Assisting the Diagnosis of Thyroid Diseases with Bayesian-Type and SOM-Type Neural Networks Making Use of Routine Test Data

Kenji Hoshi,[a] Junko Kawakami,[a] Wataru Sato,[a] Kenichi Sato,*[a] Akira Sugawara,[b] Yoshihiko Saito,[c] and Katsumi Yoshida[d]

[a] *Information Science Center, Tohoku Pharmaceutical University; 4–4–1 Komatsushima, Aoba-ku, Sendai 981–8558, Japan:* [b] *Department of Comprehensive Medicine, Tohoku University Graduate School of Medicine;* [c] *Department of Clinical Laboratory, Tohoku University Hospital; and* [d] *Department of Medical Technology, Tohoku University School of Health Sciences; 1–1 Seiryocho, Aoba-ku, Sendai 980–8574, Japan.*

Patients with hyperthyroidism sometimes take much time to receive the final diagnosis. To improve patient QOL, simple screening for hyperthyroidism by thyroid non-specialists at the physical check-up is highly expected. Therefore, we applied both Bayesian-type and SOM-type neural networks since we assured the approach useful in analysing thyroid function diagnosis in the previous work. Routine test (14 parameters) data from 66 subjects with a known diagnosis (18 patients with hyperthyroidism and 48 healthy volunteers) were adopted as learning data, and then 142 individuals who also received the same routine tests at the Tohoku University Hospital were screened to predict patients with hyperthyroidism. Both neural networks using 14 parameters predicted several patients as having hyperthyroidism with high probability, including all three hyperthyroid patients diagnosed later by the physician. Further detailed analysis of the routine test parameters that were important for classification found that screening with a set of three parameters (alkaline phosphatase, serum creatinine and total cholesterol) or plus aspartate aminotransferase allowed for quite accurate screening. These results showed that the same neural networks as previous work allows simple screening of patients for hyperthyroidism on the basis of routine test data, and that physicians not specializing in the thyroid can rapidly identify individuals suspected of having hyperthyroidism, to permit a rapid referral for examination and treatment by thyroid specialists.

Key words　thyroid function diagnosis; self-organizing map; Bayesian neural network; hyperthyroidism; screening; routine test data

Both classification problems and QSAR analyses of drugs have been better studied by the multi-layer feedforward neural networks with back-propagation learning than the classical methods in pattern recognition such as linear multiple regression or adaptive least square method, since those problems often involve nonlinear relationships between descriptors and the class (/activity).[1—3] Recently Bayesian regularized neural networks (BRNN),[4—7] which extends back-propagation learning algorithm in order to overcome its defects such as the problems of local trapping, overfitting *etc.* by introducing probabilistic treatment of the Bayesian inference technique for the synaptic weights, has been successfully applied for QSAR studies[8,9] even for massive sample data.[10]

Analytical chemistry is often concerned with the classification of samples into groups on the basis of chemical results while such classification process is called diagnosis if these groups are pathological cases. In the previous work[11] we analyzed the human thyroid data as the three class classification problem, that is, diagnosis by using two notable approaches, the Bayesian regularized neural networks and the self-organizing map.[12—14] The former presented the high classification rates and a nice soft pruning if ARD method is swichted on. The SOM enabled nicely to grasp characeric features of thyroid function from the map of well clustered thyroid patients owing to data-visualization. These results ensured that both approaches are superior to statistical analysis based on conventional multivariate analysis techniques, and would be helpful in diagnosing thyroid function based on laboratory tests. Therefore, we now try to assist the diagnosis of thyroid diseases making use of the recent clinical examination data.

Conventionally, multivariate analysis has been frequently used for analyzing data on medical care.[15] This kind of analysis is implicitly based on the assumption that the sample shows a normal distribution. However, this requirement is often not satisfied because of small sample size or other reasons. The relationship of the routine test data, disease history and physical findings to the diagnosis (disease name) for individual patients seems to assume a non-linear form in many cases. When neural networks are used for analysis, there is no need to assume a normal distribution of the sample. This kind of analysis is applicable even when the sample size is large or small. Furthermore, analysis using neural networks is valid for non-linear relationships.

In Japan, people who go to a hospital because they are complaining of some illness usually undergo routine testing as an early step in the diagnostic process. The data from these routine tests point to one or several possible diagnoses. Full-scale diagnosis is then started with the conduct of detailed tests focused on the diseases suspected on the basis of the disease history, findings from observations and data from the routine tests. However, sometimes the findings are negative for the initially suspected disease, or achieving the final diagnosis takes much time. These outcomes are much more likely when the patient suffers from hyperthyroidism, which has a low prevalence (about 0.2—0.5% of the total population).[16] Furthermore, the cost of the detailed tests is usually high. However, if data from the set of routine tests are analyzed in more detail, we can find hints concerning the possible disease the patient has. For example, patients with hyperthyroidism tend to have a set of elevated alkaline phos-

phatase (ALP) and low total cholesterol (T-Cho),[17,18] while patients with hypothyroidism tend to have a set of low ALP and elevated T-Cho. A thyroid hormone test will be performed on individuals suspected by the physician as having abnormal thyroid function. In this event, a reliable diagnosis is now possible thanks to recent marked advances in the technology for measuring T3, T4, TSH, and the anti-TSH receptor antibody.[19] If the likelihood of a disease can be suggested from the data from a set of routine test parameters, it becomes possible for specialists to start zeroing in on the correct diagnosis quickly, thus leading to an improved patient QOL.

This paper focuses on hyperthyroidism in this context, and examines the possibility of assisting in the identification of this disease by spotting patients that may be hyperthyroid on the basis of routine test data alone.

In our present study, processing the data from a set of routine test parameters using a neural network turned out to be very promising as a means of assisting the diagnosis, and it was found to be possible to discriminate patients with hyperthyroidism from among the many individuals who had received the routine tests. We used the SOM_PAK package[20] developed by Kohonen's group for calculating the SOM, and the "Software for Flexible Bayesian Modeling" package[21] by Neal for the BRNN.

Human Thyroid Data In the present study, the following 14 parameters measured for many samples at the Tohoku University Hospital were analyzed: ALP, gamma glutamyl transferase (γ-GTP), aspartate aminotransferase (AST), alanine aminotransferase (ALT), lactate dehydrogenase (LDH), serum creatinine (S-Cr), uric acid (UA), triglyceride (TG), T-Cho, white blood cell (WBC), red blood cell (RBC), hemoglobin (Hb), hematocrit (HCT), and platelet. Samples from the following individuals with known results of diagnosis were used for this study: 18 patients definitely diagnosed as having hyperthyroidism between 2000 and 2003 (4 males and 14 females) and 48 healthy volunteers (15 males and 33 females). Screening was conducted of 142 individuals (48 males and 94 females) who had received routine tests at the Department of Comprehensive Medicine of the Tohoku University Hospital throughout the 2004 year but for whom no diagnosis had been established.

In the present work we supposed the independence between routine test parameters, though γ-GTP and ALT have a little large correlation coefficient 0.72 among the parameters useful for classification.

Artificial Neural Network Models The routine test data were analyzed using two types of neural networks. One of them was a self-organizing neural network of Kohonen.[12]

This network is superior in terms of the clustering capability and the ability for visualization of data distribution. It can yield SOMs. Its algorithm is simple and universally applicable. High reliability of calculation is also well known for this neural network. If the SOM prepared from samples from individuals with known diagnosis show clustering corresponding to the diagnosis, we can estimate patients with hyperthyroidism from the individuals with unknown diagnosis by determining the location of their routine test data projected on the SOM. Furthermore, if the characteristics of data are analyzed on the component planes for each parameter, it is possible to select parameters useful for making a diagnosis.

Maps are prepared by means of calculation while setting the neuron number (to be arranged on the two-dimensional plane), the neighboring radius of the Gaussian neighborhood function, learning coefficient and learning frequency.

The other type of neural network is the Bayesian regularized neural network.[5,6,11] This is a multi-layer (three-layer) neural network, but was extended to the Bayesian probability framework of treating model parameters, to avoid defects like overfitting encountered in the traditional maximum likelihood approach.[4,5] The routine test parameters are assigned to each neuron of the input layer. The classification rate of patients with hyperthyroidism and that of healthy individuals are given by the two neurons of the output layer, respectively. Softmax function is used as an output function. The sum total of the predictive probabilities is equal to 1. The training of BRNN was carried out by Bayesian methods using the results of diagnosis for known samples. However, when this network was applied in set with the Markov Chain Monte Carlo method,[6] it was found to be a stable network suitable also for complex practical data processing and having a high capability of classification. With this technique, the Bayesian empirical synaptic weight probability spontaneously corresponded to the regularized term of the synaptic weight. Therefore, by means of automatic relevance determination (ARD), a series of inputs (routine test parameters) having large impacts on the judgment can be identified. The predictive capability of the neural network can be elevated by assigning an appropriate number of neurons to the intermediate layer (*i.e.*, model complexity) which can be determined using only the training set. In the present paper we used numbers between 5 and 15 as good models, similar to the work using back-propagation neural network by Zhang and Berardi.[22] This is actually the case for a few input parameters analyzed in the section "Elevating the accuracy of screening," and the results almost unchanged at these numbers. When we made BRNN calculation using 14 routine tests, however, the results change depending on the numbers of intermediate neurons and on switching on (or off) ARD. Rule of thumb (larger is the ratio of sample size to synaptic weights, the model works better) should be alleviated in the BRNN[9] and Roberts *et al.* reported that whether switching on ARD function is better or not depends on the category of sample data.[7,23] Then we picked up test samples as hyperthyroid patient when they are predicted by a lot of models.

Results and Discussion

Screening of Patients with Hyperthyroidism: Learning Data After the data from individual patients were preprocessed to normalize as to each parameter, the SOM shown in Fig. 1 was obtained. This SOM was calculated under the following conditions: neuron number 30×20, neighboring radius $r=30$, learning coefficient $\alpha=0.15$, and learning frequency 50000. It was checked that the calculated results remain almost unchanged when the number of neurons in the competitive layer are changed to some extent like 30×20 ($r=20$) and 40×30 ($r=30$ or 40). The data were distributed in two distinctive zones, *i.e.*, a zone for healthy individuals (label 3, red, 3-1 to 3-48, $n=48$) and a lower right zone for patients with hyperthyroidism (label 1, green, 1-49 to 1-68, $n=18$; excluding 1-57 and 1-65). It was thus possible to nicely identify patients with hyperthyroidism. Utilization of
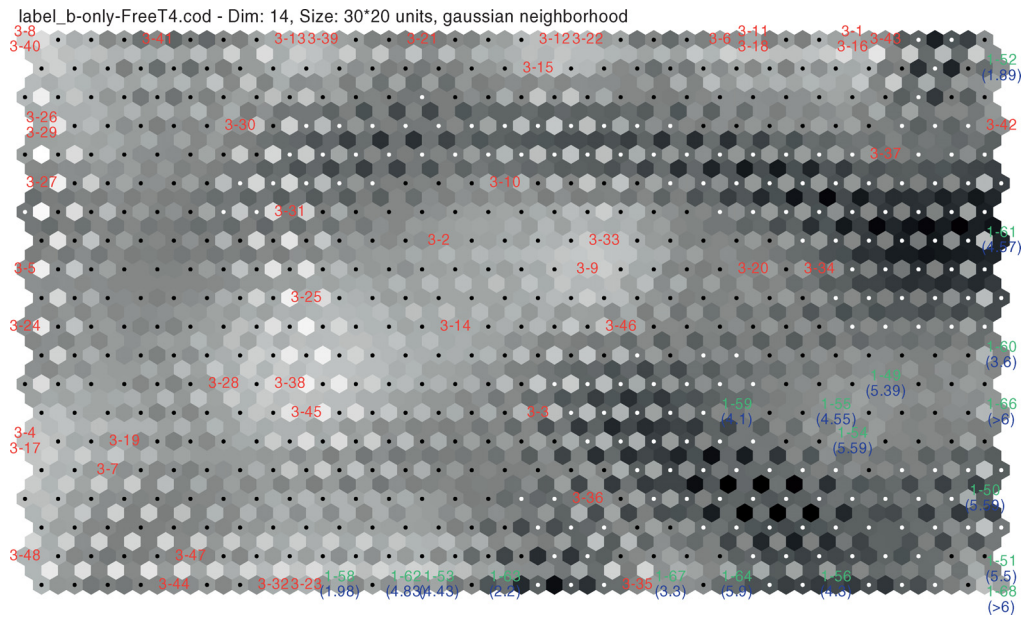
label_b-only-FreeT4.cod - Dim: 14, Size: 30*20 units, gaussian neighborhood

Fig. 1. A 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. The blue label indicates the free T4 level (ng/ml) for individual patients.
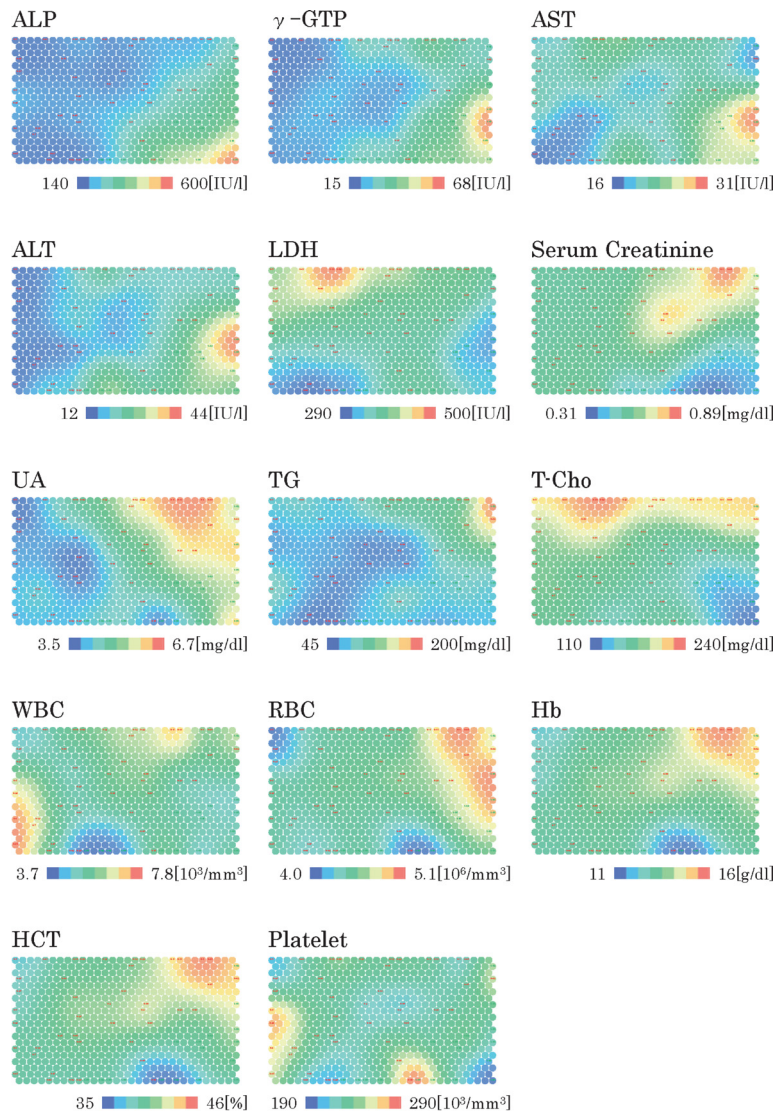

Fig. 2. Component Planes for 14 Parameters Accompanying the SOM Shown in Fig. 1

such a difference in distribution on this map makes it possible to analyze the characteristics of test parameters and to estimate patients with hyperthyroidism among the individuals tested.

When the SOM was compared with the information on medical records, it was found that patients with severe hyperthyroidism were distributed in the deeper area of the zone for hyperthyroidism, while patients with mild hyperthyroidism were distributed in the part of the hyperthyroidism zone closer to the zone for healthy individuals. Case 1-52 (free T4=1.89 ng/ml, free T3=7.3 pg/ml), who was erroneously assigned to the zone for healthy individuals, showed few abnormalities in hormones. A similar trend was also seen in case 1-58 (free T4=1.98 ng/ml), 1-63 (free T4=2.2 ng/ml) who were assigned to the borderline zone. The healthy individual 3-35, who was assigned to a point slightly entering the hyperthyroidism zone, had a slightly high ALP and slightly low S-Cr and T-Cho (characteristics akin to those of patients in the hyperthyroidism zone).

The diagnostic data for 66 individuals with known diagnosis (18 patients with hyperthyroidism and 48 healthy individuals) were entered as teacher signals into the Bayesian regularized neural network. After 400 rounds of Monte Carlo sampling, diagnosis was predicted by the calculation using the leave-one-out (LOO) method (output by the softmax function). When the number of neurons in the intermediate layer was 8, 10 and 12, the predictive accuracy was about 90%. The patients with hyperthyroidism who were erroneously predicted as healthy by BRNN were approximately identical to the patients erroneously predicted by SOM calculation.

Then, we attempted to identify routine test parameters useful in the diagnosis of hyperthyroidism. In SOM calculation, the component planes in Fig. 2 obtained for each parameter simultaneously with SOM shown in Fig. 1 allows an instantaneous judgment as to whether the level of a given parameter is high (red) or low (blue). It is thus possible to identify parameters among the 14 parameters useful for classification. Among the component planes shown in Fig. 2, parameters which show a characteristic hyperthyroidism zone in the lower right area are selected for more detailed analysis. Of these parameters, ALP, γ-GTP, AST and ALT have levels higher than the criterion levels, while S-Cr and T-Cho have levels lower than the criterion levels. The normal range in the routine tests at the Department of Comprehensive Medicine of the Tohoku University Hospital are 112—330 IU/L (ALP), 8—57 IU/L (γ-GTP), 12—30 IU/L (AST), 8—35 IU/L (ALT), 0.4—0.8 mg/dl for female and 0.6—1.1 mg/dl for male (S-Cr) and 128—220 mg/dl (T-Cho). These six parameters are judged as being important in the diagnosis of hyperthyroidism. For TG and WBC, the broad area including the hyperthyroidism zone takes a certain level and we cannot rule out some association of these two parameters with hyperthyroidism. The other parameters are judged to have low association with hyperthyroidism.

Next, we calculated the degree of relevance of 14 parameters by means of BRNN with ARD, to identify parameters useful in the diagnosis of hyperthyroidism. As an illustration, calculated relevances for females are shown in Fig. 3 (which is helpful for later discussion). Parameters ALP, γ-GTP, AST, ALT, S-Cr, T-Cho, *etc.*, were found to be useful in the diag-
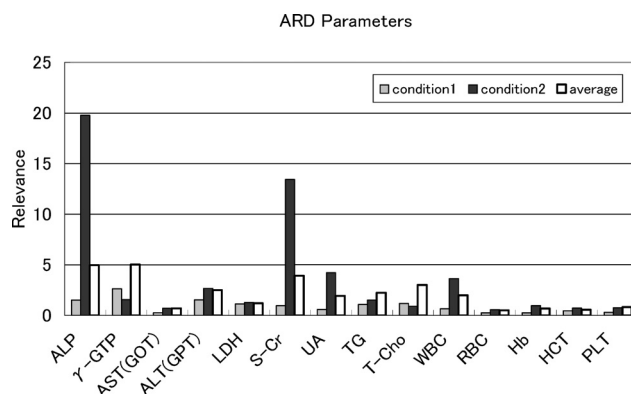

Fig. 3. Relevances for 14 Routine Test Parameters for Females Obtained by the ARD Method of BRNN Calculation

Average relevances of those at intermediate neurons 5, 8, 9, 10, 11, 12 and 15 are shown as white bars. Typical examples at intermediate neurons 5 (condition 1) and 12 (condition 2) are also shown.

nosis of hyperthyroidism, approximately consistent with the results of SOM's component plane analysis.

Screening of Patients with Hyperthyroidism: Test Data
When the 142 individuals tested (48 males and 94 females) were projected on the SOM (Fig. 1) depending on their test data as shown in Fig. 4, about 10 individuals were projected into the hyperthyroidism zone (lower right area) and many of the remaining individuals were projected into the healthy zone, although the results varied slightly depending on the models used for calculation. It seems likely that patients who were sick but had no characteristic sign specific to hyperthyroidism were projected into the border area between the hyperthyroidism zone and the healthy zone. When the data from 142 individuals with unknown disease were analyzed by the BRNN after learning of the data from 66 patients with known disease, to select patients suspected of having hyperthyroidism, there were about 10 patients for whom most of the BRNN calculations with intermediate neurons (5, 8, 10, 12, 15) predicted as hyperthyroid with probability more than 80%. (In Table 3, we will show predicted hyperthyroid patients with high probability by the BRNN calculation using 14 routine tests for 94 female test samples.) When the results of prediction with these two types of neural networks were taken together, hyperthyroidism was suspected at a high probability in 7 individuals (examinees 31, 54, 74, 167, 176, 217, 285). Of these 7 examinees, three women (examinees 167, 217, 285) were later diagnosed definitely by the physician as having hyperthyroidism.

In the SOM, all of these three patients were distributed in the deepest area of the hyperthyroidism zone, and they had been rated by BRNN as having hyperthyroidism at a high probability. The remaining 4 individuals who were erroneously estimated as having hyperthyroidism actually had hepatic dysfunction. This is probably explained as follows. Since the neural network had learned only the data from hyperthyroidism patients and healthy individuals, individuals who actually did not have hyperthyroidism but whose test data resembled the data for patients with hyperthyroidism were estimated by the neural network as having hyperthyroidism. Thus, the results of prediction with SOM were approximately identical to those with BRNN. Individuals selected by these methods of screening were found to include
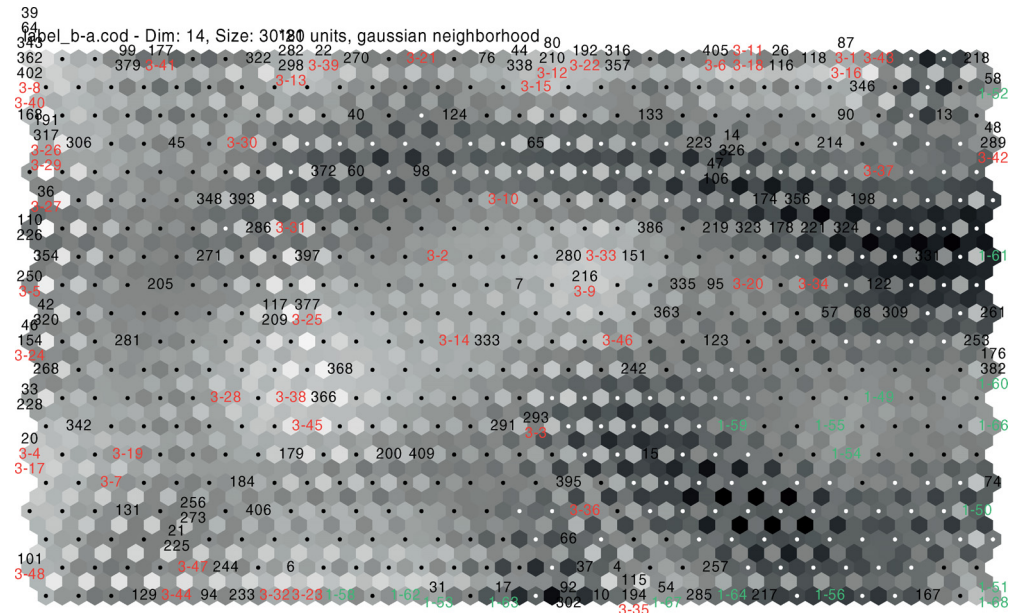
Fig. 4. The Same SOM as Fig. 1, However, 142 Examinees (Black Labels) are Projected onto It Now
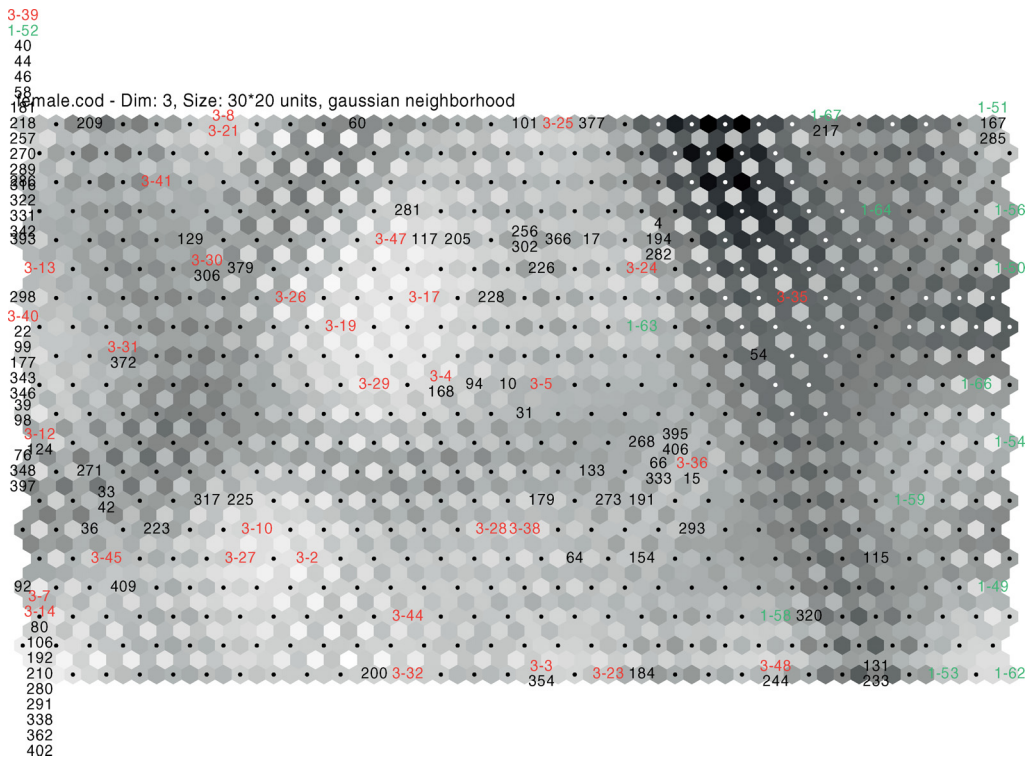

Fig. 5. Similar SOM as Fig. 1, But Now Only Taking 47 Diagnosed Females (14 Green-Labeled Hyperthyroids and 33 Red-Labeled Normals) Using Only Three Parameters ALP, S-Cr and T-Cho. Then, 94 Female Examinees (Black Labeled) are Projected onto the Map

30×20 neurons in the competitive layer, $r=20$, and 50000 times of learning cycles.

patients with hyperthyroidism at a high probability.

Elevating the Accuracy of Screening Next, we examined the possibility of elevating the accuracy of screening. Using the 6 parameters previously rated as important for the diagnosis of hyperthyroidism, we carried out BRNN calculation without ARD for sets composed of two to six parameters shown in Table 1. Because the S-Cr level differed greatly between males and females (normal range: 0.6—1.1 mg/dl for males and 0.4—0.8 mg/dl for females) and because hyperthy-

roidism is more frequent in females than in males, we analyzed the data for males separately from data for females. Table 1 shows the number of wrong predictions for the 14 females with hyperthyroidism and 33 healthy females when the prediction was based on BRNN calculation by the LOO method using data from females with known diagnosis (learning data). As shown in this table, the accuracy of prediction was low when a set of 2 parameters or a set of 5 or 6 parameters was used. For group No. 3-2 through 4-4, the

Table 1. Results of the Number of Wrong Predictions by BRNN-LOO (without ARD) Calculation of Data from Females with Known Diagnosis (Learning Data)

No.	Set of variables						Hyper-thyroid	Normal
2-1	ALP					T-Cho	5	3
3-1	ALP		AST		S-Cr		3	3
3-2	ALP				S-Cr	T-Cho	3	1
4-1	ALP		AST	ALT		T-Cho	2	2
4-2	ALP		AST		S-Cr	T-Cho	3	1
4-3	ALP			ALT	S-Cr	T-Cho	2	4
4-4		γ-GTP	AST		S-Cr	T-Cho	3	2
4-5		γ-GTP		ALT	S-Cr	T-Cho	4	3
5-1	ALP		AST	ALT	S-Cr	T-Cho	4	4
6-1	ALP	γ-GTP	AST	ALT	S-Cr	T-Cho	5	3

There were 14 hyperthyroidism females and 33 normal ones. X in the group number (X-Y) indicates the number of parameters used and Y represents the kind of parameter set. The results are nearly the same irrespective of intermediate neurons of 5, 10 and 12.

Table 2. Results of Prediction by BRNN (without ARD) Calculation of Data Collected at the Department of Comprehensive Medicine of the Tohoku University Hospital (94 Females)

No.	Set of variables						Threshold (%)	No. of wrong predictions
2-1	ALP					T-Cho	80	0
3-1	ALP		AST		S-Cr		92	0
3-2	ALP				S-Cr	T-Cho	95	0
4-1	ALP		AST	ALT		T-Cho	65	3.6
4-2	ALP		AST		S-Cr	T-Cho	89	0
4-3	ALP			ALT	S-Cr	T-Cho	80	1.0
4-4		γ-GTP	AST		S-Cr	T-Cho	80	9.8
4-5		γ-GTP		ALT	S-Cr	T-Cho	66	8.8
5-1	ALP		AST	ALT	S-Cr	T-Cho	85	0.8
6-1	ALP	γ-GTP	AST	ALT	S-Cr	T-Cho	80	5.0

Averages of the threshold (%) and the number of wrong predictions for 10 sets of parameters are shown at intermediate neurons of 5, 8, 10, 12 and 15, since the different number of neurons produce almost the same results.

error in the diagnosis of hyperthyroidism occurred at a high probability (80—90%) in cases 1-52 and 1-58. In case 1-63, the probability of error was close to 50%. Thus, the number of erroneous predictions changes to 2 or 3 as the probability of prediction is slightly varied by the use of different sets of parameters. If these findings are also taken into account, the set Nos. 3-2, 4-1 and 4-2 seem to be excellent. ALP resembles γ-GTP, and AST is akin to ALT (all of these four parameters are usually associated with hepatic dysfunction). However, since the accuracy was higher for No. 4-2 than for No. 4-4, and it was higher for No. 4-3 than for No. 4-5, ALP can be deemed as a better indicator of hyperthyroidism than γ-GTP. It is interesting to note the recent report[24] that ALP is a useful parameter in order to increase the diagnostic accuracy for Graves' disease. Furthermore, since the accuracy was higher for No. 4-2 than for No. 4-3 and it was higher for No. 4-4 than for No. 4-5, AST can be deemed as a slightly better indicator of hyperthyroidism than ALT.

Table 2 shows the results of prediction by BRNN (without ARD) calculation of data collected at the Department of Comprehensive Medicine of the Tohoku University Hospital (test data), with the threshold (%) and the number of wrong predictions serving as indicators. The data analyzed were derived from 94 females. Here, the threshold (%) means that the three females (cases 167, 217, 285) finally diagnosed as having hyperthyroidism are included in the examinees estimated as having hyperthyroidism at a probability higher than the threshold level. The number of wrong predictions means the number of examinees who were actually free of hyperthyroidism although included in the examinees estimated as having hyperthyroidism at a probability higher than the threshold level.

Table 2 indicates that variable set 3-1 (ALP, AST and S-Cr), set 3-2 (ALP, S-Cr and T-Cho) and set 4-2 (ALP, AST, S-Cr and T-Cho) allow highly accurate prediction of hyperthyroidism and highly accurate screening since individuals with hepatic dysfunction can be eliminated by these sets of variables. However, when the probability of wrong prediction was analyzed (not shown in Table 2), no wrong prediction of hyperthyroidism was seen among the individuals estimated by set 3-2 as having hyperthyroidism at a probability of 78% or higher, while some wrong predictions were observed among the individuals estimated with set 3-1 or 4-2 at a probability of 89% or higher. If this result is taken into account, set 3-2 seems to be particularly excellent. Like Tables 1 and 2 clearly shows that ALP serves as a better indicator than γ-GTP for prediction of hyperthyroidism, and that AST is slightly better than ALT. The results shown in Tables 1 and 2 allow us to say that if only ALP is selected from the known major indicators of liver function (ALP, γ-GTP, AST and ALT) and if a set of variables composed of ALP, S-Cr and T-Cho is used, highly accurate screening of patients with hyperthyroidism (capable of eliminating individuals free of hyperthyroidism but having hepatic dysfunction) is possible. Considerably accurate screening is also possible if a set composed of four variables, ALP (selected from ALP and γ-GTP), AST (selected from AST and ALT), S-Cr and T-Cho is employed.

In Table 3 we show predicted hyperthyroid patients with high probabilities by BRNN calculation with (/without) ARD for 94 female test samples. Models are of intermediate neurons 5, 8, 10, 12 and 15. Bald numbers denote hyperthyroid patients diagnosed later by the physician. Predictive capability of BRNN using three parameters ALP, T-Cho and S-Cr are much higher than those using 14 routine test parameters. Besides, results of BRNN calculation with ARD are almost similar to that without ARD for three parameters, while the former is worse than the latter for 14 routine test parameters.

Results similar to those yielded by BRNN calculation were obtained in each SOM prepared from sets of 2 to 6 variables like the sets shown in Tables 1 and 2. One example is shown in Fig. 5. The SOM shown here was prepared from set 3-2 of variables (ALP, S-Cr and T-Cho) shown in Table 1. The healthy zone (3-X, red label) is well separated beyond the black valley from the hyperthyroidism zone on the right (1-X, green label). The black label indicates the examinee number at the Department of Comprehensive Medicine (test data). It was projected on the SOM depending on the routine test data. Examinees 167, 217 and 285 are projected into the deeper area of the hyperthyroidism zone. As stated already, only these three individuals were finally diagnosed by the physician as having hyperthyroidism. No other examinee was projected into the hyperthyroidism zone, although several examinees were projected into the border area. As explained in

Table 3. Predicted Hyperthyroid Patients with a High Probability by BRNN Calculation Preprocessed to Normalize Learning Data (/Test Data), with ARD (/without ARD) Using 14 Routine Test Data for 94 Female Test Samples

Prob. of hyperthyroid	Intermediate neurons	14 parameters		ALP, T-Cho, S-Cr	
		with ARD	without ARD	with ARD	no ARD
90—100%	15		**285**	**167**, **285**	**167**, **217**, **285**
	12	**217**, 331	**285**, 331	**167**, **217**, **285**	**167**, **217**, **285**
	10	17, 58, 192, **217**, 331	**285**, 331	**167**, **217**, **285**	**167**, **217**, **285**
	8		**285**	**167**, **217**, **285**	**167**, **217**, **285**
	5	**217**, 331	331	**167**, **217**, **285**	**167**, **217**, **285**
80—90%	15	31, **217**, 331	31, **167**, **217**, 331	**217**	
	12	**167**	31, 58, **167**, **217**		
	10	10, 31, 46, **167**, 168, 223, 291, 362, 402	31, **167**, **217**, **285**, 331		
	8	331	31, 58, **167**, **217**, 331		
	5	17, 31, 58, **167**, 192	31, 58, **167**, **217**, **285**		
70—80%	15	**167**	58	115	257
	12	17, 31, 58, **285**	192		257
	10	316	58	54, 257	257
	8	17, 31, **167**, 192, **217**, 223	192		257
	5	223, **285**	131, 192	257	257

Models are of intermediate neurons 5, 8, 10, 12 and 15. Bold numbers denote hyperthyroid patients diagnosed later by the physician.

Fig. 1, Examinees 1-52, 1-58, 1-63 and 3-35 (individuals with known hyperthyroidism or healthy individuals) were erroneously assigned to the healthy zone or the hyperthyroidism zone.

There were 19 males with known diagnosis (15 healthy males and 4 males known to have hyperthyroidism). Prediction of hyperthyroidism was performed on 48 individuals tested at the Department of Comprehensive Medicine. None of these 48 males were later diagnosed by the physician as having hyperthyroidism. The results of SOM and BRNN calculations for the male group were similar to those for the female group. That is, the accuracy of screening was highest with variable set 3-2 composed of three variables (ALP, S-Cr and T-Cho). However, the accuracy with this set was lower in the male group than in the female group probably partly because too small sample size. The mean number of males who actually had hepatic dysfunction but were estimated as having hyperthyroidism at a probability over 80% was 1.5 when the neuron number of the intermediate layer was 5, 8, 10, 12 and 15.

Comparison to Empirical Model of Thyroid Specialist
In the present study, screening with a four-variable set (ALP, AST, S-Cr and T-Cho) or a three-variable set (ALP, S-Cr and T-Cho) was found to allow all patients with hyperthyroidism to be picked up and to have a considerably high accuracy. Specialists in thyroid diseases usually consider the possibility of hyperthyroidism in individuals showing both elevated ALP and reduced T-Cho.[17,18] It is also known that patients with hyperthyroidism often show elevation in AST and ALT.[25] It is noteworthy that ALP, AST, ALT and T-Cho were included in the variables rated in the present study as useful in the diagnosis of hyperthyroidism using neural networks. The results of the present study using neural networks additionally indicate a large impact of S-Cr (to which physicians usually did not pay close attention) on the accurate diagnosis of hyperthyroidism. The importance of S-Cr was also endorsed by the finding that the accuracy of prediction decreased considerably as the number of variables used for prediction was decreased from 14 (including S-Cr) to 13 (excluding S-Cr).

Conclusion
The results of the present study can be rated as being highly reliable in view of the finding that an approximately identical screening was possible by each of the two different types of neural network (an SOM type and a Bayesian regularized type). If the present method of diagnostic assistance is applied to predict individuals suspected of having hyperthyroidism among the individuals consulting hospitals because of sickness, it will allow the smooth start of appropriate diagnosis and treatment by specialists in thyroid diseases, leading to improved QOL of patients. This kind of diagnostic assistance is expected to be useful clinically.

In the future, it is desirable that the weight-incorporated neural network tool that we developed is used by many clinicians as a means of diagnostic assistance for analysis of routine test data in clinical practice. Then the much more accurate and robust screening would be possible owing to increasing sample size.

References and Notes
1) Ichikawa H., "Kaisougata Nyuraru Nettowa-ku," Kyouritsu Co., Tokyo, 1993.
2) So S. S., Karplus M., *J. Med. Chem.*, **39**, 5246—5256 (1996).
3) Sato K., Nakagawa J., *Chem. Pharm. Bull.*, **45**, 107—115 (1997).
4) Bishop C. M., "Neural Networks for Pattern Recognition," Oxford University Press, 1995.
5) MacKay D. J. C., *Network: Computation in Neural Systems*, **6**, 469 (1995).
6) Neal R. M., "Bayesian Learning for Neural Networks," Springer, New York, 1996.
7) Husmeir D., Penny W. D., Roberts S. J., *Neural Networks*, **12**, 677 (1999).
8) Sato K., Nakagawa J., Matuzaki H., *J. Tohoku Pharmaceutical University*, **44**, 187—193 (1997).
9) Burden F. R., Ford M. G., Whitley D. C., Winkler D. A., *J. Chem. Inf. Comput. Sci.*, **40**, 1423—1430 (2000).
10) Ajay, Walters W. P., Murcko M. A., *J. Chem.*, **41**, 3314—3324 (1998).
11) Hoshi K., Kawakami J., Kumagai M., Kasahara S., Nisimura N., Nakamura H., Sato K., *Chem. Pharm. Bull.*, **53**, 1570—1574 (2005).
12) Kohonen T., "Self-Organizing Maps," Springer, Berlin, 2000.
13) Tokutaka H., Kisida S., Fujimura K., "Jikososikikamappu no Ouyou," Kaibundou Co., Tokyo, 1999.
14) Kawakami J., Hoshi K., Ishiyama A., Miyagishima K., Sato K., *Chem. Pharm. Bull.*, **52**, 751—755 (2004).
15) Coomans D., Jonckheer M., Massart D. L., *Anal. Chim. Acta*, **103**,

409—415 (1979).

16) Cooper D. S., *Lancet*, **362**, 459—468 (2002).
17) 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).
18) http://thyroid.umin.ac.jp/guideline/03.html
19) Weetman A. P., *N. Engl. J. Med.*, **26**, 1236—1248 (2000).
20) Kohonen's group of Helsinki University of Technology (http://www. cis.hut.fi/research/som-research/nnrc- programs.shtml)
21) Neal R., Flexible Bayesian Modeling (http://www.cs.toronto.edu/ ~radford/fbm.2004-11-10.doc/index.html)
22) Zhang G., Berardi V. L., *Health Care Manag. Sci.*, **1**, 29—37 (1998).
23) Penny W. D., Roberts S. J., *Neural Networks*, **12**, 877 (1999).
24) Yanagisawa T., Sato K., Kato Y., Shimizu S., Takano K., *Endocr. J,* **52**, 29—36 (2005).
25) Gurlek A., Cobankara V., Bayraktar M., *J. Clin. Gastroenterol.*, **24**, 180—183 (1997).