

## 基本的検査からの甲状腺機能亢進症患者のスクリーニング

川上 準子, 星 憲司, 佐藤 渉, 佐藤 憲一, 菅原 明<sup>1</sup>, 齋藤 芳彦<sup>2</sup>, 吉田 克己<sup>3</sup>

<sup>1</sup>東北大学大学院医学系研究科, <sup>2</sup>東北大学病院, <sup>3</sup>東北大学医学部保健学科

### Screening of the patient with hyperthyroidism using routine test data

Junko KAWAKAMI, Kenji HOSHI, Wataru SATO, Kenichi SATO, Akira SUGAWARA<sup>1</sup>, Yoshihiko SAITO<sup>2</sup>,  
and Katsumi YOSHIDA<sup>1</sup>

(Received November 2, 2005)

A complementary study to our former work is performed about screening of the patient with hyperthyroidism using routine test data. The same calculation is done, however, adopting 17 routine tests in addition to former 14 tests. The result adopting 4 routine tests such as ALP, AST, ALT and T-Cho which thyroid specialist often consider as important to suspect hyperthyroidism, is evaluated. The importance of S-Cr to make screening the patients with hyperthyroidism is stressed.

**Key words** — screening; hyperthyroidism; routine test data; Kohonen's self-organizing map; Bayesian regularized neural network

体調が悪いため病院を訪れた人に対して、日本では通常基本的検査が行われるが、その段階では検査値をみても様々な疾患の可能性があり、病歴、診察所見および基本的検査のデータから疑われる病気に関連した専門検査を行うことにより診療が始まるが、疑われたものはずれることもあり、最終的な診断結果が出るまでに長い時間がかかることもある。甲状腺機能亢進症のように患者数が全人口の約0.2~0.5%と症例の少ない場合にはそのようなことが起こりやすく、消化器系疾患や心疾患などが疑われている間に1年にも及ぶ期間を要し、ようやく、最終的に甲状腺疾患が判明し治療が開始されることもよくあり、その間の患者の苦痛には大変なものがあると言われる。

また、専門検査は概して高コストでもあるので、かなりの疑いがもたれる前からその検査を実施するのは困難である。甲状腺ホルモンの専門検査データから統計処理により診断支援を行う可能性については、はじめCoomansらが判別

分析を用いて示したが<sup>1)</sup>、最近、星らはベイズ正則型ニューラルネットワーク(BRNN)<sup>2,3)</sup>と自己組織化型ニューラルネットワーク(SONN)<sup>4)</sup>の2種類を併用して用いることで優れた診断支援が可能であることを示した<sup>5,6)</sup>。ただ、現在では、最近のTSH, T4, T3や抗TSH受容体抗体測定技術の著しい進歩のおかげで甲状腺ホルモン検査を行えば、そのデータから確実な診断がほぼ可能である<sup>7)</sup>。そこで、我々は東北大学病院総合診療部で基本的検査を受けた未診断者の14項目の検査値のセットから甲状腺機能亢進症患者を高い確率で予測する診断支援の課題にニューラルネットワークの手法を応用し、その有効性を示した<sup>8)</sup>。その際、アルカリフォスファターゼ(ALP)、血清クレアチニン(S-Cr)、総コレステロール(T-Cho)の3項目のセットに注目するとスクリーニング精度の向上が図れることについても報告した。

本論文では、よく使用される検査項目として、平均赤血球容積(MCV)、平均赤血球ヘモグロビ

ン量 (MCH), 平均赤血球ヘモグロビン濃度 (MCHC) を追加した17項目を用いた場合について計算を行った. 甲状腺専門医は基本的検査の中で甲状腺疾患に関係しそうな検査として, まずALPとT-Cho, さらにAST, ALTの計4個を考えることが多いと言われるので<sup>9)</sup>, この4変数でのスクリーニング結果について詳しく述べる.

### 患者データ

東北大学病院では基本的検査として, 通常, 33項目の検査を実施対象としているが, 患者によっては未検査の項目も多い. 今回の解析では, Table 1に示す17項目で検査値の揃っている全てのサンプルを使用した. 診断結果が既知の健常者48名 (男15名, 女33名), 甲状腺機能亢進症の診断が確定した患者18名 (男4名, 女14名) の計66サンプルを使用した. また, 体調不良で来院し基本的検査を受診した病名が確定していない142名 (男48名, 女94名) を予測対象とした.

### 方法

2種類のニューラルネットワークを用いて基本的検査データを解析する.

1つ目は自己組織化マップ (SOM) を生成するKohonenのSONNであり, クラスタリング能力, データの分布を可視化する能力に優れている. アルゴリズムもシンプルで汎用性に富み, 計算の安定性が高いこともよく知られている. もし既知のサンプルを用いて作成したSOMが病名に応

じたクラスタリングを示せば, 病名不明の受検者のSOM上への投射により, 甲状腺機能亢進の患者を予測できる. また, データの特徴を解析して診断に有効な検査項目を選び出すことも出来る.

2つ目はBRNNで, 三層の階層型ネットワークである. 入力層の各ニューロンに検査項目を割り当て, 出力層の2つのニューロンはSoftmax型出力をもち, それぞれ, 健常者と甲状腺機能亢進症患者の予測確率を与える. 既知のサンプルの診断結果を用いて, ネットワークの重みをベイズ学習によって決定するが, マルコフ連鎖モンテカルロ法を用いた実装<sup>3)</sup>によりクラス分類能力も高い安定したネットワークであることが知られている. この手法ではベイズの先見的結合荷重確率が結合荷重の正則化項と自然に対応しており, ARD (Automatic Relevance Determination) により判定に大きく影響した一連の入力, 即ち, 基本的検査項目を評価することができる.

### 結果と考察

文献8の14項目に平均赤血球容積 (MCV), 平均赤血球ヘモグロビン量 (MCH), 平均赤血球ヘモグロビン濃度 (MCHC) を追加した17項目でのSONN計算で得られたSOMマップをFig.1に示す. MCV, MCH, MCHCはヘマトクリット (HCT), 赤血球数 (RBC), 血色素 (Hb) から計算で求められ, 通常, 貧血の指標として役立っている. SOMは健常者群 (ラベル3-で赤) と亢進

Table 1. Part of patient data who received routine test at Tohoku University Hospital during 2000 ~ 2004.

sample	ALP	$\gamma$ -GTP (GOT)	AST (GPT)	ALT (GPT)	LHD	S-Cr	UA	TG	T-Cho	WBC	RBC	Hb	HCT	MCV	MCH	MCHC	PLT	diagnosing
3-1	171	21	24	17	447	1.1	6.2	131	211	4	5.35	16.6	47.5	88.8	31	34.9	162	normal
3-2	94	17	14	13	409	0.6	4.4	35	172	6.4	4.55	13.7	41.3	90.8	30.1	33.2	201	normal
3-3	219	10	16	11	340	0.7	4.6	90	134	5	4.31	13.6	39.8	92.3	31.6	34.2	260	normal
1-49	244	61	33	47	309	0.3	6.1	85	135	3.4	4.81	13.8	40.5	84.2	28.7	34.1	304	hyper
1-50	524	72	27	34	373	0.3	5.3	85	122	4	4.95	14.1	42.7	86.3	28.5	33	165	hyper
1-51	793	49	35	26	369	0.3	7.5	52	108	4.1	3.94	11.2	33	83.8	28.4	33.9	159	hyper
4	355	27	26	28	427	0.5	3	88	204	6.3	3.95	12.6	37.6	95.2	31.9	33.5	235	?
6	159	33	17	15	316	0.7	4.7	46	153	4.2	4.39	13.8	39.7	90.4	31.4	34.8	162	?

症患者群（ラベル1-で緑）の2つのエリアに棲み分けられており、ほぼ9割の亢進症患者を正しく分類できていた。これを見ると結果は14項目の場合とほぼ同じであるが、SOMのクラスター分布はより明瞭でマップは見やすい。これはMCVとMCHの要素平面において亢進症患者ゾーンのみに特徴的な値の低下が見られるため、MCVとMCHの追加により、14項目の場合よりは2つのゾーンを分離する能力が上がっているものと解釈できる。Fig.1はニューロン数30x20、近傍半径 $r=30$ 、学習回数50000回、学習係数 $\alpha=0.15$ で作成したが、計算条件を変えても結果は安定していた。マップとカルテを照合してみると、機能亢進ゾーンの患者でも、重い機能亢進の患者はゾーンの奥に、軽い機能亢進の患者は健常ゾーンの近くに配置されていることがわかった。誤って健常ゾーンに配置された1-52は非常に健常に近い患者であり、境界域に配置している1-58などもその傾向があることがわかった。

次に、142名の受検者をこのSOMに投射した結果をFig.2に示すが、多くの受検者は健常ゾーンに投射されるものの、10名近くは亢進ゾーンに投射されている。計算条件の違いによらず、亢進ゾーンの一番奥に配置された3名の患者167, 217, 285は、医師による診断の結果、最終的に確認された甲状腺機能亢進症患者であった。したがって、甲状腺機能亢進症患者を洩れなく予測する点では素晴らしいが、スクリーニング精度の点では17項目での予測計算も14項目の場合と似て不十分であることがわかる。14項目の場合、BRNNに診断結果が既知の（48名の健常者と18名の甲状腺機能亢進）66名のデータを教師信号として学習させ、モンテカルロサンプリング数400回でLeave-one-out計算による診断結果の予測精度はほぼ9割であり、病名不明の受検者142名に対して健常か甲状腺機能亢進かを予測させたところ、80%以上の確率で亢進と予測された患者が10名弱で、最終的に甲状腺機能亢進症であると判明した3人の患者は、いずれも9割以上の高い確率で亢進と予測されたが<sup>8)</sup>、この結論は17項目の計算でもほぼそのまま成立することがわかった。

BRNNはARD機能をもち判定に重要であった検査項目は何かを答えることができる。Fig.3に17項目に対する関与度（relevance）のグラフを示すが、条件1は中間層のニューロン数は8でモンテカルロサンプリング数400回の場合、条件2は中間層のニューロン数は10でモンテカルロサンプリング数400回の場合、条件3は中間層のニューロン数は12でモンテカルロサンプリング数400回の場合、でありいずれもARDありのBRNN計算による。2つ以上の計算条件でも比較的関与度の大きな検査項目としては、ALP、 $\gamma$ -GTP、AST、ALT、血清クレアチニン、中性脂肪、総コレステロール、白血球数を挙げるができるが、これは文献8でSOMの要素平面の特徴分布から判明した分類に重要な検査と中性脂肪、白血球数を除いてほぼ一致している。

甲状腺専門医は基本的検査の中で甲状腺機能亢進症に関係しそうな検査として、まずALP値の上昇とT-Cho値の下降に着目し、さらに場合によってはASTとALTの変化の計4個を考慮することが多い<sup>9)</sup>。以下では、それについてのニューラルネットワークを用いた今回の研究による評価について述べる。この4項目、および、S-Crを加えた5項目でのBRNN計算によるスクリーニングの結果をTable 2に示す。数字は甲状腺機能亢進症と予測された総合診療部で基本的検査を受けた人の番号であり、太い数字は実際に医師により甲状腺機能亢進症であることが確認された、つまり、正しく予測された患者である。この4項目でのスクリーニングは難しく、S-Crを加えた5項目でのARDなしBRNN計算ならスクリーニングに使えるようである。Table 2には比較のために17項目と文献8でスクリーニング精度が一番優れていると考えられた3項目（ALP、S-Cr、T-Cho）のセットでの結果も載せてある。5項目でのARDなし計算の結果は、17検査項目の場合より優れているが、3項目（ALP、S-Cr、T-Cho）のセット（この場合、実際には90%以上の確率である）に比べてスクリーニング精度が低いことがわかる。5項目でのBRNN計算による予測結果が、ARDなしで良い結果を与えるのは、この5項目は既にブルーニングの結果（ARDあ

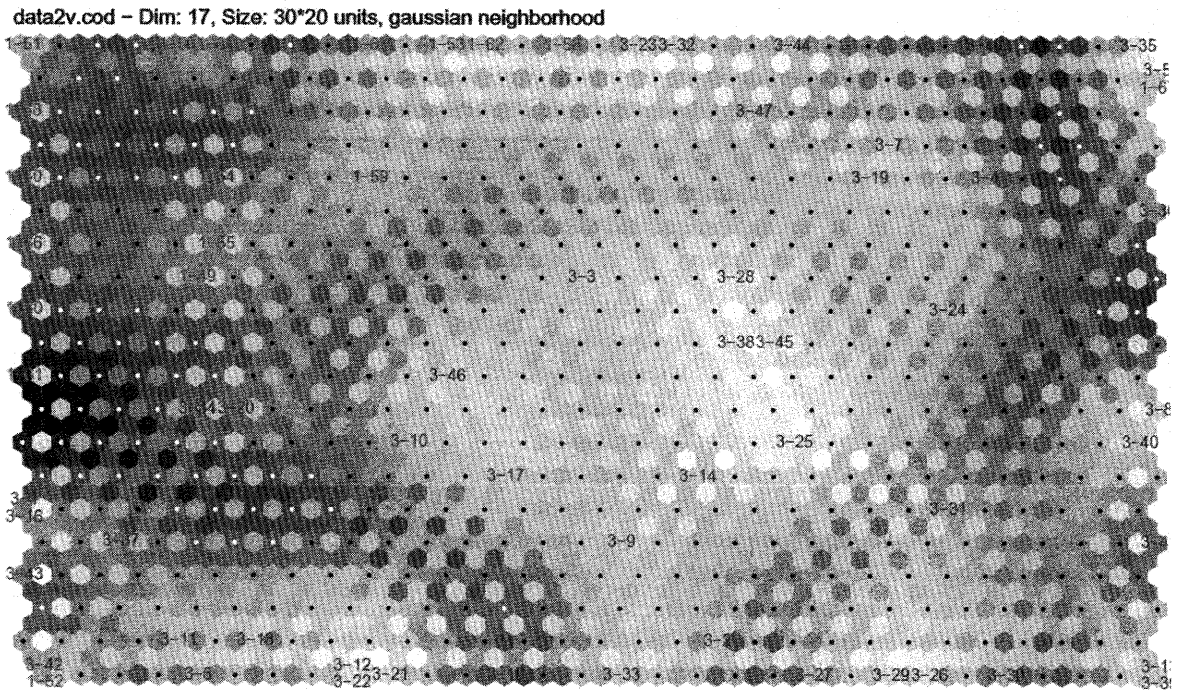


Fig. 1. The SOM for known patients with hyperthyroidism (label 1, green, 1-49 to 1-68, n=18; excluding 1-57 and 1-65) and healthy individuals (label 3, red, 3-1 to 3-48, n=48). Each group is composed of males and females.

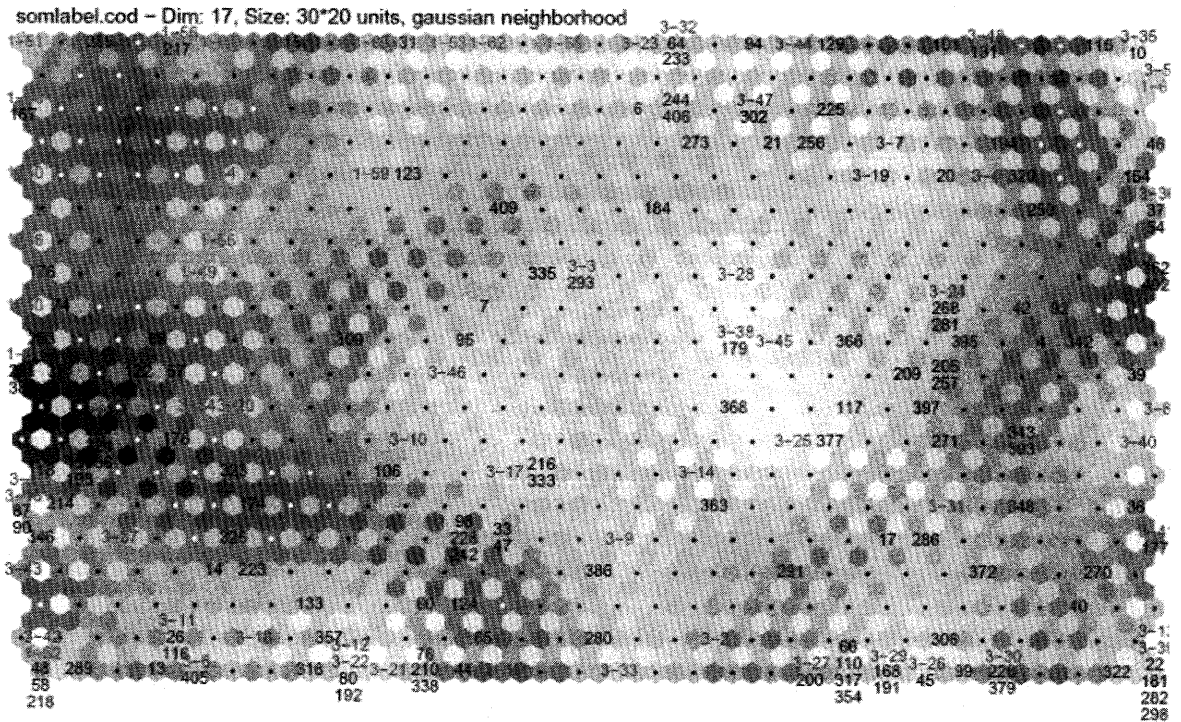


Fig. 2. The same SOM as Fig. 1, however, 142 examinees (black labels) are projected onto it now.

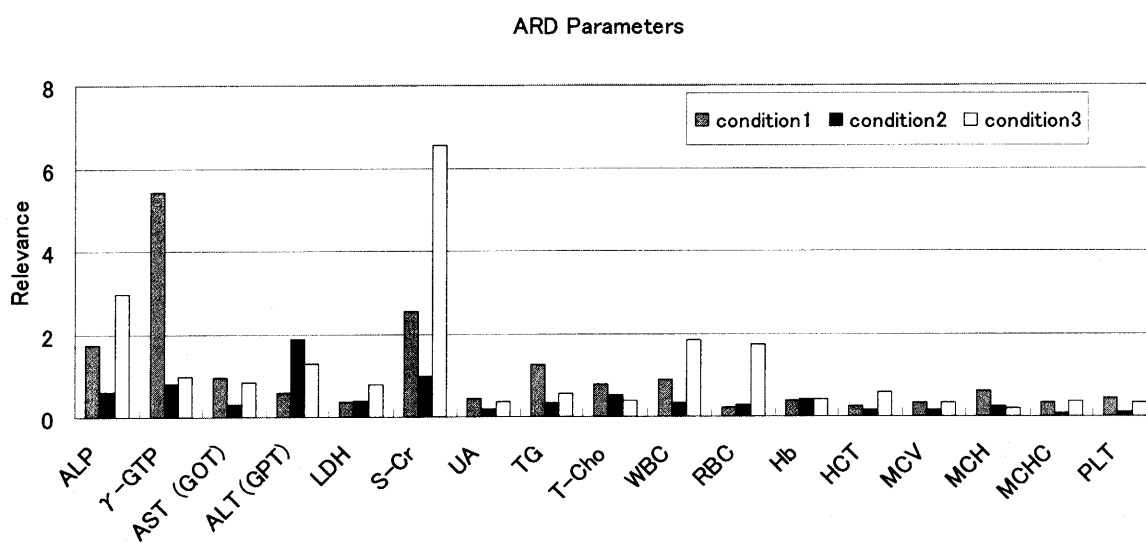


Fig. 3. Relevances for 17 routine tests obtained by the ARD method of BRNN.

Table 2. The result of screening for the female patients with hyperthyroidism using 4 routine tests which a medical specialist often takes. Other results using 5 and 17 routine tests, and 3 best routine tests are also shown.

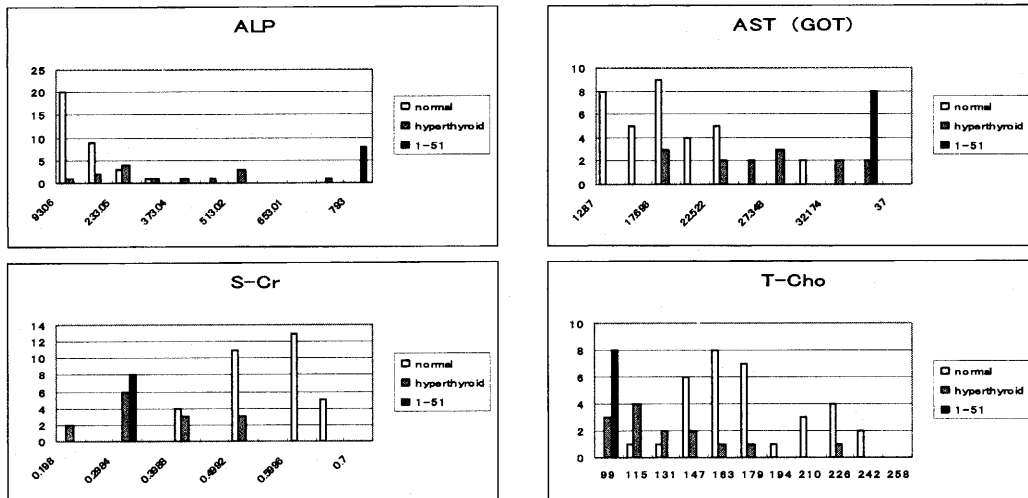
prob.of hyperthyroid	intermediate neurons	ALP, T-Cho, AST (GOT), ALT (GPT)		ALP, T-Cho, AST (GOT), ALT (GPT), S-Cr		17 parameters	ALP, T-Cho, S-Cr	
		with ARD	no ARD	with ARD	no ARD		with ARD	with ARD
more than 80%	15			31, 217	31, 167, 217, 285	31, 167, 217, 331	167, 217, 285	167, 217, 285
	12	31, 92	31	31, 217	31, 167, 217, 285	167, 217, 331	167, 217, 285	167, 217, 285
	10	31, 92, 223	31, 92	31, 217	31, 167, 217, 285	17, 31, 58, 168, 192, 217, 223, 291, 331, 402	167, 217, 285	167, 217, 285
	8	31, 92	31, 92, 285	31	31, 167, 217, 285	217, 331	167, 217, 285	167, 217, 285
	5			31, 217	31, 167, 217, 285	17, 31, 217, 331	167, 217, 285	167, 217, 285
70% ~ 80%	15			4, 115	92	17, 58, 192	115	257
	12	17, 223	92	4, 115		58, 285, 289		257
	10	17, 217, 285	167, 223, 285		92	10, 46, 167, 362	54, 257	257
	8	223	167, 223	217	92	31, 58, 167, 285, 289		257
	5			60, 92, 285		58, 167, 192, 223, 285	257	257

り)として多くの項目から選択された項目群に属するので、その段階からさらにARDを行う必要はないためと考えられる。

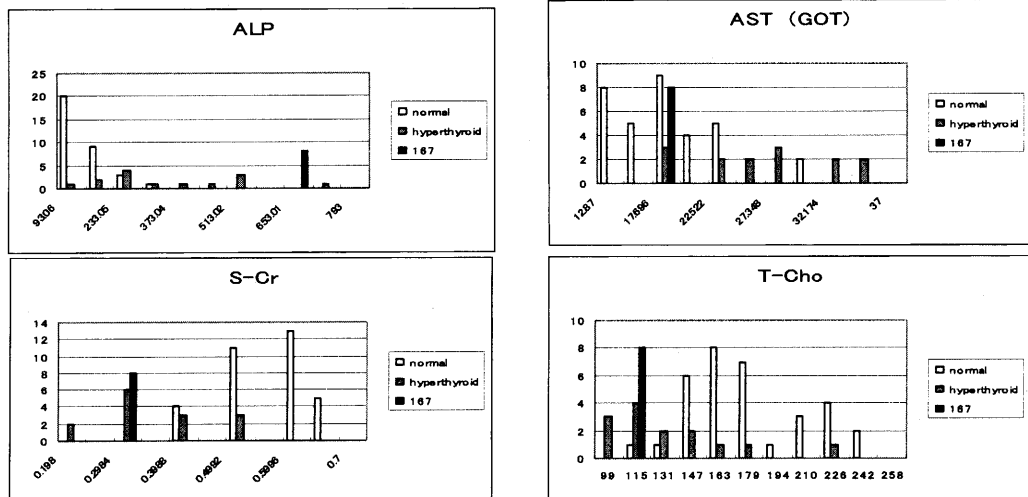
Fig.4にALP, AST, S-Cr, T-Choの各検査項目につき既知の亢進症患者(灰色)と健常者(白)の

度数分布を示す。検査項目によってはきれいに2分できるが一般には重なる部分も多い。これに病名が確定していない総合診療部受検者(黒)を検査値に対応する位置に書き込むと(受検者が目立つように、スケールは何倍かしている)

1-51



167



257

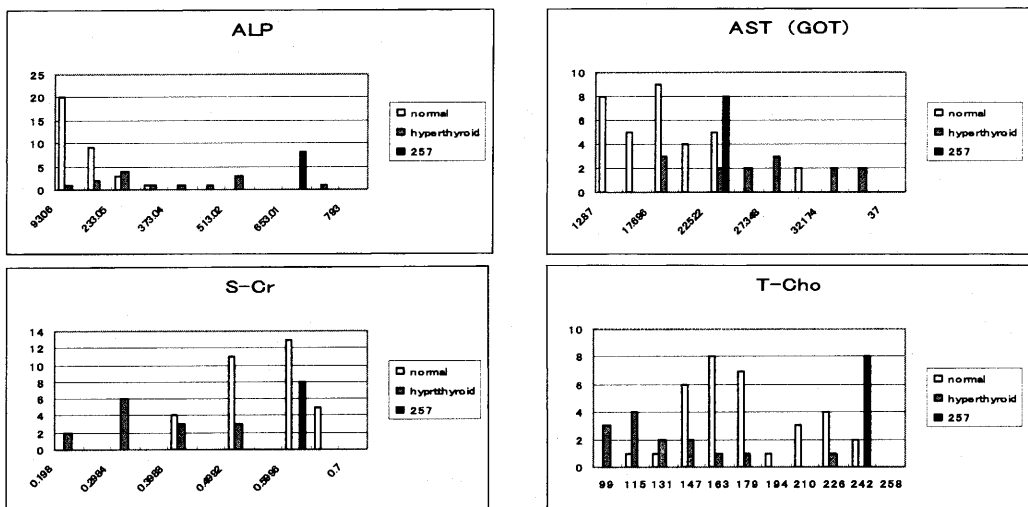


Fig. 4. Frequency distribution of 4 routine tests (ALP, AST, S-Cr, T-Cho) for hyperthyroid patients 1-51 and 167, and for not hyperthyroid patient 257.

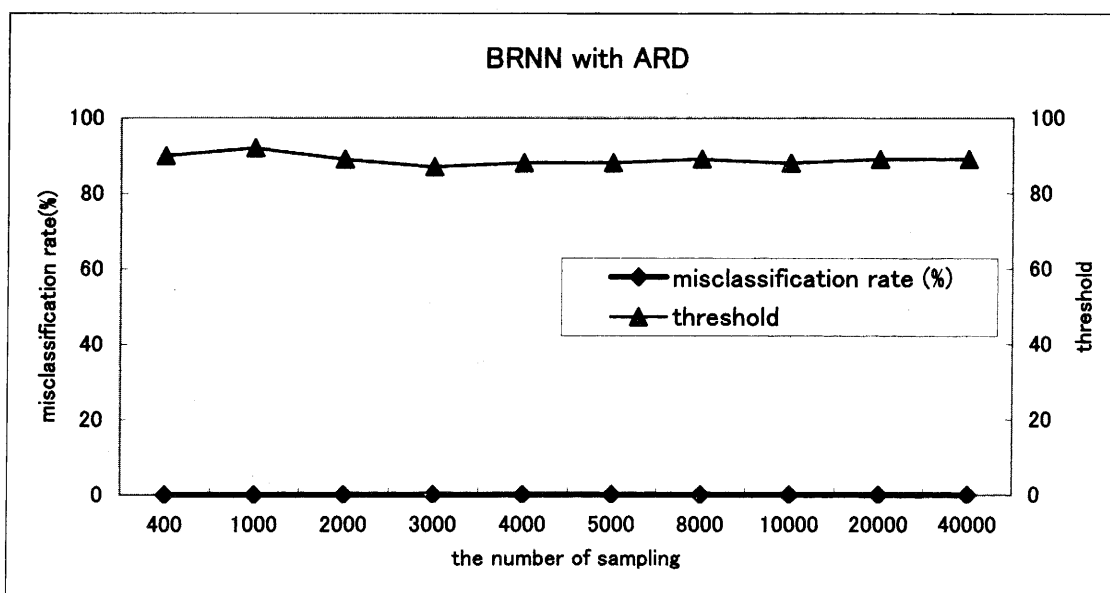


Fig. 5. The misclassification rate (%) and threshold are calculated for female patients using ALP, AST (GOT), S-Cr and T-Cho by BRNN with ARD, and their values remain almost the same.

SOMでなぜそのように分類されたのかなどの解釈をすることができる場合も多い。既知の亢進症患者1-51, 病名未確定の総合診療部受検者167, 257の例を示すが, 167は後に医師の診断により亢進症, 257はそうでないことがわかっている。Fig.2のSOMから示唆されるように1-51は重い亢進症でALPとASTが高く, S-CrとT-Choが低く, スクリーニング精度が高い検査項目セットで見られた条件と一致している。167はASTが高くないもののALPが高く, S-CrとT-Choが低く, スクリーニング精度が高い検査項目セットで見られた条件と一致している。257はS-CrもT-Choも高く亢進症の条件と合わないことが度数分布からわかる。

BRNN計算ではマルコフ連鎖モンテカルロ法を使用するので, サンプル数が十分であることを, 一般的には問題ごとに確認しておく必要がある。計算時間はサンプル数にほぼ比例する。Husmeirらは標準的な問題につき予測精度とサンプル数の関係を調べたが<sup>10)</sup>, サンプル数400程度で十分な場合も多い。文献8と本論文ではサンプル数400で解析を行っている。Fig.5に100倍までのサンプル数で誤診率と閾値<sup>8)</sup>の変化について女性のみを対象と

してALP, AST, S-Cr, T-Choのセットで調べた結果の例を示す。中間層のニューロン数などの計算条件を変えても大きな変動は見られなかった。これらの結果により本研究ではサンプル数400で十分であることが確認できた。

## REFERENCES

- 1) Coomans D, Jonckheer M, Massart D.L., *Analytica Chimica Acta*, **103**, 409-415 (1978).
- 2) MacKay D.J.C., *Neural Computation*, **4**, 448-472 (1992).
- 3) Neal R.M., "Bayesian Learning for Neural Networks", Springer, New York, 1996.
- 4) Kohonen T. "Self-Organizing Maps", Springer, Berlin, 2000.
- 5) Kawakami J, Hoshi K, Ishiyama A, Miyagishima K, Sato K., *Chem. Pharm. Bull.*, **52**, 751-755 (2004).
- 6) Hoshi K, Kawakami J, Kumagai M, Kasahara S, Nisimura N, Nakamura H, Sato K. *Chem. Pharm. Bull.*, **53**, 1570-1574 (2005).
- 7) Weetman AP., *N Engl J Med*, **26**, 1236-48 (2000).
- 8) Hoshi K., Kawakami J., Sato W., Sato K., Sugawara

- A., Saito Y., Yoshida K., submitted
- 9) Amino N and Japan Thyroid Association Working Group of Guidelines for the Diagnosis of Thyroid Disease. Guideline for the diagnosis of thyroid disease. *Clinical Endocrinology (Japanese)*. **50**, 3-13 (2002).
- 10) Husmeir D., Penny W.D., Roberts S.J., *Neural Networks*, **12**, 677-705 (1999).