研究ノート

日赤・健康管理センターの人間ドックで蓄積されたデータを活性化する情報システムの研究と開発（中間報告その３）
－データマイニング法と多変量解析法とによるデータ分析－

研究グループ　幹事　総合管理学部　教　授　野村　武
〃　〃　市村　憲治
〃　〃　藤尾　好則
〃　助教授　税所　幹幸
〃　〃　津曲　隆
〃　講　師　宮園　博光
共同研究者　日赤・熊本健康管理センター
所　長　小山　和作
企画課長　松尾　芳昭

A Design Study of Analyzing Systems concerning the data saved by Japanese Red-cross Kumamoto Health-care Center (Interim Report 3)

[概要]

日本赤十字社・熊本健康管理センターでは，人間ドックで行った測定・診断・判定などの諸データを，受診者ごとに年度別にマスターファイルに記録し保存している。熊本県立大学の上記教員グループは，受診データごとの正常／異常の分布を詳細に分析・調査する「情報処理システムの研究・開発」を，同センターの協力を得て平成６年度から２カ年間の研究テーマに設定した。これらのデータはプライバシー保護のために人名を削除してセンターから預かり，種々の実験・考察に活用してきた。

「第一回の中間報告」では，研究の進め方について報告した。

「中間報告その２」では，①汎用集計プログラムを活用することによって，通常集

計から三重集計までの集計作業，②熊本県内を地域細分化した上での集計によって，データ分布についての地域差の有無の分析，③他地域の人間ドック・データとの比較分析－他地域の人間ドックの年間報告書に記載の数字と，先方の設定条件に合わせて，集計し直した当方の数字とを比較分析する試行－などを報告した。

今回の「中間報告その3」では，項目数が大量で，多くの要素が絡んでいる蓄積データについて，共通する因子を探るとか，少ない数で全体を説明する変数を求めるとかを狙って，いわゆる'Reduction of Data'＝少ない変数で全体を表す工夫＝を行うことにし，①データマイニング法による分析，②多変量解析法による分析の二通りを実施したので報告する。

1．前書き

熊本県立大学・総合管理学部と日本赤十字社・熊本健康管理センターとが取り組んだ地域貢献共同研究事業（二カ年計画）は，平成6年4月に発足してこの平成8年3月に一応の終了となった。

この間の研究活動は，蓄積データの集計・分析システムの設計や試行を主な狙いに定め，健康管理センターの協力を得て，熊本県立大学の情報系教員グループが実作業を担当した。そして月に1～2回の会議を開いて情報の交流や研究の報告を継続して，以下のような成果を得た。なお一部のテーマについては，作業の関連から平成8年8月現在でも継続して研究作業を続けている。

実施ずみあるいは進行中の研究内容はつぎの三つに大別される。そして今回の「中間報告その3」では(2)の高次の分析手法によるデータ分析の結果を報告する。

［研究報告の項目］

(1)蓄積データの項目別集計分析　……　「中間報告その2」で報告ずみ
（アドミニストレーション学会誌）

①準備作業

a. データ項目の概要

b. データベース構築に関する基礎的な調査

c. 集計プログラムの導入，チェック

d. 正常／中間／異常の境界値

②市町村別の受診率の分布

③個別の項目ごとの集計分析

④三項目を多重集計した分析（三次元集計）

⑤データを地域別に細分化した分析

付属資料：地域細分化に関する考察

⑥他地域の人間ドック・データとの比較分析

(2)高次の分析手法によるデータ分析 …… 今回の「中間報告その3」で報告する事項（アドミニストレーション学会誌）

①データマイニング法による分析

②多変量解析による分析

(3)遡及分析 …… 完成次第報告予定
（アドミニストレーション学会誌）
人間ドック学会を学会で発表予定

①死因と生前の健康管理データとの関連分析

②まとめ

この一連の研究については，すでに平成7年7月に「日赤・健康管理センターの人間ドックで蓄積されたデータを活性化する情報システムの研究と開発（中間報告）」（アドミニストレーション第2巻1号）を出して，まず研究に関する考え方・進め方を報告した。

次いで「中間報告その2」として，上記項目の中の(1)についての研究結果をアドミニストレーション学会誌第3巻1号に報告した。

続く今回は「中間報告その3」で，上記項目の中の「(2)高次の分析手法によるデータ分析」について報告する。

そして最終的には「(3)遡及分析」については，完成しだいアドミニストレーション学会誌に報告し，総合的にまとめる予定である。加えて8月末に開かれる日本人間ドック学会で，それまでにまとまった部分を発表する予定にしている。さらに人間ドック学会誌に研究内容のまとめを報告する作業を進めている。

2．データマイニング法の適用

人間ドックの蓄積データに最近話題になっているデータマイニング法を適用し，蓄積データ項目間の関連性を調べたいと考えた。この分析に当たっては，日立製作所九州支社ならびに同システム研究所に多面的な協力をいただいた。

①データマイニング法の出現

近年，巨大データベースから知識獲得を簡便に行う技術として，データマイニング法が脚光を浴びている。一口でいえば，マイニング＝採鉱という名のとおり，鉱脈や水脈を探索するイメージである。

例えば，巨大な顧客データベースをもつ商店が商品のダイレクトメールを顧客に出すときに，商品ごとにどのような属性を持つ顧客グループに焦点を合わせると利益を最大にできるかを調べたいという場合である。つまり顧客を絞り込むためのルールを発見しようというわけである。[*1),*2)]

具体的には，パソコンにマイニング法の汎用ソフトをセットし，データを投入して分析を行う。基本的には，大量のデータの中からそれぞれ範囲を設定された三個までの変数項目と一つの結論項目との間の出現頻度を，丹念にカウントして度数を表示する仕組みになっている。

②人間ドックの蓄積データへの適用

人間ドックの蓄積データにこの方法を適用する目的は，沢山の測定データ項目はどのようなグループに分けることができるか，あるいはいくつかに絞った数値表現で言い換えられないか…を探ることにある。

人間ドックに蓄積されたデータは，それぞれの測定項目ごとに正常／軽度の異常／異常の境界値が設定されており，三つの範囲に分けられている。今回の適用は，多くの組み合わせの中から汎用プログラムでできる「3項目の範囲別組み合わせ」をカウントさせ，結論項目の「範囲別肥満度」との関連を分析した。俗な表現をすれば「肥満度が正常なグループ」の中ではどの測定項目の範囲の組み合わせが多いか少ないか，

---

*1）福田剛志ほか：データマイニングの最近の傾向－巨大データからの知識発見技術，情報処理，Vol.37,No.7,pp.597-603(1996).

*2）日立製作所・データマイニング法（解説書）

あるいは逆に「肥満度が異常なグループ」の中ではどの組み合わせが多いか少ないか…というように，すべての組み合わせについて頻度集計を行わせた。

今回のテストでは，日立製作所が開発した汎用ソフトウェアを使用させてもらった。

③ '93年度、'87年度のデータを分析

後述の３．多変量解析法との整合性をとるために，データは '93年度，'87年度の年度の同じ蓄積データを用いた。基本的には健康管理センターのデータ整備の都合に合わせたもので，年度について特別な指定はしていない。

また各年度のデータについては，まずデータを性別に分け，さらに40才から69才までの範囲のデータを抽出して，分析を進めた。

測定項目は，これまでの作業の経験から次の２１項目を取り上げた。

1. 肥満度	2. 血圧１高	3. 血圧１低	4. γＧＴＰ
5. ビリルビン	6. ＬＤＨ	7. ＧＰＴ	8. ＧＯＴ
9. 総蛋白	10. 血糖（空腹時）	11. コレステロール	12. 白血球
13. 尿酸	14. ヘモグロビン	15. 中性脂肪	16. ＨＤＬＣ
17. クレアチニン	18. 尿素窒素	19. 肺（努力）	20. 血清アミラーゼ
21. 赤血球			

それぞれについての正常／軽度／異常の境界値は，健康管理センターが現在採用している数値と同じにした。なお，アルブミンはデータ不備のため除外した。

結論項目については種々の設定の方法があると想定されたが，ここでは「肥満度」だけを採用した。

表　1　データマイニング法による分析

´93年度・男性・40～69才・［結論→条件］

設　定　条　件			肥満度	比　率	評価尺度
ＧＰＴ＝異,	総蛋白＝正,	赤血球＝正	異　50%	231/ 454	0.027
〃	〃	ﾍﾓｸﾞﾛﾋﾞﾝ＝正	〃　50〃	232/ 460	0.027
〃	赤血球＝正,	〃	〃　50〃	244/ 488	0.028
ｱﾐﾗｰｾﾞ＝正,	〃	尿素窒素＝正	〃　28〃	1028/3573	0.028
〃	〃	白血球＝正	〃　28〃	1040/3620	0.028
〃	〃	ﾍﾓｸﾞﾛﾋﾞﾝ＝正	〃　28〃	1092/3830	0.028
ＧＰＴ＝軽,	尿素窒素＝正,	尿酸＝正	軽　20〃	119/ 591	0.007
〃	ＬＤＨ＝正,	〃	〃　20〃	125/ 624	0.007
〃	総蛋白＝正,	〃	〃　19〃	123/ 616	0.007
〃	ﾍﾓｸﾞﾛﾋﾞﾝ＝正,	〃	〃　19〃	126/ 635	0.007
〃	ＬＤＨ＝正,	尿素窒素＝正	〃　19〃	135/ 697	0.007
〃	総蛋白＝正,	〃	〃　19〃	133/ 687	0.007
〃	〃	ＬＤＨ＝正	〃　19〃	139/ 722	0.007
〃	ﾍﾓｸﾞﾛﾋﾞﾝ＝正,	〃	〃　19〃	142/ 744	0.007
ＧＰＴ＝正,	中性脂肪＝正,	血糖値＝正	正　71〃	2696/3791	0.061
〃	〃	尿酸＝正	〃　71〃	2831/3982	0.064
〃	〃	ＧＴＰ＝正	〃　71〃	2710/3816	0.061
〃	〃	ＨＤＬＣ＝正	〃　70〃	2912/4107	0.065
〃	〃	血圧１低＝正	〃　70〃	2844/4013	0.063
〃	ＨＤＬＣ＝正,	血圧１高＝正	〃　70〃	2825/3992	0.062
〃	〃	ＧＴＰ＝正	〃　70〃	2777/3931	0.060
〃	〃	尿酸＝正	〃　70〃	2942/4166	0.063
〃	〃	血圧１低＝正	〃　70〃	2670/4227	0.062
〃	〃	ｸﾚｱﾁﾆﾝ＝正	〃　70〃	3008/4297	0.061

［注］異＝異常、軽＝軽度の異常、正＝正常　の略

表　2　データマイニング法による分析

´87年度・男性・40～69才・［結論→条件］

設　　定　　条　　件	肥満度	比　率	評価尺度
ＧＰＴ＝異，赤血球＝正，尿素窒素＝正	異　42%	150/ 349	0.027
〃　，ﾍﾓｸﾞﾛﾋﾞﾝ＝正，　〃	〃　42〃	151/ 355	0.026
〃　，　〃　，赤血球＝正	〃　42〃	162/ 381	0.028
〃　，総蛋白＝正，　〃	〃　42〃	161/ 379	0.028
〃　，　〃　，ﾍﾓｸﾞﾛﾋﾞﾝ＝正	〃　42〃	162/ 384	0.028
血圧１低＝軽，ＧＯＴ＝正，ｸﾚｱﾁﾆﾝ＝正	軽　20〃	54/ 264	0.008
〃　，　〃　，白血球＝正	〃　20〃	58/ 287	0.008
〃　，　〃　，赤血球＝正	〃　20〃	58/ 288	0.008
〃　，　〃　，総蛋白＝正	〃　19〃	58/ 292	0.008
〃　，　〃　，ﾍﾓｸﾞﾛﾋﾞﾝ＝正	〃　19〃	58/ 292	0.008
ｱﾐﾗｰｾﾞ＝正，赤血球＝正，ｸﾚｱﾁﾆﾝ＝正	〃　14〃	218/1485	0.009
〃　，　〃　，ＧＯＴ＝正	〃　14〃	213/1458	0.008
血圧１低＝正，ＧＰＴ＝正，中性脂肪＝正	正　74〃	1298/1732	0.072
血圧１高＝正，　〃　，　〃	〃　74〃	1312/1770	0.068
ＧＴＰ＝正，　〃　，　〃	〃　73〃	1353/1839	0.067
血圧１低＝正，　〃　，ＨＤＬＣ＝正	〃　73〃	1362/1859	0.065
中性脂肪＝正，　〃　，尿酸＝正	〃　72〃	1374/1883	0.064
ＨＤＬＣ＝正，　〃　，中性脂肪＝正	〃　72〃	1408/1935	0.065
ｸﾚｱﾁﾆﾝ＝正，　〃　，　〃	〃　72〃	1373/1887	0.063
白血球＝正，　〃　，　〃	〃　72〃	1406/1937	0.063
ＧＴＰ＝正，　〃　，ＨＤＬＣ＝正	〃　72〃	1406/1939	0.063

［注］異＝異常、軽＝軽度の異常、正＝正常　の略

表　3　データマイニング法による分析

´93年度・女性・40～69才・［結論→条件］

設　　定　　条　　件	肥満度	比　率	評価尺度
ｱﾐﾗｰｾﾞ＝正，赤血球＝正，ﾍﾓｸﾞﾛﾋﾞﾝ＝正	異　23%	238/1033	0.040
〃　　白血球＝正，　〃	〃　20〃	263/1253	0.034
〃　　赤血球＝正，ＨＤＬＣ＝正	〃　20〃	209/1008	0.026
〃　　〃　　中性脂肪＝正	〃　20〃	199/ 982	0.023
〃　　ＨＤＬＣ＝正，ﾍﾓｸﾞﾛﾋﾞﾝ＝正	〃　19〃	230/1174	0.023
〃　　中性脂肪＝正，　〃	〃　19〃	223/1152	0.022
ＧＴＰ＝軽，ＧＰＴ＝正，白血球＝正	軽　17〃	36/ 205	0.009
〃　　ｸﾚｱﾁﾆﾝ＝正，血圧１低＝正	〃　17〃	40/ 232	0.009
〃　　〃　　ﾍﾓｸﾞﾛﾋﾞﾝ＝正	〃　16〃	41/ 242	0.009
〃　　〃　　白血球＝正	〃　16〃	41/ 247	0.009
ＧＯＴ＝正，ｱﾐﾗｰｾﾞ＝正，赤血球＝正	〃　11〃	122/1042	0.010
ﾍﾓｸﾞﾛﾋﾞﾝ＝正，　〃　　〃	〃　11〃	120/1033	0.010
血圧１低＝正，ＧＯＴ＝正，　〃	〃　10〃	190/1764	0.009
〃　　ﾍﾓｸﾞﾛﾋﾞﾝ＝正，　〃	〃　10〃	188/1748	0.009
ＧＯＴ＝正，　〃　　〃	〃　10〃	189/1758	0.009
ＧＴＰ＝正，血糖値＝正，ＨＤＬＣ＝正	正　80〃	1432/1772	0.040
ＧＰＴ＝正，　〃　　尿酸＝正	〃　80〃	1493/1853	0.040
中性脂肪＝正，　〃　　〃	〃　80〃	1497/1860	0.040
ＨＤＬＣ＝正，　〃　　〃	〃　80〃	1496/1860	0.039
ＧＴＰ＝正，中性脂肪＝正，　〃	〃　80〃	1540/1916	0.040
ＧＰＴ＝正，　〃　　〃	〃　80〃	1635/2035	0.042
〃　　血糖値＝正，ＨＤＬＣ＝正	〃　80%	1529/1904	0.039
〃　　ＧＴＰ＝正，尿酸＝正	〃　80〃	1566/1951	0.039
ＨＤＬＣ＝正，　〃　　〃	〃　80〃	1532/1909	0.039
ＧＴＰ＝正，　〃　，中性脂肪＝正	〃　80〃	1592/1986	0.040

［注］異＝異常、軽＝軽度の異常、正＝正常　の略

表　4　データマイニング法による分析

'87年度・女性・40～69才・［結論→条件］

設　定　条　件	肥満度	比　率	評価尺度
ｱﾐﾗｰｾﾞ＝正，赤血球＝正，尿素窒素＝正	異　26%	115/ 442	0.042
〃　，　〃　，中性脂肪＝正	〃　24〃	100/ 407	0.032
〃　，　〃　，ＨＤＬＣ＝正	〃　24〃	101/ 413	0.032
〃　，ﾍﾓｸﾞﾛﾋﾞﾝ＝正，尿素窒素＝正	〃　23〃	116/ 490	0.034
〃　，白血球＝正，　〃	〃　23〃	120/ 517	0.033
〃　，　〃　，ﾍﾓｸﾞﾛﾋﾞﾝ＝正	〃　23〃	117/ 506	0.033
血圧１高＝軽，ｺﾚｽﾃﾛｰﾙ＝正，HDLC＝正	軽　37〃	11/ 29	0.011
〃　，　〃　，ＧＴＰ＝正	〃　36〃	11/ 30	0.010
ＧＯＴ＝正，ｺﾚｽﾃﾛｰﾙ＝軽　，総蛋白＝正	〃　16〃	36/ 225	0.010
〃　，　〃　，ｸﾚｱﾁﾆﾝ＝正	〃　16〃	36/ 225	0.010
ＧＴＰ＝正，ｱﾐﾗｰｾﾞ＝正，ﾍﾓｸﾞﾛﾋﾞﾝ＝正	〃　13〃	62/ 446	0.011
〃　，　〃　，総蛋白＝正	〃　13〃	66/ 477	0.011
〃　，　〃　，ＧＯＴ＝正	〃　13〃	65/ 472	0.011
〃　，ｱﾐﾗｰｾﾞ＝正，尿素窒素＝正	〃　13〃	63/ 460	0.010
血圧１高＝正，血圧１低＝正，ｱﾐﾗｰｾﾞ＝異	正　82〃	398/ 484	0.041
〃　，ｺﾚｽﾃﾛｰﾙ＝正，ＨＤＬＣ＝正	〃　80〃	503/ 627	0.041
〃　，ＧＰＴ＝正，血糖値＝正	〃　78〃	658/ 835	0.045
〃　，ＨＤＬＣ＝正，　〃	〃　78〃	677/ 862	0.044
〃　，中性脂肪＝正，　〃	〃　78〃	659/ 840	0.042
〃　，ＧＴＰ＝正，　〃	〃　78〃	656/ 837	0.042
〃　，ＧＯＴ＝正，　〃	〃　77〃	686/ 880	0.041
〃　，白血球＝正，　〃	〃　77〃	708/ 909	0.041
〃　，血圧１低＝正，　〃	〃　77〃	705/ 906	0.041
〃　，ｸﾚｱﾁﾆﾝ＝正，　〃	〃　77〃	707/ 909	0.041

［注］異＝異常、軽＝軽度の異常、正＝正常　の略

④分析結果について

a.上述の数表のなかから,「肥満度正常」について正常範囲で関連が高い項目を上げると,次の項目となる。

'93 男性…ＧＰＴ　中性脂肪　血糖値　ＧＴＰ　ＨＤＬＣ　血圧１低　血圧１高
　　尿酸　クレアチニン

'87 男性…血圧１低　血圧１高　中性脂肪　ＧＰＴ　ＧＴＰ　ＨＤＬＣ　尿酸
　　白血球　クレアチニン

'93 女性…ＧＴＰ　ＨＤＬＣ　ＧＰＴ　尿酸　中性脂肪　血糖値

'87 女性…血圧１低　血圧１高　中性脂肪　ＧＰＴ　ＧＴＰ　ＧＯＴ　ＨＤＬＣ
　　血糖値　クレアチニン
　　白血球　コレステロール　アミラーゼ（異）

なお,〔肥満度　軽度〕,〔肥満度　異常〕については,おおむね評価尺度値が低かったので,取り上げなかった。

b.こんごの課題

①マイニング法は四項目以上の多元化ができないので,今回の分析には不向きとも考えられたが,操作が簡便なので実行した。こんご別項の多変量解析との対比を行ってみる。

②データ区分を健診の判定境界ではなく,データの分布を三分するような境界でとらえて,グループ分けを変えてみることも検討したい。

③結論項目として肥満度だけを取り上げたが,こんごなにを設定すべきか考察する必要がある。

④医学的な見地を含む総合的な分析・判断はこれからの課題とした。

3．多変量解析法の適用

大学の学部選択の際の適性検査,商品の売れ行き調査,野球やサッカーなどのスポーツの作戦の立て方から政治問題まで,この世のほとんどのことは単純に割り切ることができず,多元的な見方をしていかねばならない。このようにいろいろな要素がからみあった複雑な問題を解析するよい方法はないものだろうか。このような要望に応じることのできる妙手－これが統計学の一分野であり,最近重要視されだした「多変量

解析法」である。[*3),*4)]

基本的な問題認識としては，この方法を人間ドックの蓄積データに適用して，多くの複雑な測定項目の間になにか単純でユニークな関係を発見できないか…と考えたことにある。

また，この分析の次に予定している遡及分析の準備作業として，死因と測定項目間の関連分析に役立てたいとも考えている。

① '87年度、'93年度のデータに適用

2．に述べたデータマイニング法による分析との整合性をとるために，'87年度，'93年度のデータを用いた。基本的には健康管理センターのデータ整備の都合に合わせたもので，特別に指定したものではない。各年度のデータについては，まず男女の性別に分け，さらに40才から69才までの範囲のデータを抽出して分析を進めた。

測定項目については，次のような21項目を取り上げた。(アルブミンは除外した)

1. 肥満度	2. 血圧１高	3. 血圧１低	4. γＧＴＰ
5. ビリルビン	6. ＬＤＨ	7. ＧＰＴ	8. ＧＯＴ
9. 総蛋白	10. 血糖（空腹時）	11. コレステロール	12. 白血球
13. 尿酸	14. ヘモグロビン	15. 中性脂肪	16. ＨＤＬＣ
17. クレアチニン	18. 尿素窒素	19. 肺（努力）	20. 血清アミラーゼ
21. 赤血球			

それぞれについての正常／軽度の異常／異常の境界値は，健康管理センターが現在採用している数値と同じにした。

②多変量解析法の詳細

a.使用したソフトウェアは，ＳＰＳＳ社製のＳＰＳＳ・Professional Statistics の因子分析プログラムである。

b.分析手順は，汎用ソフトに組み込まれた手順に従った。

---

*3) 柳井晴夫，岩坪秀一共著：多変量解析入門－複雑さに挑む科学，ブルーバックス（1990年）

*4) ＳＰＳＳ（マニュアル）：ＳＰＳＳ Inc.

ア．データのチェック

イ．相関行列の検討…すべての変数を含む相関行列を計算する
　　全分散をチェックする　主成分を大きい方から押さえる

ウ．因子の抽出　…データを表現するために必要な因子数を設定する
　　因子数を少なくする方向で，説明しやすい因子パターンを選ぶ

エ．斜交軸回転　…因子をより解釈しやすくするために変容を行う
　　パターン行列，構造行列を読む

＜イ．からエ．を繰り返す＞

③分析に関連しての設定事項

a.因子数の設定

一般的に全分散（Eigenvalue）が 1.00までを残し，それ以下を切り捨てる方法が行われているので，その方法を原則として採用した。

図1

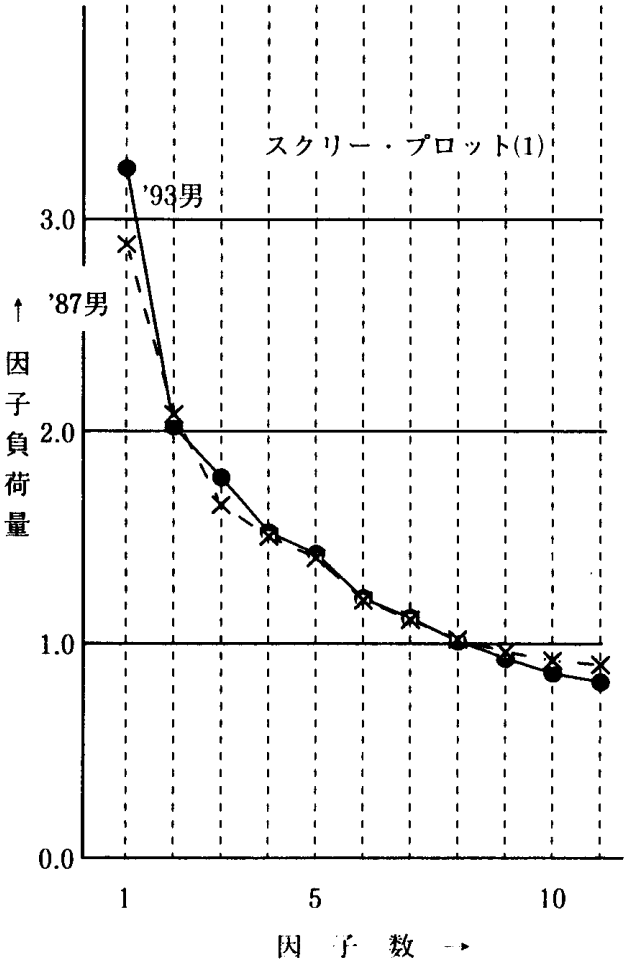
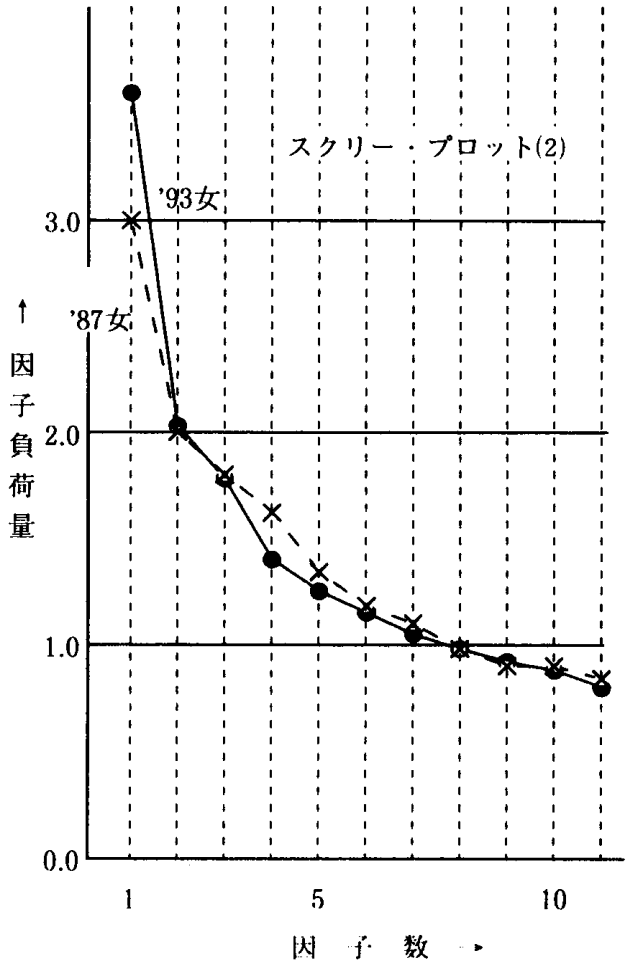


図2



いま一つ，**図1，2**のように各因子が説明している全分散をプロットし，山の麓のくずれ石（スクリー）のように下降線の緩やかになる所で判断するスクリー・プロット図法も試行したが，図に見られるように判定が難しかった。

b.因子数を削減

各年度別で，性別，年齢範囲（40～69才）のデータについて，a.の判定を行ったところ因子数はほぼ7となった。次いで因子数を1個ずつ5まで削減していき，もとの項目との関係がどのように変化していくかを調べた。

因子ごとに各データ項目の負荷量が表示されるが，因子数を変えるたびにその関連ぐあいが変化し，負荷量が±0.3以下に小さくなって項目として消滅することも出てくる。

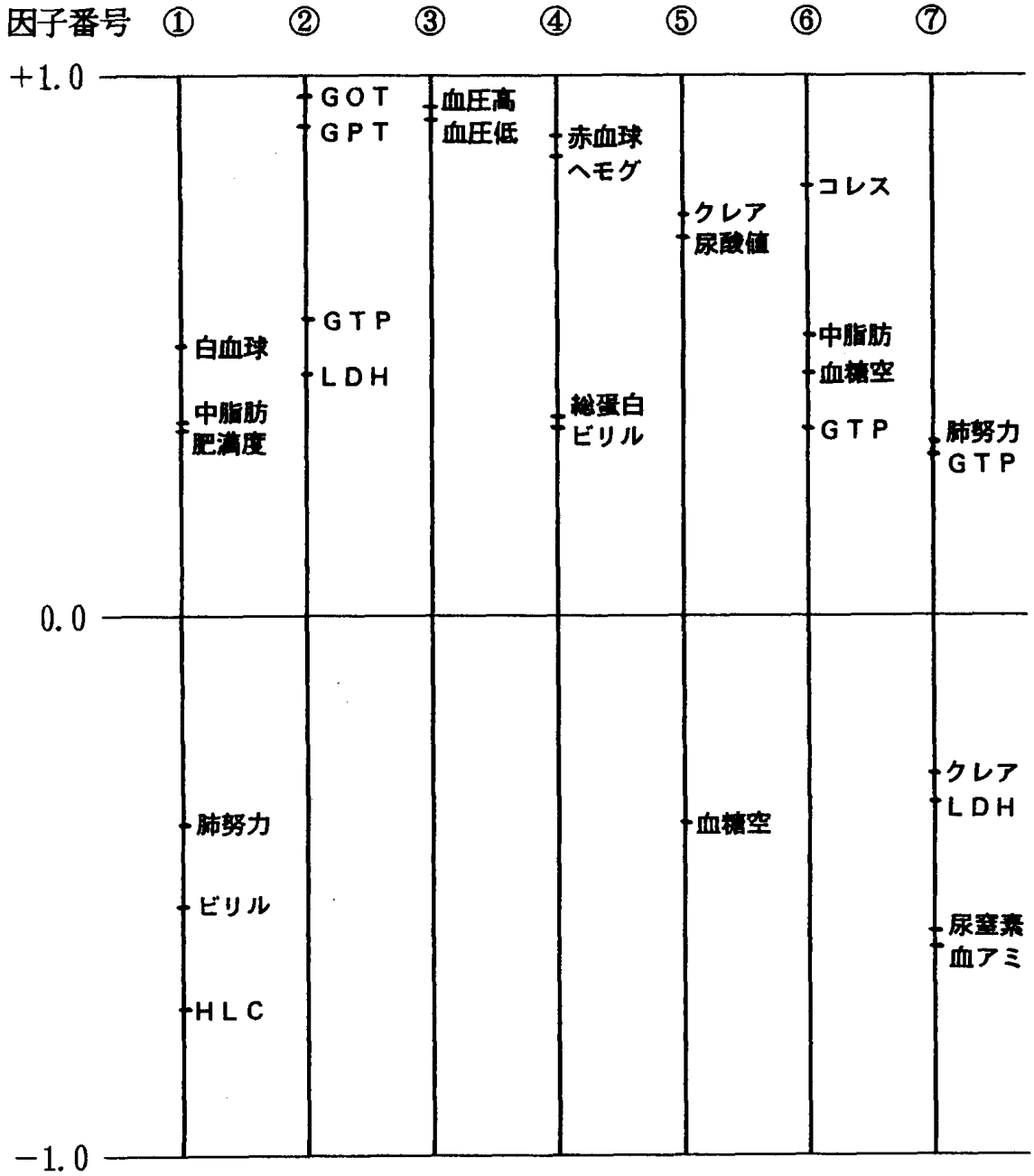
因子数が少ない方が好ましいことには違いないが，データ項目との関連性も大切で，この二つの要素の兼ね合いを吟味しなければならない。とくに医学的な見地からの吟味が重要である。

c.パターン行列の採用

斜交軸回転後はパターン行列と構造行列とが表示されるが，経験的な判断から読みやすいパターン行列を採用した。

図　3　'93年度・男性・40～69才・21項目の パターン・マトリックス［7因子指定］

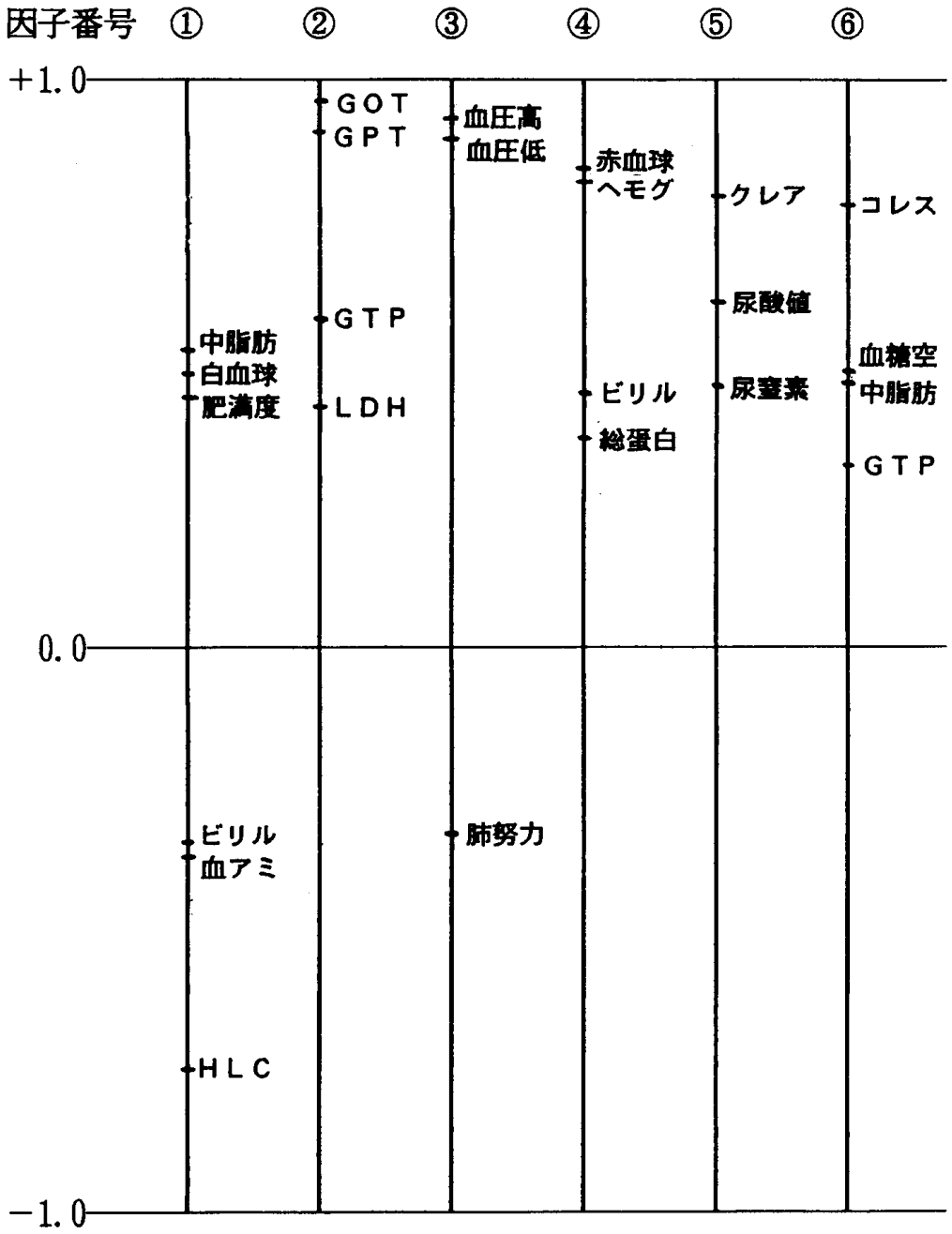
6,235 件　Com.Pct.＝58.7%　8因子指定は収斂せず



(注) 相関行列に強い相関が見える…血圧1高と血圧1低、GOTとGPT、赤血球とヘモグロビン

図 4 ´93年度・男性・40～69才・21項目のパターン・マトリックス［6因子指定］

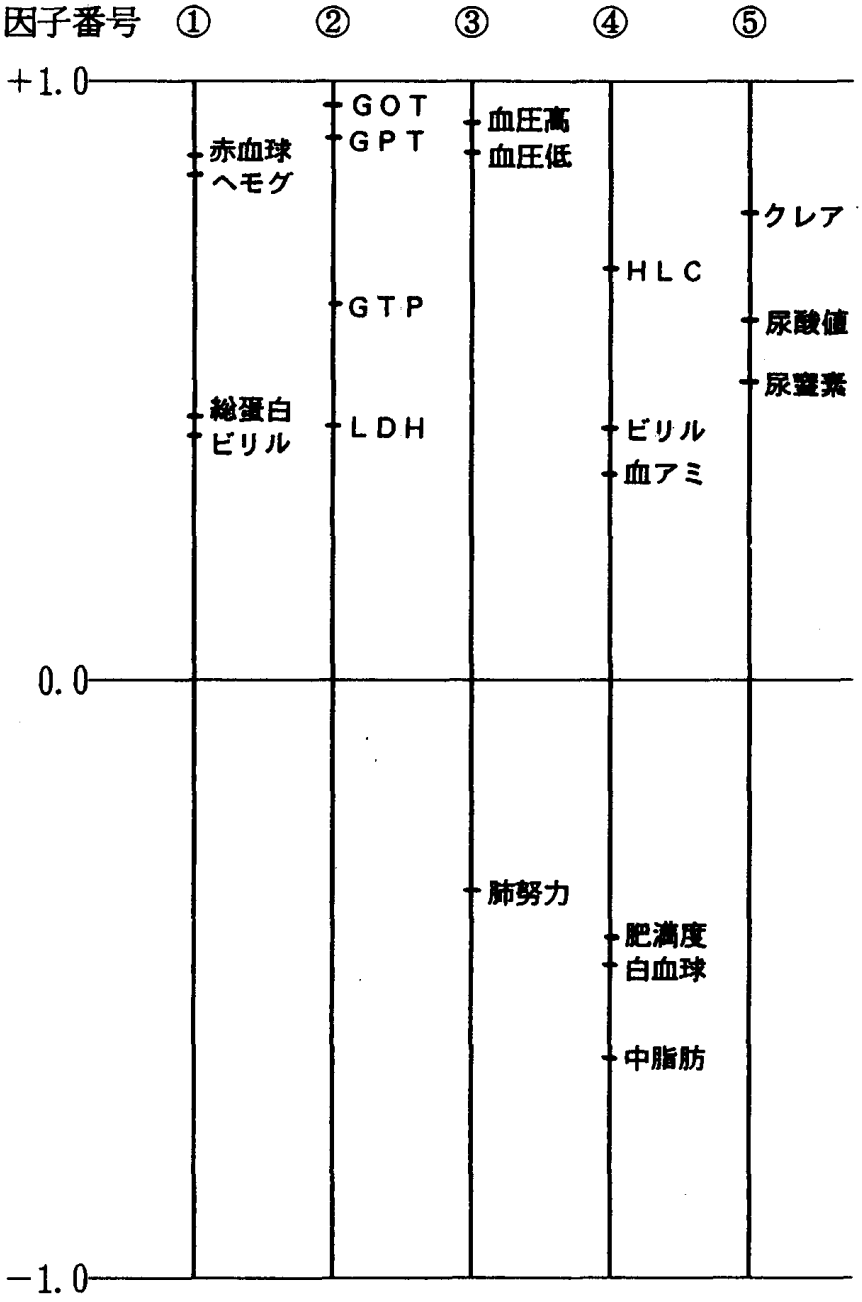
6,235 件 Com.Pct.＝53.4%



(注)

図　5　´93年度・男性・40～69才・21項目のパターン・マトリックス［5因子指定］

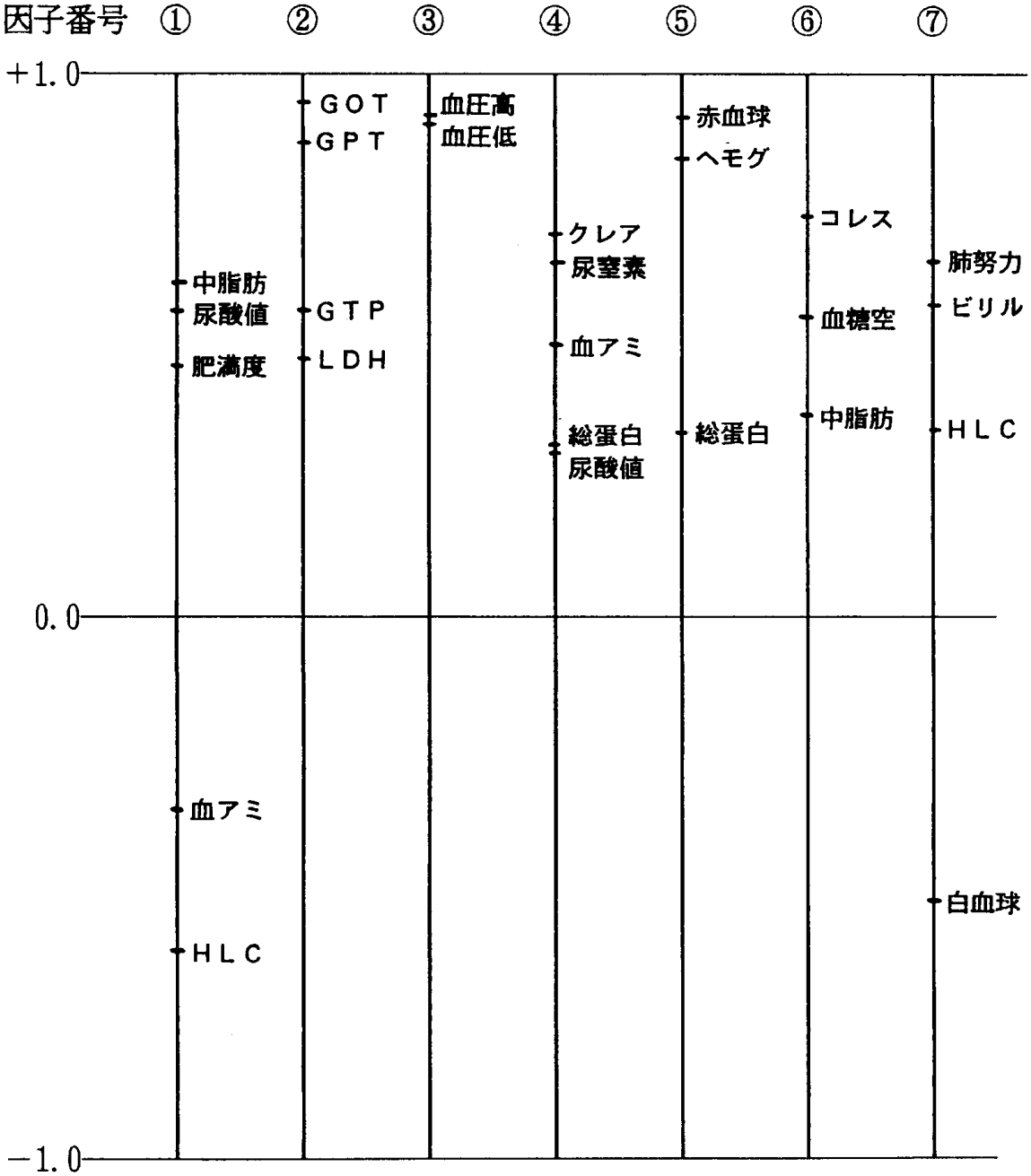
6,235 件　Com.Pct.＝47.7%



（注）コレステロール、血糖値が消えた

図 6 ´87年度・男性・40～69才・21項目のパターン・マトリックス［7因子指定］

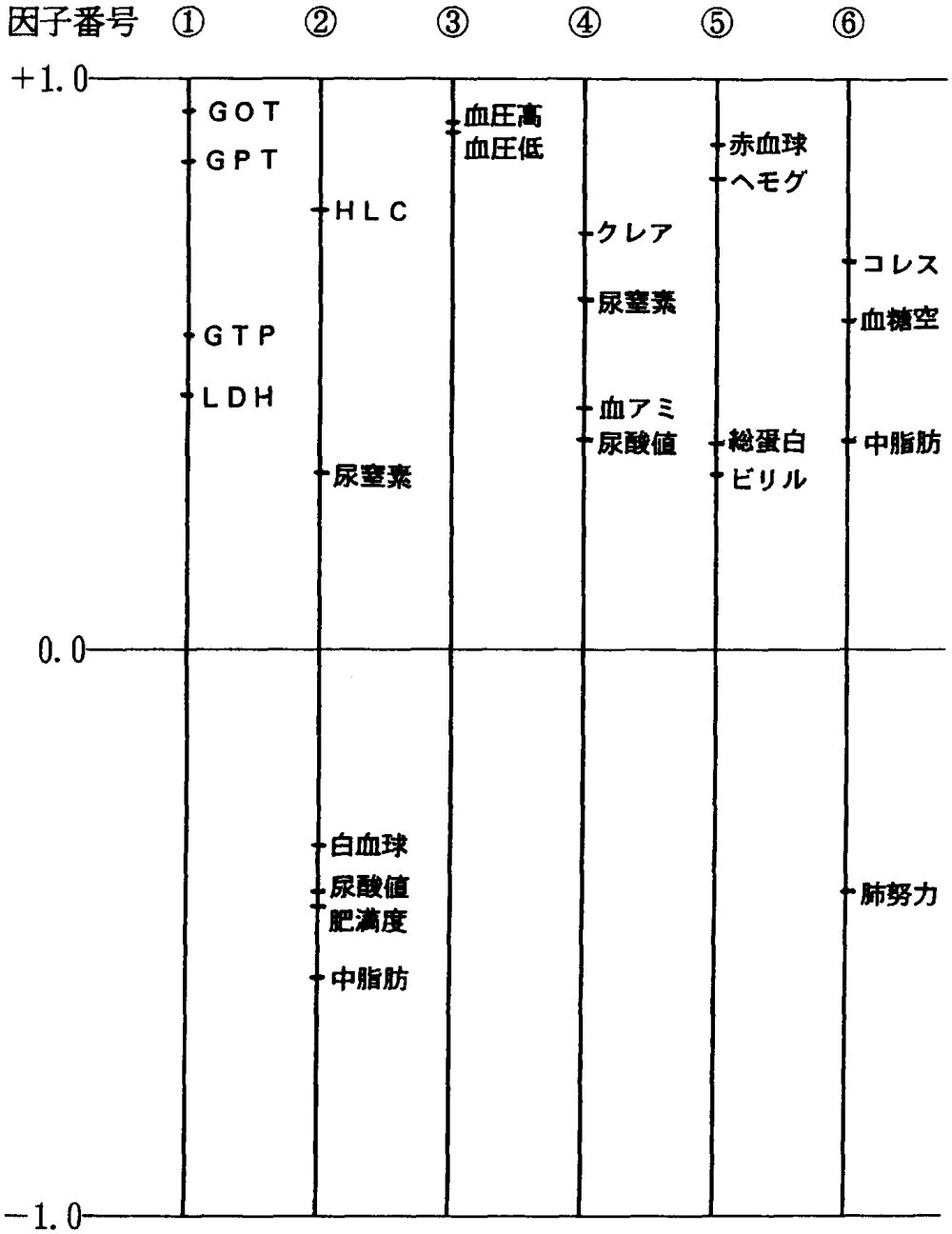
3,199 件 Com.Pct.＝57.0% 8因子は収斂せず



´93男7因子比：⑦？ ② ③ ⑤ ④ ⑥ ①逆

図　7　´87年度・男性・40～69才・21項目のパターン・マトリックス［6因子指定］

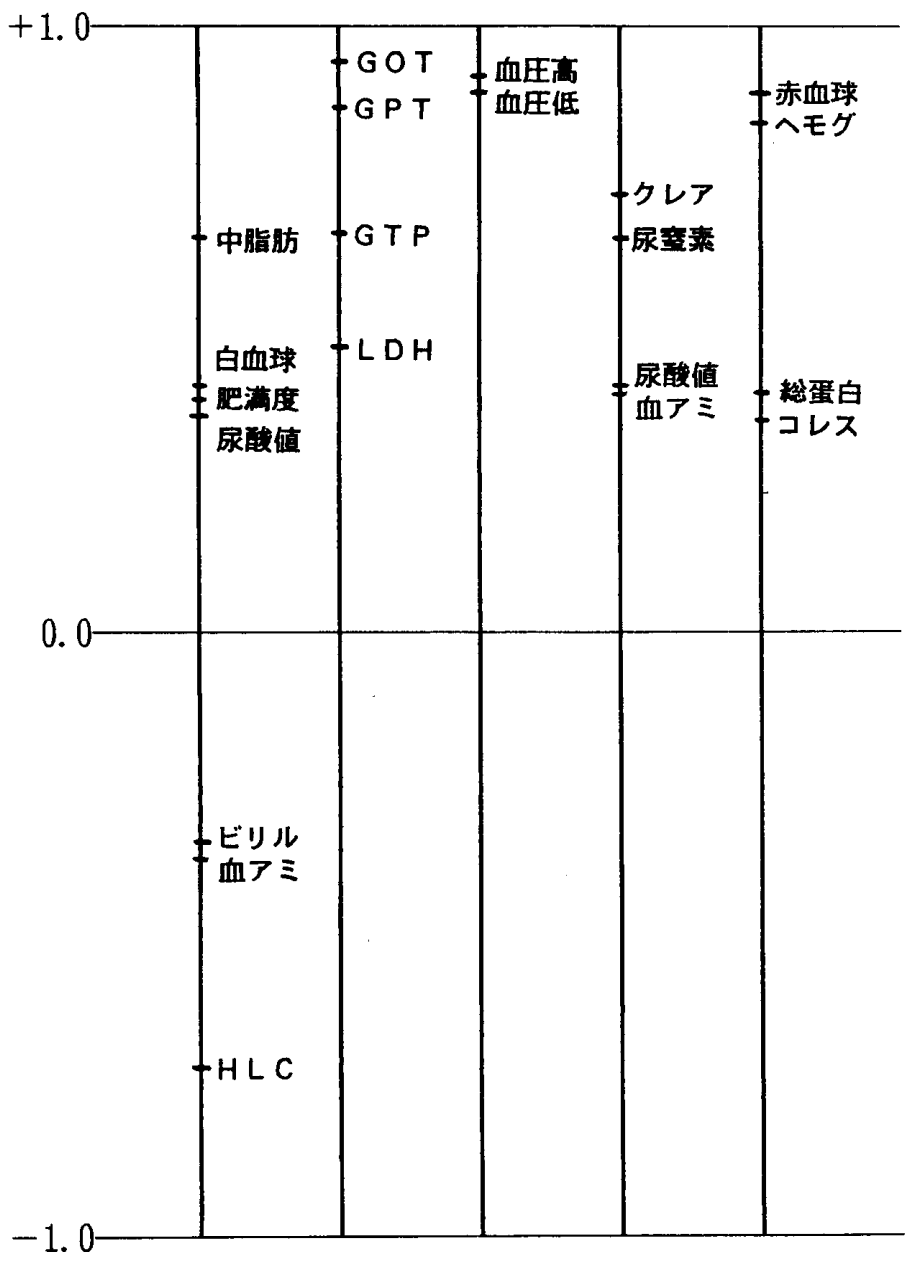
3,199 件　Com.Pct.＝51.6%



´93男6因子比：②　①逆　③　⑤　④　⑥

図 8 ´87年度・男性・40～69才・21項目のパターン・マトリックス［5因子指定］

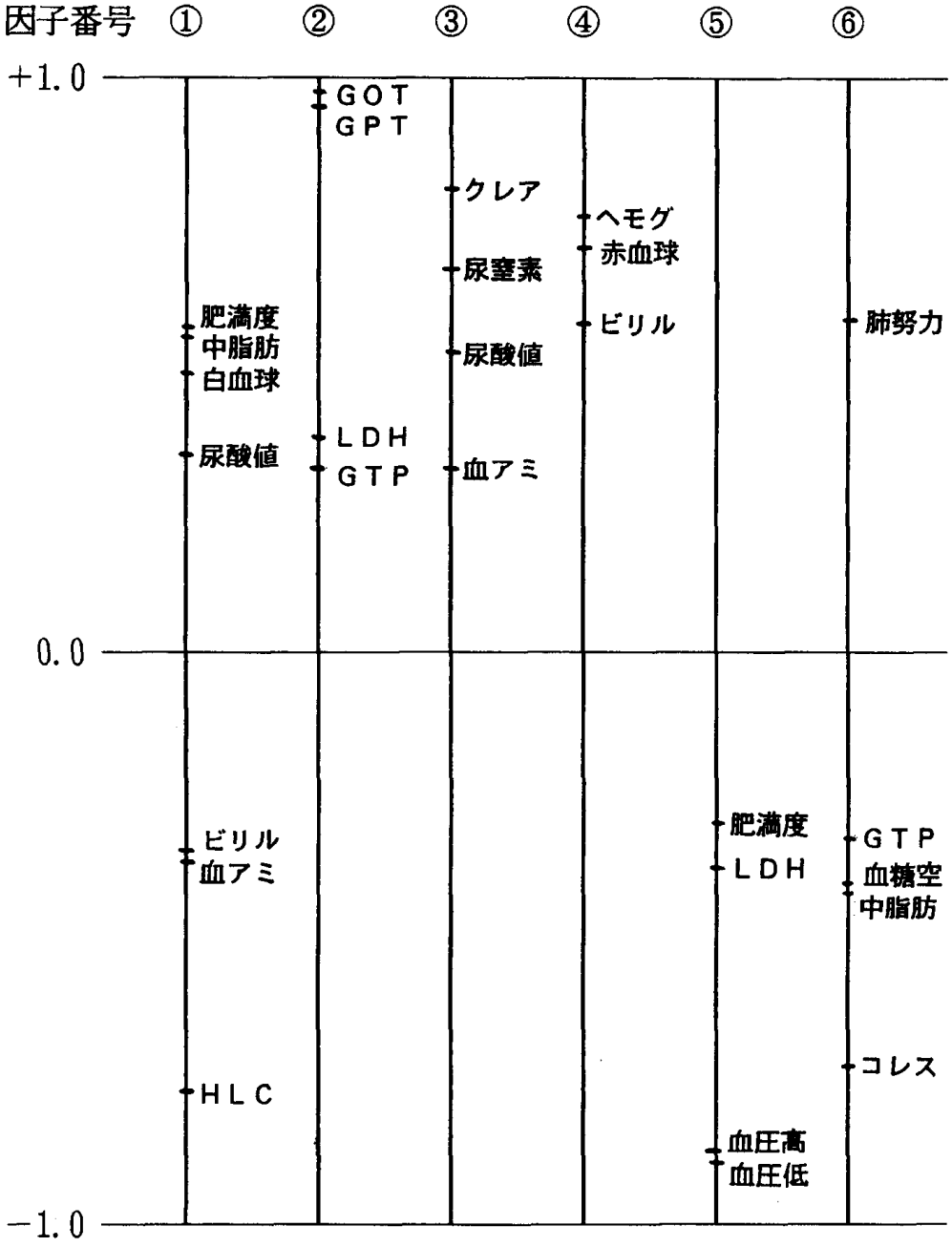
3,199 件 Com.Pct.＝46.0%



´93男5因子比：④逆 ② ③ ⑤ ①

図　9　´93年度・女性・40～69才・21項目のパターン・マトリックス［6因子指定］

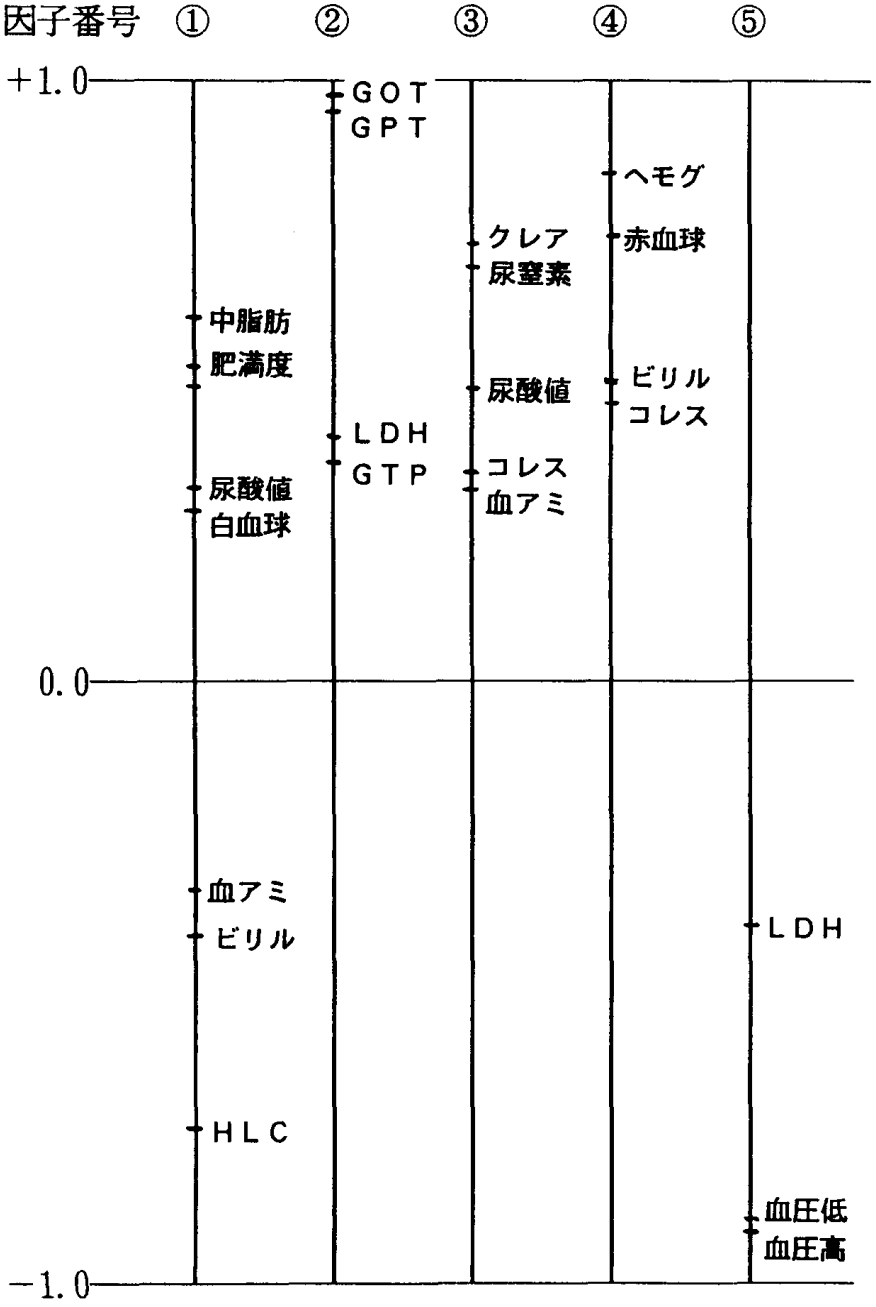
2,511 件　Com.Pct.＝53.4%　7因子指定は収斂せず



（注）

図 10 ´93年度・女性・40～69才・21項目のパターン・マトリックス［5因子指定］

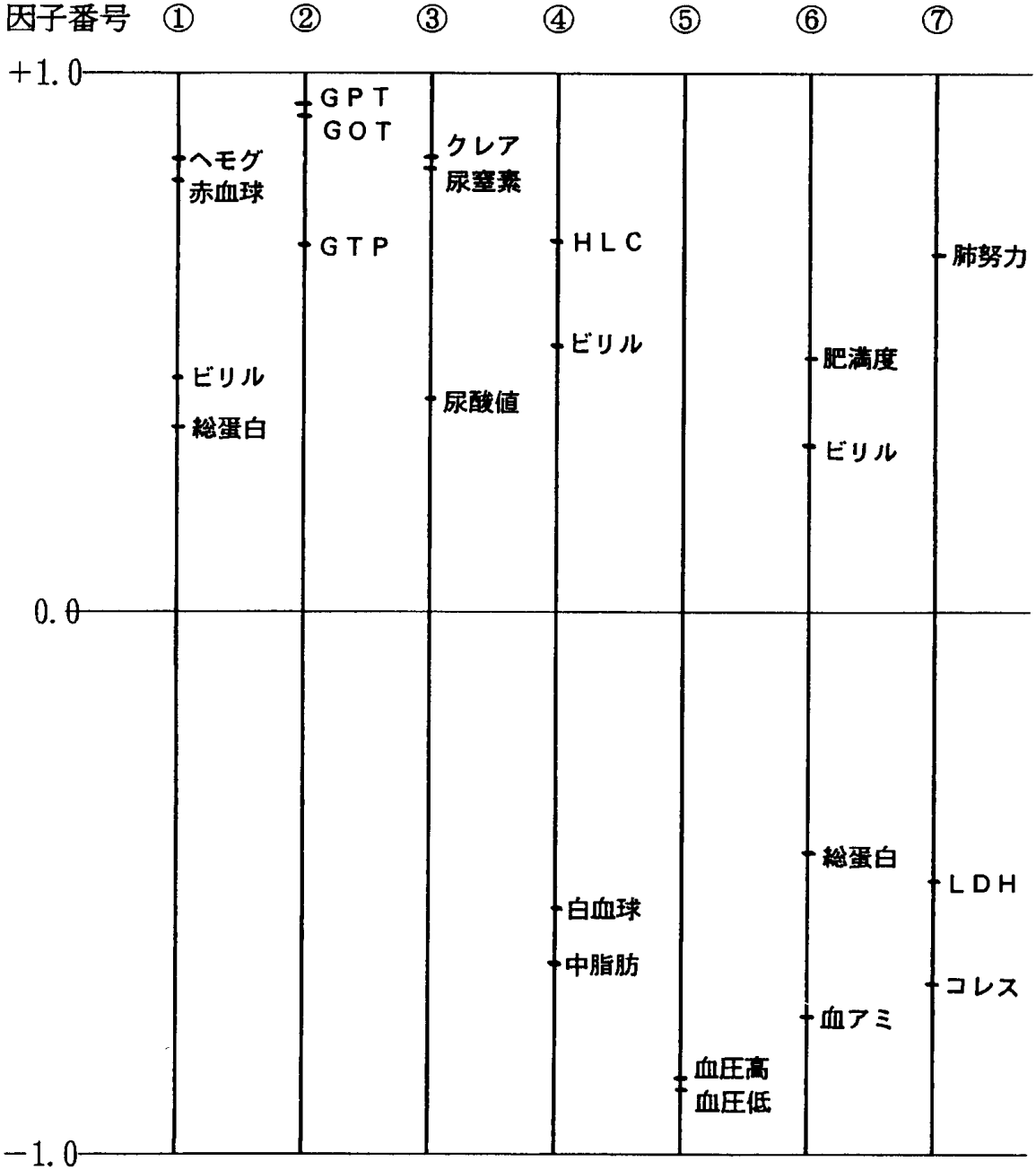
2,511 件 Com.Pct.＝47.9%



（注）血糖、蛋白、肺努力の３項目が消えた

図 11 ´87年度・女性・40～69才・21項目の パターン・マトリックス［7因子指定］

1,253 件 Com.Pct.＝57.7%

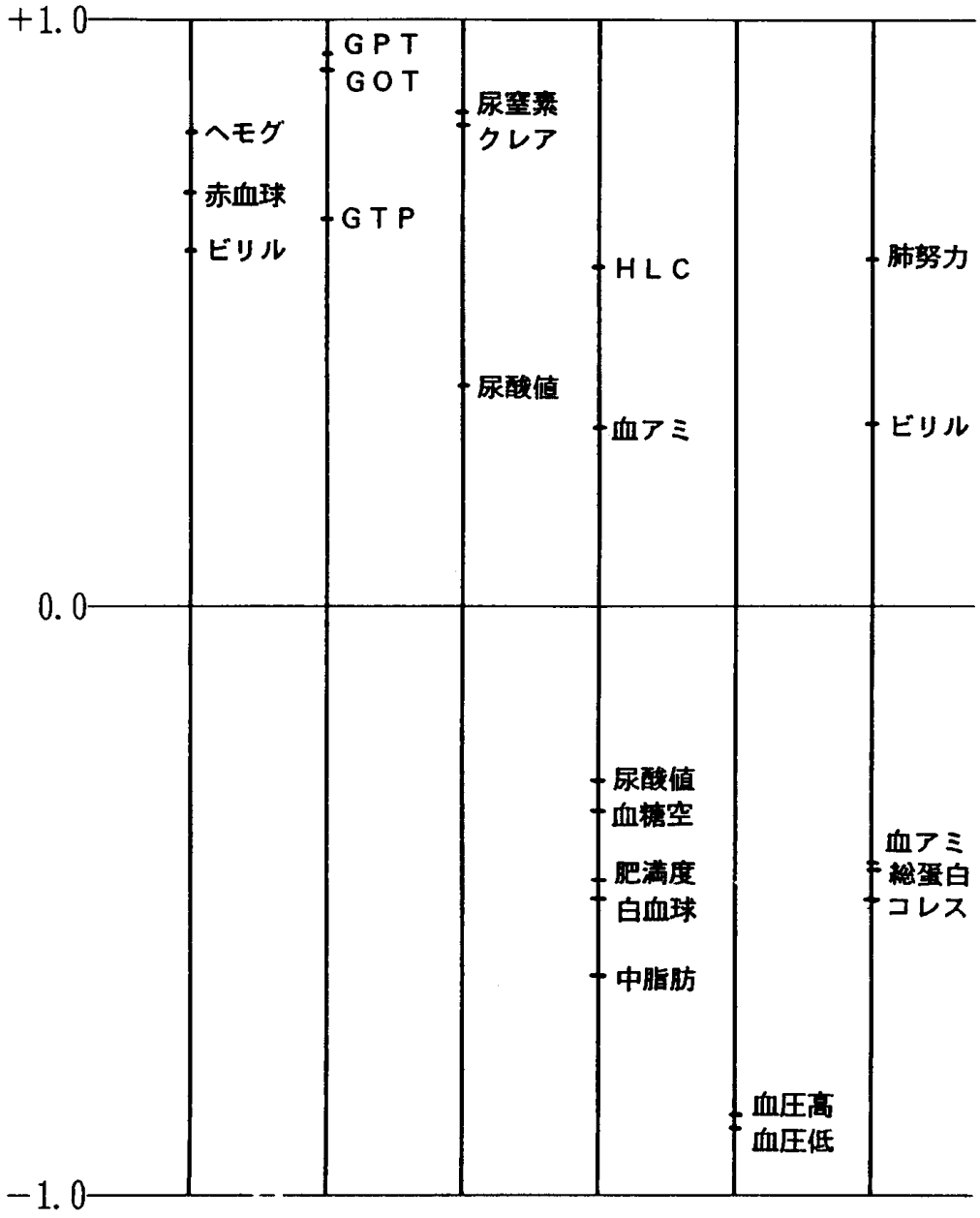


´93女7因子比：

（注）血糖の1項目が消えた

図 12 ´87年度・女性・40～69才・21項目のパターン・マトリックス［6因子指定］

1,253 件 Com.Pct.＝52.6%

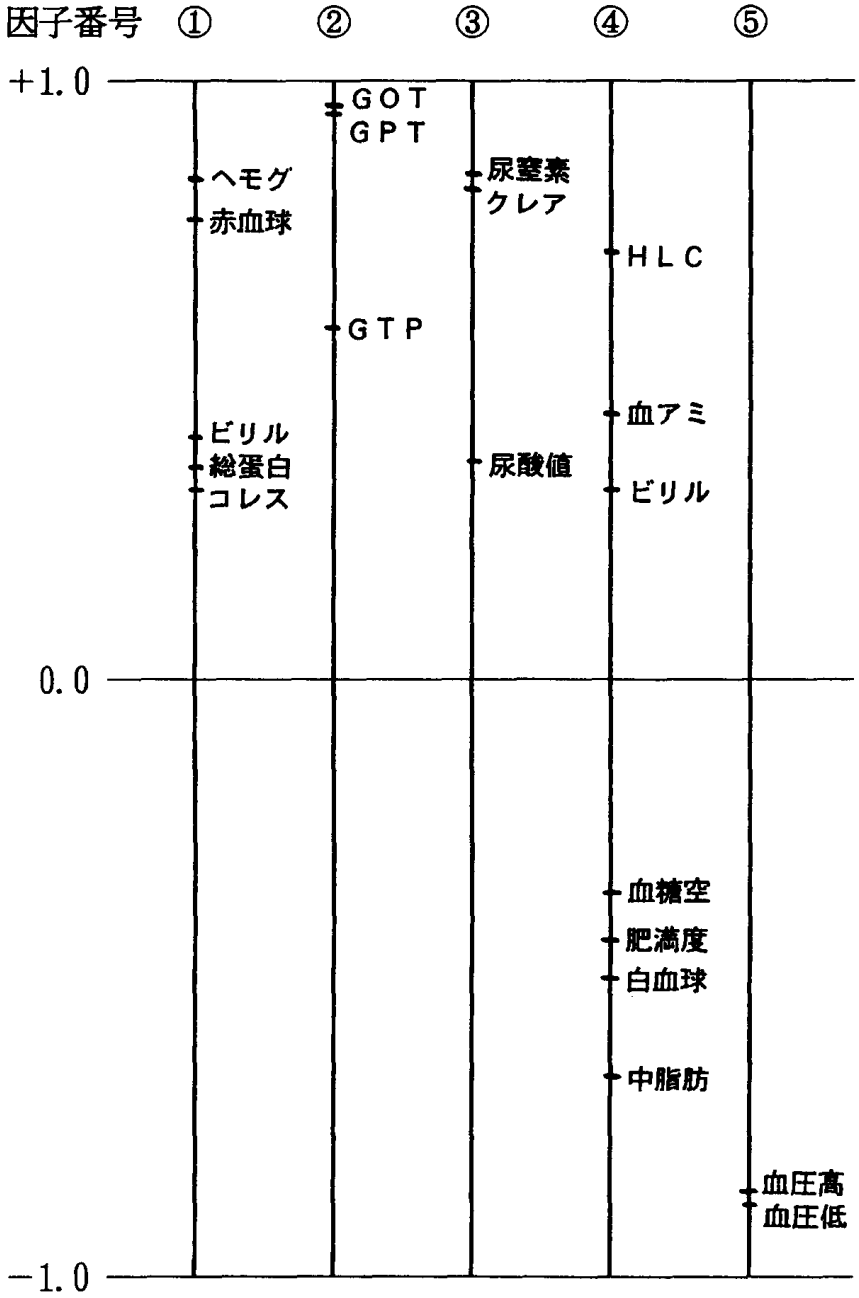


´93女6因子比：

（注）ＬＤＨの１項目が消えた

図 13 ´87年度・女性・40～69才・21項目のパターン・マトリックス［5因子指定］

1,253 件 Com.Pct.＝47.0%



´93女5因子比：④ ② ③ ①逆 ⑤

（注）ＬＤＨ、肺努力の２項目が消えた

d. 総合的判断はこれから

今回の分析は各種手法の適用までとし，医学的な立場を含む総合的な判断はこれからの課題とした。そこで試行した因子数の設定に伴うパターン行列はすべて掲載している。

ここまでの作業は，それぞれが個別の分析・研究であることに加えて，こんごに予定している遡及分析の準備作業であると言える。遡及分析では蓄積データをできるだけ簡潔にまとめておくことが必要であり，その工夫を多変量解析によって行おうとしているわけである。

以　上