

胃癌患者の新しい予後予測法としてのニューラル ネットワークの有用性について ロジスティック回帰分析との比較

岐阜大学医学部第2外科

熊澤伊和生 平岡 敬正 川口 順敬
国枝 克行 梅本 敬夫 佐治 重豊

はじめに：胃癌治療に対するテーラーメイドの医療を目指し、パターン認識に優れた数理的モデルであるニューラルネットワークの予後予測法としての有用性を評価した。**方法：**術後胃癌患者672例に対しニューラルネットワーク(NN)とロジスティック回帰分析(LR)を用いて術後1年3年時点での生死の予測を retrospective に行い精度につき比較検討した。予後因子は腹膜転移、肝転移、深達度、リンパ節転移、根治度、郭清度、年齢、組織型、INF, ly, v の臨床病理学的因子で、21の2値変数(0, 1)にカテゴリー化したものを予測モデルの共変量とし、適中率とROC解析(Az値)で評価し、leave-one-out法にて検証した。**結果：**術後1年時点の適中率で学習検証結果ともにNNは有意に高値を示した(学習 NN 90.0% vs LR 86.8%, 検証 NN 88.1% vs LR 85.3%; とともに $p < 0.01$)。術後3年でも比較的良好であった(学習 NN 85.3% vs LR 83.9%, 検証 NN 83.0% vs LR 82.7%)。Az値では両モデルは同等の数値となった。**考察：**ニューラルネットワークはロジスティック回帰と比べて同等かそれ以上の予後予測能力を示した。今後、ニューラルネットワークは個々の術後患者のリスクを判定する指標としての有用性が期待される。

はじめに

癌患者の治療効果は臨床病理学的予後因子を用いた統計学的解析¹⁾、あるいは癌関連遺伝子変異や増殖・浸潤・転移関連因子などの分子生物学的手法を用いた解析²⁾などにより癌の悪性度や宿主要因の評価がなされ、治療の標準化と個性化が試みられている。

そこで、教室では各種予後因子の統計学的処理によるオーダーメイド医療の確立を目指し、予後の判明している胃癌術後患者を対象にニューラルネットワークを構築し、個々の患者に対する予後予測の可能性を検討してきた³⁾。本法の有用性に関しては臨床医学の分野で文献的に散見されるが、癌患者を対象とした予後予測法としての臨床応用や従来の統計学的処理法との間で予測精度を比較した報告はいまだ少ない。今回著者らは、ニューラルネットワークと従来法であるロジ

スティック回帰モデルを用い、胃癌患者を対象に予測精度などを比較検討し、本法の有用性を評価した。

研究対象

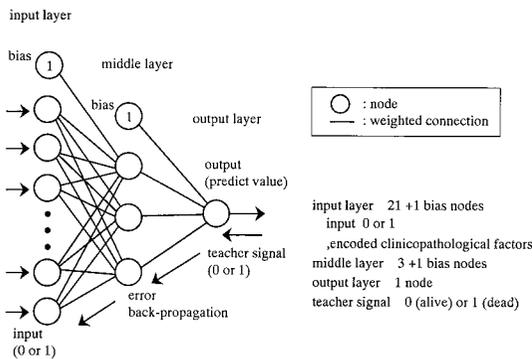
対象は1977年から1995年までの間に当科で胃切除術を施行した胃癌797例中、胃癌登録用紙に予後因子が記載され、術後3年以上を経過し予後が判明している672例(84.3%)である。内訳は男性427例、女性245例で、組織学的進行度はstage I が309例、II 74例、III 120例、IV 169例である。なお、術後3年の他病死亡例を含む全死亡例は271例で、内訳は原病死(胃癌死)206例(死亡例の76.0%)、他癌死5例、他病死31例、術後30日以内死亡16例、その他不明死13例である。

研究方法

1. 評価に用いた臨床病理学的因子

胃癌取り扱い規約⁴⁾に記載されている臨床病理学的因子の中から単変量解析で有意差がみられた、腹膜転移(以下、P)、肝転移(以下、H)、胃壁深達度(以下、t)、リンパ節転移(以下、n)、総合的根治度、リンパ節

Fig. 1 Feedforward neural network design



清(以下, D), 年齢, 組織型, INF, ly, v の11因子を用いた。

2. ロジスティック回帰モデルの作成法

ロジスティック回帰モデル(以下, LR)は上記因子を2から6にカテゴリー化した21個の共変量を用いて術後1年次と3年次の予後への重みをオッズ比(単変量, 多変量)により推定した。

3. ニューラルネットワークの作成法

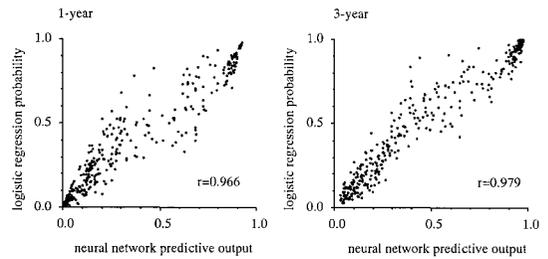
ニューラルネットワーク(以下, NN)はフィードフォワード型で, 中間層が1層からなる階層型モデルを用いた。入力層は21個のnode(疑似ニューロン)からなり, 入力信号は上記各因子を2値(0と1)にコード化し, 閾値調整目的で常に1を出力するバイアス項を入力層, 中間層に各1を加えた。中間層node数と学習回数は既報⁵⁾の手順でLeave-one-out法(以下, LOOM)による検討から, それぞれ3つ, 50回を用いた。出力層node数は1で, 術後1年または3年次での実際の予後を2値化(生存が0, 死亡が1)し, 教師信号として用いた(Fig. 1)。

なお, 教室の平岡ら³⁾が先に作成したNNに比べ, 入力層を9から21へ増加させて予後因子の追加と細分化を図り, 中間層を9から3へ, 出力層を2から1へ減らして学習負担を軽減した。

4. アプリケーションの作成法

NNは誤差逆伝搬法^{6,7)}のプログラムを, LRは最尤法⁸⁾によるプログラムを作成した。開発アプリケーションはCode Warrior IDE version 3.0J(Metrowerks Corporation)で, C言語を用いて作成した。機種はPower Macintosh(G3 300MHz)を使用した。なお, NNのプログラムの検証は文献上で公開されている臨床データ⁹⁾の学習が可能であることで確認した。

Fig. 2 Relationship between predictive outcomes of neural network and logistic regression



5. NNとLRによる予後予測法

術後1年と3年の時点でNNとLRの予測をretrospectiveに比較検討した。すなわち, 各臨床病理学的因子を患者データごとに2値(0, 1)変数にコード化した, 21のダミー変数を共変量として用い, また実際の予後は2値化(0; 生存, 1; 死亡)した。判定は2×2分割表を用い, 実際の予後と各予測結果(0から1の実数)が同じ群(cut-off値=0.5で, 0から0.5未満を生存群, 0.5から1を死亡群)となった場合を正解, 異なった場合を不正解とした。

1) 予測結果の相関の散布図の検索法

両モデルの予測値の散布図を描出し, 予測値間の相関係数と2×2分割表における一致率を算出した。

2) 予測精度の検索法

実際の生死と各モデルでの予測結果の2×2分割表で, とともに生存の群をA, 実際に生存しているが死亡と予測した群をB, 実際に死亡しているが生存と予測した群をC, とともに死亡の群をDとし, 両モデルの適中率(accuracy; $(A+D)/\text{全症例数}$), 感度(sensitivity; $A/(A+B)$), 特異度(specificity; $D/(C+D)$), 陽性反応適中度(positive predictive value; $A/(A+C)$), 陰性反応適中度(negative predictive value; $D/(B+D)$)を算出した。

3) 組織学的進行程度と予後因子別適中率

組織学的進行程度とP, H, t, n, 根治度, Dの各因子別に適中率を算出した。

4) モデルの適合度の検索法

Receiver operating characteristic(受信者動作特性; 以下, ROC)曲線を描出し, Area under the ROC curve(以下, Az)値を算出した。ROC曲線はcut-off値を0から1まで20等分し, ポイント毎にtrue positive fraction(sensitivity), false positive fraction(1-

Table 1 Distribution of clinicopathological factors and survival rate

factor categories	No. of patients	1-year survival		3-year survival		No. of patients	1-year survival		3-year survival	
		No.	%	No.	%		No.	%	No.	%
total	672	520	77	401	60					
stage						curability				
Ia	246	240	98	226	92	A	367	356	97	325
Ib	97	91	94	84	87	B	180	131	73	70
II	64	57	89	38	59	C	125	33	26	6
IIIa	70	53	76	31	44	lymph node dissection				
IIIb	43	33	77	13	30	D0 + 1	194	119	61	89
IVa	29	17	59	4	14	D2 over	478	401	84	312
IVb	123	29	24	5	4	age				
peritoneal dissemination						59	333	266	80	215
P -	575	494	86	397	69	60 69	191	143	75	107
P +	97	26	27	4	4	70	148	111	75	79
liver metastasis						histology*				
H -	632	514	81	399	63	High	340	278	82	224
H +	40	6	15	2	5	Low	332	242	73	177
depth of invasion						INF				
m	159	157	99	148	93	$\alpha + \beta$	521	427	82	351
sm	109	104	95	97	89	γ	151	93	62	50
mp	56	51	91	40	71	ly				
ss	158	114	72	84	53	0	254	247	97	228
se	132	76	58	30	23	1 + 2	211	167	79	119
si	58	18	31	2	3	3	207	106	51	54
lymph node metastasis						v				
n0	352	334	95	304	86	0	500	427	85	352
n1	110	87	79	55	50	1 + 2	143	82	57	44
n2	98	64	65	31	32	3	29	11	38	5
n3 + 4	112	35	31	11	10					

* : category " high "includes papillary, well and moderate differentiated adenocarcinoma, " low "includes poorly differentiated adenocarcinoma and signet-ring cell carcinoma.

specificity) を算出し plotting して描出した .

6. 検証と検定の方法

検証は LOOM による内部妥当性で行い, 全症例の適中率と Az 値で比較した .

検定は StatView 5.0 (SAS Institute Inc.), SPSS 6.1 J (SPSS Japan Inc.) を用い, 両モデル予測値間のピアソンの積率相関係数と LR のオッズ比, 標準誤差を算出して行った . また, 適中率是对応のある 2 群間の比率の差を McNemar 検定で, 2 群間の Az 値の差 (two-tailed test) を Hanley & McNeil の方法^{10,11)}で行った .

結 果

1. 対象症例の特徴

組織学的進行程度および各臨床病理学的因子別に症例数と術後 1 年および 3 年の生存数, 生存率を Table 1 にまとめたが, 組織学的進行度は全国胃がん登録調

査報告¹²⁾ (stage I が 49.9% , II 11.7% , III 19.7% , IV 18.7%) に比べ自験例 (順に 46.0% , 11.0% , 17.9% , 25.1%) では進行例が有意に多かった ($p=0.003$, Mann-Whitney の U 検定) .

2. 臨床病理学的因子の予後への重み

1) 単変量解析

術後 1 年と 3 年の時点で年齢や深達度 (sm) を除き, カテゴリー化された各群の予後因子は基準となる群に比べオッズ比で有意 ($p < 0.05$) 差を認めた (Table 2) .

2) 多変量解析

術後 1 年と 3 年で数値が異なり, 単変量に比べ有意因子が減少した . 各因子別オッズ比は術後 1 年の時点では深達度 (si で 7.07) , 根治度 (C で 6.90) , H (+) , ly3 , P (+) の順に高く, 術後 3 年の時点では深達度 (si で 29.23) , 根治度 (C で 4.31) , n3+4 , H (+) , P

Table 2 Summary of univariate and multivariate logistic regression analysis

factor category	univariate				multivariate						
	1-year		3-year		1-year			3-year			
	OR	p =	OR	p =	OR	95%CI	p =	OR	95%CI	p =	
peritoneal dissemination											
P +	16.65	*	51.86	*	2.27	1.08 4.73	0.03	3.49	1.02 12.02	0.05	
liver metastasis											
H +	24.68	*	32.54	*	4.36	1.48 12.79	*	3.66	0.65 20.55		
depth of invasion											
sm	3.77		1.66		2.27	0.36 14.39		1.07	0.39 2.97		
mp	7.70	0.02	5.38	*	1.48	0.19 11.28		1.94	0.66 5.69		
ss	30.30	*	11.85	*	4.21	0.66 26.94		2.38	0.83 6.85		
se	57.84	*	45.75	*	4.53	0.64 32.01		6.61	1.87 23.32	*	
si	174.44	*	376.73	*	7.07	0.94 53.37	0.06	29.23	4.21 202.73	*	
lymph node metastasis											
n1	4.91	*	6.33	*	0.60	0.24 1.53		1.32	0.66 2.64		
n2	9.86	*	13.69	*	0.90	0.34 2.39		2.49	1.00 6.24	0.05	
n3 + 4	40.82	*	58.15	*	1.54	0.56 4.24		3.54	1.15 10.86	0.03	
curability											
B	12.11	*	12.16	*	3.53	1.17 10.64	0.03	1.42	0.59 3.41		
C	90.23	*	153.47	*	6.90	1.80 26.47	*	4.31	1.03 18.00	0.05	
lymph node dissection											
D0 + 1	3.28	*	2.22	*	1.79	0.94 3.41	0.08	1.67	0.88 3.17		
age											
60 69	1.33	0.18	1.43	0.05	1.59	0.87 2.89		1.73	0.98 3.04	0.06	
70	1.32	0.23	1.59	0.02	1.20	0.60 2.39		2.05	1.09 3.86	0.03	
histology											
low	1.67	0.01	1.69	*	1.08	0.60 1.95		0.72	0.41 1.25		
INF											
y	2.83	*	4.17	*	1.06	0.57 1.99		1.98	1.03 3.80	0.04	
ly											
1 + 2	9.30	*	6.78	*	2.00	0.66 6.05		1.73	0.81 3.69		
3	33.62	*	24.85	*	3.04	0.96 9.60	0.06	2.07	0.88 4.89	0.10	
v											
1 + 2	4.35	*	5.35	*	0.78	0.42 1.47		0.93	0.49 1.77		
3	9.57	*	11.42	*	0.96	0.32 2.85		1.48	0.35 6.15		

* : $p < 0.01$, p value blank means $p > 0.1$. CI : Confidence interval OR : Odds ratio P(-) H(-) m, n0, curability A, D2 over, age under 59, histology high, INF $\alpha + \beta$, ly0, v0 : OR values of above based categories are all 1.00.

(+) の順に高かった (Table 2).

3. ニューラルネットワークとロジスティック回帰モデルによる予測

1) 予測結果の相関と散布図

予測結果は NN と LR 間で有意の相関関係を示し (術後 1 年で $r=0.966$, 3 年で $r=0.979$), 2×2 分割表での一致率は術後 1 年で 95.8% (644/672), 3 年で 93.6% (629/672) と極めて類似した予測結果であった (Fig. 2).

2) 予測精度

予後の適中率は術後 1 年次で NN (90.0%) が LR (86.8%) に比べ有意に高値で ($p < 0.001$), 3 年時次は NN (85.3%) が LR (83.9%) に比べ若干良好であった.

感度, 特異度, 陽性反応適中率, 陰性反応適中率は全般的に NN で良好な値を示したが, 術後 3 年次の感度は LR (76.0%) が NN (71.2%) に比べ若干良好であった (Table 3).

3) 組織学的進行程度と因子別適中率

術後 1 年の時点で各因子とも NN で高い適中率を示し, stage IVa, IVb, P(-) & (+), H(-), t3,

Table 3 Performances on neural network and logistic regression by training data

%	accuracy	sensitivity	specificity	positive predictive value	negative predictive value
1-year survival					
logistic regression	86.8	94.6	59.9	89.0	76.5
neural network	90.0	97.1	65.8	90.7	87.0
3-year survival					
logistic regression	83.9	89.3	76.0	84.6	82.7
neural network	85.3	94.8	71.2	83.0	90.2

* : p < 0.001 ** : p = 0.222

Table 4 Predictive accuracy on neural network(NN)and logistic regression(LR)by each factor

		1-year			3-year				
		survival %	accuracy %		p =	survival %	accuracy %		p =
			NN	LR			NN	LR	
total		77.4	90.0	86.8	< 0.01*	59.7	85.3	83.9	0.22
stage	Ia	97.6	97.6	97.6	1.00	91.9	91.9	91.9	1.00
	Ib	93.8	93.8	93.8	1.00	86.6	86.6	86.6	1.00
	II	89.1	92.2	89.1	0.48	59.4	62.5	57.8	0.58
	IIIa	75.7	78.6	78.6	1.00	44.3	64.3	60.0	0.61
	IIIb	76.7	79.1	72.1	0.25	30.2	81.4	76.7	0.75
	IVa	58.6	82.8	62.1	0.04*	13.8	82.8	82.8	1.00
	IVb	23.6	82.9	74.0	0.02*	4.1	96.7	95.9	1.00
peritoneal dissemination	P -	85.9	91.0	89.2	0.01*	69.0	83.3	82.3	0.43
	P +	26.8	84.5	72.2	0.01*	4.1	96.9	93.8	0.25
liver metastasis	H -	81.3	90.0	86.7	< 0.01*	63.1	84.5	83.1	0.21
	H +	15.0	90.0	87.5	1.00	5.0	97.5	97.5	1.00
depth of invasion	t1	97.4	97.4	97.4	1.00	91.4	91.4	91.4	1.00
	t2	77.1	89.7	87.4	0.13	57.9	77.6	74.8	0.26
	t3	57.6	78.0	70.5	0.02*	22.7	80.3	78.0	0.68
	t4	31.0	84.5	72.4	0.02*	3.4	96.6	96.6	1.00
lymph node metastasis	n0	94.9	95.2	94.6	0.48	86.4	88.9	88.1	0.55
	n1	79.1	88.2	84.5	0.22	50.0	73.6	71.8	0.77
	n2	65.3	78.6	71.4	0.07	31.6	75.5	70.4	0.30
	n3 + 4	31.2	85.7	77.7	0.01*	9.8	93.8	94.6	1.00
curability	A	97.0	97.0	97.0	1.00	88.6	88.6	88.6	1.00
	B	72.8	81.1	72.2	< 0.01*	38.9	71.7	66.7	0.22
	C	26.4	82.4	77.6	0.15	4.8	95.2	95.2	1.00
lymph node dissection	D2 over	83.9	90.8	87.9	< 0.01*	65.3	83.7	82.4	0.39
	D0 + 1	61.3	88.1	84.0	0.03*	45.9	89.2	87.6	0.51

* : p < 0.05

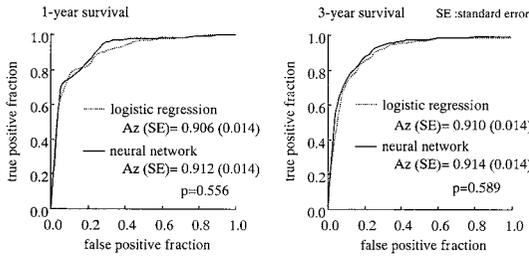
t4, n3+4,根治度 B で有意差がみられた(p < 0.05).術後 3 年の時点では n2と根治度 B で NN は軽度良好であったが有意差はみられず, 全般的に LR と同程度の

適中率であった (Table 4).

4) モデルの適合度

術後 1 年次の Az 値は NN が 0.912 (Standard Error

Fig. 3 Comparison of ROC curve and Az between neural network and logistic regression



=0.014 ; 以下, SE と略記) で LR の0.906 (SE=0.014) に比べ若干高値を示したが有意差はみられなかった . 3年次の Az 値は NN が0.914(SE=0.016)で LR の0.910 (SE=0.019) と近似した値を示した (Fig. 3) .

4. 検証

適中率は術後1年次では NN が88.1%, LR が85.3% で, 学習データと同様 NN が有意に良好であった (p =0.001) .術後3年次では NN が83.0%, LR が82.7% と同等の値であった .

Az 値は術後1年次で NN が0.892(SE=0.016)と LR の0.884 (SE=0.015) に比べ若干高値で, 3年次は NN が0.891 (SE=0.013)で LR の0.893 (SE=0.013) に比べ若干低値であったが, とともに有意差はみられなかった (Table 5) .

考 察

悪性腫瘍に対する治療効果の予測は心血管障害などの良性疾患に比べ予後因子が複雑で, 多様性に富み, 煩雑なため従来の統計的手法のみでは満足できる解析結果が得られていない .そこで, 教室ではパターン認識に優れた人工知能的手法である NN を構築して予後因子の解析に応用し, その有用性を評価してきた . NN は認知科学者の Rumelhart ら¹³⁾が1986年に NN の

階層に hidden unit を加えた誤差逆伝搬法を開発して以来汎用されており, 近年医学分野全般で応用が進み臨床的トピックとして注目されている¹⁴⁾⁻¹⁷⁾ .

1. 予後因子について

著者らは予後因子として腹膜転移, 肝転移, 壁深達度, リンパ節転移, 根治度, リンパ節郭清, 年齢, 組織型, INF, ly, n の11因子を選択・採用した .通常予後因子の選択法として stepwise 法を用いた報告¹⁸⁾⁻²⁰⁾が多いが, 今回の単変量解析のオッズ比の結果から選択した方法は, 一応の規準を満たすものと考えている .なお, 予後因子のカテゴリー化には LR におけるオッズ比を参考にして因子別に適宜設定した .

2. ニューラルネットワークを用いた予後予測について

今回の検討で, 胃癌患者の予後予測の適中率は術後1年の時点では NN が LR に比べ有意に良好で, 3年の時点では NN が若干高値を示した .検証結果でも学習結果と同様, 術後1年の時点では NN が有意に良好であったが, 3年の時点では両者間の有意差は消失し, ほぼ同等の結果を示した .それゆえ, 胃癌患者の予後予測法として NN の有用性が示唆されたと推察している .

一般に NN と LR の予測値には強い相関関係が存在すると考えられ, 冠動脈バイパス手術での予後予測において強い相関 (r=0.95) が報告されている²¹⁾ .

悪性腫瘍においても階層型 NN による予後予測の有用性を支持する報告が多い .すなわち, Jefferson ら¹⁸⁾は620例の肺癌患者の術後経時的予後予測で, LR に比べ NN が検証結果で有意に優れていたと報告し, 濱本ら⁹⁾は肝癌肝切除例の予後予測モデルを作成し, 耐術残肝体積の cut-off 値の予測にも本法の応用を試みている . Bottaci ら²²⁾は334例の結腸直腸癌切除例の予後予測を行い, 術後2年の時点で80%の適中率を得, また

Table 5 Performances on neural network and logistic regression by test data

%	accuracy	sensitivity	specificity	positive predictive value	negative predictive value	Az (SE)
1-year survival						
logistic regression	85.3	93.5	57.2	88.2	71.9	0.884(0.015)
neural network	88.1	96.0	61.2	89.4	81.6	0.892(0.016)
3-year survival						
logistic regression	82.7	88.0	74.9	83.8	80.9	0.893(0.013)
neural network	83.0	93.0	68.3	81.3	86.9	0.891(0.013)

* : p=0.001 ** : p=0.877 *** : p=0.537 **** : p=0.862 SE : standard error

Table 6 Comparison between Odds ratio(OR) and Effective odds ratio(EOR)

clinicopathological factor	category	1-year				3-year			
		OR	EOR			OR	EOR		
			mean	min	max		mean	min	max
peritoneal dissemination	P +	2.27	4.45	1.21	7.45	3.49	4.79	1.07	14.07
liver metastasis	H +	4.36	5.61	1.21	11.50	3.66	1.31	1.01	1.78
depth of invasion	sm	2.27	0.71	0.54	0.96	1.07	0.66	0.30	0.92
	mp	1.48	0.53	0.29	0.92	1.94	1.07	0.51	1.45
	ss	4.21	1.29	1.02	1.55	2.38	1.09	1.01	1.19
	se	4.53	1.68	0.99	2.45	6.61	1.72	1.05	2.85
	si	7.07	2.38	1.08	3.65	29.23	8.82	1.08	21.45
lymph node metastasis	n1	0.60	0.58	0.32	0.95	1.32	1.45	1.01	1.76
	n2	0.90	1.40	0.47	3.43	2.49	1.70	1.03	2.75
	n3 + 4	1.54	2.48	1.08	3.86	3.54	1.62	1.01	2.91
curability	B	3.53	3.06	1.04	5.40	1.42	2.24	1.10	4.16
	C	6.90	4.60	1.23	7.22	4.31	17.43	1.14	41.19
lymph node dissection	D0 + 1	1.79	2.34	1.11	3.24	1.67	1.48	0.80	2.05
age	60 69	1.59	2.97	0.98	5.39	1.73	1.39	0.88	2.84
	70	1.20	1.04	0.52	1.48	2.05	2.14	0.88	3.62
histology	low*	1.08	1.03	0.82	1.20	0.72	0.92	0.44	1.22
INF	γ	1.06	1.13	0.99	1.30	1.98	2.02	1.02	4.29
ly	1 + 2	2.00	2.01	0.53	3.83	1.73	1.86	1.04	2.74
	3	3.04	4.85	1.00	9.02	2.07	1.63	1.03	2.53
v	1 + 2	0.78	0.97	0.45	2.13	0.93	1.24	0.90	2.12
	3	0.96	1.05	0.94	1.19	1.48	1.29	0.96	2.08

* : category " low " includes poorly differentiated adenocarcinoma and signet-ring cell carcinoma.

prospective に検索した結果、臨床医師の予測に比べ NN が有意に優れていたと報告している。Burke ら²³⁾ は5,000例の術後5年時点での予後予測を行い、乳癌で78.4%、結腸直腸癌で86.9%の適中率を得たと報告している。

ところで、胃癌患者の予後予測はモデル作成が不成功に終わったと Drew ら²⁴⁾ が報告しているが、教室の平岡ら³⁾ の予測結果でも、胃癌切除症例の術後5年次の適中率は、130例の既知例で72.3%、未知例で68.7%と若干低値であった。

本研究では両学習モデル間の因子別適中率を比較したが、さらに術後1年次における根治度別の NN の有用性について検討した。全症例において予測結果が異なった28例(NN 正答25例、LR 正答3例)の中で、NN 正答例に根治度 B 16例(生存6例、死亡10例)が含まれたのに対して、LR 正答3例はすべて根治度 C 症例の死亡例であり、根治度 B 症例において LR で予測で

きた症例はすべて NN で予測可能であった。また根治度 C 症例の1年生存33例のうち、LR 正答11例であったのに対して NN は18例(P(+)群6例、t3群7例、t4群4例、n3+4群7例を含む)の正答を得た。これらのことから NN によって臨床上困難であった根治度 B、C での予後予測の実現可能性が示唆された。なお根治度 A 症例では両モデルとも死亡予測は不能であり、術後早期予後の予測は現時点では困難であった。

3. ROC 解析について

ROC 解析²⁵⁾は統計的決定理論から誕生し1960年頃より医学的意思決定の評価に利用され、主として放射線診断で応用されてきた²⁶⁾⁻²⁸⁾。Az 値は ROC 曲線で囲まれた面積で、モデルの適合度の指標とされるが、この Az 値で NN と LR の有用性を各疾患における予測結果で比較すると、NN が優れているもの^{19,20)}、NN と LR が同等であるとするもの^{21,29)}、学習結果では NN で優れていたが検証結果では同等であったとする

もの³⁰⁾など評価が異なる。今回の研究では、学習結果で NN が若干高値を示したが有意差はみられず、検証結果でも術後1年の時点で NN がわずかに高値を示した。それゆえ、NN は LR と同等もしくは若干良好なモデルと考えている。

4. 妥当性の検証について

NN は学習データに対し優れた回帰能力を示すが、未知のデータに対しては最適のモデルである保証はなく、妥当性の検証 (cross validation) が必要となる。LOOM は内部データを利用した妥当性の検証法であるが、今回の検証で適中率は術後1年次で学習データと同様 NN が有意に良好で、術後3年次で NN と LR はほぼ同等の値を示した。この LOOM による検証は画像診断^{31,32)}などで用いられているが、症例が多くなると計算に時間を要するため、短時間で至適モデルを作成することが今後の課題と思われる。

5. 予後への重みと今後の展開について

予後因子の重みの定量化には Az 値²⁸⁾、NN の平均出力比³³⁾、effective odds ratio (以下、EOR)³⁴⁾などが試みられている。本研究では EOR を算出し LR のオッズ比と比較検討した。その結果、臨床病理学的因子の平均 EOR で、上位5因子は術後1年次で H (+)、ly3、根治度 C、P (+)、年齢が60歳台であった。しかし、LR のオッズ比では深達度 (si) が最も高値を示したが、EOR では低値で、この相違が適中率の差に影響した可能性が推察される。一方、術後3年次は EOR が根治度 C、深達度 (si)、P (+)、年齢70歳以上、INF の順で高値を示した。LR のオッズ比では H (+)、n3+4 が高値を示したが、EOR では低値であった。根治度 C と P (+) が両モデルで高値を示した (Table 6)。

これらの結果から、NN は LR に比べ短期予後の予測に優れていたが、中期予後の予測は同等か若干良好で、総じて NN の有用性が示唆できたと考えている。NN はパーソナルコンピュータでも解析可能で、一般臨床で新たな知見の発見 (データマイニング) が可能となり、根拠に基づいた治療 (EBM) に役立つものと期待される。今後は治療効果に対する解析、転移形式や転移時期などの予想及び治療効果に対する解析、即ち術後補助療法の継続期間や追加投与時期の特定などへの応用が検討課題である。また NN は個々の術後患者のリスクを判定する指標としての有用性が期待される。

文 献

1) 中島聰總, 太田恵一郎, 石原 省ほか: 生存に及ぼ

す胃癌予後要因の多変量解析. 癌と化療 21: 1768-1774, 1994

- 2) Yasui W, Tahara H, Tahara E et al: Molecular-pathological diagnosis of gastrointestinal tissues and its contribution to cancer histopathology. *Pathol Int* 49: 763-774, 1999
- 3) 平岡敬正, 梅本敬夫, 佐治重豊ほか: ニューラルネットを用いた胃癌術後患者の新しい予後予測法の試み. *日消外会誌* 32: 2064-2071, 1999
- 4) 胃癌研究会編: 胃癌取扱い規約. 改訂第12版. 金原出版, 東京, 1993
- 5) 熊澤伊和生, 川口順敬, 佐治重豊ほか: leave-one-out method による階層型ニューラルネットワークの学習回数と中間層の node 数の決定方法. 術後胃癌患者の予後予測への臨床応用. *医療情報学19回連大会論集*: 944-945, 1999
- 6) 平野広美: C でつくるニューラルネットワーク. パーソナルメディア, 東京, 1991, p15-22
- 7) 熊澤逸夫: 学習とニューラルネットワーク. 森北出版, 東京, 1998, p51-75
- 8) 丹後俊郎, 山岡和枝, 高木晴良: ロジスティック回帰分析, SAS を利用した統計解析の実際. 朝倉書店, 東京, 1996, p181-183
- 9) 濱本 勲, 岡田節雄, 前田 肇: ニューラルネットワークを用いた肝癌肝切除術の予後予測. *日消外会誌* 28: 1030-1036, 1995
- 10) Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143: 29-36, 1982
- 11) Hanley JA, McNeil BJ: A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 148: 839-843, 1983
- 12) 三輪 潔: 全国胃がん登録調査報告 第55号. 三輪胃がん登録研究所, 東京, 1999, p56
- 13) Rumelhart DE, Hinton GE, Williams RJ: Learning representations by back-propagating errors. *Nature* 323: 533-536, 1986
- 14) Cross SS, Harrison RF, Kennedy RL: Introduction to neural networks. *Lancet* 346: 1075-1079, 1995
- 15) Baxt WG: Application of artificial neural networks to clinical medicine. *Lancet* 346: 1135-1138, 1995
- 16) Penny W, Frost D: Neural networks in clinical medicine. *Med Decis Making* 16: 386-398, 1996
- 17) Duh MS, Walker AM, Ayanian JZ: Epidemiologic interpretation of artificial neural networks. *Am J Epidemiol* 147: 1112-1122, 1998
- 18) Jefferson MF, Pendleton N, Lucas SB et al: Comparison of a genetic algorithm neural network

- with logistic regression for predicting outcome after surgery for patients with nonsmall cell lung carcinoma. *Cancer* 79 : 1338-1342, 1997
- 19) Dybowski R, Weller P, Chang R et al : Prediction of outcome in critically ill patients using artificial neural network synthesised by genetic algorithm. *Lancet* 347 : 1146-1150, 1996
- 20) Bryce TJ, Dewhirst MW, Floyd CE Jr et al : Artificial neural network model of survival in patients treated with irradiation with and without concurrent chemotherapy for advanced carcinoma of the head and neck. *Int J Radiat Oncol Biol Phys* 41 : 339-345, 1998
- 21) Tu JV, Weinstein MC, McNeil BJ et al : Predicting mortality after coronary artery bypass surgery : what do artificial neural networks learn? *Med Decis Making* 18 : 229-235, 1998
- 22) Bottaci L, Drew PJ, Hartley JE et al : Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions. *Lancet* 350 : 469-472, 1997
- 23) Burke HB, Goodman PH, Rosen DB et al : Artificial neural networks improve the accuracy of cancer survival prediction. *Cancer* 79 : 857-862, 1997
- 24) Drew PJ, Bottaci L, Monson J et al : Artificial neural networks. *Lancet* 350 : 1175-1176, 1997
- 25) 山下一也 : 信号検出理論・ROC解析小史 . 日本放射線技術学会専門委員会デジタル画像のROC解析検討班編 . ROC解析の基礎と応用 . 日本放射線技術学会 , 京都 , 1994, p1-3
- 26) Floyd CE Jr, Lo JY, Yun AJ et al : Prediction of breast cancer malignancy using an artificial neural network. *Cancer* 74 : 2944-2948, 1994
- 27) Ikeda M, Ito S, Ishigaki T et al : Evaluation of a neural network classifier for pancreatic masses based on CT findings. *Comput Med Imaging Graph* 21 : 175-183, 1997
- 28) Arana E, Marti-Bonmati L, Bautista D et al : Calvarial eosinophilic granuloma : diagnostic models and image feature selection with a neural network. *Acad Radiol* 5 : 427-434, 1998
- 29) Lapuerta P, L'Italien GJ, Paul S : Neural network assessment of perioperative cardiac risk in vascular surgery patients. *Med Decis Making* 18 : 70-75, 1998
- 30) Duh MS, Walker AM, Pagano M, Kronlund K et al : Prediction and cross-validation of neural networks versus logistic regression : using hepatic disorders as an example. *Am J Epidemiol* 147 : 407-413, 1998
- 31) El-Kwae EA, Fishman JE, Bianchi MJ et al : Detection of suspected malignant patterns in three-dimensional magnetic resonance breast images. *J Digit Imaging* 11 : 83-93, 1998
- 32) Ashizawa K, Ishida T, MacMahon H et al : Artificial neural networks in chest radiography : application to the differential diagnosis of interstitial lung disease. *Acad Radiol* 6 : 2-9, 1999
- 33) 高橋英孝, 飯田行恭, 岸本 剛 : ニューラルネットワークによる疫学モデルのリスクの定量化 . 医療情報学 16回連大会論集 : 250-251, 1996
- 34) Lippmann RP, Shahian DM : Coronary artery bypass risk prediction using neural networks. *Ann Thorac Surg* 63 : 1635-1643, 1997

Usefulness of Neural Network as a Novel Method of Predicting Outcome for
Gastric Cancer Patients Compared with Logistic Regression

Iwao Kumazawa, Takamasa Hiraoka, Yoshihiro Kawaguchi, Katsuyuki Kunieda,
Takao Umemoto and Shigetoyo Saji
Second Department of Surgery, Gifu University School of Medicine

Introduction : To establish tailor-made therapy for gastric cancer patients, we evaluated the usefulness of a neural network (NN) a computer-based mathematical model superior in pattern recognition. **Materials and Methods :** Predictions of 1- and 3-year survival were compared between the NN and logistic regression (LR) retrospectively using 672 gastrectomized patients with stomach cancer. As prognostic factors, we selected peritoneal metastasis (P) hepatic metastasis (H) invasion depth, lymph node metastasis (n) curability, lymph node dissection (D) age, histology, INF, ly and v, and categorized them into 21 dichotomous (0, 1) variables to suit each model. We then evaluated accuracy using a 2×2 matrix and Az (area under the receiver operating characteristic curve) Both models were tested using the " leave-1-out " method. **Results :** The accuracy of 1-year survival predicted by the NN was significantly better than that of LR (training data : NN 90.0 % , LR 86.8% , test data : NN 88.1% , LR 85.3% ; $p < 0.01$) The accuracy of 3-year survival predicted by the NN was relatively better than that of LR (training data : NN 85.3% , LR 83.9% , test data : NN 83.0% , LR 82.7%) The Az of the NN was statistically similar to that of LR. **Conclusion :** The neural network showed superior or similar predictions for gastric cancer patients compared logistic regression. The neural network may thus be used as an index for deciding the risk of individual postoperative patients.

Key words : Prediction of postoperative survival, gastric cancer, neural network, logistic regression, receiver operating characteristic analysis

[Jpn J Gastroenterol Surg 34 : 449 458, 2001]

Reprint requests : Iwao Kumazawa Second Department of Surgery, Gifu University School of Medicine
40 Tsukasamachi, Gifu, 500 8705 JAPAN
