

胃X線像の各種判別分析

新村秀一・鈴木隆一郎・中西克己

1. はじめに

本稿で紹介する事例は、昭和49年度から3年間にわたって実施された厚生省がん研究助成金「コンピュータによるがん診療の総合研究」のなかで、筆者らが分担した「がん診断の定量化」研究である。その内容は、医学的には胃がん診断を目的とする胃X線像の計量診断であるが、統計学的には名義尺度の離散変量を説明変数とする判別分析の事例といえよう。

この事例の特徴は次の3点である。

第1点は、材料がある期間にある検診機関が実施した全例のプロスペクティブなデータであり、そのためサンプリングに関する問題から解放されていた。

第3点は、この研究が、医師によって企画されたため、当初、枝分れ論理、学習法などいわば原始的な方法から始まり、よりよい成績を求めて数量化Ⅱ類やベイズ

診断など統計学的に洗練された方法へ移ってきたことである。そのため同一のデータに各種の判別手法が適用され、見方を変えれば各種判別手法の限界を探ることにもなった。

第3点は前項と同じ経緯から、各疾病例の分布の形について考慮することなく判別分析が行なわれたことである。途中からこの研究に参加した筆者は統計学を専門とする立場から、この点に興味をもち、特に今度の事例研究において、主成分分析による疾病分布モデルを導入し、このモデルのもとで判別分析を試みた。しかしながら、医師側がその専門知識（一般的には、固有知識とよんでよい）から作成したモデルと比べて、必ずしもよい成績は得られなかった。

以上述べた流れにしたがって、この事例を紹介し、検討を加える。

2. 「X線像による胃がん診断の定量化」研究の意義

がんのなかで胃がんが最も大きな部分を占めているわが国では、胃X線診断はその最初の検査手技として欠く

しんむら しゅういち 住商コンピュータサービス(株)
すぎき たかいちろう、なかにし かつみ 大阪府立
成人病センター

ひとくちコメント

X線やその他の症状から胃がんか否かを判別するという、いわゆる診断は、集団検診などで多数を扱う場合には、統計学にもとづくコンピュータによる自動診断が有効であることはかなり前から主張されてきたことであるが、この事例はまさにこの自動診断を扱ったものである。

統計学では、ある個体のもついく種かの指標からその個体がどのようなカテゴリに属するかを判別する判別関数の理論や数量化Ⅱ類などの方法が開発されている。しかしこうした技法を現実の問題に形式的に安易に適用しようとしてもうまくゆかないことが多い。これはまさに原住民の心を無視した植民地政策のような

ものである。

臨床医学の場で古くから行なわれている、枝分れ理論による判別とか、学習法による判別など医学固有のものは、それなりの良さがあるはずであり、それらをよく研究したうえで、上記のような統計的方法といかに調和させてゆくかという多くの苦闘なしには真に現実には則した解決は望めないであろう。

この事例は著者らのこうした苦闘から生れてきたものと思われる。こうした精神構造はOR手法の現実への応用のときに最も重要な点と思われる。

なおこの初期の著者らの原稿はこの2倍ほどの分量であったものを、頁数の制限から、現在のものにまで情報圧縮していただいたことを付記しておこう。(T)

胃全変	下壁	壁上	過状	反転	SI	Nema	全壁	右壁	上へ	圧排	憩室	癒着	胃角開大	小腸拡張	内腔狭少	内腔拡大	胃泡縮小	通過障礙	食物残渣	腸管過長	その他				
部位①	全体	噴門	噴下	胃泡	体上	体中	体下	胃角	前庭	前部	幽門	後胃	辺縁	小窩	大窩	吻合部	Niche	欠損	硬直	不整	胃入	痙攣	圧排	圧排	その他
													深さ	深	中	浅	不明	大きさ	大	中	小	小	不明		
Niche	Prff	血状	散在	欠損内	顔門	(A)U	その他	不明	Enf	顔門	不整	瘰癧	不明	高さ	高	中	低	不明	大きさ	大	中	小	小	不明	
Ba	Prff	血状	散在	欠損内	顔門	(A)U	その他	不明	Enf	顔門	不整	瘰癧	不明	高さ	高	中	低	不明	大きさ	大	中	小	小	不明	
欠透	Prff	血状	散在	欠損内	顔門	(A)U	その他	不明	Enf	顔門	不整	瘰癧	不明	高さ	高	中	低	不明	大きさ	大	中	小	小	不明	
性状	Prff	血状	散在	欠損内	顔門	(A)U	その他	不明	Enf	顔門	不整	瘰癧	不明	高さ	高	中	低	不明	大きさ	大	中	小	小	不明	
診断	Ca	推定	I	IIa	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba	Ba

図 1 上部消化管X線診断用紙

べからざるものであり、胃がん診断システムの効率を左右する重要な位置にある。胃がんに対するX線診断を定量化しうれば、診断精度の向上と安定化とが期待され、胃がん診断システム全体の効率向上に貢献しうる。

また、この研究により、どのような画像情報が診断決定に有用であるかを定量的に示しうれば、胃X線像の自動処理・自動診断の研究にとって不可欠な基礎知識を提供しえようと考えた。

3. 研究材料

データとして、「X線ファイル」と「イベントファイル」を使用した。

3.1 X線ファイル

センターの胃集検で、X線診断が行なわれると、図1に示す上部消化管X線診断報告が医師により作成される。これを光学マーク読取装置を介して磁気テープに集録したものがX線ファイルである。表1に昭和45年4月

表 1 検査区分・装置別 胃X線診断件数 (昭和45年4月～昭和47年12月)

検査区分	装置	直接撮影			合計	
		間接撮影	ロールTV	シートTV		暗室
スクリーニング		52,013	15,750	1	0	67,764
精検		12	9,837	12	11	9,872
特撮		0	31	122	0	153
再検		5	56	0	0	61
経過観察		1	2,507	38	1	2,547
合計		52,031	28,181	173	12	80,397

から昭和47年12月までの件数を検査区分・装置別に示した。

より情報量の多い直撮例を用いるのが適当と考え、間接撮影のスクリーニング群52,013例中ロールTVによる精検を受けた9,837例を内部標本に用いて診断モデルの設定を行ない、ロールTVスクリーニング群15,750例を外部標本としてそのモデルの検定に用いた。

表 2 X線医師診断と総括診断 (内部標本と外部標本(カッコ内))

総括診断	早期がん	進行がん	がん疑	悪性否定しえず	胃潰瘍	ポリープ	胃潰瘍癒痕	胃炎	その他	著変なし	合計
医師診断											
早期がん	7(0)	5(1)	0(0)	0(0)	3(0)	0(0)	3(1)	0(1)	0(0)	0(0)	18(3)
進行がん	4(1)	29(6)	0(0)	0(0)	2(1)	1(0)	2(0)	1(0)	0(0)	1(0)	40(8)
がん疑	12(3)	5(0)	19(2)	4(4)	19(4)	5(0)	21(10)	16(12)	4(3)	12(6)	117(44)
悪性否定しえず	2(2)	12(1)	0(0)	67(26)	57(24)	17(5)	103(56)	93(62)	11(2)	64(53)	426(231)
胃潰瘍	0(0)	3(0)	0(0)	3(5)	340(117)	4(1)	98(51)	51(37)	4(2)	47(55)	550(268)
ポリープ	0(2)	0(1)	0(0)	5(3)	4(5)	111(51)	13(5)	24(29)	4(5)	30(24)	191(125)
胃潰瘍癒痕	6(3)	4(2)	1(1)	4(14)	143(91)	22(10)	888(607)	290(217)	37(17)	289(267)	1,684(1229)
胃炎	0(1)	0(0)	0(0)	2(0)	13(8)	8(9)	45(40)	583(733)	8(7)	130(96)	789(894)
その他	1(0)	0(0)	0(0)	0(4)	9(2)	6(3)	14(10)	50(57)	26(36)	43(43)	149(155)
著変なし	0(0)	0(5)	0(0)	0(13)	16(51)	3(29)	66(219)	76(443)	32(56)	5,680(11977)	5,873(12793)
合計	32(12)	58(16)	20(3)	85(69)	606(303)	177(108)	1,253(999)	1,184(1591)	126(128)	6,296(12521)	9,837(15750)

表 3 各種判別分析の結果の四分表による表示

総括診断・自動診断	悪性：悪性	良性：悪性	悪性：良性	良性：良性	判別の中率
(1)	76	525	14	9,222	0.95
(2)	90	599	0	9,148	0.94
(3)	83	111	7	9,636	0.96
(4)	1	87	27	15,635	0.99
(5)	68	391	13	1,759	0.82
(6)	67	505	14	1,645	0.77
(7)	8	314	12	1,209	0.79
(8)	65	429	16	1,721	0.80
(9)	9	203	11	1,320	0.86
(10)	12	217	8	1,306	0.85
(11)	63	389	18	1,761	0.82
(12)	8	212	12	1,311	0.85
(13)	64	367	17	1,783	0.83

3.2 イベントファイル

センターでは胃集検申込者に対し、1人終身1番号を付して識別しているため、X線、内視鏡、細胞診、組織診、病理診などの診断情報や追跡情報の結合が可能である。

また、くり返しこれらの情報が発生してもその間を結合しうる。そこで、スクリーニング受診者が要精検とされた時に始まり、原則として診断が確定した時に終了する「イベント」という考え方を定義した。前述の諸情報を総合して1イベントに1つの診断名を与え、これを総括診断とよぶことにした。識別情報としては、集検カルテ番号、スクリーニング年月日を用いた。これに総括診断を得るに至った根拠、時点、ならびに追跡情報を加えたものがイベントファイルである。

3.3 X線医師診断と総括診断の対応

前述の両ファイルを結合させて、X線医師診断と総括診断とを対応させたのが表2である。総括診断の「がん疑い」や「悪性否定しえず」の大部分は、検査中、検査中断などの例である。

総括診断の「早期がん確定」と「がん確定」とをまとめて「悪性」とし、以外を仮に「良性」として判別することを、以降では「悪性診断」とよぶことにする。医師診断の「悪性否定しえず」以上を「悪性」とまとめて、内部標本での成績を四分表にまとめたものが表3(1)である。

4. 疾病の分布モデルを仮定しない場合

4.1 学習法・モデル1A・モデル1B

本節では、筆者らの当初の研究内容の簡単な紹介を行なう。ORの立場からは、このような摸索段階、失敗例等の概念形式時期にスポットをあてることのほうが重要と思われるが、紙面の都合で概略のみ報告する。

最初に「学習法」とよぶ方法を用いた。すなわち、内部標本で総括診断ががん確定であった90例の所見を記憶し、同一所見を「悪性」それ以外を「良性」とした。結果は表3(2)に示すように、見落としは当然ゼロであるがとり過ぎ率は86.9%となった。そこで、各所見ごとの症例数をかぞえ、その多い順に第5位

までの所見はきわめて情報に乏しいものであったので、この5所見を良性(7悪性例が含まれている)とした成績が表3(3)である。医師診断の結果と比較して、きわめて満足のものであるが、表3(4)に示すように外部標本に対しては、まったく診断力をもっていなかった。

この原因としては、多変量的な考察を欠いていたためと考えられるので、数量化II類を用いることとした。

総括診断を外的基準に、各X線所見をアイテム・カテゴリとして数量化II類を用いることにした。表4に内部標本の診断の確定度を示した。直撮X線診断を問題にしているため、それより上位の手技で確定診断名のついている2231例(表中の中枠でかこんだ部分の合計)をサンプルとした。

このサンプルにつき医師によるX線診断の判定的中率をみたものが表5である。外的基準は、早期がん、進行

表 4 総括診断とその診断手技(昭和45~47年 内部標本)

		診断手技						合計
		直撮X線	内視鏡	細胞診	組織診	手術	病理診	
総括診断								
確定	著変なし	5853	443					6296
	早期胃がん	2				30		32
	進行胃がん	7	4	5		5	37	58
	胃ポリープ	20	96	10	23		1	150
	胃潰瘍	167	324	20	61	2		574
	胃潰瘍瘢痕	464	560	30	51	1		1106
	胃炎	609	492	14	22			1137
その他	22	17	4	6	1		50	
疑診		392	41	1				434
合計		7536	1977	84	143	9	68	9837

表 5 X線医師診断と総括診断 (内部標本より抽出分)

医師診断	総括診断		がん疑	悪性否定しえず	胃潰瘍	ポリープ	胃潰瘍癒痕	胃 炎	その他	著変なし	合計
	早期がん 確	進行がん 確診									
早期がん 確診	6	4			2		3				15
進行がん 確診	4	24			2	1	2	1		1	35
が ん 疑	12	5			18	5	24	17		9	90
悪性否定しえず	2	11			53	17	93	86		57	319
胃 潰 瘍		4			183	4	90	52		41	374
ポ リ ー プ					4	75	19	24		29	151
胃 潰 瘍 癒 痕	5	3			129	16	379	265		251	1048
胃 炎					8	7	24	70		38	147
そ の 他	1				8	5	8	13		17	52
著 変 な し											
合 計	30	51			407	130	642	528		443	2231

医師診断の判定的中率 = (6+24+183+75+379+70) ÷ 2231 = 0.33

がん、胃潰瘍、ポリープ、胃潰瘍癒痕、胃炎、著変なしの7グループとなり、判定的中率は33%となった。しかし、実際にX線診断にたずさわる医師の立場からは、表3(5)に示した悪性度診断が意味をもってくる。

外的基準が7グループあるため、カテゴリー数量は第6次元まで与えられる。この各次元上で、7群の累積百分率を求め、判別境界点を設定した。これを用いて枝分れ論理を構成し判別を行なった。数量化後の枝分れ論理による「モデル1A」は、陥凹性病変と隆起性病変というX線所見の形態に着目したロジックであり、表3(6)に結果を示す。外部標本1543例の結果を表3(7)に示す。

モデル1Bは、早期がんと進行がんとを分けてゆく方向にロジックを組み立てた。内部標本の成績を表3(8)、外部標本の成績を表3(9)に示す。

表3(10)に、外部標本の医師診断の成績を示す。

4.2 数量化II類の外的基準の単純化(モデル2)

表 6 電算機診断の判別点と悪性診断力 (モデル2：内部標本)

判別点の サンプル 数 量	正 診 (True Positive)		誤 診 (False Positive)	
	累 積 症例数	累 積 百分率	累 積 症例数	累 積 百分率
7.20	7例	8.6%	0例	0.0%
2.60	40	49.4	31	1.4
1.40	47	58.0	57	2.7
0.20	60	74.1	208	9.7
0.00	63	77.8	389	18.1
-0.20	70	86.4	954	44.3
-0.40	78	96.3	1722	80.1
-0.60	81	100.0	2015	93.7

前述のモデル1は多群判別のモデルであった。このモデルから得た成績を使用した悪性度診断は医師診断に匹敵する診断力を示したが、X線所見の選択、内部標本の再検討を行ない、モデルを改訂して正診率の向上を企てるためには、状況を複雑化しすぎたきらいがあった。

そこで、悪性度診断に関して状況を単純化することを目的とし、同じ内部標本を悪性群(81例)、良性群(2150例)の2群に判別することを試み「モデル2」とする。表3(11)に内部標本の、表3(12)に外部標本での判別結果を示した。内部標本、外部標本とも判別の中率では医師診断とほぼ一致する成績をえたが、見落とし率ではおよばなかった。

この状況をさらに明らかにするため、内部標本において、この判別境界点を移動させることにより、悪性群のどれだけを悪性と正診(True Positive)しえるか。また、同時に良性群のどれだけを悪性と誤診(False Positive)してしまうかを表6に示した。この考え方を医師診断にも応用して、X線診断上で悪性疾患を考慮している表現の強さの順に、同様の数値を算出して表7に示した。これらをグラフ化すると、図2のように、ROC曲線がえ

表 7 医師診断の悪性度表現と悪性診断力 (内 部 標 本)

X線診断の 悪性度表現	正 診 (True Positive)		誤 診 (False Positive)	
	累 積 症例数	累 積 百分率	累 積 症例数	累 積 百分率
早期がん } 確診	38例	46.9%	12例	0.6%
が ん 疑	55	67.9	85	4.0
悪性否定しえず	68	84.0	391	18.2
良 性 疾 患	81	100.0	2150	100.0

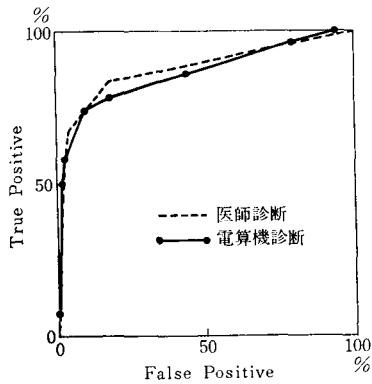


図 2 電算機診断と医師診断の ROC 曲線

られる。TPを増加させてゆくと、ある点から急にFPが増加してしまう傾向が、医師診断に比べて、モデル2でやや強いことが明らかである。

4.3 ベイズ診断

ベイズの定理が数量化II類等に比べ、わが国で利用される実績が少ないのは、そのアルゴリズムの単純さにもかかわらず疾病Dの事前確率P(D)や、所見Sの尤度P(SID)の値を推測するのが困難な点にある。しかし幸いにも、当研究では母集団を忠実に反映すると考えるに足る十分な数の標本集団が得られているので、P(D)やP(SID)を内部標本で得られる頻度比で代用できる。さらにベイズ診断の利点は、単純頻度を基礎としているため、経験との乖離が少ないと考えられる。

ベイズ診断の概略は次のとおりである。疾病群 D_i は早期がん、進行がん等を表わす。所見 S_j は数量化II類に使われたものと同じである。診断は式(1)に示すようにベイズの定理の拡張形を使う。

$$P(D_i|a_1, \dots, a_n) = \frac{\text{分子}}{\text{分母}} \quad (1)$$

$$\text{分子} = P(D_i) \prod_{j=1}^n \{P(S_j|D_i)^{a_j} (1 - P(S_j|D_i))^{1-a_j}\}$$

$$\text{分母} = \sum_{i^*=1}^k [P(D_i^*) \prod_{j=1}^n \{P(S_j|D_i^*)^{a_j} (1 - P(S_j|D_i^*))^{1-a_j}\}]$$

表 8 ベイズ診断による7群判別

疾病群 確率	著変なし	早期 胃がん	胃がん	胃 ポリープ	胃潰瘍	胃潰瘍 癒痕	胃 炎
0.0 ~ 0.20	77	23	27	46	197	147	101
0.20 ~ 0.40	334	5	5	1	7	302	319
~ 0.60	32		1			175	106
~ 0.80			1	2	2	18	2
~ 1.00		2	17	81	201		
合 計	443	30	51	130	407	642	528

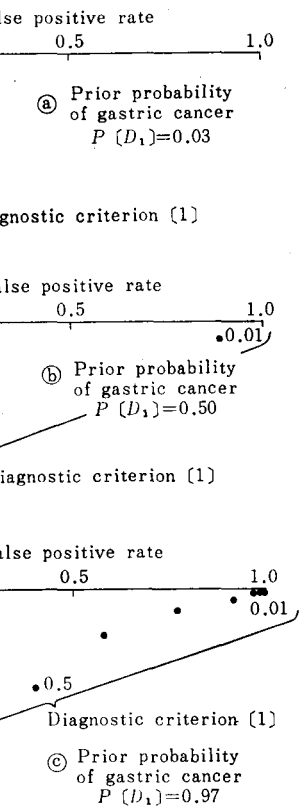


図 3 Receiver operating characteristic curve of X-ray diagnosis for gastric cancer using Bayes' theorem

ただし、

$$a_j = \begin{cases} 1 \cdots \text{所見 } S_j \text{ があり} \\ 0 \cdots \text{所見 } S_j \text{ がなし} \end{cases}$$

n : 所見の数

k : 疾病群の数

ここでは一応疾病群同士が互いに排反的であり、所見同士が互いに独立であることを仮定している。この点の当否に関しては、バッドール分布による修正やシミュレーションによる最適化も提案されているが、この手法が

表 9 ベイズ診断による2群判別

疾病群 確率	悪性群	良性群
0.0 ~ 0.20	24	172
~ 0.40	1	55
~ 0.60	2	46
~ 0.80	3	81
~ 1.00	51	1796
合 計	81	2150

他の手法に比べ操作性が悪いことを考えればこの条件は認めざるをえないと考える。

ベイズ診断で、7群判別、5群判別（早期がんとがん、胃潰瘍癒痕と胃炎をグループとする）、3群判別（早期がんとがん、ポリープと胃潰瘍と胃潰瘍癒痕と胃炎をグループとする）、2群判別を行なった。表8に7群判別の結果を示す。

結果は極端に悪かった。5群判別と3群判別の結果も同様に悪かった。表9に2群判別の結果を示す。また、事前確率を種々に変えてその影響をみることとした。結果を図3にROC曲線で示す。

5. 疾病の分布モデルを仮定した場合

医学における疾病群は、正常群より偏位したおどきのようなものであるから、正常群を地球に、疾病群を山にたとえて「地球モデル」とよぶことをここで提案する。しかし、4章でわれわれが過去に行なった研究では、このような分布にもとづく検討を欠いていた。そこで、本章ではこの点に関して新しく検討を加えた。

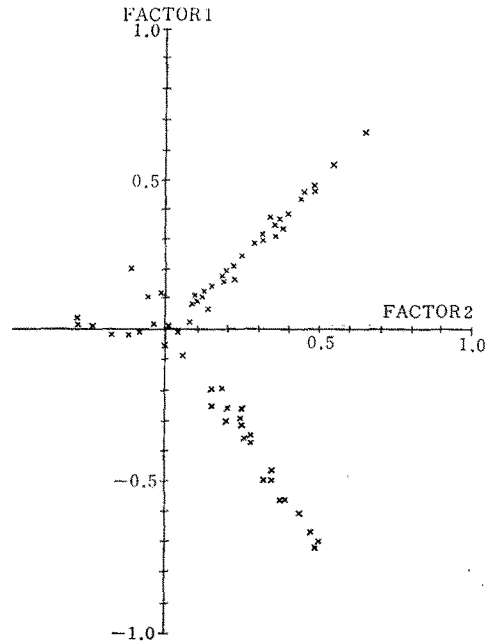


図4 因子負荷量のプロット

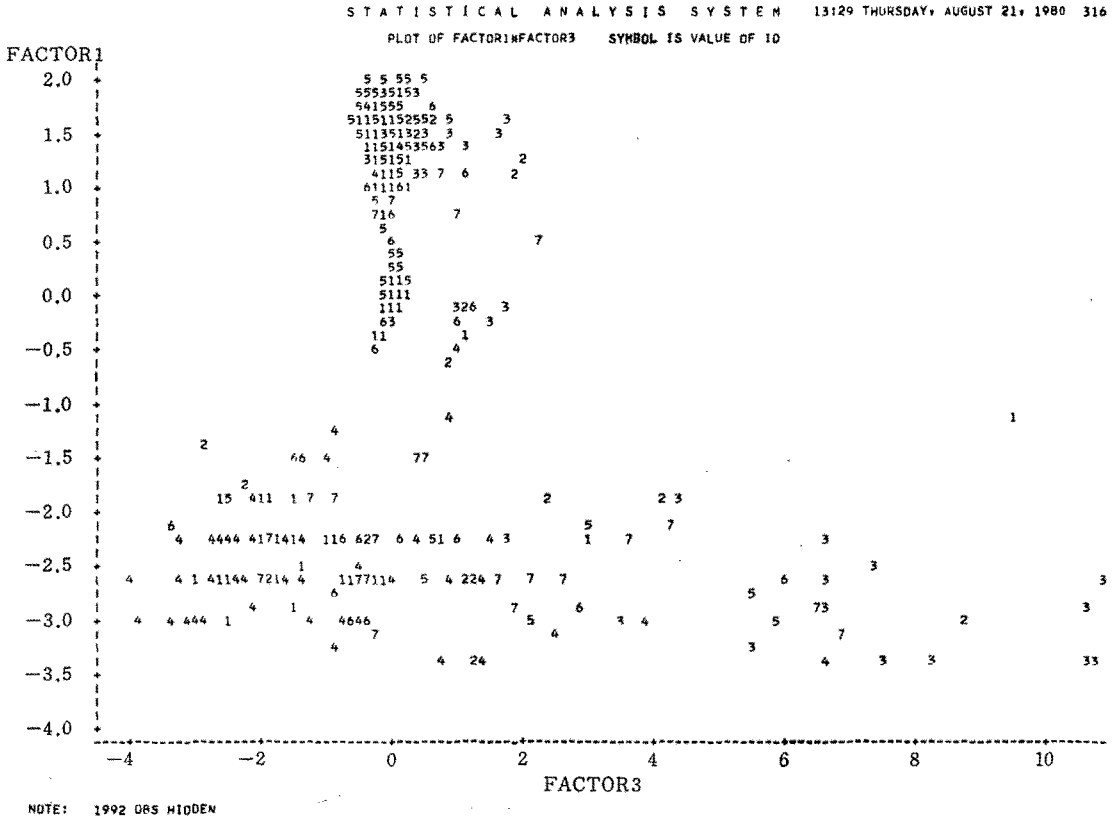


図5 サンプルスコアのプロット

表 10 2次元クロス表

頻 度	第1因子軸 <-0.5	-0.5 ≤ 第1因子軸 ≤0.5	0.5 < 第1因子軸	合計
著変なし	26	381	36	443
早期胃がん	12	12	6	30
進行胃がん	15	16	20	51
胃ポリープ	84	43	3	130
胃潰瘍	11	192	204	407
胃潰瘍癒痕	25	521	96	642
胃炎	36	441	51	528
合 計	209	1606	416	2231

5.1 主成分分析による疾病モデルの導入

4章で用いたアイテム・カテゴリーのすべてを2値型(0,1)のアイテム・カテゴリーに変換し、もとの各アイテムのカテゴリー1に相当するものを省いたデータを作成した。このデータを用いて主成分分析を行なった。変量数は63であり、固有値が1以上のものは第22因子までであり累積寄与率は65.5%であった。

第3因子軸までの累積固有値は、25.1%と少ないが、因子負荷量、サンプル・スコアに有効と考えられる情報が得られたのでこの3軸を中心に分析を行なった。図4に第1、第2因子負荷量を示す。第1象限と第4象限で1時22分の型の翼型の分布をしていることがわかる。各翼は悪性所見を、中心は良性所見とよく対応している。われわれが「地球モデル」とよぶデータの因子負荷量は、このような特徴をもつものを意味している。

この結果を反映して、サンプルスコアの第1、第2因子軸上の分布も翼型となる。図5は、第1因子軸と第3因子軸上のサンプルスコアの分布を示すが、第1因子軸の-0.5と0.5を境に3つのクラスターに分れていることがわかる。第1、第2因子軸上にもどって考えれば、-0.5と0.5の間にあるクラスターは中心に、それ以外は翼に対応していることがわかる。そこで、データをこの3クラスターに分け、7群との2重クロス表(表10)を作成した。

著変なし、胃潰瘍癒痕、胃炎はクラスター2(-0.5 < 第1因子軸 ≤ 0.5)に対応する割合が80%を越えている。残りの4群もそれぞれクラスター2に対応する割合が30%から40%と高いが、早期がんはクラスター1、がんはクラスター3とし、ポリープはクラスター1、胃潰瘍はクラスター3とよく対応している。このことから、クラスター1はポリープをよく表わしているが早期がんとかんの割合も多い。クラスター3は胃潰瘍とよく対応しているが、がんの割合も多い。クラスター2は著変な

表 11 事前確率を変えた場合の判別成績の結果
(左側：電算機，上側：総括診断)

疾病群 事前確率	クラスター1		クラスター2		クラスター3	
	(0.5, 0.5)	25	7	8	11	17
(0.7, 0.3)	2	175	20	1567	9	367
	25	9	13	234	19	36
(0.9, 0.1)	2	173	15	1344	7	354
	25	13	24	1033	19	53
(0.98, 0.02)	2	169	4	545	7	337
	26	21	28	1563	25	112
	1	161	0	15	1	278

しとよく対応しているが他の疾病群の症例も多く、軽度な疾病がデータの著変なしと見分けのつきにくいことをうかがわせる。以上から、2峰性の地球モデルを考えることができるが、割合の多少はあれ、どの疾病群もすべてのカテゴリーに表われるので判別の困難性が予見できる。

5.2 判別関数による結果

各7群とも翼形の分布をしているので、仮りに第1因子軸でカテゴライズした3グループごとに判別を行なうことにした。

判別成績は、Fisherの線形判別関数による結果を掲載し、最近隣法による判別や多群判別の結果を割愛する。

層別したため、症例数の関係で、クラスター1(209例)では、 x_1, x_2, x_4 から $x_{16}, x_{20}, x_{25}, x_{45}$ から x_{63} までの36変数を用いた。クラスター2(1606例)では、 x_1, x_2, x_4 から x_{17} までの16変数、クラスター3(416例)では、 x_1 から x_{44} までの44変数を用いた。

表11に、上から順に、悪性群と良性群の事前確率を(0.5, 0.5), (0.7, 0.3), (0.9, 0.1), (0.98, 0.02)とした場合の結果を示す。上段の結果は尤度方式、それ以外はリスク方式とよばれる結果である。

表3(11)のモデル2と比較するため、クラスター1では事前確率を(0.98, 0.02)、クラスター2では(0.7, 0.3)を、クラスター3では(0.98, 0.02)とすれば、表3(11)の成績が得られる。この成績は、モデル2に比べFPもFNもともに少なく、判別の中率は83%となり1%改善されたことになる。

6. 医学知識にもとづくモデルを設定した場合

本章では、5章でデータ解析的に導入した分布モデルを考える以前に、医師がその専門知識にもとづいて作りあげたモデルに沿って話を進める。

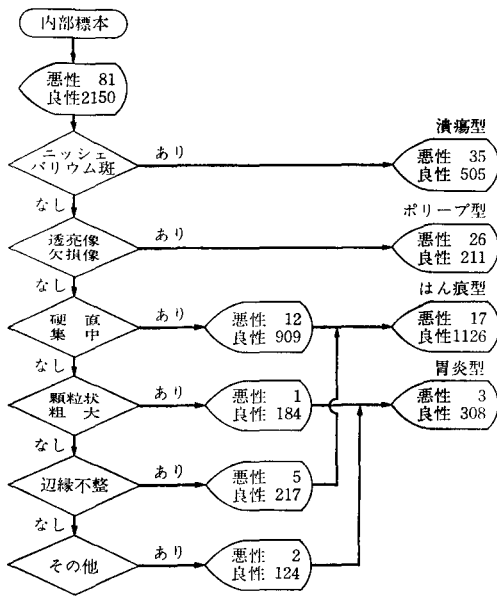


図 6 内部標本を 4 型に分ける流れ図

すなわち、4 章までは、がんを病型別に分けなくて一括してとり扱ってきたが、本章では、標本症例を X 線所見が示す形態上の特徴にもとづいて、いくつかの層に分類し、その各層内で、さらに悪性群と良性群の 2 群に分けるために林の数量化 II 類を使用した。

形態上の特徴により分類する方法としては、「学習法」を行なう以前に枝分れ論理を用いた「病型診断モデル」を開発した。このモデルは、X 線所見の各組合せと疾病症例数を直接に検討することにより、特定の所見群がきわめて強く特定の診断に対応しており、その症例数の 80% 以上が 1 つの診断に結びついているものを拾ってゆくと、顆粒状レリーフは胃炎と、レリーフ集中や硬直は胃潰瘍癒痕と、透亮はポリープと、ニッセンバリウム貯留像は胃潰瘍との関係があることがわかった。図 6 に、その論理を流れ図にして示した。簡単な枝分れであるが、良性疾患での病型診断については、X 線読影医師の診断を 70% 以上再現しうることがすでに確かめられている。

さらに図 6 には、これまでも使用してきた内部標本 2,231 例を流した場合の症例の分け方を、総括診断の悪性・良性別に記載した。このようにして、内部標本を 4 層に分け、おのおの「潰瘍型」「ポリープ型」「癒痕型」「胃炎型」と仮りに名づけた。この各層を、さらに数量化 II 類により悪性群と良性群の 2 群に判別する仮定のもとに算出されたカテゴリー数量を表 12 に示した。表 12 の空欄部分は、その所見の存在頻度が 2% 以下と少なかったため、その項目をモデルから削除したことを示している。

おのおのを通常の判別境界点で 2 群に判別した成績

を、「モデル 3 A」として、表 13 の中欄に示した。判定的中率は、癒痕型を除き、一応、十分な値となっているが、左欄の X 線読影医師診断に比べて、悪性症例の見逃しがわずかに多くなっている。また、癒痕型ではとりすぎが約 2 倍も多くなっており、内部標本合計の判定的中率を低下させる原因になっている。

潰瘍型とポリープ型で見逃し症例を減じ、癒痕型でとりすぎ症例を減ずる方向に、適当に判別点を調節した成績を「モデル 3 B」として、表 13 の右欄に示した。このモデルでの内部標本合計は、医師診断より、わずかに 1 例多く見逃したのみで、とりすぎは 32 例少なくなり、医師診断をしのぐ判定的中率を示した。

医師ならびに各モデルによる外部標本各層の診断成績も表 13 に併記した。外部標本合計の判定的中率も良好な成績で、医師診断をしのぐものになっている。その反面、見逃し症例は、医師診断に比べてやや多くなっている。しかし、その原因が、癒痕型にあることは明らかである。

7. 考 察

医療の日常業務の場でプロスペクティブに集められた全例調査によるデータを使用した胃 X 線像の各種判別分析の事例を紹介し、疾病の分布モデルを仮定した場合に関して、特に追加検討した。

用いた手法としては、枝分れ法、学習法、数量化 II 類、ベイズ診断、多重ロジスティックモデル、主成分分析、線形判別関数、最近隣法による判別分析、クラスター分析、クロス集計等である。結果について特にここで述べる意味のないものは割愛した。

結論としては、表 14 に示すように、①多群判別における成績が悪い。②データの分布モデルを仮定したほうが仮定しないよりは成績はよいが、医学知識にもとづく成績にはおよばなかった。

判別分析手法を開発する側での今後の課題としては、多群判別の成績の改善を計ることと、固有知識にもとづく援助が得られない場合でもデータ解析からの分布モデルを使用することにより成績の改善を計ることにあると思われる。

(本事例研究の一部は厚生省がん研究助成金研究課題番号 55-17「映像によるがん診断技術の開発」(班長飯沼武)の助成を受けた。)

文 献

- [1] 中西克己, 他: コンピュータによるがん診療の総合研究会議資料, 昭和 49 年 6 月。
- [2] 同 上 昭和 49 年 12 月。

表12 林の数量化第Ⅱ類「モデル3」によるカテゴリー数量

	アイテム	カテゴリー	潰瘍型	ポリープ型	癩痕型	胃炎型		アイテム	カテゴリー	潰瘍型	ポリープ型	癩痕型	胃炎型
全 般 変 化	1. 胃角開大	1. なし	.0232	-.0106	-.0941	-.0243		Niche, Ba たまりの 性状分析	20. E n f	1. 不明	.1434		
		2. あり	-.1639	.3472	.6743	.4790							
	2. 小弯短縮	1. なし	.2016	-.0037	-.3014	-.0254			21. R 集中	1. 不明	-.1390		
		2. あり	-.1384	.1212	1.5811	.4401							
見 所 見 辺 縁 所 見 レ リ ー 所 見	3. 内腔狭少	1. なし						22. 深 さ	1. 不明	.1404			
		2. あり											
	4. 残胃液	1. なし	.0024	.0468	.0665	-.0072		23. 数	1. 不明	.8846			
		2. あり	-.0351	-1.8030	-1.4536	.1424							
	5. Niche	1. なし	-.0778					24. 大 小 じ	1. 不明	-.9708			
		2. あり	.0750										
	6. 欠 損	1. なし	-.0474	.2229					25. P r f	1. 不明	-.0430		
		2. あり	1.6582	-1.2864									
	7. 硬 直	1. なし	-.1411	.0582	.1415			26. 表 面	1. 不明	.0593			
		2. あり	.7881	-1.0904	-.1193								
8. 不 整	1. なし	-.0606	-.0394	-.2278	.2395		27. E n f	1. 不明	.0373				
	2. あり	.2636	.4048	.3090	-1.1146								
9. 弯 入	1. なし	-.0232		-.0313	.1820		28. 芽	1. 不明	-.0807				
	2. あり	.9100		.8010	-1.9947								
10. けいれん	1. なし	.0055			-.0842		29. 高 さ	1. 不明	.3120				
	2. あり	-.1702			.5873								
11. B a 貯	1. なし	-.1246					30. 数	1. 不明	.0403				
	2. あり	.0613											
12. 透 亮	1. なし	-.0231	.1776				31. 大 小 じ	1. 不明	-.2418				
	2. あり	.8689	-.0129										
13. 粗 大	1. なし				-.0501		相 関 比		.5663	.7735	.1143	.1369	
	2. あり				.1447								
14. 不 整	1. なし	-.0064	-.0073	-.0820	.2713								
	2. あり	.0435	.2101	.2456	-.3187								
15. 顆 粒 状	1. なし	-.0038	.0109	-.1118	-.3742								
	2. あり	.1662	-.3585	1.4853	.6837								
16. 集 中	1. なし	.0738	.0127	.0153									
	2. あり	-.1593	-.5898	-.0193									
17. 消 失	1. なし			-.0233	-.0087								
	2. あり			.6953	.3283								
Niche, Ba たまりの 性状分析	18. P r f	1. 不明	-.0376										
		2. 皿状	.0571										
		3. 嵌込	2.3195										
		4. 欠損内	1.6900										
		5. 類円	-.1071										
		6. くさび	-.0725										
		7. その他	-.0688										
19. 周 囲 R	1. 不明	-.1641											
	2. 顆粒状	-.0545											
	3. 不 整	.1311											
	4. 透 亮	.6149											
	5. 消 失	.2604											
	6. N P	-.0088											
	7. その他	-.3130											

表 13 層化 2 群判別分析の成績 (判定的中率)

総括診断 電算機 診断		X線読影医師診断				モデル 3 A				モデル 3 B			
		内部標本		外部標本		内部標本		外部標本		内部標本		外部標本	
		悪性	良性	悪性	良性	悪性	良性	悪性	良性	悪性	良性	悪性	良性
潰瘍型	悪性	32	121	6	59	29	73	5	35	32	124	6	57
	良性	3	384	0	208	6	432	1	232	3	381	0	210
		(0.770)		(0.784)		(0.854)		(0.868)		(0.765)		(0.791)	
ポリ- プ型	悪性	25	53	1	25	24	21	2	15	26	36	2	23
	良性	1	158	3	108	2	190	2	118	0	175	2	110
		(0.772)		(0.796)		(0.903)		(0.876)		(0.848)		(0.818)	
はん痕 型	悪性	10	179	3	98	11	330	2	235	7	156	0	87
	良性	7	947	4	751	6	796	5	614	10	970	7	762
		(0.837)		(0.881)		(0.706)		(0.720)		(0.855)		(0.890)	
胃炎型	悪性	1	38	2	35	2	43	1	28	2	43	1	28
	良性	2	270	1	239	1	265	2	246	1	265	2	246
		(0.871)		(0.870)		(0.859)		(0.892)		(0.859)		(0.892)	
合 計	悪性	68	391	12	217	66	467	10	313	67	359	9	195
	良性	13	1759	8	1306	15	1683	10	1210	14	1791	11	1328
		(0.819)		(0.854)		(0.784)		(0.791)		(0.833)		(0.866)	

- [3] 同 上 昭和50年 6月.
- [4] 同 上 昭和50年12月.
- [5] 同 上 昭和51年 6月.
- [6] 同 上 昭和51年12月.
- [7] 稲本一夫, 他: 電算機による上部消化管X線診断情報処理, 日本医放会誌, 30(9), 1-10, 1970.
- [8] 中西克己, 他: 電算機を利用した胃X線悪性度診断の試み, 成人病, 15(3), 16-30, 1975.
- [9] 鈴木隆一郎, 他: 胃X線像自動診断論理の開発, 成人病, 14(3), 72-79, 1974.
- [10] 新村秀一, 他: 自覚症問診項目の統計的評価, 第15回日本ME学会大会論文集, 28(1976).
- [11] 平岡力, 他: 自覚症状の組合せによる胃がん判別の限界, 癌の臨床, 23巻14号, 1321/1327(1977).

表 14 林の数量化第II類による悪性度診断モデルの総括

モ デ ル	内部標本			外部標本			
	見逃し 率	取過ぎ 率	判定的 中率	見逃し 率	取過ぎ 率	判定的 中率	
多 群 判 別 分 析	モデル 1 A	.173	.235	.767	.600	.206	.789
	モデル 1 B	.198	.213	.801	.550	.133	.861
2 群 判 別 分 析	モデル 2	.222	.181	.818	.600	.139	.855
層化 2 群判別分析	モデル 3 A	.185	.217	.784	.500	.206	.791
	モデル 3 B	.173	.167	.833	.550	.128	.878
2 群 判 別 分 析	地球モデル	.210	.171	.828			
X 線 読 影 医 師 診 断		.160	.182	.819	.400	.142	.854

- [12] 新村秀一, 他: 胃X線像の電算機診断の試み, 第16回日本ME学会大会論文集, 300/301(1977).
- [13] 鈴木隆一郎, 他: 乳房X線像の計量診断, 第17回日本ME学会大会論文集, (1978).