

Predicting Thyrotoxicosis in Patients Using a Set of Routine Tests: Adding their Rate of Annual Time-Series Variations to Self-Organizing Map-Based Predictive Model Improves Diagnostic Accuracy

Sorama Aoki, Sono Nishizaka, Kenichi Sato, Kenji Hoshi, Junko Kawakami
 Medical and Pharmaceutical Information Science, Tohoku Pharmaceutical University,
 4-4-1 Komatsushima, Aoba-ku, Sendai 981-8558, Japan.
 Email: s-aoki@tohoku-pharm.ac.jp, 21452111@is.tohoku-pharm.ac.jp,
 {ksato, hoshi, jnaka}@tohoku-pharm.ac.jp

Kouki Mori
 Center for Health Promotion,
 JR Sendai Hospital
 Sendai, Japan.
 kouki-mori@jreast.co.jp

Yoshinori Nakagawa
 Sendai Thyroid Clinic
 Sendai, Japan.
 na@sendaikojosen.com

Wataru Hida, Katsumi Yoshida
 Department of Health Supervision,
 Tohoku Kosai Hospital
 Sendai, Japan.
 wa-hida@tohokukosai.com,
 kayosimd@beetle.ocn.ne.jp

Abstract— Difficulties have been associated with accurately diagnosing patients with thyroid dysfunction (PTD); however, measuring thyroid hormone levels in all individuals is challenging. We successfully constructed a prediction model for PTD by adopting pattern recognition methods using a combination of six routine laboratory tests, and identified 21 new PTD using our screening method, which was executed at two health check-up centers. In the present study, we newly introduced time-series variations in routine tests as additional parameters in order to develop the model by eliminating the influence of individual differences in routine tests. We constructed self-organizing maps (SOM) using the time-series traceable data of 13 PTD and 45 healthy individuals. We then investigated the locations of 140 projected false positives in our previous study on SOM and found that the number of false positives markedly decreased, thereby demonstrating the progression of our new model.

I. INTRODUCTION

PATIENTS with thyroid dysfunction (PTD) are often overlooked and misdiagnosed. For example, thyrotoxicoses have been misdiagnosed as a heart disease or malignant tumor of the digestive tract, while hypothyroidism has been misdiagnosed as muscle, heart, or liver disorders or hyperlipidemia in general screening by doctors or internists. Therefore, these patients frequently receive the wrong treatment [1]–[4]. Thyroid dysfunction progresses slowly as a result of PTD being overlooked or misdiagnosed in addition to patients often not being aware of and tolerating their

symptoms, which results in many patients remaining unexamined.

Thyroid specialists previously reported difficulties in identifying PTD based on physical findings alone; therefore, the measurement of TSH levels is considered indispensable [5], [6]. However, TSH levels cannot be measured in all individuals who visit a clinic due to the associated costs, such that receiving a full health check-up in this regard is difficult in terms of cost-effectiveness and remains a topic of debate [7].

Since an excess or lack of thyroid hormones influences the whole body and produces abnormal routine test results, we attempted to construct a predictive model using an appropriate set of routine tests, which are measured in the health check-up system, in order to detect PTD using a low-cost method. In our previous studies, we analyzed the routine test data of PTD and healthy individuals using three types of pattern recognition methods (PRM) [8]–[10] in addition to medical statistics. We found that PRMs with three parameters (an elevation in alkaline phosphatase (ALP) and decreases in serum creatinine (S-Cr) and total cholesterol (TC)) [11]–[13], or four parameters (the previous three parameters in addition to an elevation in heart rate (HR)) [14] allowed accurate screening for hyperthyroidism (thyrotoxicosis). We also reported that PRMs with another set of four parameters (elevations in lactate dehydrogenase (LDH), S-Cr, and TC, and a lower red blood cell count (RBC)) allowed accurate screening for hypothyroidism [15]. We applied these predictive models to the screening of 4,355 Japanese people whose routine test data had already been measured in the general health check-up system, referred to as “the Ningen Dock”, at JR Sendai Hospital between July 2008 and December 2011, and identified 7 overt PTD (2

This work was supported in part by a Grant-in-Aid for Scientific Research (23590811, 26460777) from the Japan Society for the Promotion of Science.

patients with Graves' disease, 2 with painless thyroiditis, and 3 with hypothyroidism), who were subsequently treated by thyroid specialists [14], [16]. None of the 7 PTD had expressed concerns regarding their health. We also applied these predictive models to the screening of 8,831 examinees of the Ningen Dock at Tohoku Kosai Hospital, which is a larger scale institution than JR Sendai Hospital, between September 2011 and March 2013, and successfully identified 14 overt PTD (8 patients with Graves' disease, 2 with painless thyroiditis, and 4 with hypothyroidism) [17]. Although we identified 21 undetected PTD, there were 218 false positives (91 at JR Sendai Hospital and 127 at Tohoku Kosai Hospital) in our screening who were subsequently found to be normal after measuring serum levels of thyroid hormones (free thyroxin (FT4) and thyroid stimulating hormone (TSH)). Most of these false positives originated from the low threshold of the predicted probability (60%) in the screening to prevent false negatives. An analysis of the screening results at the Ningen Dock of Tohoku Kosai Hospital revealed that setting a higher threshold value for probabilities (85%) that satisfied the condition of not overlooking patients with Graves' diseases and with overt hypothyroidisms, who were found in the screening, decreased the number of false positives to 45 [17]. Although this threshold level caused two false negatives, both were patients with painless thyroiditis who did not require medical treatment; therefore, overlooking these cases did not yield any clinical problems.

We found that many of the time-series variations in the routine tests of false positives did not always change toward a pattern characteristic of a thyroid dysfunction. In order to quantitatively evaluate these time-series changes, we parsed time-series variations by calculating the average rates of the annual time-series variation (RATV) in each of the routine test data between previous and current visitations to the Ningen Dock. We only selected cases suspected of thyrotoxicosis in our screening with a previous visitation record to the Ningen Dock of within three years. In each time-series traceable subject, we significantly decreased the number of false positives from 140 to nine, while maintaining the true positives of Graves' disease, by simultaneously plotting both the predicted probability and average RATV in a scattergram [18].

In the present study, we attempted to construct an advanced predictive model with PRMs, especially Self-Organizing maps (SOM), by introducing the 4 RATVs of routine tests (ALP, S-Cr, TC, and HR) in subjects suspected of having thyrotoxicosis in our screening to analyze RATV more non-linearly than our former analysis. We then determined whether introducing RATV into the SOM model improved diagnostic accuracy.

This study was approved by the Ethical Review Boards of Tohoku Kosai Hospital and JR Sendai Hospital. Data of all subjects were handled using linkable anonymization.

II. SUBJECTS AND METHODS

A. Subjects

At the Ningen Docks of JR Sendai Hospital (between July 2008 and December 2014) and Tohoku Kosai Hospital (between September 2011 and March 2013), we identified 13 true positives (6 at JR Sendai Hospital and 7 at Tohoku Kosai Hospital) and 140 false positives (104 at JR Sendai Hospital and 36 at Tohoku Kosai Hospital) of thyrotoxicosis in our screening, whose previous visitation records to the Ningen Dock were within three years. Furthermore, 45 healthy female controls were collected from the Ningen Dock of JR Sendai Hospital to train normal data for the SOM model; therefore, 198 subjects were parsed in the present study. In those screened, we used the predictive screening tool shown in Figure 1 and obtained the predictive probability for examinees after inputting six routine tests and the presence/absence of medication with dyslipidemic therapeutic drugs. The threshold level of the predicted probability for the need for hormone measurements was set to 60%. The mean \pm SD interval days of visitation was 383 ± 68 days for true positives, 379 ± 79 days for false positives, and 399 ± 100 days for healthy controls; every interval day was nearly equal to a year.



Fig. 1. The interface of our simple screening tool based on a predictive model that instantaneously yields classification results for examinees if their routine test (ALP, S-Cr, TC, LDH, RBC, and HR) data are input [17].

Table 1 shows the 13 patients (true positives) with thyrotoxicosis identified by our screening. The clinical and laboratory features of each patient shown in the table are the current visitation values. The diagnostic results of patients who consulted the Thyroid Outpatients Service of JR Sendai Hospital or Tohoku Kosai Hospital according to our encouragement after our screening are shown for hyperthyroidism as Graves' disease or destructive thyroiditis (painless or sub-acute).

B. Rates of annual time-series variations in routine tests

In order to handle each time-series variation in the routine tests quantitatively, we calculated the rates of annual time-series variations in each subject's routine test data between previous and current visitations to the Ningen Dock [18]. In this calculation, we divided the interval days of visitations in each subject by 365 days in order to normalize inconsistencies in the interval day; therefore, the rate of the annual time-series variation (RATV) was obtained by the following equation:

$$v_R = \frac{R_c - R_p}{R_p} \times \frac{1}{\frac{D_c - D_p}{365}} \quad (1)$$

where v and R denote the RATV and measured value of each routine test (ALP, TC, S-Cr, and HR), and D denotes the visitation date to the Ningen Dock. Subscript c denotes the current visitation, while p denotes the previous visitation to the Ningen Dock. Therefore, v_{ALP} , v_{TC} , v_{S-Cr} and v_{HR} were obtained for each subject.

C. Analyses using pattern recognition methods

As we reported previously [11]-[13], [15], routine test data were analyzed using three types of PRMs together in order to make our predictive model more robust. The first method, SOM [8], is a simplified model that reflects biological neural networks that realize character extraction in the cortex, and is also a non-linear extension of a principal component analysis in statistics. This method is superior in terms of classification capability and visualization of data distribution. If the SOM prepared from the samples from individuals with definite diagnoses shows clustering reflecting these diagnoses, we have the ability to predict which patients have thyrotoxicosis

by determining the locations of individuals with unknown diagnoses projected on the SOM according to their routine test data. Furthermore, by analyzing the characteristics of the data within the component plane for each parameter, it is possible to select parameters that are useful for making a diagnosis [11]-[13], [15]. In the present study, we used the SOM_PAK package [19] after adding a facility of our own. We then adopted Bayesian regularized neural networks (BRNN) [9]. This is a multi-layer neural network, but was extended to include the Bayesian probability framework for treating model parameters in order to avoid defects such as overfitting, which are encountered in the conventional maximum likelihood approach. We used the Software for Flexible Bayesian Modeling package [20] by Neal for the BRNN. We then adopted the support vector machine (SVM) [10] with the LIBSVM package [21] and selected radial basis function kernels. We showed in Figure 1 the interface of our predictive screening tool [17], which displays the three types of predictive outputs for both hyperthyroidism and hypothyroidism in an examinee. The calculated probabilities by BRNN and SVM are given by %. We judged through visual observations whether each examinee had a thyroid dysfunction as well as the degree of severity by noting the location projected on the SOM (labeled by a yellow star).

In the present study, we added four RATV parameters (v_{ALP} , v_{TC} , v_{S-Cr} , and v_{HR}) to the set of four routine tests (ALP, TC, S-Cr, and HR) used to date; therefore, we adopted a set of eight parameters (ALP, TC, S-Cr, HR, v_{ALP} , v_{TC} , v_{S-Cr} , and v_{HR}) to construct the predictive model by using the SOM. We initially constructed the SOM using the training data set (13 true positives and 45 healthy controls) to determine whether these eight parameters had the ability to classify these two groups. One hundred and forty

TABLE 1.
CLINICAL AND LABORATORY FEATURES OF 13 TIME-SERIES TRACEABLE PATIENTS WITH OVERT THYROTOXICOSIS IDENTIFIED IN OUR SCREENING, WHOSE PREVIOUS VISITATION RECORDS TO THE NINGEN DOCK WAS WITHIN THREE YEARS

	Subject No.	Sex	Age	Diagnosis	FT4 (ng/dL)	TSH (μ IU/mL)	ALP (IU/L)	S-Cr (mg/dL)	TC (mg/dL)	HR (/min)
Tohoku Kosai Hospital	K1	Female	41	Painless Thyroiditis	1.76 ^{*1}	<0.01	216	0.47	164	81
	K2	Female	54	Graves' Disease	3.24 ^{*1}	<0.01	154	0.37	131	89
	K3	Male	54	Graves' Disease	3.56 ^{*1}	<0.01	376	0.65	148	70
	K4	Female	57	Graves' Disease	3.35 ^{*1}	<0.01	396	0.57	134	81
	K5	Female	61	Graves' Disease	3.65 ^{*1}	<0.01	375	0.33	128	58
	K6	Female	47	Graves' Disease	2.25 ^{*1}	<0.01	392	0.45	146	92
	K7	Female	41	Graves' Disease	3.78 ^{*1}	<0.01	306	0.53	128	80
JR Sendai Hospital	JR1	Male	48	Painless Thyroiditis	1.8 ^{*2}	<0.005	198	0.56	184	102 ^{*3}
	JR2	Male	48	Graves' Disease	5.6 ^{*2}	<0.005	494	0.56	106	74 ^{*3}
	JR3	Female	35	Graves' Disease	5.35 ^{*2}	<0.005	468	0.33	162	85 ^{*3}
	JR4	Female	42	Thyrotoxicosis ^{*4}	2.2 ^{*2}	0.008	135	0.44	171	92 ^{*3}
	JR5	Female	49	Thyrotoxicosis ^{*4}	1.99 ^{*2}	0.092	122	0.51	132	80 ^{*3}
	JR6	Male	51	Graves' Disease	3.28 ^{*2}	<0.005	391	0.67	165	85 ^{*3}

*1: The reference range of FT4 in the Tohoku Kosai Hospital Ningen Dock was 0.70-1.48[ng/dL]

*2: The reference range of FT4 in the JR Sendai Hospital Ningen Dock was 0.90-1.70[ng/dL]

*3: Heart rate in JR Sendai Hospital was the pulse rate measured by an automated sphygmomanometer (sitting position), while heart rate in Tohoku Kosai Hospital was determined by ECG (recumbent position).

We corrected the pulse rate to heart rate using a regression formula [14] when calculating predictive probabilities.

*4: No diagnosis was given to patients who did not consult the JR Sendai Hospital Thyroid Outpatient Service.

false positives were then projected onto the SOM and the number of false positives who were located inside the zone of true positives, that is, the number of unavoidable false positives, was counted. We then evaluated the accuracy of the present extended model by comparing the number of our present SOM to that of the SOM in our previous study.

III. RESULTS

A SOM called a grey map that was constructed using a set of 8 parameters is shown in Figure 2A. In the following SOM, true positives (patients with Graves' disease were represented by red numbers, while patients with painless thyroiditis (/thyrotoxicosis without diagnosis) were represented by orange numbers) and healthy individuals (represented by green numbers) were distributed in two distinctive zones.

True positives were located on the left, and healthy individuals were located on the right. Patients with severe Graves' disease (high FT4 level) were more likely to be distributed in the deeper area of the left patient zone, whereas those with painless thyroiditis (/thyrotoxicosis who had a mild FT4 level) were more likely to be distributed inside the left patient zone, but closer to the healthy individual zone, such as JR1 (PT), JR4, and JR5. The characteristic distribution in this SOM was similar to our previous findings in which the SOM was constructed using a set of three or four routine tests [13] [14]. These distributional differences made it possible to analyze the characteristics of parameters specific to thyrotoxicosis and use them to differentiate patients.

We showed the component planes (Fig. 2B) accompanying

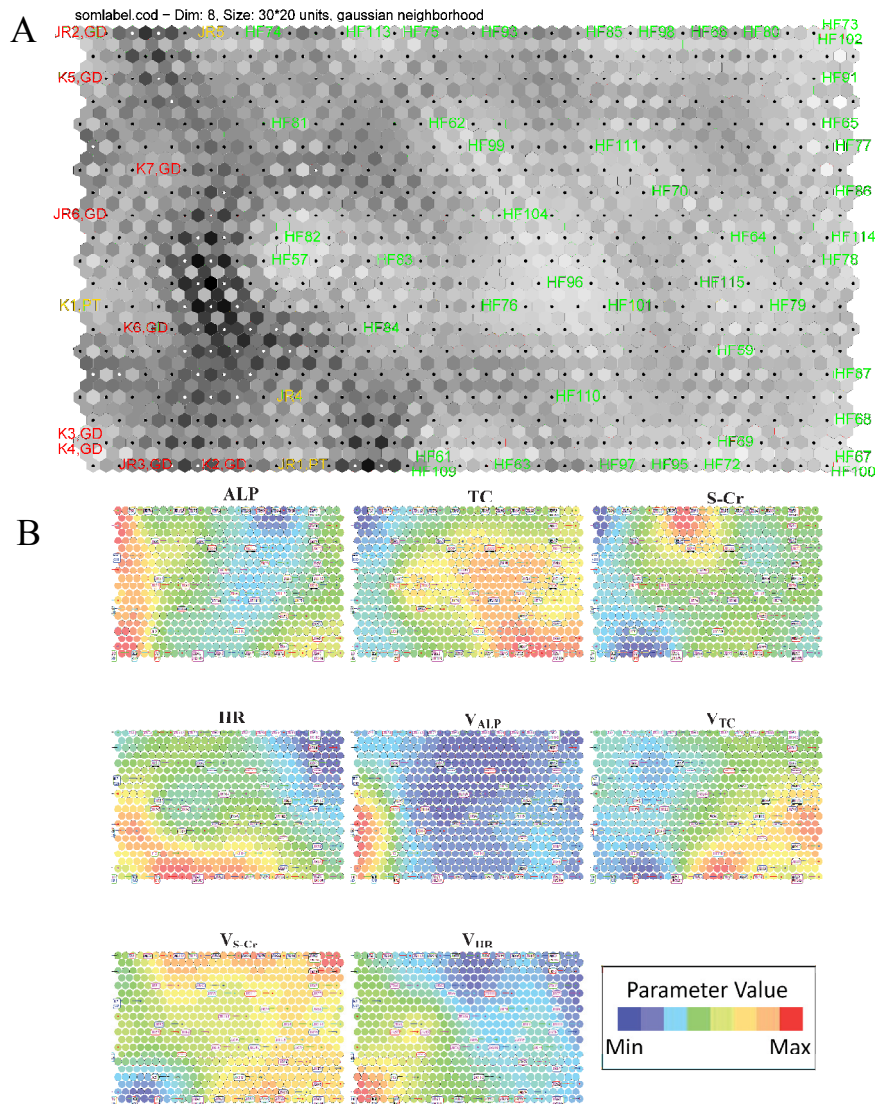


Fig. 2. Features of true positives of thyrotoxicosis were prominently characterized by variations in 8 parameters such as routine tests (ALP, TC, S-Cr, and HR) and their RATV. Red labels show 9 true positives with Graves' disease, orange labels show 4 true positives with painless thyroiditis (/thyrotoxicosis without diagnosis), and green labels show healthy controls.

- (A) Grey map of SOM in which true positives and controls were distributed according to the rule that samples that have a similar set of parameters lie nearer, while samples that have different test results lie further away. The calculation was performed under the following conditions: neuron number: 30×20 , neighboring radius $r = 20$, learning coefficient = 0.15 and learning frequency: 50,000.
- (B) Eight component planes of SOM for each parameter, in which locations of true positives and healthy subjects are all the same as in A, but those with high or low values of these parameters are represented with colors.

the grey map (Fig. 2A), which gave the same location of all samples in Figure 2A, but indicated the level of a given parameter for each sample from high (red) to low (blue). The component planes enabled an instantaneous judgment to identify valuable routine tests in order to discriminate patients with thyrotoxicosis from healthy individuals. The four component planes (ALP, TC, S-Cr, and HR) showed the characteristic ups and downs of the 4 routine tests (slightly elevated ALP and HR and slightly lower TC and S-Cr) in patients with thyrotoxicosis who were distributed over the left area, and another four component planes concerning RATV showed ups and downs in accordance with time-series

variations in routine tests. The level distribution of each RATV was almost similar to that of the routine test itself; however, some different distributions were observed in places within the zone. The distribution of v_{HR} was the highest (red) in the lower-left area in which K2 and K4 were located, while the distribution of HR was mild-high (green to yellow) in that area. In contrast, the distribution of HR in the lower-middle area in which healthy individuals HF61 and HF109 were located was the highest (red), whereas that of v_{HR} was mild-low (green) in that area.

Figure 3 shows the same SOM as that shown in Figure 2A; however, 140 false positives (normal of both FT4 and TSH,

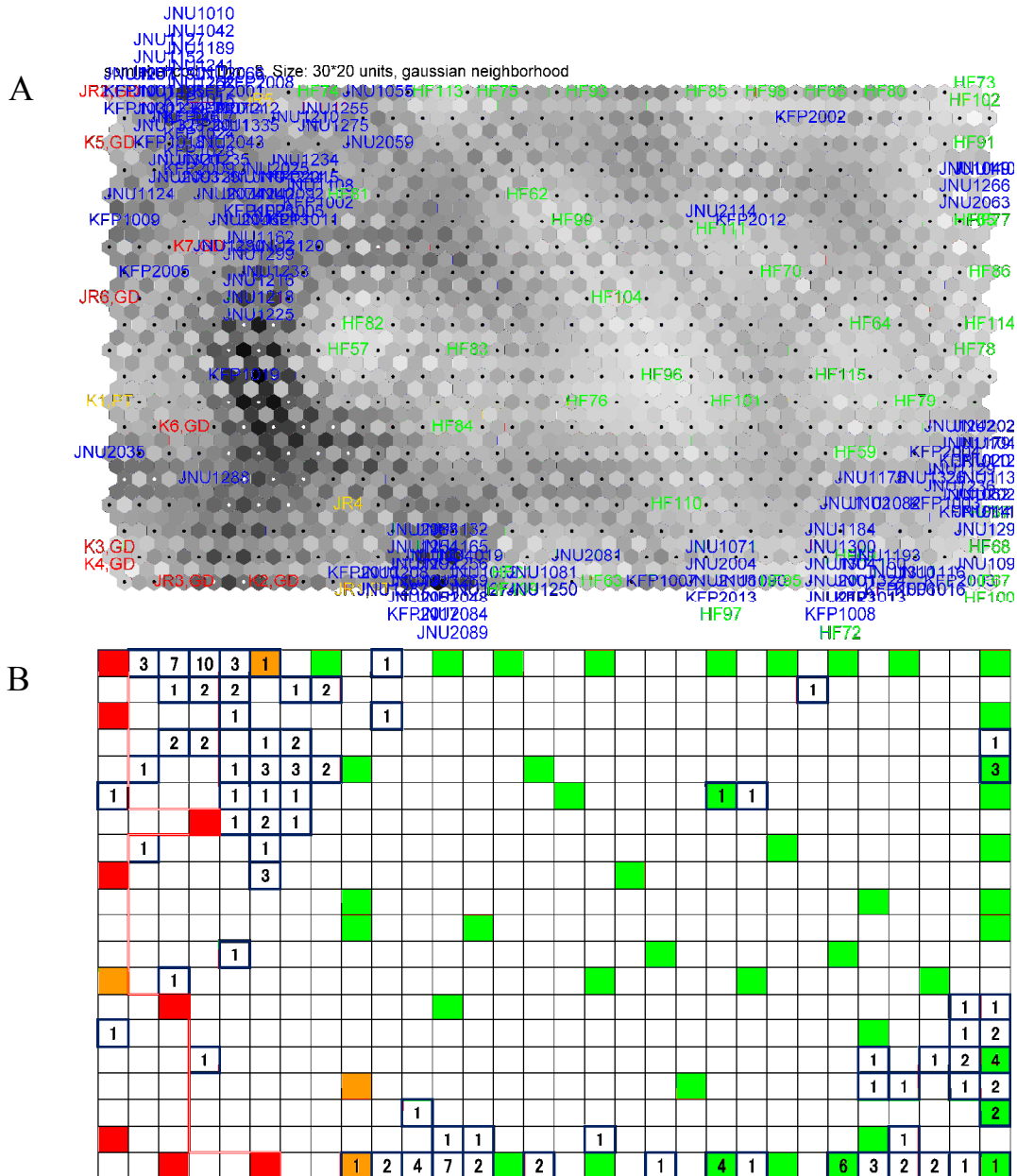


Fig. 3. The same SOM as that shown in Figure 2, but with 140 false positives (with normal FT4 and TSH, represented by blue labels) projected onto it. (A) Gray map of the original SOM by Kohonen. (B) A chart-type SOM on the sheet of Microsoft Excel using the same codebook vector obtained in the original SOM shown in 3A. Each of the red, orange, and green cells denotes that the cell has subjects with Graves' disease, painless thyroiditis (/ thyrotoxicosis without diagnosis), and healthy controls, respectively. Numbers in the blue-boldd cell denote the number of false positives projected onto that cell.

represented by blue labels) were projected onto it. In Figure 3A (grey map of the SOM by Kohonen), approximately half of the subjects were projected around the border region of two groups (thyrotoxicosis (left) and healthy individuals (right)), and only several subjects were located in the deeper left area of the left zone, while the others were dispersedly located in the right healthy region. In order to assist in evaluating these numbers quantitatively, we introduced a chart-type SOM on a sheet of Microsoft Excel, as shown in Figure 3B, using the same codebook vector obtained in the original SOM shown in Figure 3A. Each of the red, orange, and green cells in this chart denoted that the cell was allocated subjects with Graves' disease, painless thyroiditis / thyrotoxicosis without diagnosis, and healthy controls, respectively. Numbers in the blue-bolded cell denoted the number of false positives projected onto the cell. We then drew a threshold line along the cells having patients with Graves' disease who required treatment (indicated by the red line in Fig. 3B), and counted the number of false positives projected on the area surrounded by the line. As a result, the number of unavoidable false positives was significantly reduced from 140 to two while maintaining the true positives of Graves' diseases.

IV. DISCUSSIONS

In the present study, we successfully confirmed the usefulness of considering time-series variations more non-linearly with the SOM by using a set of 8 parameters that included four RATV in addition to the routine tests ALP, S-Cr, TC, and HR. Two groups, patients with thyrotoxicosis and healthy individuals, were separately distributed. Such separated distribution was found in the SOM constructed using a set of three (/four) routine tests, as we previously reported [11]-[14]. It should be noted that a complementary up and down distribution was found among the component planes of each routine test and its RATV. Regarding such a complementary distribution between HR and v_{HR} , most of the patients with Graves' disease had a higher HR, whereas those with a lower HR had a higher v_{HR} . This result suggested that using RATV together with the routine test itself helps to prevent false negatives. In contrast, most of the healthy individuals had normal-lower HR, whereas those with a higher HR had a lower v_{HR} . This result suggested that using RATV together with the routine test itself also helps to prevent false positives. These results may be interpreted as originating from inter-individual variations in heart rate [22]. For example, someone with an athlete's heart will have a lower personal normal range for HR than others; therefore, when they develop Graves' disease, elevations in HR are masked in its absolute value, whereas a relative difference between before and after its onset is expected. In our previous study, the addition of HR to a set of three routine tests yielded many advantages, but an unignorable disadvantage; the HR-added model sometimes avoided false positives and false

negatives of Graves' disease, but sometimes caused false positives and false negatives [17]. This limitation was successfully avoided in our present model by considering the RATV of HR in addition to HR itself. We expect these effects in the other routine laboratory tests. Furthermore, by comparing the number of unavoidable false positives in the present study with our previous findings [18], we further decreased this number, which suggests that fully considering the non-linear effect of RATV by adopting a pattern recognition method such as the SOM may be more effective.

Individual differences are known to exist in routine tests, thereby making statistical analyses insufficient. However, we considered the influence of individual differences to be markedly reduced by the introduction of RATV. It became possible to markedly reduce the number of false positives while maintaining that of true positives of Graves' diseases. It is considered easy to track these time-series variations in examinees who visit the Ningen Dock because approximately 70% of visitors are repeaters [23], [24] and missing values are rare; thus, our time-series analysis in the screening demonstrated in this study may be realistic and useful. However, the amount of training data obtained to date is small, and further studies are needed in order to confirm the robustness and versatility of this model by increasing samples hereafter. In addition, it is desirable to perform similar analyses using another PRM such as BRNN and SVM in future research.

REFERENCES

- [1] G.J. Canaris, N.R. Manowitz, G. Mayor, E.C. Ridgway: The Colorado thyroid disease prevalence study. *Arch Intern Med*, 160, pp. 526-534, 2000.
<http://dx.doi.org/0.1001/archinte.160.4.526>.
- [2] S.J. Landers: Overlooked, underdiagnosed? Thyroid disease poses a challenge. *American Medical News*, 47, pp. 25-26, 2004.
- [3] S. Reza, A. Shaukat, T.M. Arain, Q.S. Riaz, M. Mahmud: Expression of osteopontin in patients with thyroid dysfunction. *PLoS One*, 8, e56533, 2013.
<http://dx.doi.org/10.1371/journal.pone.0056533>
- [4] M.C. Mosher: Amiodarone-induced hypothyroidism and other adverse effects. *Dimens Crit Care Nurs*, 30, pp. 87-93, 2011.
<http://dx.doi.org/10.1097/DCC.0b013e3182052130>
- [5] M. Beniko, S. Sato, K. Iida, H. Ikawa, H. Kudo, A. Matsumoto, M. Sekiguchi, A. Imamura, Y. Mashio, K. Kawasaki: Evaluation of serum TSH levels in Ningen Dock, Ningen Dock, 24, pp. 885-890, 2009 (In Japanese).
<http://dx.doi.org/10.11320/ningendock.24.885>
- [6] K. Kasagi, N. Takahashi, G. Inoue, T. Honda, Y. Kawachi, Y. Izumi: Thyroid function in Japanese adults as assessed by a general health checkup system in relation with thyroid-related antibodies and other clinical parameters. *Thyroid*, 19, pp. 937-944, 2009.
<http://dx.doi.org/10.1089/thy.2009.0205>.
- [7] P.W. Ladenson, P.A. Singer, K.B. Ain, N. Bagchi, S.T. Bigos, E.G. Levy, S.A. Smith, G.H. Daniels: American Thyroid Association guidelines for detection of thyroid dysfunction. *Arch Intern Med*, 160, pp. 1573-1575, 2000.
<http://dx.doi.org/10.1001/archinte.160.11.1573>
- [8] T. Kohonen: *Self-organizing maps* 3rd edn. Springer-Verlag, Berlin, 2000.
- [9] D.J.C. Mackay: A practical Bayesian framework for backpropagation networks. *Neural Computation*, 4, pp. 448-472, 1992.
<http://dx.doi.org/10.1162/neco.1992.4.3.448>

- [10] V.N. Vapnik: *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, 1995.
- [11] K. Hoshi, J. Kawakami, W. Sato, K. Sato, A. Sugawara, Y. Saito, K. Yoshida: Assisting the Diagnosis of Thyroid Diseases with Bayesian-Type and SOM-Type Neural Networks Making Use of Routine Test Data. *Chem Pharm Bull*, 54, pp. 1162-1169, 2006. <http://doi.org/10.1248/cpb.54.1162>
- [12] W. Sato, K. Hoshi, J. Kawakami, K. Sato, A. Sugawara, Y. Saito, K. Yoshida: Assisting the diagnosis of Graves' hyperthyroidism with Bayesian-type and SOMtype neural networks by making use of a set of three routine tests and their correlation with free T4. *Biomed Pharmacother*, 64, pp. 7-15, 2010. <http://dx.doi.org/10.1016/j.biopha.2009.02.007>
- [13] S. Aoki, K. Hoshi, J. Kawakami, K. Sato, K. Satoh, K. Mori, A. Sugawara, Y. Saito, K. Yoshida: Assisting the diagnosis of Graves' hyperthyroidism with pattern recognition methods and a set of three routine tests parameters, and their correlations with free T4 levels: Extension to male patients. *Biomed Pharmacother*, 65, pp. 95-104, 2011. <http://dx.doi.org/10.1016/j.biopha.2010.10.005>
- [14] S. Aoki, K. Sato, K. Hoshi, J. Kawakami, S. Suzuki, K. Mori, A. Sugawara, Y. Saito, K. Yoshida: improvement of New Low-cost Method for Detecting Abnormal Thyroid Function Making Use of Set of Routine Tests -Addition of Heart Rate, Influence of Dosing and Time Series Drifting are Very Effective-. *Ningen Dock*, 27, pp. 87-96, 2012. <http://doi.org/10.11320/ningendock.27.87>
- [15] S. Aoki, K. Hoshi, J. Kawakami, K. Sato, W. Sato, K. Satoh, K. Mori, A. Sugawara, Y. Nakagawa, K. Yoshida: Assisting the diagnosis of overt hypothyroidism with pattern recognition methods making use of a set of routine tests, and their multiple correlation with total T4. *Biomed Pharmacother*, 66, pp. 195-205, 2012. <http://doi.org/10.1016/j.biopha.2011.11.018>
- [16] S. Aoki, K. Satoh, K. Hoshi, J. Kawakami, K. Sato, Y. Saito, K. Mori, K. Yoshida: New Method for Aiding Detection of Abnormal Thyroid Function in Patients Making Use of Set of Routine Tests- Applying It in Screening During General Health Check-ups. *Ningen Dock*, 26, pp. 9-16, 2011. <http://doi.org/10.11320/ningendock.26.9>
- [17] S. Aoki, K. Sato, K. Hoshi, J. Kawakami, K. Mori, Y. Nakagawa, W. Hida, K. Yoshida: New Low-cost Method for Detecting Abnormal Thyroid Function in Patients Making Use of a Set of Routine Tests- Testing Many More Ningen Dock Examinees and Studying Appropriate Threshold Levels. *Ningen Dock International*, 2, pp. 19-26, 2014.
- [18] S. Aoki, S. Nishizaka, K. Sato, K. Hoshi, J. Kawakami, K. Mori, Y. Nakagawa, W. Hida, K. Yoshida: New clinical decision support system using a set of routine tests to detect thyrotoxicosis in patients: Adding the average rate of annual time-series variations improves diagnostic accuracy. to be submitted.
- [19] T. Kohonen, J. Hynninen, J. Kangas, J. Laaksonen: SOM_PAK: The Self-Organizing Map Program Package. <http://www.cis.hut.fi/research/som-research/nrc-programs.shtml>
- [20] R.M. Neal: Software for Flexible Bayesian Modeling and Markov Chain Sampling. <http://www.cs.toronto.edu/~radford/>
- [21] C.C. Chang, C.J. Lin: LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*, 2, pp. 1-27, 2011. <http://dx.doi.org/10.1145/1961189.1961199>
- [22] H. Kobayashi: Inter and intra individual variations in Heart Rate Variability of Japanese males. *J Physiol Anthropol*, 26, pp. 173-177, 2007. <http://doi.org/10.2114/jpa2.26.173>
- [23] Committee for Statistical Survey of Ningen Dock Examination: Report on National Aggregate Results of Ningen Dock in 2012. *Ningen Dock*, 28, pp. 678-690, 2013(In Japanese). <http://doi.org/10.11320/ningendock.28.678>
- [24] Committee for Statistical Survey of Ningen Dock Examination: Report on National Aggregate Results of Ningen Dock in 2013. *Ningen Dock*, 29, pp. 623-635, 2014(In Japanese). <http://doi.org/10.11320/ningendock.29.623>