equation which included BACK, VITAL, SBP and GRIP was selected the best equation.

$$\hat{y} = 53.07 - 0.219 \times \text{BACK} - 0.00398 \times \text{VITAL} + 0.1573 \times \text{SBP} + 0.253 \times \text{GRIP}.$$

However, this equation was a problem that the regression coefficient of GRIP was

positive. Therefore, another forward stepwise regression analysis was conducted after GRIP were deleted. Consequently, following equation was the better one and its multiple correlation was 0.612.

$$\hat{y}=57.93-.171\times\text{BACK}-.00327\times\text{VITAL}+.1468\times\text{SBP}$$

4) When VO₂ (estimated maximal aerobic power by Margaria's method) in replace of HR1 and HR2 was added as the 10th independent variable, at the 1st BACK was entered. VO₂ at the 2nd step, and VITAL at the 3rd step were respectively entered and the equation was described as follows:

$$\hat{y}=89.0-0.161\times\text{BACK}-0.042\times\text{VO}_2-0.00278\times\text{VITAL} \quad (R=.6408)$$

However, the calculated VO₂ was unequal to other variables as for the age changes because this variables were essentially designed to correlate to chronological age.

5) The specific problems of these equations was not obtained from residuals and the co-ordinates of *e*is (standardized residuals) and predicted age, so the high reliability was verified and the validity of this equation was verified by the results of preceding studies.

(Journal of Health Science, Kyushu University, 9: 147-157, 1987)

緒 言

われわれは他人の顔ぼう、姿勢、運動動作などを手懸りに、「あの人は見かけでは20代にしか見えない」とか、「とても50代とは思えない身のこなしだ」とかの様式をとることがある。すなわち、年齢に相応の平均的な容姿や動作を基準にして、年齢そのものに立ち帰るといふ思考様式が存在している。このような観点は論理学にいう逆問題の解を求める手続と符合し、加齢の結果としての容姿や動作あるいは測定値から年齢に逆戻りする過程が存在している。

因果関係の「結果から原因の探索を行う」逆問題もしくは不適切問題の近似解を得ようとするこうした思考実験様式においては、真の解や誤差についての事前情報を利用して解の範囲を限定することが必要であり、有限なデータを対象に関した不完全な知識から解の構造を帰納するプロセスを繰り返すことになる(田辺, 1976)⁴²⁾。

ここには統計的モデル(族)を積極的に導入せねばならない意義が本質的に存在する。その基本的な方略は広義の重回帰モデルに帰着することが多いが、これまでの年齢予測研究では計量データを扱うことが多く、カテゴリーカル回帰分析(一般にいわれる数量化理論I類)よりも通常の重回帰分析が広く応用されている。たとえば法医学・人類学の領域では頭蓋縫合や歯の咬耗度を独立変数として年齢の予測式の作成を意図した研究¹¹⁾¹²⁾²¹⁾³⁹⁾⁴¹⁾があり、内科学的な循環器系や体力的な変数により説明を求めようとする研究¹⁰⁾¹¹⁾¹⁸⁾²⁵⁾²⁶⁾⁴⁰⁾⁴⁴⁾⁴⁵⁾などでもその多くが重回帰分析を主たる手続として用いている。筆者のうちの一人、吉川(1985)は

女子の大学教職員や大都市に在住する主婦を対象に簡便な体力テストの結果をもとに年齢予測の重回帰式を作成している¹⁸⁾、中村(1975)、Nakamura(1985)も多変数からなる同様の回帰式の提示を行っている²⁵⁾。

本研究ではいわゆる暦年齢に生理的機能を対応させ、後者の結果から暦年齢を予測するための重回帰式を作成することが主たる目的であるが、「母教節約的」に比較的少数の変数からの予測式を定立すること、予測式の誤差への反応が過剰となりすぎないように配慮することを課題の一側面とする。

すなわち、予測式の精度は重回帰分析での重相関係数(multiple correlation coefficient, *R*)で評価できるものの、その値は独立変数個数とともに増昇していくので好ましいものではないことは容易に理解できる。ここでは自由度とパラメータ数を加味し安定性と偏りを妥協させた最良な回帰式を選択するために、最近の統計学の知見¹⁾²⁾³⁾⁴⁾⁶⁾⁸⁾⁹⁾¹²⁾³⁴⁾³⁶⁾³⁸⁾を援用しながら予測式の定立をめざすものである。

研究 方 法

1. 対 象 者

大都市近郊に存在し、特に重要な疾病を有しない住民のうち、自治体からの健康診断のよびかけに応募した住民、男子55名である。その年齢は最小32歳から最大59歳、平均年齢45.2、標準偏差8.58歳となっている。

2 独 立 変 数

年齢にともなって線型な変化を有すると考えられる次のような身体計測値を独立変数とし、所定の手続によって測定を行った。

体組成の指標として上腕背部および肩甲骨下部の皮脂厚値を用い長嶺の方法により体密度を推定し, Keys & Brozcek の式によって体脂肪率 (コード名: FAT, 以下同様) を計算したものをとりあげた。

筋力・柔軟性の面からは握力 (GRIP) 背筋力 (BACK), 立位体前屈 (FLEX) をそれぞれ一般的な方法で測定した⁴³⁾。また運動調節機能として平衡性をとりあげ, 具体的には閉眼片脚立ち (BALANCE) を同様に測定した⁴³⁾。呼吸機能の指標として肺活量 (VITAL)⁴³⁾, また循環機能の指標として Margaria et

al.²³⁾ の方法で最大酸素摂取量を推定しようとした時の第1負荷時および第2負荷時の心拍数をそれぞれ測定 (HR1, HR2) し, 安静時の心拍数 (HR) とともに独立変数とした。収縮期・拡張期の安静臥位血圧値 (SBP, DBP) も同様に循環機能の指標として採択した。

結果ならびに考察

1. 基本統計量について

表-1には暦年齢 (AGE, 従属変数) および独立変数についての平均値, 分散, 分布の程度など基本統計量を示している。BALANCEの歪度・尖度を除くと特に問題とすべきデータ固有の問題はないが, 今後に行う重回帰分析はその適用が「このデータの得られた」範囲に限られるもの⁸⁾として考えておく必要がある。

2. 相関行列の検討

表-2は従属変数, 独立変数相互間の相関係数を示したものである。独立変数間の相関が高いと係数値がわずかのデータ変動によって影響されるという多重共線性 (Multi-colinearity)⁸⁾¹³⁾¹⁹⁾²⁰⁾²⁷⁾²⁸⁾³⁴⁾の問題がある。チャタジーとプライスによれば多重共線性の基準として (1) 行列の固有値が0.01以下の小さいものがあらわれる場合, (2) 固有値の逆数和が独立変数個数の5倍を越える場合となることなどがあげられている⁶⁾。この点が實際上, 問題となるのは回帰式を得た後においてであって仮に全変数を普く投入した時の

Table 1. Descriptive statistics.

(MALE) VARIABLE	MEAN	SD	SKEWNESS	KURTOSIS
AGE, y.o	45.2	8.58	0.531	-0.650
FAT, %	16.1	3.24	0.109	-0.538
GRIP, kg	45.4	7.06	0.262	-0.776
BACK, kg	116.7	20.56	-0.031	0.185
BALANCE, sec	34.9	29.6	1.531	1.618
FLEX, cm	3.53	8.03	-0.687	-0.429
VITAL, cc	3458.4	733.8	-0.040	-0.622
HR 1, beat	120.3	13.2	0.234	-0.996
HR 2, beat	156.9	15.8	-0.075	-0.586
SBP, mmhg	126.3	11.9	0.241	-0.666
DBP, mmhg	82.7	9.58	0.160	-0.073
HR, beat	63.6	7.56	-0.044	-0.427

Table 2. Correlation coefficients.

(MALE) VARIABLE	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
AGE	1.00											
FAT	112	1.00										
GRIP	-301	-186	1.00									
BACK	-498	008	626	1.00								
BALANCE	-208	-097	184	155	1.00							
FLEX	-028	-142	032	-013	060	1.00						
VITAL	-412	-345	439	260	354	114	1.00					
HR 1	-008	164	104	123	-123	-172	005	1.00				
HR 2	072	157	-043	034	-212	-278	-123	830	1.00			
SBP	269	158	-147	-074	-258	006	-125	296	342	1.00		
DBP	235	257	-076	-043	-203	112	-181	231	223	692	1.00	
HR	082	-082	135	-031	072	067	061	468	414	218	336	1.00

decimals omitted.

ことを考えると最小の固有値は0.1442, 固有値の逆数
和は23.14となり一応, 多重共線性の問題は考えなく
ともよいことになる。

3. 重回帰方程式の作成

実際の重回帰式の作成にあたっては11個の独立変数
を一個ずつ, 回帰式に投入していくステップワイズ方
式を採用した⁷⁾⁸⁾¹³⁾²⁷⁾²⁸⁾³⁴⁾。ただし, この方式でも独
立変数の寄与の程度が変化していく様態について検討
しながら分析するのが好ましいので次のような方略を
とることにした。ステップごとに, 未投入変数が仮に
投入されたと考え, その時の残差平方和に基づく F 値
を F ENTER 値とし, 同じくすでに投入された変数
が仮に, 除去された時の残差平方和に基づく F 値を F
 $REMOVE$ 値とした。したがって(1)この統計量の
うち $F REMOVE$ 値を手懸りに全変数によって成る
回帰式から寄与度の低い変数を順次除去していく変数
減少法, (2)この統計量のうち $F ENTER$ を手懸りに
寄与度の高い変数を順次投入していく変数増加法,
(3) $F ENTER$ 値, $F REMOVE$ 値を並行的に利用し,
一度投入された変数でも寄与の程度が一定以下
ならば方程式から除去していく変数増減法, (4)同
様に変数減増法が考えられる⁷⁾⁸⁾¹³⁾¹⁷⁾²⁸⁾³⁴⁾。ここでは
筆者なども実際に行ってきた方法であり, 他にも適用
例が多い³⁹⁾ことを鑑みて変数増減法を用いることにした。
ただし, F 値そのものの変動についてはいわゆる単調
な増加法や減少法が適切であることも考えられるので
状況に応じてこれらを利用することもあり得る。

資料は省くが, 単調な増加方式を行った場合のステ
ップごとの F 統計量の変化を概観するとまず BACK
が投入される。これは他の変数に比較して第IIステッ

プ以降でも圧倒的に高い F 値を有し続けている。VI-
TAL は第IIステップで投入され, 以下 III. SBP, IV.
GRIP, V. HR, VI. DBP, VII. HR2, VIII. HR1, IX.
FAT, X. BALANCE, XI. FLEX といった投入順
位になる。上位で投入された3変数の F 統計量は明ら
かに残り8変数のそれと異なる。

変数増減法では変数の採択と除去の基準が問題とな
り, F 値が小さいと多くの変数が投入され, F 値が大
きいと少数の変数しか取り込まれないことになる。基
本的にはパラメータ数の多少によって「推定や予測
という本来の目的の信頼度を低下させる」か, 逆に
「対象の構造を十分に反映せずに情報に偏りを生ず
る」³⁶⁾かという問題が存在する。

「適合度」と「信頼度」のトレード・オフを加味し
た基準として赤池弘次(1974)は Kullback-Leibler
情報量に依拠して次式 [1] によって定義される情報量
基準を導入している¹⁾²⁾³⁾⁴⁾³³⁾³⁴⁾。 $AIC = -2 \ln(\text{尤度})$
 $+ \text{パラメータ数} \cdots [1]$ 。赤池(1981)はまた, 変数増減
法による変数の選択は AIC 最小化 ($MAICE$, Mini-
mum AIC Estimates) を達成するものであることを
述べているし²⁾, 佐和(1979)は F 統計量と $MAICE$
ならびに C_p 基準間の符合の根拠について論じてい
る³⁴⁾。

以上の点や適用度などを総合的に勘案した結果, こ
こでの選択方式は AIC を中心に種々の統計量基準に
依拠しながら変数増減方式によって行うことにした。

表-3, 図-1には単調な増加方式を行った場合の
ステップごとの投入変数名や残差平方和および情報量
基準として一般によく知られている AIC (An Infor-
mation Criterion) や Mallows による C_p 基準⁶⁾⁹⁾

Table 3. Entered/removed variables and several criteria statistics of their equation.

	± VARIABLES	RSS	R^2	R^{*2}	C_p	AIC	Schwarz
I	+BACK	2989.6	0.248	0.234	0.828	444.159	444.167
II	+VITAL	2649.1	0.333	0.308	-3.303	439.509	439.52
III	+SBP	2487.4	0.374	0.337	-4.214	438.045	438.06
IV	+GRIP	2397.5	0.397	0.348	-3.833	438.02	438.04
V	+HR 1	2394.0	0.398	0.336	-1.897	439.94	439.96
VI	+HR	2391.0	0.398	0.323	0.050	441.87	441.90
VII	+DBP	2390.0	0.399	0.309	2.032	443.85	443.88
VIII	+FLEX	2388.9	0.399	0.294	4.012	445.82	445.86
IX	+HR 2	2388.7	0.399	0.279	6.009	447.82	447.85
X	+FAT	2388.5	0.399	0.262	8.005	449.81	449.85
XI	+BALANCE	2388.3	0.399	0.245	10.001	451.81	451.85

34), Schwarz の基準³⁸⁾, 重相関係数 (R), 自由度調整済みの重相関係数 (R^*) など基準統計量を示している。

このうち, R については変数の増加とともに増昇していくものであるが, R^* も含めて他の統計基準は自由度との間に一定の傾向はない。これまでの筆者ほかの類同した研究¹⁷⁾¹⁸⁾を概観すると R^* は変数の節約という点では甘い基準であり, C_p はその点では厳しいものと考えられる。ために AIC を折衷的に利用してきたというのではなく, 本来の目的上, 総合的に予測式が最良のものとなり得るか否かが最終的な判断のよりどころと考えなければならない。実際に得られた基準統計量の変化をみるとパラメータ数とあてはまりの良さを加味した時に採択される回帰式の変数個数は

AIC と R^* , Schwarz の基準で一致し, C_p 基準はより少数の変数から成る回帰式を最良と判断することになる。以上のように AIC を判断の基礎に置くと 4 変数から成る次式 [2] が予測のために好ましい回帰式として考えられることになる。

$$\hat{Y} = 53.07 - 0.219 \times \text{BACK} - .00398 \times \text{VITAL} + 0.1573 \times \text{SBP} + 0.253 \times \text{GRIP} \quad (R = .630) \dots\dots [2]$$

ところでこの式 (2) では係数の符号が問題となる。つまり, BACK, VITAL は一般的には青・壮年期の男子では加齢とともに低下していくこと³⁰⁾⁴³⁾, SBP は同様に増昇²²⁾³²⁾していくことが知られているのでここで得た符号に照合して問題はないが, GRIP の係数値が正の値を示すことになり一般的な知見に反するものである。この様な場合, 大西によると当該変数を除去

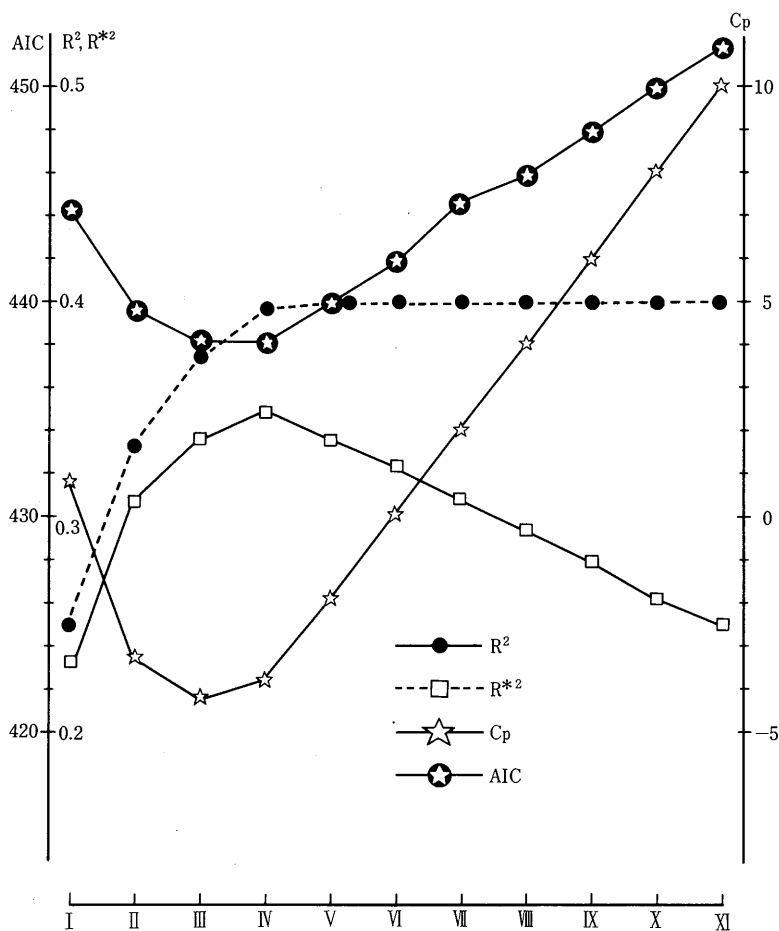


Fig. 1 The changes of criteria of goodness of fitting.

していくことも一つの方略と考えられている²⁹⁾。

GRIP を除去した10独立変数群について同様のステップワイズ選択方式で重回帰分析を行った結果、すでに述べたと同じ順序で変数の投入/除去が行われた。この結果、AIC, Schwarz の基準はこの場合、BACK, VITAL, SBP の3独立変数からなる以下の式 [3] を最良と評価することになった。

$$\hat{Y} = 57.93 - .171 \times \text{BACK} - .00327 \times \text{VITAL} \\ + .1468 \times \text{SBP} \quad (R = .612) \dots \dots [3]$$

以上から、好ましいと考えられるのは統計基準によるれば [2] 式、変数の係数値符号について考慮すると [3] 式を挙げることができる。ここに投入された変数群から多重共線性を検討するために相関係数行列から固有値を求めたところ4変数群の場合、 $\lambda_1 = 1.924$, $\lambda_2 = 0.9711$, $\lambda_3 = 0.7674$, $\lambda_4 = 0.3374$ ($\sum 1/\lambda_i = 5.8$) となり、また、3変数群の場合、 $\lambda_1 = 1.289$, $\lambda_2 = 0.964$, $\lambda_3 = 0.7465$ ($\sum 1/\lambda_i = 3.2$) となった。これらの値は先述のチャタジーとプライスの示唆する基準を満足することはなく、したがっていずれも多重共線性の存在は否定されることになった。以上のように、BACK [=筋力], VITAL [=呼吸機能], SBP [=循環機能] といった変数を含む重回帰式が、変数個数の面、データの対象範囲などの面から妥当なものとして考えられる。当初の独立変数は11個をとりあげたが、母数節約的にAICなどを適用すると3ないし4個で十分に説明が可能であることになった。のみならず、採択された変数は筋力、呼吸機能、循環機能をそれぞれ代表し得るものであり特に人間の活動能力を身体の健康に不可欠な要素とする最近の体力科学の一般の見解⁵⁾¹⁶⁾に照合しても興味深い。壮年・老年期には「機能によって加齢変化の起こる時期が多様であり」¹³⁾、「変化の勾配も小さく」¹⁰⁾ 度数分布にも偏りがありがちなこと³¹⁾⁴⁵⁾が知られている。したがって変数群も単一の領域に限定するだけのものでもなく、多領域的に当初は設定されるべきであろう。この点を考えると [3] 式は総合的な生理機能をもつ重回帰式を得ようとした初期の目標からして適切なものといえるであろう。

4. 残差の分析

回帰分析の信頼性は観測値と予測値との間にみられる差、すなわち残差についての検討を必要とする。残差について一般的に知られている見解は、佐和(1979)³⁴⁾、奥野ほか(1978)²⁷⁾、奥野ほか(1981)²⁸⁾、あるいは三宅(1978)²⁴⁾などに詳しく解説されている⁶⁾⁸⁾²⁰⁾。残差の分析方法としては、予測値と残差の二次元座標によって大まかな傾向を知ったり、残差をさらに標準化

($\mu = 0, \sigma = 1$) して観測値との二次元座標を分析する方法などがある。三宅(1978)によると場合によっては観測値とのものを対数・逆数などに変換したりする必要がある²⁴⁾。ここでは [3] 式として得られた重回帰式について以下の様な一般的な方法で残差分析を進めた。なお [3] 式での残差の標準差は6.79(歳)である。

残差分析の基本的な手続として予測値と標準化残差による二次元座標を作成する図-2のようになる。これをみると特定の変動傾向があるとはみなし難いことが概観できる。

さらに残差については基本統計量を求めたが、それらを予測値の基本統計とともに示すと表-4のようになる。このうちで、KURTOSIS, SKEWNESSは、それぞれ1次から3次あるいは4次までの積率によって算出したものをさし、SK, Q, SK, D,などは%ile値を手懸りにして計算した尖度、歪度である。このうち、残差のSKEWNESSは2.25と有意($P < 0.1$)に高い値を示し、分布の歪みが相当に大きいものであることを示唆した。

また標準化残差のうち2.0を越えたものは1名に過ぎなかった。総合して全体としての残差分析の結果は、やはり極限値のあたりに生じやすいということが指摘できる。こうした回帰効果への対処はまた回帰分析の宿命でもあるが、Nakamura(1985)²⁶⁾が考察している手続きなどは興味深い方法である。

残差分析の結果ではこのようないくつかの問題点を考えておかねばならないが、さらに重要なのは得られた重回帰式そのものの予測度の低さである。つまり [2] 式によるしる、 [3] 式で予測しようしるこれらが説明し得るのは全体の高々40%程度にしか過ぎない。したがって新しい変数の導入などが当面考えられる課題であろう。

5. 持久性指標の導入による重回帰式の作成

古川ほか(1976)を始めいわゆる生理学的指標を導入した同様の研究¹⁰⁾¹¹⁾¹⁸⁾²⁵⁾²⁶⁾⁴⁵⁾では動的な循環系の指標が一貫して用いられている。この背景には一般成人を対象とした場合、循環器系機能は筋力や呼吸器系を含む総合的な健康指標であること、最近における高血圧症や代謝障害などの疾病の増加がみられること⁵⁾¹⁶⁾などがある点を無視することはできない。これまでに本研究で展開した重回帰分析でも単純な変数として Margaria et al. の方法による2段階の踏台昇降テストで得られた心拍数の測定値 (HR1, HR2) を独立変数として用いている。これまでの結果をふまえ、ここでは Margaria et al. たちの方法によって、これら

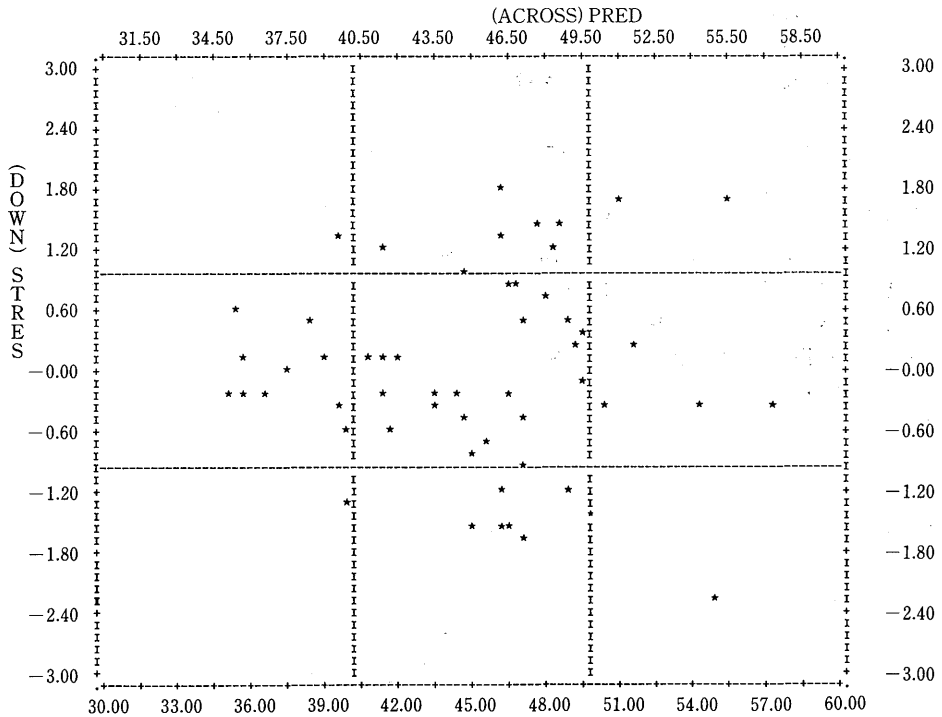


Fig. 2 Plotting of standardized residuals (down, STRES) and predicted age (across, PRED) from equation [3].

Table 4. Descriptive statistics of predicted age and residuals from equation [2].

	MEAN	SD	25%ILE	50%ILE	75%ILE	SK.QT	SK.DQ	KUQ	SKEW	KURT
PREDICT	45.2	5.25	41.5	46.3	48.6	-.352	-.299	.504	2.51	-.002
RESID	0.0	6.79	-4.0	-1.2	4.6	.349	.115	.430	2.25	-.111

Table 5. Entered/removed variables and several criteria statistics of their equation.

	± VARIABLES	RSS	R ²	R ^{*2}	C _p	AIC	Schwarz
I	+BACK	2989.6	0.248	0.234	5.990	444.159	444.167
II	+VO 2	2613.4	0.342	0.317	0.567	438.762	438.773
III	+VITAL	2417.4	0.392	0.356	-1.301	436.475	436.489
IV	+SBP	2342.3	0.411	0.364	-0.783	436.739	436.757
V	+GRIP	2269.1	0.429	0.371	-0.227	436.993	437.015
VI	+HR	2257.0	0.432	0.361	1.534	438.698	438.724
VII	+BALANCE	2245.8	0.435	0.351	3.313	440.425	440.454
VIII	+FLEX	2234.3	0.438	0.340	5.086	442.143	442.176
IX	+DBP	2232.2	0.438	0.326	7.045	444.091	444.128
X	+FAT	2230.1	0.439	0.312	9.004	446.039	446.079

HR1, HR2 から得た最大酸素摂取量 [$\dot{V}O_2$, ml/kg · min] を HR1, HR2 の代替変数 [VO2] として採用し、合計10個の独立変数から始まるステップワイズ重回帰分析を試みた。

表-5にはこれまで行ってきたと同様にステップワイズ重回帰分析で投入された変数名と、それらの重回帰式ごとの重相関係数 (R)、その他諸基準統計量を示している。これによると、AICのほか C_p , Schwarz の各統計量基準はいずれも第4ステップで最小値を示し、 R^* のみは第5ステップで最大値を示すことになる。これらの過程での投入変数は BACK, VO2, SBP, VITAL, GRIP と進んできている。第3ステップまでの変数群による重回帰式は、以下の式 [4] のようになる。

$$\hat{Y} = 89.0 - 0.161 \times \text{BACK} - 0.442 \times \text{VO2} - 0.00278 \times \text{VITAL} \quad (R = .6408) \dots \dots [4]$$

諸基準統計量のうち、 R^* 以外のものは第4ステップで最小値となり、したがってこの重回帰式が最良の方程式として選択されることになる。ここでの標準化変回帰係数は、BACK = -.384, VO2 = -.213, SBP = .198, VITAL = -.210 となっている。BACK, SBP,

VITAL については、すでに3の項で、その一般的な年齢変化については指摘した。新たに投入された VO2 については、都立大学の資料など⁵⁾¹⁶⁾³⁸⁾によるまでもなく、年齢によつての低下が汎く認められるものでもあり、ここでの標準化係数の符号なども妥当なものであろう。

この VO2 を含んだ重回帰式 [4] より、残差の計算を行ったところ、残差の標準偏差は6.60となった。また、これをもとに残差を標準化したものと、予測値とで二次元在標を構成したところ、図-3のようになった。この図には特定の変動パターンを認めることができず、この点からいえば得られた重回帰式の信頼性については確認できたことになろう。

[4] 式として採択された重回帰式に含まれる独立変数は 軀幹筋力 (BACK), 呼吸機能 (VITAL), 循環機能 (SBP), 有酸素的作業能力 (VO2) といった変数である。そして小筋の筋力 (GRIP) や柔軟性 (FLEX), 平衡性 (BALANCE) や体組成 (FAT) の貢献度は必ずしも大きいとはいえないことになる。生理的に加齢を考えるひとつの資料として $\dot{V}O_2 \text{ max}$, 肺活量, 収縮期血圧などの呼吸・循環機能や背筋力を代表とする

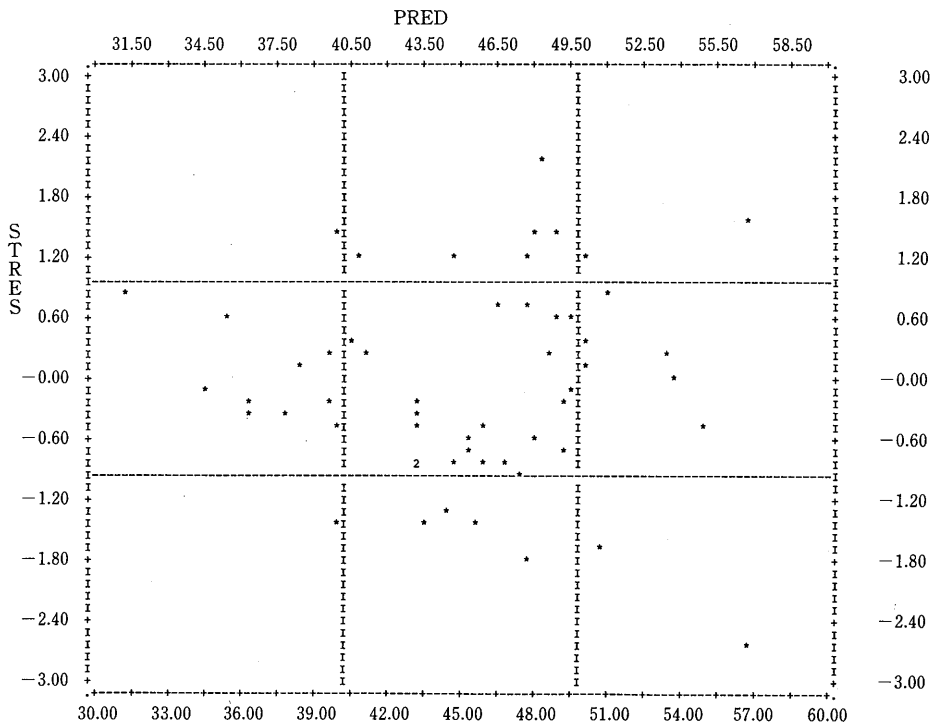


Fig. 3 Plotting of standardized residuals (down, STRES) and predicted age (across, PRED from equation [4]).

大筋の筋力」, 「有酸素的作業能力」などがより重要であると考えられるのではないだろうか。吉川 (1985) は主婦や教職員についての年齢予測重回帰分析を試み SBP, BALANCE, VITAL, FAT の4変数が有意に貢献し寄与率51%となったことを報告している¹⁷⁾。本研究で採択された変数に比較しモデルの性差は体脂肪率 (FAT), BALANCE の寄与に求めることができ、一方、血圧値や呼吸機能は共通して有意に寄与していると考えられる。

これまでの年齢予測に関する研究では単一の学問領域で得られた変数(群)を独立変数(群)として扱うことが多かった。無論、それはそれぞれの科学自体に課された要求にこたえるべき成果として重要である。一方では多くの変数の塊として「ヒト」という生物は存在し、また種々の機能の系(システム)³¹⁾としての役割についても考えないわけにはいかない。そのための一次的なアプローチとして本研究を位置付けることもできよう。

また決定係数は0.4106となり、全分散の約41%がこの重回帰式によって説明可能ということになる。ただし、ここで新しく導入した VO₂ については、最高心拍数 (MH) を年齢そのものから求め、その値を用いて Margaria et al. (1965)²³⁾ に基づいて計算されたものであり、すでに年齢と一次的な従属関係あるものとして得られている。したがって有意な変数として投入されるのはむしろ当然なのかも知れない。この点については、今後の課題となるものである。

VO₂ を含まない式 [2], [3] についても述べたが同様に寄与率が式 [4] の場合41%にしか過ぎず予測という本来の目的からすれば甚だ不完全なものである。被験者の特性についての偏り、population-specific な問題を考慮した結果でサンプリングが新しい変数の導入とともに必要であろう。

結論と討論

大都市近郊に在住している特に傷病は有していない健康な男子 (年齢45.8才±8.1才) 55名を対象として、歴年齢を簡単なパフォーマンステストや運動負荷テストあるいは健康診断の測定で得られた測定値から予測するための重回帰分析を行った。

結果を要約すると以下の様になる。

(1) 独立変数としたのは皮下脂肪厚から推定した体脂肪率 (FAT), 握力 (GRIP)・背筋力 (BACK), 立位前屈 (FLEX), 閉眼片足立 (BALANCE), 肺活量 (VITAL), Margaria の方法により踏台昇降運動時

の心拍数 (HR1, HR2), 安静時の心拍数 (HR) および収縮期と拡張期の血圧 (SBP, DBP) の合計11変数であった。これらの変動係数や正規分布に比した尖度・歪度などには特に問題点はみあたらなかった。

(2) 全体の独立変数11個についての相関行列からは固有値を中心にした検討結果では多重共線性は示唆されなかった。

(3) ステップワイズ方式によれば BACK→VITAL→SBP→GRIP→HR→DBP→HR2→HR1→FAT→BALANCE→FLEX の順に回帰式に投入されていた。

(4) 赤池の情報量基準 (AIC) や Mallows の Cp 基準, Schwarz の基準, 自由度調整済み重回帰係数 (R^*) など幾つかの統計的を評価の尺度とし、ステップごとに得られた重回帰式を検討した結果、意味ある変数として BACK, VITAL, SBP, GRIP が含まれた以下の回帰式が最良とみなされることになった。

$$\hat{Y}=53.07-0.219\times\text{BACK}-.00398\times\text{VITAL}+0.1573\times\text{SBP}+0.253\times\text{GRIP}$$

この場合、重回帰係数 (R) は0.630, 寄与率39.7%となった。しかし、このうち、GRIP の回帰係数は負の値となりこれは一般的な見解とは符合しないことになる。

(5) そのため当該変数 GRIP を除去した上でステップワイズ重回帰分析を進め同様に AIC, Mallows や Schwarz の基準によってステップごとの回帰式を評価したところ、AIC や Schwarz の基準などは BACK, VITAL, SBP を含む以下の式を最良と評価することになった。

$$\hat{Y}=57.93-.171\times\text{BACK}-.00327\times\text{VITAL}+.1468\times\text{SBP}$$

この回帰式の重回帰係数は $R=612$, 寄与率 37.5%となった。

(6) BACK, VITAL, SBP までを含む回帰式からの残差の標準偏差は5.55となった。モーメント系の尖度・歪度のうち歪度は有意に大きいものとなった。しかし2標準偏差を越すケースは1例しか見当たらなかった。因みに第4変数として GRIP を含む回帰式についても残差分析では特定の変動パターンは認めることができなかった。

(7) 同様に該当する回帰式に含まれた変数群間の相関行列からは多重共線性は示唆されなかった。

(8) 運動負荷時の心拍数 (HR1, HR2) の代替変数として Margaria et al. の方法で推定された最大酸素摂取量 ($\dot{V}O_2$, ml/kg·min) を新たな独立変数とし

て同様のステップワイズ重回帰分析を行ったところ、BACK, VO₂, VITAL, SBP の順に投入された場合の回帰式が AIC その他の統計基準によって最良であることが確認でき、これまでと同様に残差分析によっても信頼度の高いものであると考えられた。

筆者らはこの種の「総合的研究」の中ですでに勤労婦人や主婦を対象にして同様の重回帰式の作成を試みてきた。ここまでは記述的に一応の予測式を求めることができたと考えられるが、あくまで記述に過ぎないし、また寄与率からみても不完全の域を脱するものではない。

またこの問題の基本には一般的な健康測定値から得られた実質的な生理的「老化度」を知ろうとするところに動機があった。したがって今後はこうした実質的な「老化度」、逆の面からは「若さ度」という個体差をもたらしているのはどのような要因であるのか、運動や生活の習慣などを中心に「老化制御」の問題が検討されるべき¹⁴⁾であろう。

付 記

本研究は昭和58年度九州大学特定研究「生活形態と健康度に関する総合的研究」として実施されたものの一環である。本論文の執筆には著者のうち吉川があたり、データ解析も九州大学大型計算機センターで吉川が担当した。ここで用いたデータの収集には健康科学センターの全スタッフがかわった。いうまでもなく本研究のデータは健康科学センター共有のものであり、たずさわられたスタッフに深甚の謝意を表したい。

文 献

- 1) Akaike, H.: A new look at the statistical model identification, *IEEE Trans. Autom. Contr.*, AC9(6): 716-723, 1974.
- 2) 赤池弘次: 統計的検定の新しい考え方, *数理科学* 17(12): 51-57, 1979.
- 3) 赤池弘次: 情報量規準 AIC とは何か—その意味と将来への展望—, *数理科学*, 14(3): 5-11, 1976.
- 4) 赤池弘次: モデルによってデータを測る, *数理科学*, 19(3): 7-10, 1981.
- 5) 朝比奈一男編: 日本人の体力と健康, 社会保険新報社, 1979, 1-30 および 61-150.
- 6) チャタジー・プライス: 回帰分析の実際 (佐和隆光, 加納悟訳), 新曜社, 1981, p.249.
- 7) Dixon, W.J. and M.E. Brown (Eds.): *BMD—Biomedical computer programs P-series*, Univ Carifor. Press: Berkley, 1979, 399-417.
- 8) ドレーパー・スミス, 応用回帰分析 (中村慶一訳): 森北出版, 1968, 163-216.
- 9) Forsythe, A. B., L. Engelman, R. Jenrich, and P.R. May: A stopping rule for variable selection in multiple regression, *J. Amer. Assoc.*, 68: 75-77, 1973.
- 10) 古川俊之・刈田全世・吉川博通・稲田 紘・森脇健・阿部 裕・宮崎 学・梶谷文彦・宮脇一男: 多変量理論による老年者の生物学的年齢推定, *高齢医学*, 6: 308-317, 1968.
- 11) Furukawa, T., M. Inoue, F. Kajiya, H. Inada, S. Takasugi, S. Fukui, H. Takeda, and H. Abe: Assessment of biological age by multiple regression analysis, *J. Gerontol.*, 30: 422-434, 1975.
- 12) Hanihara, K. and T. Suzuki: Estimation of age from the pubic symphysis by means of multiple regression analysis, *Am. J. Phys. Anthropol.*, 48: 233-240, 1978.
- 13) 芳賀敏郎・橋本茂司: 回帰分析と主成分分析, *日科技連*, 1980, p.228.
- 14) 旗野脩一・芳賀 博・荒尾静代: 老化と予後の関連因子, *応用統計学*, 11(1): 49-58, 1982.
- 15) 菱沼從尹: 年齢の断面, *発達*, 1(3): 10-18, 1980.
- 16) 池上晴夫: 運動処方, 朝倉書店, 1982, 1-54.
- 17) 吉川和利・小宮秀一・小室史恵: 体内総水分量 (TBW) 予測式作成の試み (1), *体力科学*, 32: 39-48, 1983.
- 18) 吉川和利: 生理的年齢予測の重回帰分析, *健康科学*, 7: 1-10, 1985.
- 19) ケンドール: 多変量解析 (奥野忠一・大橋靖雄訳): 培風館, 1981, p.225.
- 20) 小林龍一: 相関・回帰分析入門 (新訂版), *日科技連*, 1982, p.259.
- 21) Koizumi, K: The estimation of age from the cranial sutures by means of multivariate analysis methods, *J. Anthropol. Soc. Nippon*, 90(2): 109-118, 1982.
- 22) 萬代 隆・田淵義勝・矢野敦雄・小沢秀樹・池田正男: 食塩—食塩と高血圧について, *内科*, 48(5): 749-752, 1981.
- 23) Margaria, R., P. Agehemo, and E. Revelli: Indirect determination of maximal O₂ consumption in man, *J. appl. Physiol.*, 20(5): 1070-1073, 1965.
- 24) 三宅章彦: 解説重回帰分析, *最新医学*, 33(1): 68-71, 1978.
- 25) 中村栄太郎: 体力年齢の推定, *日本体育学会第31回大会号*: 1980, p.551.
- 26) Nakamura, E.: The assessment of physiological age based upon a principal component analysis of various physiological variables, *J. Kyoto Pref. Univ. Med.*, 94(8): 757-769, 1985.
- 27) 奥野忠一・久米 均・芳賀敏郎・吉澤 正: 多変量解析法, *日科技連*, 1978, 7-76.

- 28) 奥野忠一・久米 均・芳賀敏郎・吉澤 正: 多変量解析法(改訂版). 日科技連, 1981, 52-157.
- 29) 大西正和: 需要予測とコンピュータプログラム. 日刊工業新聞社, 1982, p.259.
- 30) 小野三嗣: 健康と体力の科学. 不昧堂, 1971, 189-204.
- 31) 大島正光: ヒト——その未知へのアプローチ. 同文書院, 1982, 226-239.
- 32) 小沢利男: 老年者の循環機能. 日本臨床, 39(3): 510-514, 1981.
- 33) 斎藤乾二郎: 予測—未来の読み方—. 朝日出版, 1981, 355.
- 34) 佐和隆光: 回帰分析. 朝倉書店, 1979, p.187.
- 35) Shephard, R.J.: Physical activity and aging. Croom Helm, London, 1978, 52-145.
- 36) 柴田里程: 統計的推測方式とモデル選択. 数理科学, 19(3): 45-50, 1981.
- 37) 白石 透・毛利昌史・本田憲業: 老年者の呼吸機能—加齢による変化と老年者における評価—. 日本臨床, 39(3): 510-514, 1981.
- 38) Stone, J.: Comments on model selection criteria of Akaike and Schwarz, J. Roy. Statist. Soc. B 41: 1979, 276-278.
- 39) 杉山高一・尾崎 公・牛沢賢治・清水政利: 歯の咬耗度による年齢推定にする重回帰分析. 応用統計学, 5(3): 123-128, 1979.
- 40) 高杉成一・武田 裕・吉川博通・古川俊之: 老化の標準値. 最新医学, 33(1): 74-75, 1978.
- 41) 竹井哲司: 歯の咬耗による年齢の推定. 日法医誌, 24(1): 4-17, 1970.
- 42) 田辺國士: 不適切問題への統計的アプローチ. 数理科学, 14(3): 60-64, 1976.
- 43) 東京都立大学身体適性学研究室編: 日本人の体力標準値第三版. 不昧堂, 1980, p.458.
- 44) 柳井晴夫・岩坪秀一: 複雑さに挑む科学. 講談社, 1976, 156-176.
- 45) 吉川博通: 多変量理論による生物学的年齢の推定. 阪大医誌, 22(1, 2): 113-136, 1976.
- 46) 吉川政己: 老化度・健康度をどう測定するか. 内科, 48(5): 729-735, 1981.