

Diagnostic Criteria, Clinical Features, and Incidence of Thyroid Storm Based on Nationwide Surveys

Takashi Akamizu,¹ Tetsurou Satoh,² Osamu Isozaki,³ Atsushi Suzuki,⁴ Shu Wakino,⁵ Tadao Ihuri,⁶
Kumiko Tsuboi,⁷ Tsuyoshi Monden,⁸ Tsuyoshi Kouki,⁹ Hajime Otani,¹⁰ Satoshi Teramukai,¹¹
Ritei Uehara,¹² Yosikazu Nakamura,¹² Masaki Nagai,¹³ and Masatomo Mori,²
for the Japan Thyroid Association

Background: Thyroid storm (TS) is life threatening. Its incidence is poorly defined, few series are available, and population-based diagnostic criteria have not been established. We surveyed TS in Japan, defined its characteristics, and formulated diagnostic criteria, FINAL-CRITERIA1 and FINAL-CRITERIA2, for two grades of TS, TS1, and TS2 respectively.

Methods: We first developed diagnostic criteria based on 99 patients in the literature and 7 of our patients (LIT-CRITERIA1 for TS1 and LIT-CRITERIA2 for TS2). Thyrotoxicosis was a prerequisite for TS1 and TS2 as well as for combinations of the central nervous system manifestations, fever, tachycardia, congestive heart failure (CHF), and gastrointestinal (GI)/hepatic disturbances. We then conducted initial and follow-up surveys from 2004 through 2008, targeting all hospitals in Japan, with an eight-layered random extraction selection process to obtain and verify information on patients who met LIT-CRITERIA1 and LIT-CRITERIA2.

Results: We identified 282 patients with TS1 and 74 patients with TS2. Based on these data and information from the Ministry of Health, Labor, and Welfare of Japan, we estimated the incidence of TS in hospitalized patients in Japan to be 0.20 per 100,000 per year. Serum-free thyroxine and free triiodothyroine concentrations were similar among patients with TS in the literature, Japanese patients with TS1 or TS2, and a group of patients with thyrotoxicosis without TS (Tox-NoTS). The mortality rate was 11.0% in TS1, 9.5% in TS2, and 0% in Tox-NoTS patients. Multiple organ failure was the most common cause of death in TS1 and TS2, followed by CHF, respiratory failure, arrhythmia, disseminated intravascular coagulation, GI perforation, hypoxic brain syndrome, and sepsis. Glasgow Coma Scale results and blood urea nitrogen (BUN) were associated with irreversible damages in 22 survivors. The only change in our final diagnostic criteria for TS as compared with our initial criteria related to serum bilirubin concentration >3 mg/dL.

Conclusions: TS is still a life-threatening disorder with more than 10% mortality in Japan. We present newly formulated diagnostic criteria for TS and clarify its clinical features, prognosis, and incidence based on nationwide surveys in Japan. This information will help diagnose TS and in understanding the factors contributing to mortality and irreversible complications.

¹The First Department of Medicine, Wakayama Medical University, Wakayama, Japan.

²Department of Medicine and Molecular Science, Gunma University Graduate School of Medicine, Maebashi, Japan.

³Department of Medicine 2, Tokyo Women's Medical University, Tokyo, Japan.

⁴Division of Endocrinology and Metabolism, Department of Internal Medicine, Fujita Health University, Toyoake, Japan.

⁵Department of Internal Medicine, Keio University, Tokyo, Japan.

⁶Department of Clinical and Molecular Endocrinology, Graduate School, Tokyo Medical and Dental University, Tokyo, Japan.

⁷Division of Diabetes, Metabolism, and Endocrinology, Department of Medicine, School of Medicine, Toho University, Tokyo, Japan.

⁸Department of Endocrinology and Metabolism, Dokkyo Medical University, Mibu, Japan.

⁹Department of Endocrinology and Metabolism, University of the Ryukyus, Nishihara, Japan.

¹⁰Second Department of Internal Medicine, Kansai Medical University, Moriguchi, Japan.

¹¹Translational Research Center, Faculty of Medicine, Kyoto University, Kyoto, Japan.

¹²Department of Public Health, Jichi Medical University, Shimotsuke, Japan.

¹³Department of Public Health, Saitama Medical University, Moroyama, Japan.

Introduction

THYROID STORM (TS) IS A LIFE-THREATENING CONDITION requiring emergency treatment (1–3). The condition arises in thyrotoxic patients and is manifested by the decompensation of multiple organs, which is often triggered by severe stress. Since its pathophysiologic mechanisms have not been clarified, the diagnosis of TS is based on clinical manifestations. Some of the characteristic clinical manifestations include unconsciousness, high fever, heart failure, diarrhea, and jaundice. The older literature suggests that TS occurs in no more than 1%–2% of hospitalized thyrotoxic patients (3,4), but its incidence in a large population has not been fully investigated.

With regard to diagnostic criteria for TS, few have been published other than those by Burch and Wartofsky (3,5). Their criteria are useful, but the approach taken, by utilizing the summation of multiple clinical manifestation scores, may often reach the threshold for the diagnosis of TS in thyrotoxic patients with severe nonthyroid illness, but not necessarily with TS. Furthermore, the scores that are allocated to signs and symptoms in this diagnostic scheme are complex and have not been validated. In the context of this information, the Japan Thyroid Association organized a committee that developed diagnostic criteria for TS and surveyed its incidence in Japan, linking the research activities of both the Japan Endocrine Society and the Ministry of Health, Labor, and Welfare of Japan. Here, we the members of this committee, report our findings regarding the clinical features of TS and propose diagnostic criteria for TS. Our criteria were based on an analysis of the literature, followed by a survey of the thyrotoxic patients in Japan who had the features of TS as reported in the literature. We finalized our criteria based on the clinical features and course of these patients. In addition, we sought to provide data regarding the incidence of TS in Japan.

Methods

The present study was approved by the ethics committee of the Jichi Medical University.

Development of the literature diagnostic criteria for TS1 and TS2

Initially, we developed tentative diagnostic criteria for TS (LIT-CRITERIA1 for TS1 and LIT-CRITERIA2 for TS2) based on the patients reported in the literature (Appendix A). Since TS is rare, a prospective study that develops criteria was considered very difficult and time consuming. Therefore, the criteria were mainly based on information obtained from the literature. In April 2006, the PubMed (for English literature) and Ichushi (for Japanese literature) databases were searched using the terms “TS” or “thyroid crisis” or “thyrotoxic crisis.” Valid publication dates were from 1992 to 2006 for PubMed (www.ncbi.nlm.nih.gov/pubmed) and from 1983 to 2006 for the Ichushi database (<http://login.jamas.or.jp>), respectively. Ninety-three original case reports (see Supplementary Data, available online at www.liebertonline.com/thy) and their abstracts were carefully evaluated by three independent researchers. Seven unpublished cases of TS diagnosed in the researchers’ facilities were also included in the analysis. The individual clinical parameters extracted from these cases were analyzed. By analyzing these data, we created LIT-CRIT-

ERIA1 for TS1 and LIT-CRITERIA2 for TS2. TS1 and TS2 were envisioned as indicative of two grades of severity for TS. In our questionnaires for surveys and/or study documents, TS1 was referred to as “Definite” TS, and TS2 was referred to as “Suspected” TS. LIT-CRITERIA1 and LIT-CRITERIA2 were referred to as the First Edition of the Diagnostic Criteria for TS.

In developing LIT-CRITERIA1 and LIT-CRITERIA2, we made a consensus decision that thyrotoxicosis would be considered an absolute criteria for the diagnosis of TS. Next, we examined the prevalence of various clinical features in these patients and compared these to the corresponding prevalence in our patients with thyrotoxicosis but without TS (Tox-NoTS). We also analyzed the patterns of combinations of clinical manifestations. Patterns were stratified by the presence or absence of central nervous system (CNS) manifestations, because these were most frequent and appeared to be very specific to TS.

Thyrotoxic patients without TS

For a comparison with patients who met LIT-CRITERIA1 and LIT-CRITERIA-2 for TS1 and TS2, respectively, we also collected data from Tox-NoTS patients ($n=133$). We recruited them from either our outpatient clinics or inpatient wards in a serial manner over a period of several months. We did not make an effort to age and gender match Tox-NoTS patients with TS1 and TS2 patients.

The First Nationwide Survey for cases of TS1 and TS2 in Japan

SURVEY-1 for cases of TS1 and TS2 in Japan was conducted in 2009. In our records, SURVEY-1 was referred to as the First Nationwide Survey. An English translation from the Japanese of the SURVEY-1 is presented in Appendix B. The first part of the survey shows the questionnaire that the respondents were asked to fill in (Appendix B). The second part gives LIT-CRITERIA1 and LIT-CRITERIA2 for the diagnosis of TS1 and TS2, respectively (Appendix A). TS1 is referred to in the survey questionnaire as “definite,” and TS2 is referred to as “suspected.” Respondents were asked to report the number of their cases of “definite” and “suspected” TS for the years 2004, 2005, 2006, 2007, and 2008. SURVEY-1 also collected information regarding gender and the year patients were seen but did not ask for clinical details.

In accordance with the Nationwide Epidemiologic Survey Manual (6–8), we selected hospital departments of internal medicine, endocrinology, thyroidology, cardiology, and emergency medicine as the targets for SURVEY-1. Study hospitals were randomly selected from a list of all the hospitals in Japan. The selection rate was based on a stratification of Japanese hospitals by the number of beds; 100% for university hospitals and those with ≥ 500 beds, whereas only 5% of hospitals with fewer than 100 beds were selected at random (eight-layered random extraction) (6) (see Supplementary Table S1). The average extraction rate was $\sim 20\%$ of hospitals or medical care institutions in all the hospitals in Japan.

The Second Nationwide Survey: the clinical features of TS1 and TS2 in Japan

After the completed questionnaires for SURVEY-1 had been received, the respondents were sent a second questionnaire, SURVEY-2, shown in Appendix C. In our documents,

this was referred to as the Second Nationwide Survey. In SURVEY-2, the respondents were asked to provide detailed clinical and laboratory information regarding the cases of TS1 and TS2 that they had reported in SURVEY-1.

Analysis of the responses to SURVEY-2 for TS1 and TS2 in Japan

After the responses to SURVEY-2 had been received, we organized and analyzed the information provided, initially entering this into a database. To fill gaps in the data provided for individual patients, we directly contacted responders to obtain this information. We then validated the assigned diagnoses that were TS1 and TS2 based, respectively, on LIT-CRITERIA1 and LIT-CRITERIA2. Using these validated data, we analyzed the content of the database, including information regarding clinical features and the number of patients from reporting centers.

Development of the final diagnostic criteria for the diagnosis of TS1 and TS2

Using the information regarding patients from our surveys who met LIT-CRITERIA1 and LIT-CRITERIA2 for TS1 and TS2, respectively, we developed our final criteria, FINAL-CRITERIA1 and FINAL-CRITERIA2, for the diagnosis of TS1 and TS2. In our study records, we referred to them as the Second Edition of the Diagnostic Criteria for TS. As in the creation of LIT-CRITERIA1 and LIT-CRITERIA2, we aimed at distinguishing TS from other disorders, including Tox-NoTS and severe nonthyroid illness.

Statistical analyses

Differences in clinical manifestations among patients with TS1, TS2, and Tox-NoTS patients were analyzed by the analysis of variance or the chi-squared test, as appropriate. We compared this information with the criteria by Burch and Wartofsky (3,5) for TS (BWC-TS). The comparison between the BWC-TS and our diagnostic criteria for TS were assessed with logistic regression, Spearman rank correlation, and the chi-squared test (Fisher's exact test), respectively. To identify the factors independently associated with clinical outcomes, logistic regression analysis or multiple regression analysis with the stepwise method was used as appropriate after the possible relevant factors had been selected by simple regression analysis. The clinical outcomes that were evaluated were death, irreversible complications, and severity of thyroid crisis. Two-sided $p < 0.05$ was regarded as being statistically significant. All statistical analyses were done using JMP version 8 (SAS Institute, Cary, NC).

Results

Development of LIT-CRITERIA1 and LIT-CRITERIA2 for TS1 and TS2, respectively

Twenty-two cases of TS were in PubMed, and 77 cases were in the Ichushi database (71 citations) (see Supplementary Data). Information regarding these 99 cases and 7 of our unpublished cases are summarized in Table 1 (see first column). Table 1 also contains information on the clinical characteristics of Tox-NoTS patients. The most prominent clinical characteristics observed in the patients in the literature with TS, and

TABLE 1. CHARACTERISTICS OF PATIENTS REPORTED IN THE LITERATURE WITH THYROID STORM, IN PATIENTS WITH THYROTOXICOSIS WITHOUT TS, AND IN JAPANESE PATIENTS WITH THYROID STORM

	TS reported in the literature	Tox-NoTS patients	TS reported, with validation, in the SURVEY-2 of Japanese patients	
			TS1	TS2
Number	106	133	282	74
Age (years old)	42.1 ± 14.9 (7–73)	43.2 ± 15.5 (14–80)	44.7 ± 16.7 (6–87)	44.6 ± 14.6 (20–80)
Male:female	30:76	34:99	74:204	15:59
Basic thyroid diseases				
Graves' disease	95.2%	97.7%	98.9%	97.30%
Others	4.8%	2.3%	1.1%	2.70%
Free T4 (ng/dL)	6.76 ± 3.17	6.35 ± 5.13	6.38 ± 3.40	6.18 ± 2.56
Free T3 (pg/mL)	15.9 ± 7.9	16.5 ± 8.2	19.70 ± 12.70	17.81 ± 8.78
Fever ≥ 38°C	55.7%	3.0%	41.5%	41.9%
Pulse rate				
≥ 120/min	82.1%	24.0%	84.0%	75.7%
≥ 130/min	67.9%	7.5%	76.2%	60.8%
CNS symptoms ^a	64.2%	Not frequent	84.4%	2.7%
GI/hepatic symptoms ^b	51.9%	Not frequent	69.5%	63.5%
CHF ^c	38.7%	Not frequent	39.4%	37.8%
NYHA classification IV	20.1%	Rare	24.1%	9.5%
Killip classification ≥ III	20.1%	Rare	22.7%	17.6%
Precipitating factors	76.4%	Not applicable	71.3%	64.9%
Mortality rate	17.0%	Very low	11.0%	9.5%

Systeme International (SI) units for free T4 to picomoles per liter (conversion factor, 12.87); for free T3 to picomoles per liter (0.0154).

^aCNS symptoms with agitation, restlessness, delirium, mental aberration/psychosis, somnolence/lethargy, convulsion or coma.

^bGI/hepatic symptoms with abdominal pain, diarrhea, nausea/vomiting, or jaundice with liver dysfunction.

^cCHF with pedal edema, bibasilar rales, or pulmonary edema.

TS, thyroid storm; Tox-NoTS, thyrotoxicosis without TS; CNS, central nervous system; GI, gastrointestinal; CHF, congestive heart failure; NYHA, New York Heart Association; T4, thyroxine; T3, triiodothyroine.

our seven patients with TS, were fever, tachycardia, CNS and gastrointestinal (GI)/hepatic signs and symptoms, and congestive heart failure (CHF). In Tox-NoTS patients, the most prominent clinical characteristics were tachycardia, goiter, weight loss, and finger tremor. All patients with TS that had CNS manifestations exhibited at least one or other of the clinical characteristics of TS which were prominent in the literature. These were fever, tachycardia, GI/hepatic, and cardiac manifestations. Since none of these manifestations showed any evident correlation with one another, it was considered that each manifestation could be dealt with independently.

Based on the findings just mentioned, we attempted to establish cut-off values for each manifestation to obtain better sensitivity, specificity, and predictive values. Since TS occurs only rarely, it was supposed that high specificity, >90% for each parameter, would be needed to obtain a good positive predictive value. In this context, we tentatively set cut-off values of $\geq 38^{\circ}\text{C}$ and ≥ 130 beats/min (bpm) for fever and tachycardia, respectively. With regard to CNS manifestations, the mild symptom, agitation, was excluded. In addition to signs and symptoms, our criteria included specific quantitative scores on the Glasgow Coma Scale (GCS) (9,10) and the Japan Coma Scale (JCS) (11). Since GI/hepatic symptoms appeared to be less specific than CNS manifestations, the mild symptom, abdominal pain, was excluded. Although the prevalence of CHF in patients with storm was <40%, CHF when present tended to be very severe. Fifty-four percent of cases of TS1 could be classified as New York Heart Association (NYHA) class IV and/or Killip class III/IV. Therefore, the criteria for CHF were limited to relatively severe manifestations: pulmonary edema, moist rales in more than half the lung fields, or cardiogenic shock. These findings were applied to create the First Edition of the Diagnostic Criteria (LIT-CRITERIA-1 and LIT-CRITERIA-2) for TS1 and TS2. These were arrived at by a consensus and are presented in Appendix A, after the questionnaire.

Cases of TS in Japan as obtained from SURVEY-1

One of the purposes of SURVEY-1 for cases of TS1 and TS2 among patients admitted for the treatment of thyrotoxicosis was to obtain a 5-year estimate of their incidence in Japan during the years 2004 through 2008. In a group of 1463 randomly selected hospitals, there was a 52.5% response rate to SURVEY-1 with 541 patients being reported, or 1030 patients if the proportion of TS in nonrespondents had been the same. Based on this correction, the total number of patients in these 1463 hospitals in Japan during the 5-year period was estimated to be 1283 [95% confidence interval: 1077–1489]. The Japanese general population in 2008 was reported to be ~127 million (12), and the number of total and admitted thyrotoxic patients were 118,000 and 4800, respectively (13). If we simply calculate using the estimated number of TS patients and then apply this information to the Ministry data regarding the incidence of thyrotoxicosis in Japan, then the incidence rate of TS is estimated to be ~0.20 persons per 100,000 Japanese population per year, and the condition occurs in 0.22% of all thyrotoxic patients and in 5.4% of admitted thyrotoxic patients.

Of the patients who met the criteria for TS1 and TS2, 77% were women. Patients with TS1 and TS2 were seen in the

departments of internal medicine, emergency medicine, and cardiology, in 67%, 21%, and 12% of cases, respectively.

SURVEY-2 of the clinical features of TS1 and TS2 in Japan

Among the 541 patients with TS1 and TS2 identified in SURVEY-1, detailed clinical information, for what was said to be 437 patients, was obtained in the SURVEY-2 responses. However, 7 patients were reported twice. Furthermore, 24 patients were seen outside the study period, and 50 patients did not meet either LIT-CRITERIA1 or LIT-CRITERIA2 for the diagnosis of TS1 or TS2. Therefore, data from 356 patients were available for further analyses from SURVEY-2. Amongst these patients, 282 had TS1, and 74 had TS2 based on LIT-CRITERIA1 and LIT-CRITERIA2, respectively.

Table 1 presents the characteristics of patients with TS as obtained in SURVEY-2. The ages of patients with TS1 and TS2 were 44.7 ± 16.7 (median: 42, range: 6–87) and 44.6 ± 14.6 (median: 44, range: 20–80), respectively. The ages of men and women with TS1 and TS2 were similar (data not shown). The male-to-female ratio for TS1 and TS2 was ~1:3 (89:263). Graves' disease (GD) was the most common cause of thyrotoxicosis among the patients with TS1 and TS2. This was followed by very rare cases of destructive thyroiditis (five cases). In 45% of the cases, the duration between the onset of TS1 or TS2 and the initial diagnosis of GD was less than one year, and about 20% of patients developed TS1 or TS2 before they received antithyroid treatment (Fig. 1). The factors classically considered to trigger TS were present in 70% of the patients with TS1 and TS2 (Table 2). The highest trigger of TS was the irregular use or discontinuation of the patient's antithyroid drug, and the second highest trigger was infection, particularly that of the upper respiratory tract (data not shown). The combined mortality rate of patients with TS1 and TS2 was 10.7% (38/356). The mortality rates of TS1 and TS2 were 11.0% and 9.5%, respectively. The age of patients who died from TS1 and TS2 was 49.56 ± 16.96 years.

The results of thyroid function tests in patients with TS1 and TS2 are presented in Table 1. The mean serum free thyroxine (FT4) concentration in patients with TS1 was 6.38 ± 3.40 ng/dL, and in patients with TS2 was 6.18 ± 2.56 ng/dL. The mean serum free triiodothyronine (FT3) concentration in patients with TS1 was 19.70 ± 12.70 pg/mL, and in patients with TS2 was 17.81 ± 8.78 pg/mL. In our Tox-NoTS patients the serum FT4 and FT3 concentrations were 6.35 ± 5.13 ng/dL and 16.5 ± 8.2 pg/mL, respectively. Serum FT4 or FT3 concentrations were not significantly different among patients with TS1, patients with TS2, and Tox-NoTS patients. Figure 2 presents information regarding the simultaneous determination of thyrotropin (TSH), and the FT4/FT3 ratio. Several cases of TS1 had normal serum FT3 levels with increased serum FT4 levels and suppressed serum TSH levels. Approximately two-thirds of patients with TS1 and TS2 had body temperatures of $\geq 38.0^{\circ}\text{C}$ (Table 1). The range of body temperatures was 32.2°C – 41.2°C (median: 38.0°C) in patients with TS1 and 35.4°C – 41.8°C (median: 38.1°C) in patients with TS2 (Fig. 3).

In 76.2% of the patients with TS1 and 60.8% of the patients with TS2, the heart rate was ≥ 130 bpm (Table 1). The distribution of the heart rate is presented in Figure 4. In patients with TS1 the median heart or pulse rate was 144 bpm, and the

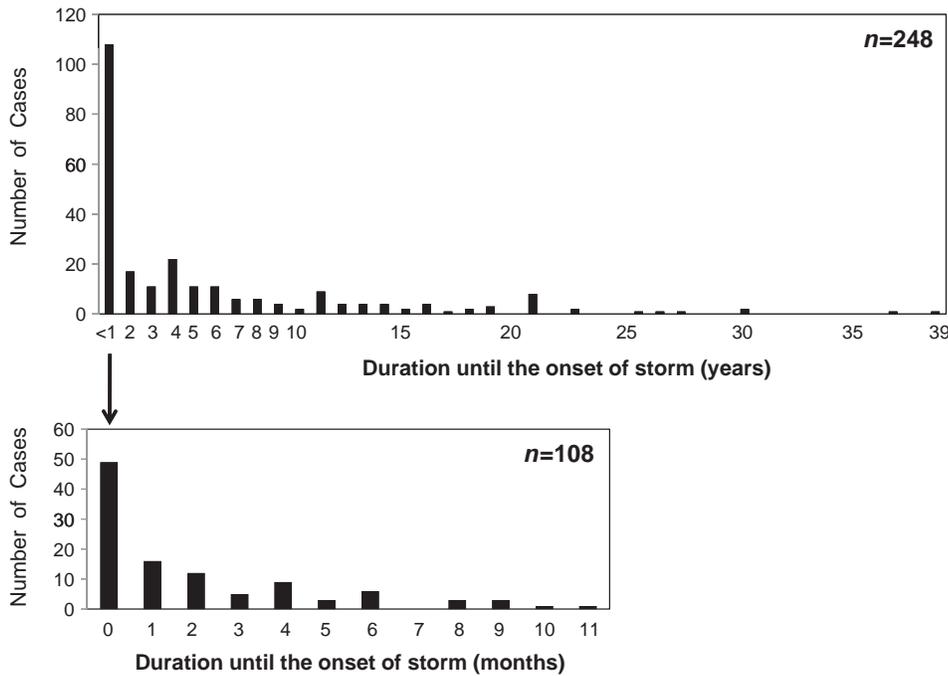


FIG. 1. Duration until the onset of thyroid storm in SURVEY-2. The duration between the initial diagnosis of Graves' disease and the onset of thyroid storm is summarized in the upper panel. The number of patients who developed thyroid storm within one year after the initial diagnosis of underlying thyroid disease is shown in the lower panel. SURVEY-2, The Second Nationwide Survey.

TABLE 2. TRIGGERS OF THYROID STORM IN JAPANESE PATIENTS

Triggers	Cases	Triggers	Cases
Irregular use or discontinuation of antithyroid drugs	122	Pregnancy/delivery	5
Infection	87	Cerebrovascular disease	3
Diabetic ketoacidosis	12	Intense exercise	2
Severe emotional stress	12	Ischemic heart disease	1
Trauma	12	Adrenocortical insufficiency	1
Nonthyroid surgery	8	Administration of iodinated contrast medium	1
Radioiodine therapy	6	Extraction of teeth	1
		Others	31

Data are from the SURVEY-2 of TS with validation. SURVEY-2, Second Nationwide Survey.

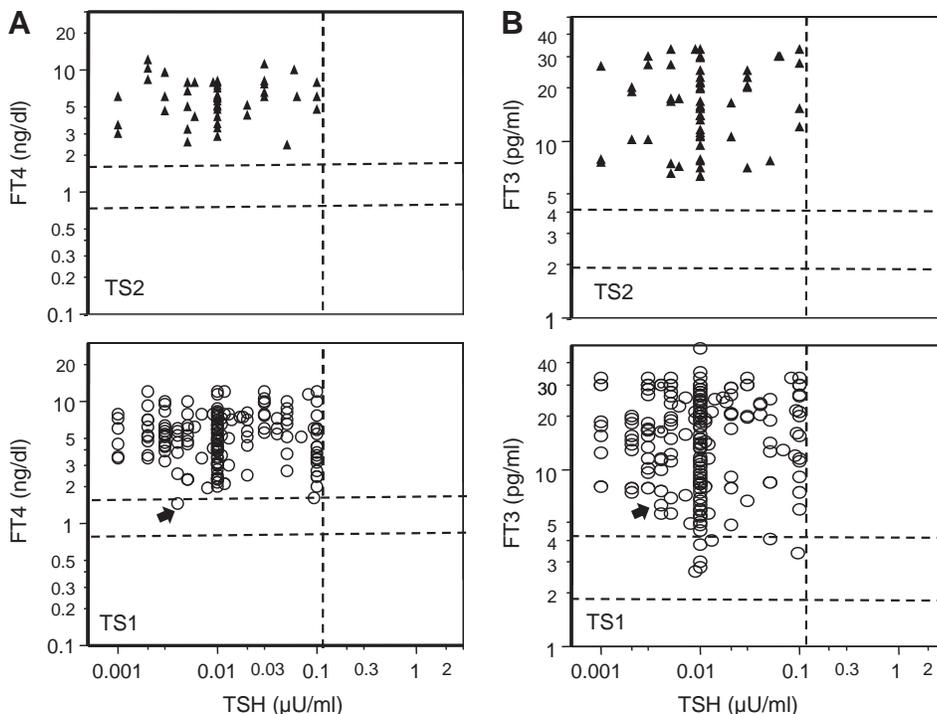


FIG. 2. Relationship between (A) FT4 and (B) FT3 levels and TSH in patients with TS1 and TS2 as obtained in SURVEY-2. Dashed lines show the limits of reference ranges. One TS1 case whose FT4 level was normal had a suppressed TSH and an elevated FT3 level, indicating a thyrotoxic state (an arrow). Six TS1 patients with normal FT3 levels had increased FT4 levels and suppressed TSH levels, being diagnosed as thyrotoxic. FT4, free thyroxine; FT3, free triiodothyronine; TSH, thyrotropin; TS, thyroid storm.

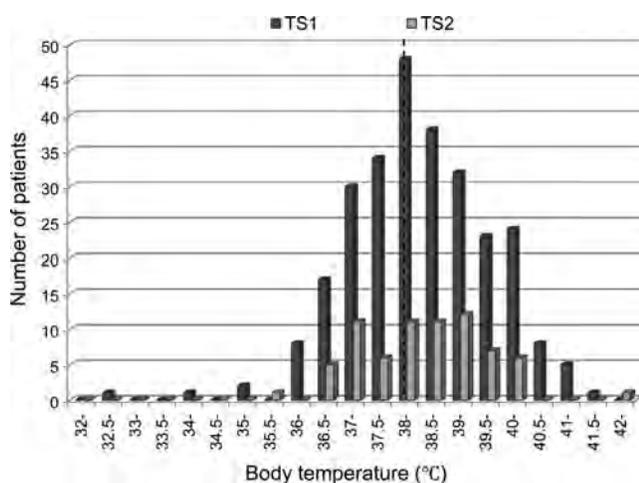


FIG. 3. Distribution of body temperature in patients with TS1 ($n=272$) and TS2 ($n=71$). The dashed line indicates the cut-off value for our diagnostic criteria ($\geq 38^{\circ}\text{C}$).

upper and lower quantiles were 171.5 and 130 bpm, respectively. In patients with TS2, the median heart rate was 140 bpm, while the upper and lower quantiles were 160 and 120 bpm, respectively.

The prevalence rates of CNS manifestations in patients with TS1 and TS2 are presented in Tables 1 and 3. Although agitation was found in 25.6% of patients with TS1 and TS2, logistic regression analysis showed that the presence of agitation as a diagnostic item did not alter these diagnoses, because the GCS, JCS, and other CNS manifestations encompassed this symptom. More than half of the patients with TS1 (Fig. 5) had abnormal values for the GCS (53.5%) and/or the JCS (62.6%).

GI/hepatic manifestations were present in 68.3% of patients with TS1 or TS2 (Table 1). We inquired if GI/hepatic manifestations were necessary for the diagnoses of TS1 or TS2. Patients with TS1 who had CNS manifestations would have

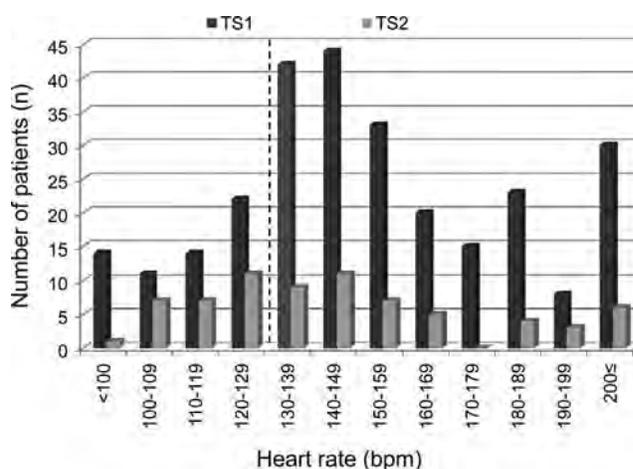


FIG. 4. Distribution of heart rate in patients with TS1 ($n=276$) and TS2 ($n=71$). The dashed line indicates the cut-off values in our criteria (≥ 130 bpm).

remained in the TS1 category even if they had not had GI/hepatic manifestations. However, 38 of the 55 patients with TS1 who did not have CNS manifestations would have been assigned to the TS2 category if they had not had GI/hepatic manifestations. The mortality in the group with TS1 who did not have CNS manifestations was 18.4% (7/38). Forty of the 59 patients with TS2 who did not have CNS manifestations would have been considered as not having any form of TS if they had not had GI/hepatic manifestations and, of these 40 patients, the GI/hepatic manifestations were nausea, vomiting, and/or diarrhea. In these 40 patients, there were 4 fatalities. However, in these TS2 patients, the presence of nausea, vomiting and/or diarrhea did not influence the prognosis as far as mortality was concerned ($p=0.807$, χ^2 test).

Thirty-five of the 55 patients with TS1 who did not have CNS manifestations would have been assigned to the TS2

TABLE 3. CENTRAL NERVOUS SYSTEM AND GASTROINTESTINAL/HEPATIC MANIFESTATIONS OF THYROID STORM IN JAPANESE PATIENTS

	TS1 (n=282)		TS2 (n=74)		Total cases (n=356)	
	Number	%	Number	%	Number	%
<i>CNS features</i>						
Restlessness	109	38.7	1	1.4	110	30.9
Delirium	35	12.4	0	0.0	35	9.8
Mental aberration/psychosis	20	7.1	0	0.0	20	5.6
Somnolence/lethargy	49	17.4	0	0.0	49	13.8
Convulsion	13	4.6	1	1.4	14	3.9
Coma	27	9.6	1	1.4	28	7.9
Total	238	84.4	2	2.7	240	67.4
<i>GI/hepatic features</i>						
Nausea	77	27.3	15	20.3	92	25.8
Vomiting	58	20.6	7	9.5	65	18.3
Diarrhea	99	35.1	28	37.8	127	35.7
Jaundice	60	21.3	13	17.6	73	20.5
GI/hepatic symptoms	166	58.9	38	51.4	204	57.3
Total	196	69.5	47	63.5	243	68.3

Data are from the SURVEY-2 of TS, with validation.

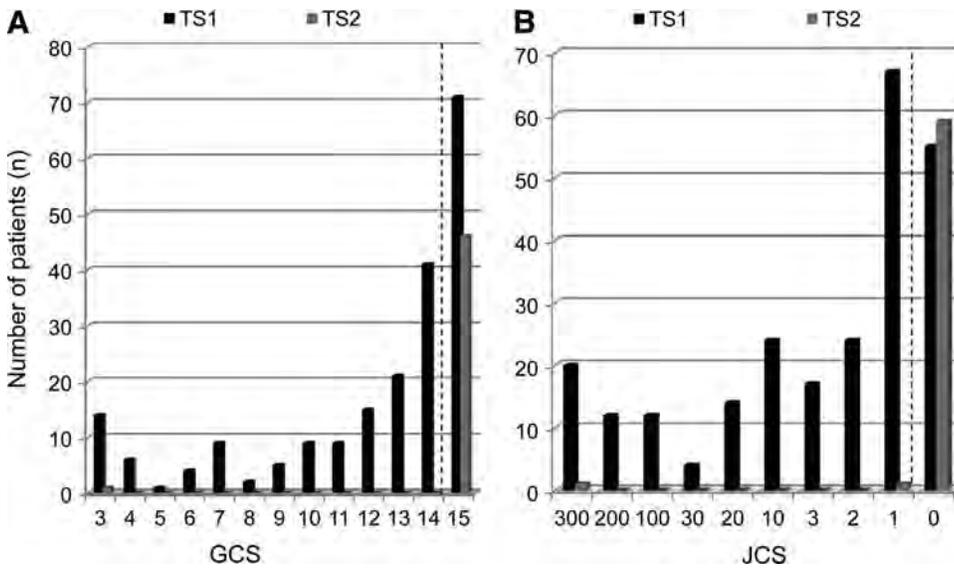


FIG. 5. Distribution of (A) GCS and (B) JCS scores in patients with TS1 (GCS, $n=207$; JCS, $n=249$) and TS2 (GCS, $n=47$; JCS, $n=61$). The dashed lines indicate the cut-off values in our criteria ($GCS \leq 14$, $JCS \geq 1$). GCS, Glasgow Coma Scale; JCS, Japan Coma Scale.

category if they had not had jaundice. Considering patients with TS1 and TS2, the mortality rate in patients with a total bilirubin level more than 3.0 mg/dL (32.3%) was much higher than that in patients with total bilirubin levels equal or lower than 3.0 mg/dL (10.5%) ($p=0.018$, χ^2 test).

Approximately 40% of patients with both TS1 and TS2 had features of CHF (Table 1). Severe features of CHF were present in 30.9% of patients with TS1 and 18.9% patients with TS2. The distributions of NYHA and Killip classifications of CHF for patients with TS1 and TS2 are shown in Figure 6. Although leg edema and pleural effusions were generally the most frequent signs of CHF, logistic regression analysis showed that their contribution to the diagnosis of TS1 or TS2 was not significant (data not shown). Atrial fibrillation (AF) was observed in 39.3% and 33.8% of patients with TS1 and TS2, respectively. Significantly, AF occurred in 52.6% of the patients who died.

Among patients with TS1 and TS2, the occurrence of any one of the five major groups of clinical manifestations (CNS,

fever, tachycardia, CHF, and GI-hepatic) did not positively correlate with the occurrence of any other group of clinical manifestations (Table 4). The occurrence of CHF and fever, and tachycardia and GI/hepatic manifestations, was negatively correlated with each other, but the correlation coefficient was low. Of the patients with TS1, 76% had more than three manifestations (Table 5), consistent with multiple organ failure (MOF). The most common association, found in 15% of patients with TS1, was concomitant CNS symptoms, fever, tachycardia, and GI/hepatic manifestations. The next most common association, found in 11% of patients with TS1, was CNS symptoms, tachycardia, and GI/hepatic manifestations.

Assigning the scores for TS (BWC-TS) used in the reports by Burch and Wartofsky (3,5), the median scores (ranges) of patients with TS1 was 70 points (15–120), and those of patients with TS2 was 52.5 (25–90) points. Table 6 presents the BWC-TS scores for the patients with TS1 and TS2. Table 6 also includes the BWC-TS scores in 50 patients who did not meet LIT-CRITERIA1 or LIT-CRITERIA2 for TS1 or TS2,

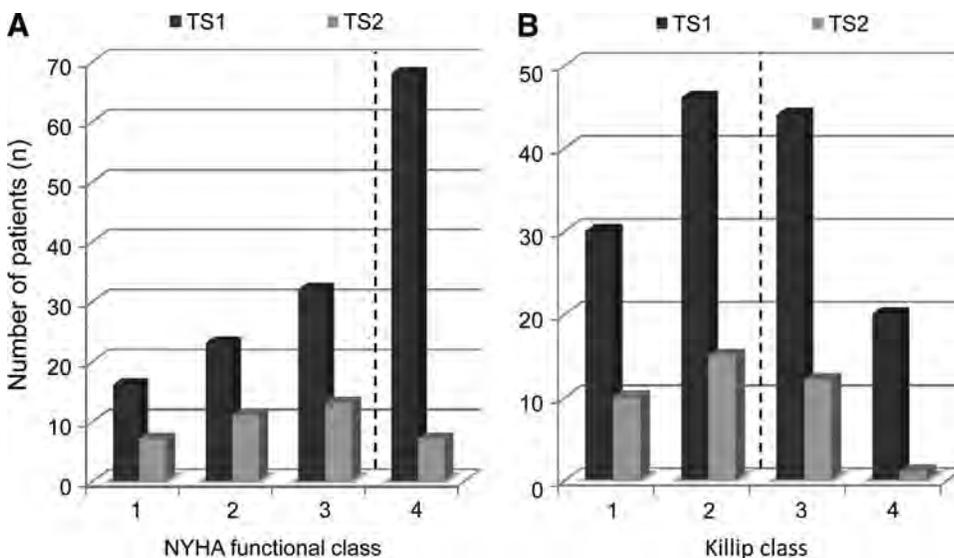


FIG. 6. (A) NYHA and (B) Killip classifications of CHF in patients with TS1 ($n=140$) and TS2 ($n=38$). The dashed bars indicate the cut-off values in our criteria ($NYHA \geq 4$, $Killip \geq 3$). NYHA, New York Heart Association; CHF, congestive heart failure.

TABLE 4. CORRELATIONS (SPEARMAN RANK CORRELATION COEFFICIENT) BETWEEN CLINICAL FEATURES IN JAPANESE PATIENTS WITH THYROID STORM

	Fever CNS ($\geq 38^\circ\text{C}$)	Tachycardia ($\geq 130\text{ bpm}$)	CHF	GI/hepatic
CNS	-0.105	-0.034	-0.028	-0.088
Fever ($\geq 38^\circ\text{C}$)		-0.080	-0.190	-0.089
Tachycardia ($\geq 130\text{ bpm}$)			0.102	-0.144
CHF				-0.066
GI-hepatic				

Data are from the SURVEY-2 of TS, with validation. *p*-Values for bold coefficients were <0.01 .

CHF, congestive heart failure.

respectively. These patients were reported to have TS in the SURVEY-1, but we determined that they did not meet LIT-CRITERIA1 or LIT-CRITERIA2 based on the information obtained in SURVEY-2. A logistic regression analysis for the diagnosis according to the new criteria as ranking variables revealed that the BWC-TS score (3,5) made a significant contribution to differentiating between our patients categorized

TABLE 5. COMBINATIONS OF CLINICAL FEATURES IN JAPANESE PATIENTS WITH THYROID STORM

	CNS symptoms		Other manifestations			Number	
TS1	+		F	T	C	G	16
	+		F	T		G	41
	+			T	C	G	18
	+		F	T	C		8
	+		F		C	G	3
	+			T		G	31
	+		F	T			20
	+		F			G	15
	+			T	C		12
	+				C	G	6
	+					G	27
	+			T			23
	+		F				16
	+				C		2
	+		F	T	C	G	2
	+		F	T		G	22
+			T	C	G	12	
+		F	T	C		5	
+		F		C	G	3	
Total							282
TS2	-			T		G	22
	-		F			G	21
	-		F	T			16
	-			T	C		7
	-				C	G	4
	-		F		C		2
	+		F	T			1 ^a
	+				C		1 ^a
Total							74

Data are from the SURVEY-2 of TS, with validation.

^aThe presence of thyrotoxicosis was not confirmed.

F, fever; T, tachycardia; G, GI/hepatic; C, cardiac.

TABLE 6. RELATIONSHIP BETWEEN BURCH AND WARTOFSKY'S CRITERIA-THYROID STORM SCORES AND THE DIAGNOSIS BASED ON OUR DIAGNOSTIC CRITERIA FOR THYROID STORM

Our diagnostic criteria	BWC-TS scores				Total
	U (0-24)	I (25-44)	L (45-60)	H (≥ 61)	
Reported, not confirmed TS ^a	9	24	16	1	50
TS2	1	19	39	15	74
TS1	1	14	80	187	282
Total	11	57	135	203	406

^aThese patients were reported as having TS by the respondents to SURVEY1, but, after their findings reported in SURVEY-2, they were reviewed by the authors if they did not meet the criteria for the diagnosis of TS1 or TS2.

BWC, Burch and Wartofsky's criteria; U, unlikely; I, impending; L, likely; H, highly likely.

as TS1 or TS2 and our Tox-NoTS patients. The contribution was estimated to be 27.7% ($p < 0.0001$).

The direct causes of death in the 356 patients with TS1 and TS2 were MOF and CHF (Table 7). In a simple regression analysis, many factors including the GCS (odds ratio [OR]: 0.863, $p = 0.0016$), the comorbidity of CHF (OR: 2.086, $p = 0.0426$), and serum creatinine levels (OR: 2.535, $p = 0.0025$) were significantly associated with patient mortality. However, a multiple logistic regression analysis revealed that the comorbidities of shock (OR: 3.90, $p = 0.006$), disseminated intravascular coagulation (DIC) (OR: 3.91, $p = 0.012$), and MOF (OR: 9.85, $p < 0.001$) were identified as independent prognostic factors for death (Table 8). Among the survivors, 22 were reported as having irreversible damage of some kind with brain damage occurring in 6, disuse atrophy occurring in 5, cerebrovascular disease occurring in 4, renal insufficiency occurring in 2, and psychosis occurring in 2. In three patients, the irreversible damage was not specified. A higher GCS and elevated blood urea nitrogen (BUN) were significantly associated with the development of irreversible deficits (GCS, OR: 0.846, $p = 0.0062$; BUN, OR: 1.01, $p = 0.0434$) (Table 8). To differentiate clinically between TS1 and TS2, we

TABLE 7. CAUSES OF DEATH IN JAPANESE PATIENTS WITH THYROID STORM

Causes	Cases	%
MOF	9	24
CHF	8	21
Respiratory failure	3	8
Arrhythmia	3	8
DIC	2	5
Perforation of the GI tract	2	5
Hypoxic brain	1	3
Sepsis	1	3
Unknown	9	23
Total	38	100

Data are from the SURVEY-2 of TS, with validation.

MOF, multiple organ failure; DIC, disseminated intravascular coagulation.

TABLE 8. FACTORS RELEVANT TO THE MORTALITY AND IRREVERSIBLE DAMAGE IN JAPANESE PATIENTS WITH THYROID STORM (MULTIPLE LOGISTIC REGRESSION ANALYSIS)

	Odds ratio for death	p-Value
Mortality		
The presence of shock	3.901	0.0055
The presence of DIC	3.914	0.0123
The presence of MOF	9.853	<0.0001
Irreversible Damages		
GCS	0.846	0.0062
BUN	1.01	0.0434

Data are from the SURVEY-2 of TS, with validation.
GCS, Glasgow Coma Scale; BUN, blood urea nitrogen.

analyzed the prognosis of each syndrome in terms of mortality and irreversible defects. Mortalities of TS1 and TS2 were not significantly different (TS1 vs. TS2; 11.0% and 9.5%, $p=0.70$). In fatal cases, the independent risk factors for mortality were not different between the two syndromes (TS1 vs. TS2, shock; 58.1% vs. 57.1%, $p=0.96$, DIC; 41.9% vs. 28.6%, $p=0.51$, MOF; 54.8% vs. 57.1%, $p=0.91$). Similarly, the rates of irreversible complications were not different between TS1 and TS2 (TS1 vs. TS2; 9.0% vs. 3.3%, $p=0.11$). However, when these analyses were limited to the irreversible neurological defects, including hypoxic brain damage, disuse atrophy, cerebrovascular disease, and psychosis, patients with TS1 had a higher prevalence of irreversible neurological defects than patients with TS2 (TS1 vs. TS2; 6.74% vs. 1.35%, $p<0.05$). As is evident from the diagnostic criteria for each syndrome, the clinical differences between TS1 and TS2 are the presence of neurological symptoms and ensuing defects.

We sought to determine how severely ill patients with TS were, or were perceived to be, based on three indices. One was the level of inpatient care they were assigned to, this being an intensive care unit, a high care unit, or the general wards. Another was their Acute Physiology and Chronic Health Evaluation (APACHE) II scores (14), and the last was the Sequential Organ Failure Assessment (SOFA) scores (15,16). The mean and SD of APACHE II and SOFA scores were 10.97 ± 0.35 (min., 0; max., 37) and 2.67 ± 0.47 (min., 0; max., 12), respectively ($n=354$). The dead exhibited significantly higher scores than those of the survivors in both score methods: 15.00 ± 1.04 vs. 10.48 ± 0.36 , $p<0.0001$ in APACHE II; 3.12 ± 0.51 vs. 2.38 ± 0.13 , $p<0.0001$ in SOFA. In a multiple logistic regression analysis, the inpatient level of care was significantly related to the GCS levels (OR: 0.857, $p=0.0242$) and serum levels of total cholesterol (OR: 0.973, $p=0.0005$) (see Fig. 7). In our TS1 and TS2 patients, the factors that were independently associated with the calculated APACHE II scores were age ($p=0.0003$), GCS ($p<0.0001$), serum albumin ($p=0.0329$), serum creatinine ($p=0.0366$), and base excess ($p=0.0011$) (see Supplementary Data). The factors independently associated with calculated SOFA scores were GCS ($p<0.0001$), Graves' ophthalmopathy ($p<0.0001$), serum albumin ($p=0.0004$), serum total cholesterol ($p<0.0001$), serum total bilirubin ($p<0.0001$), shock ($p<0.0001$), CHF ($p=0.013$), and PaCO₂ ($p=0.0059$).

Although an accurate estimation is impossible because not all possible sites for the occurrence of TS were asked to report their patients with TS, we estimate that the number of cases of TS1

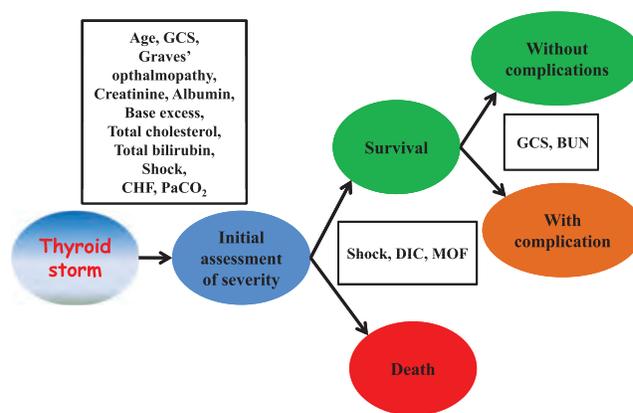


FIG. 7. Possible factors contributing to the severity and prognosis of thyroid storm. DIC, disseminated intravascular coagulation; MOF, multiple organ failure.

and TS2 in Japan would be ~165 and 43 per year, respectively, based on the response rates for SURVEY-1 and SURVEY-2.

Development of FINAL-CRITERIA1 and FINAL-CRITERIA2 for the diagnosis of TS1 and TS2

To develop the final diagnostic criteria for the diagnosis of TS1 and TS2, we considered modifying LIT-CRITERIA1 and LIT-CRITERIA2 to reflect our analysis of the features of TS patients in Japan who were identified in SURVEY-1 and characterized in SURVEY-2. First, we examined the prevalence and independence of clinical manifestations of these patients and of Tox-NoTS patients, including analyses of combinations of clinical features. Next, we investigated whether the cut-off values for each manifestation provided appropriate sensitivity, specificity, and predictive values to differentiate between TS1, TS2, and Tox-NoTS. Based on these analyses, we made only one change to generate the final diagnostic criteria for TS. This was related to the serum bilirubin concentration. As noted in the results, the mortality rate for TS1 and TS2 patients with total serum bilirubin concentration >3.0 mg/dL was significantly higher than that in those with lower serum bilirubin concentrations. Therefore, we made a serum total bilirubin concentration of >3 mg/dL a criteria for GI/hepatic manifestations. Based on this revision, three patients who were classified as TS2 by LIT-CRITERIA2, met the criteria for TS1 in FINAL-CRITERIA1. This was because by FINAL-CRITERIA1 now met the criteria for GI/hepatic manifestations. These patients were not identified as having jaundice in SURVEY-2, but their serum bilirubin was >3 mg/dL. Notably, none of these three patients died in TS. The final criteria for the diagnosis of TS1 and TS2 (i.e., The Second Edition of the Diagnostic Criteria for TS) are presented in Table 9. They differ from the criteria based on an analysis of the literature only in the fact that a serum bilirubin of >3 mg/dL was used for one of the criteria for GI/hepatic manifestations.

Discussion

The present study helps clarify the epidemiology of TS, in some regard, for the first time. The incidence of TS cases in Japan, including both TS1 and TS2, was estimated to be 1283 ± 105 [95% confidence interval: 1077–1489] per 5 years

TABLE 9. FINAL CRITERIA FOR THE DIAGNOSIS OF THYROID STORM

Grade of TS	Combinations of features	Requirements for diagnosis
TS1	First combination	Thyrotoxicosis and at least one CNS manifestation and fever or tachycardia, or CHF or GI/hepatic manifestations
TS1	Alternate combination	Thyrotoxicosis and at least three combinations of fever, or tachycardia, or CHF, or GI/hepatic manifestations
TS2	First combination	Thyrotoxicosis and a combination of two of the following: fever or tachycardia or CHF or GI/hepatic manifestations
TS2	Alternate Combination	Patients who meet the diagnostic criteria for TS1 except that serum FT3 or FT4 values are not available but whose data before or after the episode suggest that they are thyrotoxic at the time of TS.

Definitions

Thyrotoxicosis: Elevated FT3 or FT4. CNS manifestations: Restlessness, delirium, mental aberration/psychosis, somnolence/lethargy, convulsion, coma including a score of 1 or higher on the Japan Coma Scale (JCS) or 14 or lower on the Glasgow Coma Scale (GCS).

Fever: 38°C or higher.

Tachycardia: ≥ 130 beats/min (arrhythmias such as atrial fibrillation are evaluated by measuring the heart rate).

CHF: The patient presents with severe symptoms such as pulmonary edema, moist rales for more than half the lung field, or cardiogenic shock. The patient's CHF is categorized as Class IV by the NYHA classification or Class III or higher by the Killip classification.

GI/hepatic manifestations: The patient presents with nausea, vomiting, diarrhea, or a bilirubin of >3 mg/dL.

Exclusions and Provisos

Cases are excluded if other underlying diseases are clearly causing any of the following symptoms: fever (e.g., pneumonia and malignant hyperthermia), impaired consciousness (e.g., psychiatric disorders and cerebrovascular disorders), heart failure (e.g., acute myocardial infarction), and liver disorders (e.g., viral hepatitis and acute liver failure). However, some of these disorders trigger thyroid storm. Therefore, it is difficult to determine whether the symptom is caused by thyroid storm or is simply a symptom of an underlying disease that is possibly triggered by thyroid storm; the symptom should be regarded as being due to a thyroid storm that is caused by these precipitating factors. Clinical judgment in this matter is required.

(0.2 persons/100,000 Japanese population/year), $\sim 0.22\%$ of all thyrotoxic patients and 5.4% of admitted thyrotoxic patients. The age of TS1 and TS2 patients was 44.7 ± 16.7 (median: 42, range: 6–87) and 44.6 ± 14.6 (median: 44, range: 20–80), respectively. No gender differences in the age of patients with TS1 or TS2 were observed. The ratio of male-to-female patients with TS1 and TS2 was $\sim 1:3$ (89:263). These epidemiological findings concerning age and gender distribution were similar to those for the 99 cases of TS found in our literature search and the 7 cases from the researchers' facilities.

There were several striking aspects regarding the clinical features and course of TS. The mortality rate of TS is still high. Many cases of TS occur in patients who have not received treatment, and many occur within the first year of treatment for GD. An apparent trigger for the development of TS was observed in 70% of our patients. The remaining 30% of patients may have had some type of stress, one not identified by their clinicians, that occurred before they developed TS. A common feature of patients with TS is either that they had taken their antithyroid drugs irregularly or had discontinued them altogether. In the management of patients with GD, therefore, it is important to counsel patients and their families about the need for regular and reliable use of their antithyroid drugs. Other illness or events that triggered TS in this survey were similar to those described in the literature (Table 2). No cases of TS were caused by thyroid surgery, indicating an improvement in the management of patients with GD before thyroid surgery.

None of the clinical manifestations assessed in SURVEY-2 were positively associated with each other. CNS manifestations were weighted most heavily in our diagnostic criteria, as they were found to be common in TS. The exclusion of agitation as a CNS-related diagnostic item did not affect the diagnosis of TS (data not shown). Although the presence of nausea, vomiting, and/or diarrhea was not associated

with the prognosis of TS, these manifestations, which are not frequent in nonstorm thyrotoxicosis, are rather specific for TS. Moreover, if we exclude these manifestations from diagnostic criteria of TS, a significant number of patients with TS1 and TS2 would become classified, respectively, as TS2 and Tox-NoTS. Therefore, we consider these GI manifestations as being important for the diagnosis of TS.

Prompted by the Child-Pugh score, which assesses the prognosis of liver diseases (17–19), we analyzed whether elevated serum total bilirubin levels were associated with the prognosis of TS1 and TS2. We not only found that the mortality rate in the group with a total bilirubin level >3 mg/dL was significantly higher than that in the group with total bilirubin levels equal to or lower than 3 mg/dL, but we also observed that simply asking whether patients had jaundice did not identify all patients with bilirubin levels >3 mg/dL. Therefore, in formulating the final criteria for the diagnosis of TS1 and TS2, we changed the criteria for GI/hepatic manifestations from jaundice to a serum bilirubin concentration 3 mg/dL.

Tachycardia and high output-induced CHF are common cardiac manifestations of TS (20). SURVEY-2 confirmed the feasibility of a cut-off value for heart rates ≥ 130 bpm, irrespective of the presence or absence of AF. The survey also demonstrated that leg edema and pleural effusion were not specific signs for CHF in TS, but CHF categorized as NYHA functional class IV and/or Killip class III or IV that was associated with pulmonary edema and cardiogenic shock was of particular importance for the diagnosis of TS. Arrhythmias are often associated with TS, with AF being the most common. This occurred in 10%–15% of cases of TS (21). In SURVEY-2, AF was observed in 38.2% of patients with TS1 and TS2 and 52.6% of patients who died. We suggest that CHF facilitates the occurrence of AF in TS. This, in turn, exaggerates CHF, resulting in a vicious cycle that contributes to mortality.

Our criteria for the diagnosis of TS differs in two major ways from the BWC-TS (3,5). First, in our criteria, the presence of evidence for thyrotoxicosis is a prerequisite item for these diagnoses. We consider that making thyrotoxicosis a prerequisite item is needed to reduce false-positive diagnoses. Second, in contrast to the BWC-TS, there is no scoring system for our criteria for TS. It is impossible, in our opinion, to generate a scoring system in a retrospective study, as a prospective study is needed for validation. Indeed, the BWC-TS score is supposed to be neither evidence based nor validated. In addition, the BWC-TS scoring system is rather complex, making it difficult to memorize precisely. Although there was a significant correlation between our diagnostic criteria and the BWC-TS, the fact that the contribution of the BWC-TS was small indicates a discrepancy between the two diagnostic systems (Table 6). For instance, one patient who scored 75 points according to the BWC-TS was not diagnosed with TS based on our criteria. He had tachycardia (146 bpm) with AF (25+10 points), ongoing mild CHF (10 points), and a slight elevation of serum bilirubin (2.4 mg/dL) (20 points) that was exacerbated by infection (10 points), and no fever. However, mild CHF (NYHA III) and only slight bilirubin elevation without hepatic dysfunction meant that the patient did not meet our criteria for TS according to our guidelines, even though he had marked tachycardia. In contrast, a patient whose score was only 15 points by the BWC-TS (diarrhea, edema, and mild unconsciousness) was diagnosed as TS2 by our criteria. One reason for the different diagnoses in this case may be the method for judging CNS features. Mild CNS symptoms (GCS: 14, or JCS: 1) were evaluated differently by our system than by the BWC-TS. Notwithstanding these exceptional cases, however, both the BWC-TS and our criteria may be helpful in diagnosing TS.

All of our criteria for TS require elevated serum thyroid hormone levels, FT3 and/or FT4, as a prerequisite item. It is, however, well known that circulating thyroid hormones, particularly T3, tend to be relatively low in severe non-thyroidal illness. This is the case in TS as we formulated the definition and, in fact, several patients with TS1 had normal serum FT3 levels despite elevated serum FT4 and low serum TSH levels. We should keep in mind, moreover, that even serum FT4 levels become low in severe illness as do serum TSH concentrations. In addition, the administration of drugs, including antithyroid drugs, iodine, and corticosteroid, modify thyroid function tests. Therefore, we may need to expand the diagnostic criteria for TS, but further studies will be needed.

The mortality rate of TS was 10.7%. This is lower than previously reported (2,22). Improvements in the general management of patients and early diagnosis likely contributed to the reduced mortality rate. None of our 133 thyrotoxic patients without TS (i.e., Tox-NoTS patients) died. Notably also, only 1 of 50 patients identified by the respondents in the SURVEY-1 as having TS, but who did not meet the criteria for TS1 or TS2 based on the information from SURVEY-2, died. This patient had only a GI manifestation and died because of *Pneumocystis carinii* pneumonia.

With regard to patients who met our criteria for TS1 and TS2, it is notable and perhaps surprising that the mortality rate was almost as high, using LIT-CRITERIA2 and LIT-CRITERIA1 respectively, for TS2 (9.5%) as it was for TS1

(11.0%). In fact, using FINAL-CRITERIA2 and FINAL-CRITERIA1 for TS2 and TS1, the mortality rates were even more similar, being 9.9% and 10.9% respectively. Moreover, the most frequent direct causes of death, such as shock, DIC, and MOF, were very similar in the two groups. This observation strongly suggests that mortality depends on severe complications rather than the characteristic clinical features of thyroid crisis. In addition, treatments/interventions for these severe conditions, which were variably used among doctors, probably influenced the mortality rate. Consistently, the comorbidity of mortality with a CNS manifestation or CHF was found in simple regression analysis. Indeed, TS mortality has been attributed to these factors in the literature (1,2). However, in our multiple regression analysis, these were not pivotal independent factors for patient death. This is perhaps because recent developments in the management of critically ill patients reduced the mortality related to these factors. However, CNS symptoms did have a profound impact on the quality of life as shown by our analysis of irreversible damage (Table 8). In general, we evaluated many factors that are assessed in critically ill patients (Fig. 7). We recommend that, in treating TS, all these complications be carefully managed.

Since the mortality rates, and the rates of irreversible complications, were so similar in TS1 and TS2, it might be argued that there is no need for graded diagnostic criteria for TS. There was a difference, however, in the prevalence of irreversible neurological defects between the two categories with irreversible defects being higher in TS1 than in TS2. This may reflect that CNS manifestations were the central item in LIT-CRITERIA1 and FINAL-CRITERIA1. Further studies may uncover gradations of TS that are meaningfully different. For the present, it is evident that patients who meet the criteria for both TS1 and TS2 are seriously ill, and they require intense management.

There are several limitations of this study. First, this was a retrospective study. However, TS is rare and its occurrence in unpredictable, making a prospective study difficult to perform. Second, the surveys were performed in Japan only, and the response rates for SURVEY-1 and SURVEY-2 were only 52.5% and 65.1%, respectively. Therefore, our results may reflect only the characteristics of an advanced and East-Asian country. Third, the estimate of incidence rate is based on the assumption that the incidence of TS is the same in the hospitals that did not respond to the surveys. Considering that responder hospitals may have responded because they did have a case or cases to report, we may have overestimated the number of the patients. Nonetheless, we believe that TS may occur more often than endocrine and thyroid experts realize, and its high mortality rate is under appreciated. Thus, even nowadays, TS is one of the important endocrine emergencies. Fourth, in different studies, the analyses of the relative importance of factors that influence prognosis are substantially influenced by the treatment interventions used for the patients in a particular study. Since patients with TS are quite variably treated, according to their severities and clinical manifestations, it is impossible to analyze them as a homogeneous group. Such analyses can be accurately achieved by the prospective study only. At this point, we think it may be useful to grade TS into two categories, but it should be emphasized that even patients with TS2 have high mortality rates and rates of irreversible complications. Therefore, it is

probably important to “over-diagnose” TS. Ideally, one set of diagnostic criteria that are 100% sensitive and specific should be available but, at the moment, it seems impossible to diagnose TS in a simple way. TS has dynamic, heterogeneous and complex features. Moreover, it needs early treatment to prevent increasing severity and an irreversible outcome.

In summary, we propose new diagnostic criteria for TS, assigning two grades relating to signs and symptoms, though not necessarily to severity or prognosis. These were based on evidence derived from nationwide surveys in Japan using questionnaires developed from an analysis of a large number of studies in two major literature databases. We hope they will contribute to prompt and precise clinical decisions and the treatment of this disorder. Furthermore, we hope to conduct prospective evaluations of our diagnostic criteria and points of emphasis in the therapy of TS. We believe that we can ultimately achieve a better prognosis of TS by utilizing these approaches.

Acknowledgments

The authors thank the members of the Japan Thyroid Association and Japan Endocrine Society, as well as the doctors participating in Japanese hospitals and clinics for their valuable and kind cooperation in the questionnaires and nationwide surveys. This study was supported by a fund from the Ministry of Health, Labor, and Welfare of Japan.

Disclosure Statement

The authors declare that they have nothing to disclose, except for research grants for T.A. and Y.N. from the National Government of Japan.

References

1. Gavin LA 1991 Thyroid crises. *Med Clin North Am* **75**:179–193.
2. Tietgens ST, Leinung MC 1995 Thyroid storm. *Med Clin North Am* **79**:169–184.
3. Wartofsky L 2005 Thyrotoxic storm. In: Braverman L, Utiger R (eds) *Werner & Ingbar's the Thyroid*, 9th edition. Williams & Wilkins, Philadelphia, PA, pp 651–657.
4. Dillmann WH 1997 Thyroid storm. *Curr Ther Endocrinol Metab* **6**:81–85.
5. Burch HB, Wartofsky L 1993 Life-threatening thyrotoxicosis. Thyroid storm. *Endocrinol Metab Clin North Am* **22**:263–277.
6. Nagai M, Tamakoshi A, Hashimoto S, Ohno Y, Nakamura K 2006 The Nationwide Epidemiological Survey Manual, 2nd edition (in Japanese). Japanese Ministry of Health and Welfare, Tokyo.
7. Hashimoto S, Fukutomi K, Nagai M, Nakamura Y, Yanagawa H, Sasaki R, Ohno Y, Kubo N, Aoki K 1991 [A method of interval estimation for number of patients in the nationwide epidemiological survey on intractable diseases] (in Japanese). *Nippon Koshu Eisei Zasshi* **38**:880–883.
8. Nakamura Y, Matsumoto T, Tamakoshi A, Kawamura T, Seino Y, Kasuga M, Yanagawa H, Ohno Y 2000 Prevalence of idiopathic hypoparathyroidism and pseudohypoparathyroidism in Japan. *J Epidemiol* **10**:29–33.
9. Langfitt TW 1978 Measuring the outcome from head injuries. *J Neurosurg* **48**:673–678.
10. Starmark JE, Stalhammar D, Holmgren E, Rosander B 1988 A comparison of the Glasgow Coma Scale and the Reaction Level Scale (RLS85). *J Neurosurg* **69**:699–706.
11. Ono K, Wada K, Takahara T, Shirotani T 2007 Indications for computed tomography in patients with mild head injury. *Neurol Med Chir (Tokyo)* **47**:291–297; discussion 297–298.
12. Ministry of Internal Affairs and Communications. Results of the Population Estimates. Statistics Bureau, Ministry of Internal Affairs and Communications, Japan. Available at www.stat.go.jp/english/data/jinsui/2.htm (accessed May 1, 2012).
13. Ministry of Health, Labor, and Welfare 2008 Patient Survey. Statistics and Information Department, Ministry of Health, Labor, and Welfare, Japan. Available at www.mhlw.go.jp/english/database/db-hss/ps.html (accessed May 1, 2012).
14. Wagner DP, Draper EA 1984 Acute physiology and chronic health evaluation (APACHE II) and Medicare reimbursement. *Health Care Financ Rev Suppl*:91–105.
15. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonca A, Bruining H, Reinhart CK, Suter PM, Thijs LG; on behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine 1996 The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med* **22**:707–710.
16. Antonelli M, Moreno R, Vincent JL, Sprung CL, Mendoca A, Passariello M, Riccioni L, Osborn J 1999 Application of SOFA score to trauma patients. *Sequential Organ Failure Assessment. Intensive Care Med* **25**:389–394.
17. Child CG, Turcotte JG 1964 Surgery and portal hypertension. In: Child CG (ed) *The Liver and Portal Hypertension*. W.B. Saunders, Philadelphia, PA, pp 50–64.
18. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R 1973 Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg* **60**:646–649.
19. Pugh RN 1992 Pugh's grading in the classification of liver decompensation. *Gut* **33**:1583.
20. Woeber KA 1992 Thyrotoxicosis and the heart. *N Engl J Med* **327**:94–98.
21. Sawin CT, Geller A, Wolf PA, Belanger AJ, Baker E, Bacharach P, Wilson PW, Benjamin EJ, D'Agostino RB 1994 Low serum thyrotropin concentrations as a risk factor for atrial fibrillation in older persons. *N Engl J Med* **331**:1249–1252.
22. Burger AG, Philippe J 1992 Thyroid emergencies. *Baillieres Clin Endocrinol Metab* **6**:77–93.

Address correspondence to:
Takashi Akamizu, M.D., Ph.D.
The First Department of Medicine
Wakayama Medical University
811-1 Kimi-idera
Wakayama 641-8509
Japan

E-mail: akamizu@wakayama-med.ac.jp

Appendix A. The Diagnostic Criteria for Two Grades of Thyroid Storm, TS1 and TS2, Based on Our Review of the Literature

Definition

Thyrotoxic storm or crisis is a life-threatening condition requiring emergency treatment. The condition, which is often triggered by severe physical or mental stress, arises in thyrotoxic patients. These patients manifest multiple organ failure as a result of the breakdown of compensatory mechanisms.

Prerequisite for diagnosis

Presence of thyrotoxicosis (elevated free triiodothyroine [T3] and/or thyroxine [T4] levels)

Symptoms (Note 1)

1. Central nervous system (CNS) manifestations (Note 2)
2. Fever (38°C or higher)
3. Tachycardia (130 beats/min or faster) (Note 3)
4. Congestive heart failure (CHF) (Note 4)
5. Gastrointestinal (GI) and hepatic manifestations (Note 5)

Diagnostic criteria for “definite”

Patients who meet the prerequisite for thyrotoxicosis and either of the following criteria are regarded as definite cases (Note 6):

- (i) At least one CNS manifestation plus one or more of the other symptoms just listed;
- (ii) Three or more of the manifestations just listed other than CNS manifestations

Diagnostic criteria for suspected cases

Patients who meet either of the following criteria are regarded as suspected cases:

- (i) A prerequisite for diagnosis is diagnosis of thyrotoxicosis plus two of the manifestations just listed, excluding CNS manifestations;
- (ii) In cases where it cannot be confirmed whether the patient meets the prerequisite for diagnosis, it is known that the patient has a history of thyroid disease, presents with exophthalmos and goiter, and meets (i) or (ii) from the criteria for definite cases (Note 6).

Notes

Note 1. Cases are excluded if other underlying diseases are clearly causing any of the following symptoms: fever

(pneumonia and malignant hyperthermia), impaired consciousness (psychiatric disorders and cerebrovascular disorders), heart failure (acute myocardial infarction), and liver disorders (viral hepatitis and acute liver failure). However, some of these disorders trigger thyrotoxic crisis. Therefore, if it is difficult to determine whether the symptom is caused by thyrotoxic crisis or is simply a symptom of an underlying disease that is possibly triggered by thyrotoxic crisis; the symptom should be regarded as being due to thyrotoxic crisis caused by these precipitating factors.

Thus, thyrotoxic crisis often occurs in the presence of its triggers. Triggers that are directly related to thyroid diseases include the irregular use or discontinuation of antithyroid drugs, thyroid surgeries, isotope therapy of the thyroid, excessive palpation or biopsy of the thyroid, and use of thyroid hormone preparation at high doses. Triggers that are not directly related to the thyroid include infection, surgeries on organs other than the thyroid, trauma, pregnancy and delivery, adrenocortical insufficiency, diabetic ketoacidosis, administration of iodinated contrast medium, cerebrovascular disorders, pulmonary thromboembolism, ischemic heart diseases, extraction of teeth, severe emotional stress, and strenuous exercise.

Note 2. The patient presents with restlessness, delirium, mental aberration/psychosis, somnolence/lethargy, convulsion, or coma. The patient scores 1 or higher on the Japan Coma Scale (JCS) or 14 or lower on the Glasgow Coma Scale (GCS).

Note 3. Arrhythmias such as atrial fibrillation are evaluated by measuring the heart rate.

Note 4. The patient presents with a severe symptom such as pulmonary edema, moist rales for more than half the lung field, or cardiogenic shock. The patient’s CHF is categorized as Class IV by the New York Heart Association (NYHA) classification or Class III or higher by the Killip classification.

Note 5. The patient presents with nausea, vomiting, diarrhea, or jaundice.

Note 6. Elderly patients should be diagnosed carefully, because, in some cases, they do not show typical symptoms of thyrotoxic crisis, that is, high fever and hyperactivity (apathetic thyroid storm).

Appendix B. Thyroid Storm: Questionnaires for the First Nationwide Survey

Name of Doctor _____
 Department/Organization _____
 Date (Month) (Day), 2009

The number of patients with thyroid storm

2004 (year)	1. Absent	2. Present:	Male ____,	Female ____ cases
2005 (year)	1. Absent	2. Present:	Male ____,	Female ____ cases
2006 (year)	1. Absent	2. Present:	Male ____,	Female ____ cases
2007 (year)	1. Absent	2. Present:	Male ____,	Female ____ cases
2008 (year)	1. Absent	2. Present:	Male ____,	Female ____ cases

Notes

1. According to the first version of diagnostic criteria, please include both definite and suspected cases with thyroid storm which you encountered from January 1, 2004, to December 31, 2008.
2. Since the prevalence rate of thyroid storm in Japan is calculated, please return even if there is no case.
3. After the first survey, the second survey will follow. Thank you for your cooperation.

Please return until February 15, 2009.

**Appendix C. Nationwide Epidemiological Survey of Thyroid Storm:
Questionnaires for the Second Nationwide Survey (Patient Information)**

Organization: _____
 Department: _____
 Address _____

Name of Doctor _____
 Date _____ (Month) _____ (Day), 2009

I. Patient profile

Name (Initials)	Family	First	Sex	M / F	Serial no.	
Date of birth	(Year)	(Month)	(Day)		(yrs old)
Address (Pref.)						
Family history of thyroid diseases	Absent / Present, Relation (Parents • Children • Siblings • Grand parents) Name of disease ()					
Past history (Year of occurrence)	Absent / Present (Disease:) (year) (Disease:) (year) (Disease:)					
Liking	Alcohol : No • Yes (/day) Smoking : No • Yes (cigarettes/day)					
Basal thyroid diseases	1. Graves' disease 2. Others () Organization diagnosed: Yours / Others () History of treatment: None / Present					
	Organization treated: Yours / Others () Duration from the development of basal thyroid disease to storm : Years Months Treatment : MMI / PTU mg/day / Radioisotope / Thyroid Surgery / Others () Condition of control: Well /Poor (Reason:)					
1st organization diagnosed as storm	Yours (Department:) / Others (Transferred from)					
Date diagnosed	(Year)	(Month)	(Day)			

Triggers	Absent / Present
Triggers of storm	1. Irregular use or discontinuation of anti-thyroid drugs, 2. Thyroid surgeries, 3. Isotope therapy of the thyroid, 4. Excessive palpation or biopsy of the thyroid, 5. Use of thyroid hormone preparation at high doses, 6 Infection (Specify: _____), 7. Surgeries on organs other than the thyroid (Specify: _____), 8. Trauma, 9. Pregnancy and delivery, 10. Adrenocortical insufficiency, 11. Diabetic ketoacidosis, 12. Administration of iodinated contrast medium, 13. Cerebrovascular disorders, 14. Pulmonary thromboembolism, 15. Ischemic heart diseases, 16. Extraction of teeth, 17. Severe emotional stress, 18. Strenuous exercise, 19. Others (_____)

II. Clinical manifestations at the diagnosis of storm (Please fill in the maximum values of body temperature (BT), pulse/heart rate before treatment)

Chief complaints	
Height	cm BW kg Temperature °C
BP	/ mmHg Respiration rate /min
Goiter (Palpation / US)	Absent / Present : Mild (< 5cm in diameter) • Moderate (≥ 5cm in diameter) • Huge (≥100g in estimated weight)
Ophthalmopathy	Absent / Present (Exophthalmos / Other ocular manifestations)
Pulse rates (Heart rates)	beats/min (beats/min), Regular / Irregular (Atrial fibrillation: absent / present)
Consciousness	Japan Coma Scale , Glasgow Coma Scale
CNS manifestations	Absent / Present (Agitation, Restlessness, Delirium, Mental aberration/psychosis, Somnolence/lethargy, Convulsion, Coma)
CHF manifestations	Absent Present (Leg edema, Pleural effusion, Moist rales over more than half the lung field, or Cardiogenic shock) New York Heart Association Classification Class 1 / 2 / 3 / 4 Killip Classification Class I / II /III/ IV Forrester Classification Subset I / II /III/ IV
GI/hepatic manifestations	Absent Present (Nausea, Vomiting, Diarrhea, Jaundice with hepatic dysfunction)
Complications (during Management)	Absent Present (1. Shock, 2. Rhabdomyolysis, 3. DIC, 4. MOF, 5. Others (Specify: _____))
Others	'Profuse sweating': Absent / Present

III. Laboratory data at the development of thyroid storm

<u>Urinalysis ((M), (D), (days after development))</u>	Results	<u>Biochemical ((M), (D), (days after development))</u>	Results
pH		TP (g/dL)	
Gravity		Alb (g/dL)	
Glucose		T-Bil (mg/dL)	
Protein		Direct-Bil (mg/dL)	
Ketone		AST (IU/mL)	
Occult blood		ALT (IU/mL)	
Urobilinogen		LDH (IU/mL)	
<u>Blood cell counts ((M), (D), (days after development))</u>		CK (IU/mL)	
Hct (%)		ALP (IU/mL)	
Hb (g/dl)		γ-GPT (IU/mL)	
RBC (10 ⁴ / L)		ChE (IU/L)	
WBC (/ L)		BUN (mg/dL)	
Plt (10 ⁴ / L)		Cr (mg/dL)	
<u>Thyroid ((M), (D), (days after development))</u>		Na (mEq/L)	
TSH (U/mL)		K (mEq/L)	
freeT4 (ng/dL)		Cl (mEq/L)	
T4 (ng/dL)		Ca (mg/dL)	
freeT3 (pg/mL)		P (mg/dL)	
T3 (ng/dL)		T-Chol (mg/dL)	
TRAb (% or IU/L)		LDL-Chol (mg/dL)	
TSAb (%)		TG (mg/dL)	
TPO-Ab (U/ml) (Ref. value)		Glucose (mg/dL)	
MCHA (fold)		CRP (mg/dl)	
TG-Ab (U/mL) (Ref. value)		<u>Coagulation ((M), (D), (days after development))</u>	
TGHA (fold)		Fibrinogen (mg/dL)	
<u>Others ((M), (D), (days after development))</u>		PT (INR)	
ACTH (pg/mL)		FDP (µg/mL)	
Cortisol (µg/dL)		D-dimer (µg/mL)	
BNP (pg/mL)			
<u>Blood gas ((M), (D), (days after development))</u>	Results		Results
pH		BE (mEq/L)	
PaO ₂ (mmHg) (FiO ₂ : %)		A-aDO ₂ (mmHg)(FiO ₂ : %)	
PaCO ₂ (mmHg)			
HCO ₃ ⁻ (mEq/L)			

Ultrasound Cardiography

LVEF (lt. ventricular ejection fraction): _____ %

Cardiac output: _____ L/min

E/A ratio: _____
 (the ratio between early (E) and late (atrial - A) ventricular filling velocity)

LAD (lt. atrial diameter) : _____ mm

IVC diameter: _____ mm

PVs (pulmonary vein systolic flow velocity) : _____ cm/sec

PVd (pulmonary vein diastolic flow velocity) : _____ cm/sec

PVa (pulmonary vein atrial reversal flow) : _____ cm/sec

IV. Treatment of thyroid storm

Wards	General • HCU • ICU • Others ()
Antithyroid drugs	Absent Present (Initial dose: MMI oral/intravenous mg/day • PTU mg/day)
Iodines	Absent Present (Initial dose: mg/day) (KI, iodinated contrast medium)
β-blockers	Absent Present (, Initial dose: oral mg/day, intravenous mg/hr)
Glucocorticoids	Absent Present (, Initial dose: mg/day)
Cardiotonics/vasopressors	Absent Present (, Initial dose:)
Antipyretics	Absent Present (, dose: /day)
Drugs to CNS manifestations	Absent Present (, dose: /day)
Other specific treatments	Absent Present (Intratracheal tubing /Plasmapheresis / Hemodialysis / Heart-lung machine / Others [])
Other particular treatments (e.g. treatment for triggers)	

