

|||||

日本公衆衛生雜誌

Japanese Journal of
Public Health

第 26 卷 別 刷

|||||

原著

糖尿病長期治療患者に関する疫学研究

(2) 多変量解析による患者の初診時臨床像
と治療法に関する研究

高木廣文* 豊川裕之* 小坂樹徳**

I はじめに

糖尿病は、治療上インスリンを必須とするか否かによって、インスリン依存性糖尿病と非依存性糖尿病に大別されている¹⁾。前報²⁾において、厳密な意味でのインスリン依存性ではないが、初診から5年間のインスリン治療の期間により、インスリン治療群と非インスリン治療群の2群に患者を分類し、両群における発症年齢、糖尿病の家族歴および既往最大肥満度などの差を報告した。これは、いわば患者の発症までの状態の相違を検討したことになる。

糖尿病患者の初診時臨床像の統計的解析としては、因子分析を用いた森ら³⁾の報告があり、また初診時検査データにもとづく糖尿病治療剤選択の判定について、判別関数を用いた王子ら⁴⁾の報告がある。王子らによれば、少数の初診時検査項目を用いて、短期ではあるが、以後の治療法にインスリンを用いるかスルフォニール尿素剤を用いるかの判定が、約80~90%の範囲で正確に行なえることを示している⁴⁾。実際の治療法の選択が主治医の経験を背景に、患者の臨床像を総合的に判定してなされるものであり、治療効果いかんによって変更、修正されるものであることを考えると、王子ら⁴⁾の成績は注目すべきものと思われる。このことから、糖尿病治療にインスリンを必要とする者とそうでない者とは、初診時臨床像においてすでに差があり、治療にインスリンを用いるか否かという点では、長期的にみて大きな変更は少ないという仮説が考えられる。

本研究は、糖尿病患者を初診時から5年間の治療法により、インスリン治療群と非インスリン治療群に分け、初診時における諸検査項目の2群間の差を検討し、さらに多変量解析の手法である判別分析⁵⁻⁷⁾、および主成分分析⁸⁻⁹⁾を用いて、2群を判別する上で最も重要な項目、すなわち患者の初診時臨床像を多角的にとらえたとき、最も2群間の差異を示すのに適した項目を統計的に選定することを目的とした。さらに、前述した仮説が本

研究により支持され得るかを検討する。

II 資 料

資料には昭和50、51年度厚生省特別研究「糖尿病治療に用いる薬剤と心血管系障害の推移に関する研究」のデータ⁸⁾を用いた。この研究には、朝日生命成人病研究所、済生会中央病院、国立病院医療センター、虎の門病院、東京女子医大、および東京大学医学部第三内科の6施設が参加し、5年以上の治療経過の明らかな患者1,358例(男762例、女596例)の資料が収集された⁹⁾。本研究では上記資料中、初診時検査項目(表2,3)、患者の家族歴・既往歴(表4)、および初診時から5年間の治療方法を分析データとして用いた。

III 方 法

1. 治療法による対象の分類

患者を初診時以降5年間のうち80%以上の期間にわたって行なわれた治療法により分類した(表1)。表1に示した「食餌」群は食餌単独療法を意味するが、今回はインスリン治療か否かに着目し、上記基準に従い「インスリン治療群(以下、I群と略す)」と「その他の非インスリン治療群(以下、N群と略す)」とに大別した。対象者数はI群257例(男133例、女124例)、N群1,101例(男629例、女472例)となった。このうち判別分析には、空腹時血糖値および発症年齢の不明の者を除外したところ、I群は全例の257例が、N群は1,091例(男622例、女469例)が分析対象となった。線型判別関数の算出には、I群は257例の全例を、N群は電子計算機を用いて一様乱数を発生させ、1,091例から261例(男147例、女114

表1 初診時から5年間の治療方法

患者の性 治療方法*	患者の性			計
	男 性	女 性		
食 餌	138(18.1)	69(11.6)		207(15.2)
経 口 剤	339(44.5)	305(51.2)		644(47.4)
インスリン	133(17.5)	124(20.8)		257(18.9)
そ の 他**	152(19.9)	98(16.4)		250(18.4)
計	762(100%)	596(100%)		1,358(100%)

* 期間の80%以上にわたって同一療法のケース

** 治療法が複数にわたるケース

* 東京大学医学部疫学教室

** 東京大学医学部第三内科学教室

UDC: 616.633.45-036.4-085

表2 各項目の2群の平均値・標準偏差およびt検定

項 目	群	人数	Mean	S.D.	t*	P
発 症 年 齢 (歳)	I	257	41.6	13.6	8.82	0.000
	N	1099	49.7	11.5		
罹 病 期 間 (年)	I	257	5.3	6.6	5.28	0.000
	N	1100	3.0	4.9		
最 大 肥 満 度 (%)	I	238	22.5	20.2	3.44	0.001
	N	986	27.6	20.6		
最大体重時年齢 (歳)	I	200	33.0	12.0	6.65	0.000
	N	796	39.3	12.1		
空腹時血糖値 (mg/dl)	I	257	202.1	75.1	11.03	0.000
	N	1093	147.7	51.1		
血清コレステロール (mg/dl)	I	194	232.5	59.4	2.26	0.024
	N	928	222.1	51.9		
血清トリグリセリド (mg/dl)	I	54	110.7	69.5	0.18	0.861
	N	290	112.6	90.7		
血清尿素窒素 (mg/dl)	I	115	17.6	5.1	2.73	0.007
	N	621	16.0	5.7		
血清クレアチニン (mg/dl)	I	24	1.6	3.1	0.91	0.371
	N	76	1.0	0.3		
初診時肥満度 (%)	I	243	2.9	19.0	6.37	0.000
	N	1015	11.4	17.1		
収縮期血圧 (mmHg)	I	175	128.2	26.6	4.27	0.000
	N	934	137.4	23.5		
拡張期血圧 (mmHg)	I	176	77.0	13.3	4.36	0.000
	N	935	81.7	13.0		

* 等分散性が棄却された場合には Welch の方法によって計算した

表3 2群における眼底、心電図および尿たん白所見 (): %

項目	所 見	I	N	Chi-sq.	P
眼 底	{No	85 (57.8)	573 (75.0)	18.504	0.000
	{(1)*	43 (29.3)	139 (18.2)		
	{(2)**	19 (12.9)	52 (6.8)		
心 電 図	{Normal***	158 (91.3)	727 (91.6)	0.003	0.959
	{Abnormal	15 (8.7)	67 (8.4)		
尿 た ん 白	{No	139 (62.6)	751 (77.1)	20.365	0.000
	{(+)	69 (31.1)	179 (18.4)		
	{(++)以上	14 (6.3)	44 (4.5)		

* Scott I, II or Wagener I, II

** Scott III, IV, V or Wagener III, IV, V

*** Q, ST, T 変化に異常のないものを正常, 他のもを異常とした

例) を抽出して分析対象とし, N群1,091例中の残りの830例は, 判別関数のチェックのために用いた(後述, IV-2)。

2. カテゴリの設定

家族歴, 既往歴および喫煙歴は「あり」と「なし」に

2分類した。眼底所見はScott I, II度またはWagener I, II度を(1)とし, Scott III度以上またはWagener III以上を(2)とし, 異常所見のないものを「なし」とした。心電図所見は, Q, ST, T 変化に異常のないものを正常, 1つ以上あるものを異常とした。尿たん白所見は「なし」, (+), (++)以上の3分類とした。

これらの属性変数に対して, 判別分析および主成分分析を行なうさいには, 表5に示したように数値を与え, 分析を行なった。

3. 分析方法

1) 線型判別関数

初診時データから, その後の治療法を判別するために, 線型判別関数⁵⁾を用いた。線型判別関数とは, P個の説明変数 $x_1, x_2, \dots, x_P$ (本研究では各種の検査項目)を用いて,

$$y = a_1x_1 + a_2x_2 + \dots + a_Px_P$$

の線型の合成得点 y を求め, この y の値によって各患者を判別しようとするものである⁹⁾。 $a_1, a_2, \dots, a_P$ は各説明変数の重みを表わし, 判別係数 discriminant function coefficient (d.f.c.) と呼ばれ, とくに説明変数が標準化されているときには標準化判別係数 standardized d.f.c. と呼ばれる¹⁰⁾。線型判別関数に含める変数は, 前進選択法⁶⁾によって1変数ずつ選択した。選択基準はWilksの λ (ラムダ) 統計量の減少が最大となる変数を選択することにした^{6,10)}。変数選択の打ち切り基準としては, 自由度調整済みマハラノビスの汎距離⁷⁾の増加がなくなる段階までとした。

2) 主成分分析

線型判別関数に選択された変数の持つ意味を検討するために, 既報の方法¹¹⁾により主成分分析を行なった。主成分は全分散に対する寄与率が約5%のものまで抽出した。

また, 線型判別関数の算出には, 変数に欠損値のある場合, それに各群における当該項目の平均値を代入して用いたが, その他の分析では, 欠損値のあるケースは分析から除外した。実際の計算には, 東京大学大型計算機センター (HITAC 8800/8700) のSPSS¹⁰⁾を用いた。

IV 結果とその考察

1. 各項目の2群間の比較

表2にI群, N群における糖尿病の発症年齢, 罹病期間, 発症前の最大肥満度および, そのときの年齢, 初診時における主要検査の平均値とt検定の結果を示した。平均値の差が有意ではない項目は, 血清トリグリセリドおよび血清クレアチニンのみであった。また, これら2項目は有効な標本数がきわめて少なく, 以後の分析から除外した。表2に示したように, I群はN群に比べ発症

表4 2群の家族歴、既往歴および喫煙歴

家族歴・既往歴・喫煙歴		I	N	Chi-sq.	P
糖尿病：家族歴	{No	161(66.0)	650(65.0)	0.046	0.831
	{Yes	83(34.0)	350(35.0)		
虚血性心疾患：家族歴	{No	206(94.5)	789(90.7)	2.761	0.097
	{Yes	12(5.5)	81(9.3)		
脳卒中：家族歴	{No	156(66.4)	565(61.3)	1.815	0.070
	{Yes	79(33.6)	356(38.7)		
高血圧：家族歴	{No	155(70.5)	586(66.8)	0.901	0.368
	{Yes	65(29.5)	291(33.2)		
虚血性心疾患：既往歴	{No	243(99.2)	974(97.1)	2.696	0.101
	{Yes	2(0.8)	29(2.9)		
脳卒中：既往歴	{No	246(98.4)	995(98.5)	0.024	0.876
	{Yes	4(1.6)	15(1.5)		
高血圧：既往歴	{No	214(86.3)	798(78.5)	7.147	0.008
	{Yes	34(13.7)	219(21.5)		
腎炎・ネフローゼ：既往歴	{No	234(92.9)	933(92.4)	0.016	0.900
	{Yes	18(7.1)	77(7.6)		
肝炎・肝硬変：既往歴	{No	234(95.9)	947(94.3)	0.678	0.410
	{Yes	10(4.1)	57(5.7)		
胃切：既往歴	{No	243(99.2)	998(99.0)	0.013	0.91 ¹
	{Yes	2(0.8)	10(1.0)		
甲状腺疾患：既往歴	{No	236(99.2)	968(98.9)	0.001	0.97 ⁶
	{Yes	2(0.8)	11(1.1)		
喫煙歴	{No	134(73.6)	489(64.7)	4.867	0.027
	{Yes	48(26.4)	267(35.3)		

() : %

表5 属性変数に与えた数値

項目	数値	0	1	2
性	Male	Female	—	
家族歴・既往歴	No	Yes	—	
喫煙歴	No	Yes	—	
尿たん白	No	(+)	(++)	
心電図	Normal	Abnormal	—	
眼底	Normal	(1)	(2)	

年齢は若く、罹病期間は長く、最大肥満度は軽度で、そのときの年齢は若く、そして初診時肥満度も軽度であった。また、空腹時血糖値、血清コレステロール、血清尿酸窒素はI群で高く、逆に収縮期血圧と拡張期血圧はN群で高かった。

表3に初診時における眼底、心電図および尿たん白所見を示した。眼底および尿たん白の異常は、I群において高率であり、統計的にも有意であった。しかし、心電図所見での異常は両群間に差はなかった。

表4に2群での家族歴、既往歴および喫煙歴を示した。有意差の認められた項目は、高血圧の既往歴および喫煙歴のみであり、両者ともにN群で高率であった。

I群とN群の初診時検査データの単純な比較によ

て、上述のように27項目中14項目に統計的有意差が認められた。このように2群の初診時臨床像には明らかな差があるが、逆にこれらの項目から以後の治療法が判定できるか否かは別の問題であり、また項目相互の関係から、たんに2群間の差が有意でないからといって、その項目が2群を識別する上で重要でないとは言えない⁷⁾。このような点から、多変量解析の一手法である判別分析を用いて、2群の差をより明確に表わし得る変数を以下のように選定した。

2. 線型判別関数による2群の判別

初診時検査データを用いて、I群かN群かの判別を統計学的に行なったが、空腹時血糖値および発症年齢は重要な役割をもつと考えられるので、これら2項目に欠損値のあるケースは分析から除外した。この結果、前述のごとく(Ⅲ-1)、I群は257例となり、判別分析のためにN群も261例を選定したものである。

表6に前進選択法による変数の選択順位、全例での外的基準(I群に0、N群に1を与えた)との単相関係数、判別関数にもとづく重相関係数⁷⁾、マハラノビスの汎距離^{8,9)}、および各段階での判別関数にもとづいて患者を分類したとき、実際の治療群との一致の程度を示す判別率(適中率¹²⁾とも呼ばれる)を示した。表7には判別得点算出のための、素データに対する判別係数と、標

表6 治療群の判別に用いた変数 (前進選択法による)

順位	変数	r^*	R^*_{H}	D^*_{H}	$D_m/2$	判別率 (%)				D^*_{H}	D^*_{H}
						I 群	N 群	両群	理論値		
1.	空腹時血糖値	-0.354	0.394	0.731	0.427	67.6	67.7	67.6	66.5	0.722	0.713
2.	発症年齢	0.255	0.484	1.222	0.553	66.9	74.7	70.9	71.0	1.201	1.181
3.	血清尿素窒素	-0.100	0.545	1.684	0.649	68.1	77.4	72.9	74.2	1.651	1.619
4.	眼底所見	-0.138	0.580	2.018	0.710	70.4	75.1	72.8	76.1	1.972	1.926
5.	初診時肥満度	0.188	0.601	2.254	0.751	70.4	76.2	73.4	77.4	2.193	2.134
6.	罹病期間	-0.170	0.625	2.559	0.800	72.8	77.0	74.9	78.8	2.483	2.409
7.	最大体重時年齢	0.206	0.639	2.757	0.830	72.4	76.6	74.5	79.7	2.666	2.577
8.	血清コレステロール	-0.073	0.651	2.931	0.856	72.0	78.2	75.1	80.4	2.824	2.721
9.	収縮期血圧	0.139	0.664	3.140	0.886	75.1	80.1	77.6	81.2	3.065	2.897
10.	尿たん白	-0.114	0.675	3.338	0.914	73.5	80.8	77.2	82.0	3.197	3.061
11.	喫煙歴	0.075	0.684	3.513	0.937	73.2	81.2	77.2	82.6	3.353	3.200
12.	高血圧・家族歴	0.031	0.691	3.641	0.954	73.5	80.8	77.2	83.0	3.464	3.296
13.	心電図所見	-0.003	0.695	3.725	0.965	73.9	80.5	77.2	83.3	3.531	3.348
14.	糖尿病・家族歴	0.008	0.699	3.801	0.975	74.3	80.5	77.4	83.5	3.590	3.390
15.	脳卒中・既往歴	0.004	0.701	3.856	0.982	74.7	80.8	77.8	83.7	3.628	3.414

* 全ケースによる外的基準との単相関係数

** I 群257例, N 群261例における重相関係数

マハラノビスの汎距離

▲ 自由度調整済みマハラノビスの汎距離

▲▲ 自由度再調整済みマハラノビスの汎距離

表7 前進選択法による15変数の判別係数

変数	判別係数	標準化判別係数
1. 空腹時血糖値	-0.00598	-0.422
2. 発症年齢	0.01751	0.230
3. 血清尿素窒素	-0.07026	-0.248
4. 眼底所見	-0.41835	-0.217
5. 初診時肥満度	0.01204	0.225
6. 罹病期間	-0.03846	-0.217
7. 最大体重時年齢	0.01652	0.176
8. 血清コレステロール	-0.00299	-0.150
9. 収縮期血圧	0.00953	0.206
10. 尿たん白	-0.26986	-0.151
11. 喫煙歴	0.36697	0.143
12. 高血圧・家族歴	0.24617	0.106
13. 心電図所見	-0.31011	-0.074
14. 糖尿病・家族歴	0.18969	0.087
15. 脳卒中・既往歴	-0.59434	-0.073
定数項	0.38792	

標準化データに対する標準化判別係数を示した。これによって求めた判別得点が負の場合にI群, 正の場合にN群と理論上分類し, 実際の治療群との一致のよさを示すのが表6に掲げた判別率の意味である。

変数選択の順位が必ずしも判別に対する寄与順位と一致しないので, 標準化判別係数の大きさで寄与をみるものとした。最も重要な変数は空腹時血糖値であり, これのみで67.6%の判別率が得られ(表6), 測定値が高い場合にI群と理論上判定される可能性が大きくなる。つぎに血清尿素窒素, 発症年齢, 初診時肥満度, 眼底所見, 罹病期間の順に治療群判別への寄与が高く, これらのうち発症年齢は若い場合に, 他の変数は測定値が大きい場合にI群に, 逆の場合にN群に理論上判定される可能性が大きくなる。最大体重時年齢は, 外的基準との単相関係数は比較的大きいが($r=0.206$), 表8に示したように発症年齢との単相関係数が0.581あり, 内部相関の影響として, 発症年齢が最大体重時年齢の治療群判別への寄与を包含しているため, 選択順位も寄与順位もやや低下したものと考えられる。

収縮期血圧は低い場合に, 最大体重時年齢は若い場合に, 尿たん白は異常の場合に, そして血清コレステロールは高い場合に, それぞれI群と理論上判定される可能性が大きくなる。

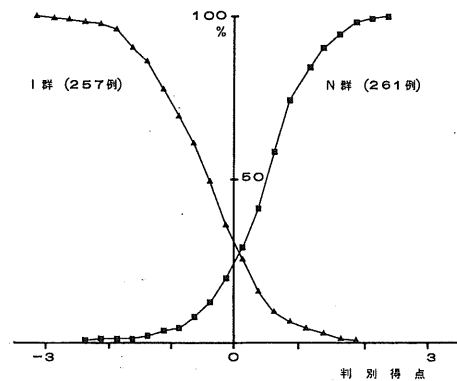
選択順位が11番目以降はすべて属性変数であり, 寄与も大体この順であり, 喫煙歴が比較的大きな標準化判別係数を有していた。

これらの15変数を用いて計算された判別得点の累積分布は図1のようになり, 両群あわせた場合の判別率は

表8 治療群の判別に用いた15変数間の相関係数

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
1. 空腹時血糖値	1.000													
2. 発症年齢	-0.139	1.000												
3. 血清尿酸素	0.049	0.038	1.000											
4. 眼底所見	0.116	-0.033	0.096	1.000										
5. 初診時肥満度	-0.129	0.157	-0.077	0.003	1.000									
6. 罹病期間	0.063	-0.242	0.037	0.250	0.002	1.000								
7. 最大体重時年齢	-0.127	0.581	-0.017	-0.070	0.244	-0.132	1.000							
8. 血清コレステロール	0.115	0.125	0.050	0.000	0.140	0.015	0.056	1.000						
9. 収縮期血圧	-0.111	0.247	-0.026	0.170	0.300	0.179	0.175	0.117	1.000					
10. 尿たん白	0.164	-0.002	0.037	0.267	0.029	0.099	-0.020	0.119	0.182	1.000				
11. 喫煙歴	-0.083	0.030	0.078	-0.054	-0.105	0.024	-0.026	-0.119	-0.048	-0.031	1.000			
12. 高血圧・家族歴	-0.034	-0.108	-0.033	-0.033	0.037	0.001	-0.042	-0.002	0.027	-0.033	0.006	1.000		
13. 心電図所見	0.024	0.139	-0.002	-0.035	0.117	0.070	0.070	0.138	0.125	0.072	-0.055	0.020	1.000	
14. 糖尿病・家族歴	0.007	-0.184	0.013	-0.006	-0.018	0.031	-0.110	-0.037	-0.059	-0.002	-0.034	0.053	-0.010	1.000
15. 脳卒中・既往歴	-0.060	0.025	-0.009	0.075	0.045	0.010	-0.031	0.009	0.087	0.033	0.090	-0.004	0.170	0.002

図1 2群の判別得点の累積分布



77.8%であった。また、N群のうち判別関数の算出に用いなかった830例について、判別得点の分割点 cutting point を0として判別したところ、81.0%の判別率が得られた。

変数追加による判別率の推移をみると(表6, 図2), I群, N群ともに変数選択の各段階で判別率が単調に増加するわけではなく, I群では9変数のときに75.1%と最も高く, N群では11変数のときに81.2%と最も高かった。両群あわせた場合には, 9変数のときに77.6%の判別率を示し, 以後14変数までの変数追加による判別率の上昇は認められない。判別得点が正規分布をすると仮定したときの理論上の判別率²⁾は変数追加の全段階で増加しているが(表6, 理論値の項), 実際の判別率は9変数のときに比べ, 15変数のときにさえもわずか0.2%, 人数にして1例の増加にすぎない。したがって, 実際上はこの9変数が最もよくI群とN群の初診時臨床像の差を表わすものと考えられる。ここで注意すべき点は, これらの変数を個別にとらえるのではなく, 多元的に評価したときに, より意味を持つということである。

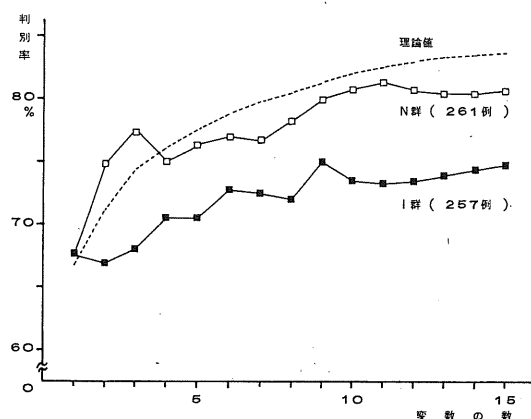
上述のように判別分析の結果から, 糖尿病患者の初診時検査データを用いて, 以後の治療にインスリンを用いるか否かの判別は約75~80%の範囲で可能であることが認められた。しかし, この結果はたんにインスリンを用いるか否かだけに対してであり, その用量や細かな修正, 変更についてのものではない。そのような点に関しては, 当然のことではあるが, 患者の主治医がより詳細に, 各種検査にもとづいて決定すべきであることは言うまでもなからう。

つぎに, 判別関数に選択された各変数の持つ意味を検討するために, 以下のように主成分分析を行なった。

3. 主成分分析による変数の分類

判別分析に用いた説明変数と治療方法および眼底異常出現に関する変数を1つ加え, 表9に示したように, 28変数による主成分分析を行なった。主成分はその寄与率

図2 変数追加による判別率の推移



が約5%である第5主成分まで求めた。以下に各主成分に対する説明を加える。

表9に示したように、5年間の治療法は第1主成分および第2主成分負荷量が大きく、この2主成分が治療法との関係が大きいことが認められた。

第1主成分は発症年齢および既往最大体重時年齢などで、その負荷量が大きいことから、年齢もしくは加齢に関する主成分と考えられる。収縮期および拡張期血圧の負荷量が大きいことも、年齢（加齢）に関する主成分であることを支持するものと考えられる。また、最大肥満度が発症年齢と関係があることはすでに報告した²⁾。第1主成分負荷量の大きな変数からは、収縮期血圧、初診時肥満度、発症年齢および最大体重時年齢の4変数が判別関数に選択され、これらの測定値が小さい場合にI群と理論的に判定される可能性が大となる。このことは、I群が厳密な意味でのインスリン依存性ではないにしても、インスリン依存性糖尿病が若年者に多いという従来の報告¹⁾に符合するものと考えられる。また、最大肥満度、拡張期血圧、高血圧の既往歴は負荷量は大きいが判別関数に選択されなかった。これは先に判別関数に導入された変数が、これらの変数の外的基準に対して説明できる部分を包含しており、単独の場合には有意差をもつにもかかわらず、判別関数に導入されなかったものと考えられる。この考えは、他の変数についてもあてはまり、多元的にデータをとらえるときの基本的な考え方であり、偏相関¹²⁾の考えにつながるものである。

第2主成分負荷量の最も大きな変数は「眼底異常悪化・5年次」であり、これは初診から5年後の時点での眼底異常と悪化に関するもので、主成分の解釈を容易にするために加えたもので、これについての詳細は次報¹³⁾で述べる。第2主成分負荷量の大きな変数は、初診時眼底所見、罹病期間、尿たん白、空腹時血糖値であり、糖尿

病性細小血管症に関する変数と考えられる^{1,14,15)}。これらの4変数はすべて判別関数に選択され、測定値が大きい場合に、I群となる可能性が理論上大きくなるが、これはインスリン治療を必要とする患者は、初診時において糖尿病性細小血管症のリスクが非インスリン治療者に比べ大きいことを示唆するものと考えられる。

第3主成分負荷量は、性および喫煙歴で大きい。一般に喫煙者は男でその割合が高く、この主成分が性差に関するものか喫煙に関するものか断じ難いが、喫煙習慣等を含めた男女差と考えることは可能であろう。判別関数には喫煙歴が選択された。

第4主成分負荷量は高血圧および脳卒中の家族歴で大きく、動脈硬化に関する家族歴についての主成分と考えられるが、判別関数には前述のごとく、高血圧の家族歴が選択された。

第5主成分負荷量は、虚血性心疾患の既往歴、心電図所見、脳卒中の既往歴で大きく、心脳血管障害に関する主成分と考えられるが¹⁵⁾、判別関数には心電図所見と脳卒中の既往歴の2変数が選択された。

これら5主成分で説明できる分散は全分散の35.9%であり、5年間の治療法については38.9%であった(表9)。これは治療法が、ここで示された5主成分ではすべて説明しきれないことを表わしている。この5主成分の負荷量が比較的小さい変数のうち、判別関数に選択されたものは、血清コレステロール、糖尿病の家族歴、血清尿素窒素の3変数であった。このうち、血清コレステロールと血清尿素窒素は、糖尿病代謝¹⁶⁾に関するものと考えられる。

ところで、前述したようにI群とN群の初診時臨床像の差を表わすのは、表6に示した変数のうち選択順位が9番までのものが重要であった。このうち、空腹時血糖値、眼底所見および罹病期間の3変数は第2主成分負荷量が大きく、発症年齢、初診時肥満度、最大体重時年齢および収縮期血圧の4変数は第1主成分負荷量が大きかった。残りの2変数は血清コレステロールと血清尿素窒素であった。先の考察から、I群かN群かの判別には、これら9変数が最重要であり、患者の年齢、糖尿病性細小血管症および糖尿病代謝に関係があると考えられる変数が、これら2群の初診時臨床像の差異を最も良く示すものと考えられる。

V 分析方法についての考察

初診時検査データを用いて、インスリン治療群と非インスリン治療群の比較を行ない、2群の初診時臨床像の差を指摘し、逆に初診時検査データから以後の治療法の判別を行ない、約75~80%の範囲で予測が可能であることを示した。ここでは、おもに本研究で用いた分析方法

表9 28変数による主成分分析 (第5主成分まで)

変 数	I	II	III	IV	V	2 乗和
△収縮期血圧	0.683	0.285	0.292	0.056	-0.283	0.716
△初診時肥満度	0.682	0.096	-0.248	0.020	0.000	0.536
最大肥満度	0.620	0.242	-0.231	-0.063	0.079	0.507
拡張期血圧	0.585	0.175	0.278	0.127	-0.392	0.620
△発症年齢	0.544	-0.308	0.035	-0.464	0.150	0.630
△最大体重時年齢	0.539	-0.289	-0.060	-0.345	0.102	0.507
高血圧・既往歴	0.472	0.020	0.178	0.295	0.032	0.343
眼底異常悪化・5年次	-0.092	0.652	0.166	-0.128	-0.073	0.483
△初診時眼底所見	0.039	0.604	0.161	-0.079	-0.060	0.402
△罹病期間	-0.031	0.514	0.287	0.177	0.071	0.384
△尿たん白	0.119	0.496	0.053	-0.088	0.116	0.284
△空腹時血糖値	-0.250	0.478	-0.221	-0.094	0.163	0.375
□5年間の治療法	0.369	-0.461	0.113	-0.024	-0.164	0.389
性	0.171	0.165	-0.722	0.180	0.070	0.615
△喫煙歴	-0.122	-0.203	0.668	-0.077	0.087	0.516
△高血圧・家族歴	0.092	-0.087	-0.033	0.686	0.030	0.489
脳卒中・家族歴	0.128	-0.149	0.008	0.591	0.057	0.391
虚血性心疾患・既往歴	0.133	-0.080	0.140	0.066	0.562	0.364
△心電図所見	0.291	0.083	0.005	0.040	0.552	0.398
△脳卒中・既往歴	0.126	0.084	0.312	0.116	0.443	0.330
△血清コレステロール	0.230	0.204	0.210	0.115	0.139	0.171
△糖尿病・家族歴	-0.132	0.042	-0.085	0.295	0.036	0.115
△血清尿酸値	-0.108	0.182	0.170	-0.232	0.098	0.137
虚血性心疾患・家族歴	0.116	-0.048	0.100	0.220	-0.036	0.075
甲状腺疾患・既往歴	-0.020	-0.038	-0.085	-0.048	0.236	0.067
肝炎、肝硬変・既往歴	-0.029	-0.065	0.149	0.100	0.202	0.078
胃切	-0.051	-0.101	0.191	0.030	0.198	0.089
腎炎、ネフローゼ	0.095	0.103	-0.008	0.101	0.121	0.045
固 有 値	2.984	2.324	1.755	1.620	1.370	10.05
寄 与 率	10.7%	8.3%	6.3%	5.8%	4.9%	35.9%

△ 判別分析において変数選択されたもの

□ 判別分析における外的基準

について検討する。

本研究では、いわゆる Fisher の線型判別関数を用いて判別を行なったが、この関数は説明変数が多変量正規分布をすることを仮定して導入されたものであり^{5,7,9)}、本研究のように属性変数を含む場合に若干問題を残す。このような場合、ダミー変数を用いる方法があり¹²⁾、これは二分法的属性に対して、各カテゴリに0, 1の数値を与えることに対応する。したがって、分析上問題となるのは、カテゴリ数が3である尿たん白と眼底所見の2変数である。そこで、これら2変数をダミー変数を用いて分析したところ、注目すべき改善とはならず、表5に示したように順次0, 1, 2の数値を与えるものとした。

またダミー変数を用いる方法とは別に、林の数量化理論Ⅱ類を用いる方法がある^{7,12)}。この方法を用いて著者らは胃がん¹⁷⁾および胃潰瘍¹⁸⁾のリスク・ファクターに関する研究を行ない、実用上の有効性を確認している。林の数量化理論Ⅱ類は、全変数が属性変数の場合には、ダミー変数による判別分析と一致するが^{5,12)}、本研究のように定量的変数を多く含む場合、連続量のカテゴリ化(区間の決め方)に一定の方法がない。試みに、各変数の平均値と標準偏差を用いて2~5区分し、試行錯誤的に分析を行なったが、線型判別関数以上の効率を上げられなかった。この理由は、変数選択の上位の変数は眼底所見を除けばすべて連続量であり、これらの変数をカテゴリ

化することは、情報量の損失を意味するためと考えられる。

線型判別関数に導入する説明変数は、Wilks の λ (ラムダ) 統計量の減少が最大となる変数を1変数ずつ選択した¹⁰⁾。このほかにも、変数選択の基準にはマハラノビスの汎距離を最大にするものなどがあるが^{6,7,10)}、各基準によって結果が大きく異なることはないようである。また選択された変数の意義は、より実際の観点から評価すべきであると考えられる⁹⁾。

線型判別関数に選択された変数の意味を検討するために主成分分析を用いたが、本来この方法は新しい総合特性値を求めるために使用されるが⁶⁾、主因子分析において共通性 communality を1とした場合に一致し^{6,7,10)}、因子分析の一種とも考えられる⁹⁾。本報告では、共通因子の探索というよりも、5年間の治療法がいかなる成分と関係するかを検討することを目的としたので、共通性を推定しない主成分分析を用いたわけである。

今回、糖尿病患者の初診時検査データから、以後5年間の治療法によりインスリン治療群と非インスリン治療群に分類した2群について、初診時臨床像の比較を行ない、明らかに両群に差異のあることを指摘した。さらに、判別分析により、逆に初診時検査データから2群の判別が約75~80%の範囲で正確に行なえることを示した。これらの結果は、糖尿病治療が主治医の経験を背景に、患者の臨床像を総合的に判定してなされるものであることを考えると、注目に値すべきことであろう。

本成績の妥当性は、他施設および新規の患者から得られたデータによって確認すべきであることは言うまでもなからう。また、判別分析において糖尿病代謝に関する項目として、血清トリグリセリド¹⁰⁾を含めずに分析を行なったが、有効な標本数を多く集めて分析を行なえば、結果に若干の差異が現われる可能性は否定できない。今後、これらの点については、さらに検討を重ねる必要があると考えられる。

Ⅶ ま と め

糖尿病患者を初診時以降5年間の治療法により、インスリン治療群と非インスリン治療群に分け、2群の初診時臨床像の差異を検討し、次いで初診時検査データによる治療群の判別を行なった。さらに、線型判別関数に導入された変数の意味を検討するために主成分分析を行なった。資料には、朝日生命成人病研究所、済生会中央病院、国立病院医療センター、虎の門病院、東京女子医大、および東京大学第三内科の6施設における5年以上の治療経験者のデータを用いた。インスリン治療群は257例、非インスリン治療群は1,091例であったが、この

うち非インスリン治療群は判別関数の算出のための261例と、判別関数チェックのための830例とに分けた。

得られた成績は以下のとおりである。

1) インスリン治療群は非インスリン治療群に比べ、発症年齢は若く、罹病期間は長く、最大肥満度は軽度で、そのときの年齢は若く、初診時肥満度も軽度であり、空腹時血糖値、血清コレステロール、血清尿素窒素は高く、逆に収縮期血圧と拡張期血圧は低かった。また、眼底および尿たん白の異常はインスリン治療群で高率であり、高血圧の既往歴および喫煙歴を有する者は、非インスリン治療群で高率であった。

2) 判別分析によって、治療群の判別に重要な初診時検査項目を選定した。それらは、①空腹時血糖値、②発症年齢、③血清尿素窒素、④眼底所見、⑤初診時肥満度、⑥罹病期間、⑦最大体重時年齢、⑧血清コレステロール、⑨収縮期血圧、⑩尿たん白、⑪喫煙歴、⑫高血圧の家族歴、⑬心電図所見、⑭糖尿病の家族歴、⑮脳卒中の家族歴の15変数である。これらを用いての判別率は、インスリン治療群74.7%、非インスリン治療群80.8%、両群あわせて77.8%であった。チェック用の非インスリン治療群830例では81.0%の判別率であった。

3) 変数追加による判別率の推移から、実際上は上記①~⑨の9変数を用いても、判別率は77.6%あり、これら9変数が最も重要であることを指摘した。さらに、主成分分析により、これらの変数が初診時年齢、糖尿病性細小血管症および糖尿病代謝に関係する変数と考えられることを示した。

4) 上記1)~3)から、インスリン治療群と非インスリン治療群では、初診時臨床像に明らかな差があり、また逆に初診時臨床像から長期的なインスリン使用を予測することが、ある程度可能であると考えられる。

稿を終るにさいし、貴重な資料を提供していただいた朝日生命成人病研究所：葛谷信貞所長ほか、済生会中央病院：堀内光院長ほか、東京女子医大：水野美淳教授ほか、国立病院医療センター：尾形安三内科医長ほか、虎の門病院：伊藤徳治部長ほか、および東京大学第三内科：糖尿病研究グループの諸先生に深謝します。

(受稿 昭53. 10. 2)

文 献

- 1) 小坂樹徳：糖尿病の臨床における2・3の基本的問題、現代医療、9(8~9)、1015~1025、1977
- 2) 高木廣文、丸井英二、他：糖尿病長期治療患者に関する疫学研究(1)インスリン治療群と非インスリン治療群の発症に関する血縁者の糖尿病の影響、日本公衛誌 25(11)、607~616、1978

- 3) 森 雄材, 他: 糖尿病患者の初診時状態像の解析, 糖尿病, 16(2), 136~146, 1973
- 4) 王子亘由, 他: 判別函数による糖尿病治療剤(インスリンおよびスルフォニール尿素剤)選択の判定, 糖尿病, 16(5), 430~434, 1973
- 5) 竹内 啓, 柳井晴夫: 多変量解析の基礎, 東洋経済(東京), 1972
- 6) 奥野忠一, 他: 多変量解析法, 日科技連(東京), 1971
- 7) 柳井晴夫, 高根芳雄: 現代人の統計2, 多変量解析法, 朝倉書店(東京), 1977
- 8) 小坂樹徳, 他: 厚生省特別研究「糖尿病治療に用いる薬剤と心血管系障害の推移に関する研究」, 総括報告書, 1977
- 9) C. R. Rao: Linear statistical inference and its applications, 2nd edit., 奥野忠一, 他訳, 統計的推測とその応用, 東京図書(東京), 1977
- 10) 三宅一郎, 他: SPSS 統計パッケージII, 解析編, 東洋経済(東京), 1977
- 11) 高木廣文, 豊川裕之: 判別関数による糖尿病治療群の判別, 第6回日本行動計量学会総会発表論文抄録集, 64~67, 1978
- 12) 安田三郎: 社会統計学, 丸善(東京), 1969
- 13) 高木廣文, 他: 未発表資料
- 14) 小坂樹徳: 糖尿病の臨床における合併症の意義, 内科, 37(3), 364~368, 1976
- 15) 小坂樹徳, 他: 糖尿病治療と血管障害——われわれの分析成績からみた考察——, 内科, 37(3), 402~407, 1976
- 16) 細谷憲政, 編: 病態栄養学 双書5, 糖尿病, 130, 第一出版(東京), 1972
- 17) 柳井晴夫, 吉本泰彦, 高木廣文, 他: 胃がんの Risk Factor に関する研究, 日本公衛誌, 24(9), 547~555, 1977
- 18) 吉本泰彦, 高木廣文, 他: 胃潰瘍に関する統計学的分析, 日本公衛誌, 25(2), 64~71, 1978

EPIDEMIOLOGICAL STUDIES
ON DIABETICS TREATED FOR A LONG-TERM
(II) A STUDY ON THE CLINICAL STATE AT INITIAL VISIT
AND THE TREATMENT OF DIABETICS
BY MEANS OF MULTIVARIATE ANALYSIS

Hirofumi TAKAGI, Hiroyuki TOYOKAWA

(Department of Epidemiology, School of Health Sciences, Faculty of Medicine, University of Tokyo)

Kinori KOSAKA

(The Third Department of Internal Medicine, University of Tokyo)

In order to elucidate the differences of medical data at initial visit between two groups of which diabetic patients were divided into insulin treated group (ITG) and non-insulin treated group (NTG) by the method of therapies during five years from initial visit, we applied discriminant analysis with statistical forward selection procedure to selecting significant variables and did principle component analysis to considering the meanings of selected variables. Used data were collected in 1,358 diabetics who were treated during five years or more at one of six institutions, which were Institution of Adult Diseases of Asahi Life Foundation, Saisei-kai Central Hospital, National Medical Center Hospital, Tokyo Women's Medical College, Tora-no-mon Hospital, and University of Tokyo. In this study, ITG and NTG consisted of 257 and 1,091 patients respectively, and further NTG was divided into two groups of which 261 patients were selected at random to calculate linear discriminant function and other 830 patients were done to check on the estimated function.

The results obtained were as follows:

1. Most of medical items had statistically significant differences between ITG and NTG, and we selected 15 variables to discriminate the two groups by discriminant analysis. Correct diagnosis rates, which mean the degree of correctness of estimation, of ITG, NTG (261 patients) and both groups were 74.7, 80.8 and 77.8 percent respectively, and that of other 830 patients of NTG was 81.0 percent by the discriminant function.
2. According to the change of correct diagnosis rates at each variable selection procedure, nine variables seemed to be most significant: fasting blood sugar, age at onset, serum urea nitrogen, eye fundi findings, degree of obesity at initial visit, duration from onset, age at maximum weight, serum cholesterol, and systolic pressure. These variables seemed to relate to age, diabetic microangiopathy, and diabetic metabolism of patients at initial visit from results of principle component analysis.
3. As before, it was suggested that there were significant differences of the clinical state at initial visit between ITG and NTG, and that we might predict by initial medical examinations of diabetic patients whether the insulin treatment for a long-term should be applied.