



## Validity and Reliability of the Korean Version of the Hyperthyroidism Symptom Scale

Jie-Eun Lee<sup>1</sup>, Dong Hwa Lee<sup>2</sup>, Tae Jung Oh<sup>2</sup>, Kyoung Min Kim<sup>2</sup>, Sung Hee Choi<sup>2</sup>, Soo Lim<sup>2</sup>, Young Joo Park<sup>3</sup>, Do Joon Park<sup>3</sup>, Hak Chul Jang<sup>2</sup>, Jae Hoon Moon<sup>2</sup>

<sup>1</sup>Department of Internal Medicine, Seoul National University Healthcare System Gangnam Center, Seoul National University College of Medicine, Seoul; <sup>2</sup>Department of Internal Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam; <sup>3</sup>Department of Internal Medicine, Seoul National University Hospital, Seoul National University College of Medicine, Seoul, Korea

**Background:** Thyrotoxicosis is a common disease resulting from an excess of thyroid hormones, which affects many organ systems. The clinical symptoms and signs are relatively nonspecific and can vary depending on age, sex, comorbidities, and the duration and cause of the disease. Several symptom rating scales have been developed in an attempt to assess these symptoms objectively and have been applied to diagnosis or to evaluation of the response to treatment. The aim of this study was to assess the reliability and validity of the Korean version of the hyperthyroidism symptom scale (K-HSS).

**Methods:** Twenty-eight thyrotoxic patients and 10 healthy subjects completed the K-HSS at baseline and after follow-up at Seoul National University Bundang Hospital. The correlation between K-HSS scores and thyroid function was analyzed. K-HSS scores were compared between baseline and follow-up in patient and control groups. Cronbach's  $\alpha$  coefficient was calculated to demonstrate the internal consistency of K-HSS.

**Results:** The mean age of the participants was  $34.7 \pm 9.8$  years and 13 (34.2%) were men. K-HSS scores demonstrated a significant positive correlation with serum free thyroxine concentration and decreased significantly with improved thyroid function. K-HSS scores were highest in subclinically thyrotoxic subjects, lower in patients who were euthyroid after treatment, and lowest in the control group at follow-up, but these differences were not significant. Cronbach's  $\alpha$  coefficient for the K-HSS was 0.86.

**Conclusion:** The K-HSS is a reliable and valid instrument for evaluating symptoms of thyrotoxicosis in Korean patients.

**Keywords:** Thyrotoxicosis; Hyperthyroidism symptom scale; Validity

### INTRODUCTION

Thyrotoxicosis refers to a clinical syndrome of excess circulating thyroid hormone, irrespective of the source, which affects many organ systems [1]. The clinical symptoms and signs of

thyrotoxicosis are fatigue, anxiety, palpitations, sweating, heat intolerance, anxiety, disturbed sleep, weight loss, tachycardia, and tremor of the extremities [2,3]. These clinical manifestations are relatively nonspecific and can vary depending on factors such as the patient's age, sex, comorbidities, and the dura-

Received: 7 October 2017, Revised: 15 November 2017,

Accepted: 26 December 2017

Corresponding author: Jae Hoon Moon

Department of Internal Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, 82 Gumi-ro 173beon-gil, Bundang-gu, Seongnam 13620, Korea

Tel: +82-31-787-7068, Fax: +82-31-787-4052, E-mail: jaemoon76@gmail.com

Copyright © 2018 Korean Endocrine Society

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

tion and cause of the disease [4,5]. This variety of nonspecific symptoms and signs makes it difficult to diagnose or assess disease status based on symptoms and signs alone. Moreover, when deciding how to treat patients diagnosed with thyrotoxicosis, the physician needs to assess the clinical severity of disease and its predicted response to the available treatments. Because the clinical severity of thyrotoxicosis does not always reflect the magnitude of the elevation in thyroid hormone levels [6], it would be useful to have a method to assess clinical status and the response to various treatments.

Previous attempts have been made to develop instruments to evaluate the clinical symptoms and signs of thyrotoxicosis. Crooks et al. [7] described a diagnostic index that scores the presence or absence of various signs and symptoms of thyrotoxicosis for the purpose of establishing a diagnosis. In 1965, Sir Edward Wayne described a scale to improve the accuracy of diagnosis of hyperthyroidism [8]. Wayne's index reports a score ranging from -25 to +45, where a score <11 is defined as "euthyroidism" and a score >19 suggests "toxic hyperthyroidism" [9]. Klein et al. [10] developed a hyperthyroidism symptom scale (HSS) and used it in 10 patients with untreated Graves' disease (GD). In that study, HSS was compared at three different phases: during hyperthyroidism at the time of diagnosis (phase 1), after 2 weeks of treatment with propranolol hydrochloride (phase 2), and in a euthyroid state after 6 months of treatment with propylthiouracil (phase 3). The mean scores decreased significantly between phases 1 and 2, paralleling the decreases in the symptoms and appeared to be sensitive to changes in both adrenergic stimulation and the metabolic effects of thyroid hormone as reflected by the sequential fall in scores between phases 2 and 3.

Because thyroid function tests (TFTs) are currently used to confirm thyrotoxicosis, the evaluation of the clinical status of patients using the HSS might be regarded as less important than previously [11]. However, it remains a helpful tool for diagnosis and assessing the response to treatment. Because of these advantages of the HSS, and because there is no Korean-language tool to evaluate the clinical status of patients with thyrotoxicosis, our aim was to develop a Korean version of the HSS (K-HSS) and evaluate its validity and reliability.

## METHODS

### HSS translation

The original version of HSS was initially translated into Korean by a Korean endocrinologist. Subsequently, this Korean version

was back-translated by a bilingual endocrinologist who was not familiar with the original version. The resulting revision was reviewed and amended by a translation committee consisting of three Korean endocrinologists. Thus, the final draft version was the linguistic equivalent of the original.

### Anthropometric and biochemical measurements

Height and weight were measured to the nearest 0.1 cm and 0.1 kg, respectively, with the subject in light clothing and without shoes. Body mass index was calculated as weight divided by the height squared (expressed in kilograms per square meter). Blood pressure and heart rate were measured on the right arm with the subject in a seated position. Serum glucose was measured using a Hitachi 747 chemistry analyzer (Hitachi, Tokyo, Japan). Serum total protein, albumin, total bilirubin, alkaline phosphatase, aspartate aminotransferase (AST), and alanine aminotransferase (ALT) were measured with an autoanalyzer (TBA-200FR, Toshiba, Tokyo, Japan). For TFT, concentrations of serum free thyroxine (T4; DiaSorin, Saluggia, Italy) and thyroid stimulating hormone (TSH; CIS Bio International, Gif-sur-Yvette, France) were measured using immunoradiometric assays. The free T4 assay had an analytical sensitivity of 0.05 ng/dL, while that for TSH had an analytical sensitivity of 0.04 mIU/L and a functional sensitivity of 0.07 mIU/L. The reference ranges for free T4 and TSH were 0.89 to 1.79 ng/dL and 0.3 to 4.0 mIU/L, respectively. Thyrotoxicosis was defined based on the results of the TFT: i.e., overt thyrotoxicosis was defined as high free T4 and low TSH, and subclinical thyrotoxicosis as normal free T4 and low TSH.

### Participants

Patients with thyrotoxicosis were recruited from the outpatient clinic of the endocrinology department at Seoul National University Bundang Hospital (SNUBH). We included 30 patients aged 15 to 60 years with newly diagnosed or recurrent thyrotoxicosis. Two patients were excluded from the study because of poor compliance, so only the data from the other 28 patients were used for the final analysis. Of these, 25 had GD and three had transient thyrotoxicosis due to thyroiditis. Patients with GD were prescribed a specific dose of antithyroid drugs (ATDs) as decided by the physician. Patients with transient thyrotoxicosis were reassured that their symptoms and signs were benign and self-limiting. Patients taking a  $\beta$ -blocker for control of symptoms such as palpitations and tremor were advised to reduce or discontinue the dose of  $\beta$ -blocker if the symptoms improved. Patients with thyrotoxicosis caused by a toxic nodular goiter

were also excluded because this condition requires treatment options other than ATDs, which have an inadequate therapeutic effect. Ten healthy volunteers who were checked to confirm that they had no history of thyroid disease were also included as a control group. All subjects participating were informed about the study and completed written informed consent. The study was approved by the Ethical Review Board of SNUBH (IRB No. B-1709/418-102).

### Procedures

Anthropometric data such as height, weight, blood pressure, and heart rate were measured at the first visit of those patients with newly detected or recurrent thyrotoxicosis who agreed to participate in this study. They also underwent blood tests including TFT and tests to identify the cause of the thyrotoxicosis (e.g., autoantibodies, thyroid scan). To assess clinical hyperthyroidism, the endocrinologist in charge of this study evaluated the patients using the K-HSS [12]. The information needed to rate each item was obtained by the endocrinologist through the history taking and physical examination. It was possible to recruit all patients from outpatient clinics of one endocrinologist with a relatively small study of 40 participants. Therefore, one endocrinologist, who was in charge of this study, was responsible for the evaluation of all patients. After 1 to 2 weeks, the patients visited the clinic to confirm the results of the previous tests and to start treatment based on the cause of their thyrotoxicosis. All patients were followed-up every 1 to 2 months, at which times they underwent blood tests including TFT. In the patients with GD, the ATD dose was adjusted as necessary. At each visit, anthropometric data were collected and vital signs were measured. The same endocrinologist evaluated the clinical thyrotoxicosis status using K-HSS.

Healthy volunteers were recruited as a control group through an official announcement in SNUBH. The controls visited the hospital on the same schedule as the patients with thyrotoxicosis. They also underwent the same blood tests and met the endocrinologist at each visit and participated in the evaluation of K-HSS.

### Assessment items in K-HSS

The original HSS gives scores from 0 to 40, where higher scores indicate more severe disease. It consists of a 10-item scale that rates nervousness, diaphoresis, heat intolerance, motor activity, tremor, weakness, hyperdynamic precordium, diarrhea, weight loss/appetite, and overall function. This structure was retained in the Korean version. The information needed to score each

item was obtained through history taking and physical examination. The intensity of each item can be rated on a subscale from 0 to 4 points and the item scores were summed to obtain the overall score.

### Statistical analysis

Data were expressed as mean  $\pm$  standard deviation (SD). To compare variables between the patient and control groups, we used the Mann-Whitney test (non-parametric data such as TSH, TSH receptor antibody [Ab] and K-HSS) or Student *t* test (for the remaining parametric data) for continuous variables and the chi-square test for categorical variables. The associations of free T4 or TSH levels and K-HSS scores at baseline were evaluated by ordinary least-square regression analysis. To compare the variables at baseline and follow-up, we used the Wilcoxon

**Table 1.** Baseline Characteristics of Participants

Characteristic	Control group (n=10)	Patient group (n=28)	P value
Age, yr	34.1 $\pm$ 5.9	34.9 $\pm$ 10.9	0.787
Male sex, %	30.0	35.7	0.744
BMI, kg/m <sup>2</sup>	20.7 $\pm$ 1.6	20.6 $\pm$ 4.7	0.932
SBP, mm Hg	124.6 $\pm$ 11.4	130.2 $\pm$ 14.3	0.274
DBP, mm Hg	70.6 $\pm$ 9.6	78.1 $\pm$ 10.6	0.057
HR, bpm	81.9 $\pm$ 14.8	101.6 $\pm$ 14.5	0.001
Free T4, ng/dL	1.4 $\pm$ 0.1	3.1 $\pm$ 1.1	<0.001
TSH, $\mu$ U/mL	1.3 (1.1–2.3)	0.0 (0.0–0.0)	<0.001
TSH receptor Ab, IU/L	0.0 (0.0–0.0)	9.5 (5.6–18.3)	<0.001
K-HSS	0.5 (0.0–2.8)	12.5 (8.0–18.0)	<0.001
Glucose, mg/dL	96.6 $\pm$ 23.1	105.8 $\pm$ 18.6	0.243
Total protein, g/dL	7.2 $\pm$ 0.4	7.1 $\pm$ 0.5	0.372
Albumin, g/dL	4.4 $\pm$ 0.3	4.3 $\pm$ 0.3	0.179
Total bilirubin, mg/dL	0.7 $\pm$ 0.3	0.7 $\pm$ 0.3	0.497
ALP, IU/L	53.2 $\pm$ 12.6	147.8 $\pm$ 160.9	0.074
AST, IU/L	18.3 $\pm$ 3.1	26.4 $\pm$ 10.4	0.001
ALT, IU/L	15.1 $\pm$ 7.9	33.1 $\pm$ 22.0	0.017
WBC, $\times 10^3/\mu$ L	5.8 $\pm$ 1.4	5.5 $\pm$ 1.7	0.610
Hemoglobin, g/dL	13.6 $\pm$ 1.4	14.0 $\pm$ 1.4	0.424
Platelets, $\times 10^3/\mu$ L	276.6 $\pm$ 59.1	243.0 $\pm$ 59.0	0.149

Values are expressed as mean  $\pm$  SD or median (interquartile range).

BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; HR, heart rate; T4, thyroxine; TSH, thyroid stimulating hormone; Ab, antibody; K-HSS, Korean version of the hyperthyroidism symptom scale; ALP, alkaline phosphatase; AST, aspartate aminotransferase; ALT, alanine aminotransferase; WBC, white blood cell.

signed rank test or paired *t* test. One-way analysis of variance or Kruskal-Wallis test was used to compare thyroid function and K-HSS scores between three groups: posttreatment patients who remained subclinically thyrotoxic, patients who became euthyroid after treatment and the control group at follow-up. Cronbach's  $\alpha$  coefficient was calculated to evaluate the internal consistency of K-HSS. A two-tailed *P* value of  $<0.05$  was considered significant. IBM SPSS for Windows version 22.0 (IBM Co., Armonk, NY, USA) was used for the data analyses.

## RESULTS

### Baseline characteristics

The baseline characteristics of the study subjects are displayed in Table 1. Mean age, sex ratio, and mean BMI were similar in both groups. As expected, the results of TFT including free T4, TSH, and TSH receptor Ab levels differed significantly between the groups. Although blood pressure did not differ significantly between the two groups, heart rate was higher in the patient group. AST and ALT levels were higher in the patient group but other laboratory results were similar for both groups.

### Correlation between K-HSS scores and baseline thyroid function test results

At the first visit, all subjects participating in this study underwent a TFT and completed the K-HSS on the same day. The data collected on this first day indicated that the K-HSS scores were positively associated with free T4 ( $B=4.31$ ; standard error [SE], 0.76;  $P<0.001$ ), and negatively associated with log TSH ( $B=-4.30$ ; SE, 0.84;  $P<0.001$ ) as calculated by ordinary least-square regression for all subjects. However, there was no significant linear correlation between free T4 or log TSH and K-HSS in thyrotoxic patients ( $B=1.91$ ; SE, 1.04;  $P=0.079$  or  $B=-0.01$ ; SE, 2.37;  $P=0.962$ , respectively). For triiodothyronine (T3), the analysis of linearity with K-HSS revealed similar results with that of free T4 in the thyrotoxic group ( $B=0.01$ ; SE, 0.02;  $P=0.618$ ).

### Changes in K-HSS scores after treatment of thyrotoxicosis

Patients were followed up for about 3 months ( $3.12\pm 0.96$ ). Six patients improved to a euthyroid state, 19 patients remained subclinically thyrotoxic, and the remaining three patients were still thyrotoxic according to the follow-up TFT results. The results of the comparison between baseline and follow-up/post-

**Table 2.** Comparison of Thyroid Function Test and K-HSS between Baseline and Posttreatment/Follow-up in Patient and Control Groups

	Patient group <sup>a</sup> (n=25)			Control group (n=10)			P value <sup>c</sup>
	Baseline	Posttreatment	P value <sup>b</sup>	Baseline	Follow-up	P value <sup>b</sup>	
Free T4	3.0±1.1	1.5±0.5	<0.001	1.4±0.1	1.4±0.1	NA	0.950
TSH	0.0 (0.0–0.0)	0.0 (0–0.4)	0.051	1.3 (1.1–2.3)	1.6 (1.0–3.2)	0.445	<0.001
K-HSS	12.0 (8.0–17.8)	4.0 (2.0–7.0)	<0.001	0.5 (0.0–2.8)	0.0 (0.0–2.0)	0.496	0.001

Values are expressed as mean±SD or median (interquartile range).

K-HSS, Korean version of the hyperthyroidism symptom scale; T4, thyroxine; TSH, thyroid stimulating hormone.

<sup>a</sup>Those whose serum free T4 returned to reference range were analyzed (n=25); <sup>b</sup>Baseline vs. posttreatment/follow-up; <sup>c</sup>Posttreatment patient group vs. control group at follow-up.

**Table 3.** Comparison of Thyroid Function Test and K-HSS between Subclinical Thyrotoxic and Euthyroid Posttreatment Patients, Euthyroid Posttreatment Patients and Control Group at Follow-up

	Posttreatment of patient group <sup>a</sup> (n=25)		Control group at follow-up (n=10)	P value <sup>b</sup>
	Subclinical thyrotoxic (n=19)	Euthyroid (n=6)		
Free T4	1.4±0.1	1.2±0.4	1.4±0.3	0.227
TSH	0.0 (0.0–0.0)	1.5 (0.5–3.9)	1.6 (1.0–3.2)	<0.001
K-HSS	4.0 (2.5–7.5)	3.0 (0.5–10.8)	0.0 (0.0–2.0)	0.004

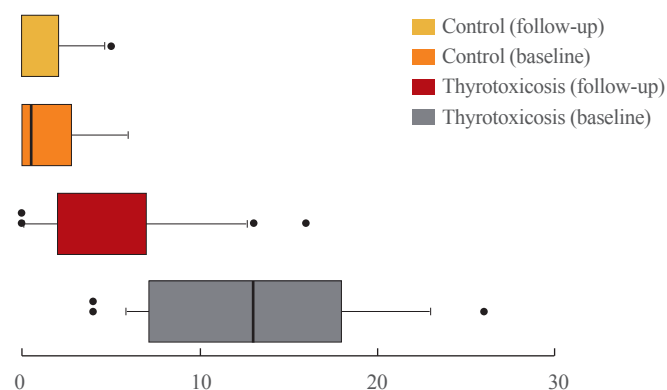
Values are expressed as mean±SD or median (interquartile range).

K-HSS, Korean version of the hyperthyroidism symptom scale; T4, thyroxine; TSH, thyroid stimulating hormone.

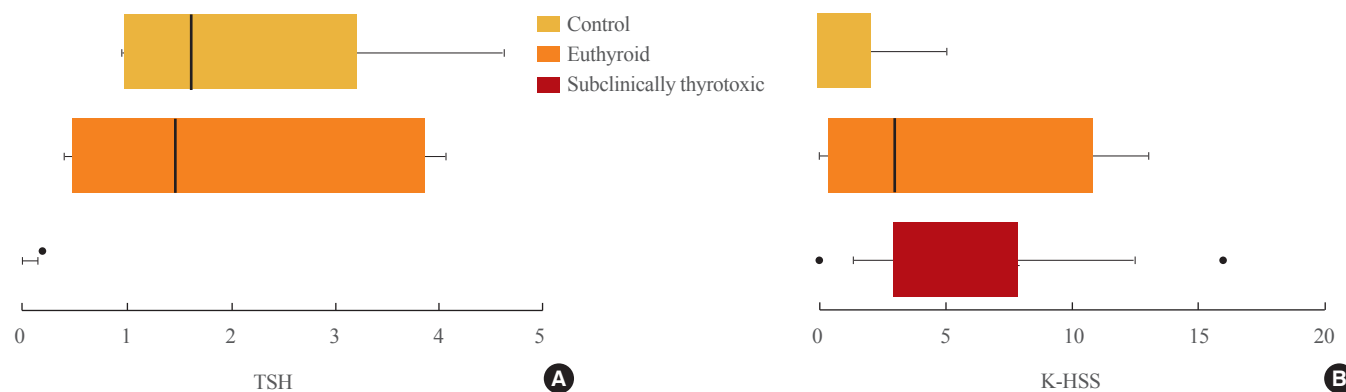
<sup>a</sup>Those whose serum free T4 returned to reference range were analyzed (n=25); <sup>b</sup>P value in one-way analysis of variance or Kruskal-Wallis test.

treatment for control and patient groups are summarized in Table 2. Because two patients who remained overtly thyrotoxic had not improved at all and one was in fact worse despite treatment with ATDs, only those patients whose serum free T4 returned to the reference range were analyzed (six euthyroid and 19 subclinically thyrotoxic). K-HSS scores at follow-up were significantly decreased from that at baseline when free T4 improved to within the reference range with ATD treatment. Free T4 did not differ significantly between the posttreatment patient group and the control group at follow-up, but TSH and K-HSS scores were significantly higher and lower, respectively, in the control group (Table 2). Within the posttreatment patient group, TSH differed significantly between subclinically thyrotoxic and euthyroid subjects, but the K-HSS scores showed no difference between these subgroups (Table 3).

The K-HSS scores for each group are graphed as boxplots in Fig. 1. The boxes represent the 10 to 90 percentile range of the distribution, the line in the middle of the box denotes the medi-



**Fig. 1.** Distribution of Korean version of the hyperthyroidism symptom scale.



**Fig. 2.** (A) Distribution of thyroid stimulating hormone (TSH) and (B) Korean version of the hyperthyroidism symptom scale (K-HSS) in subclinically thyrotoxic and euthyroid posttreatment patients and control group at follow-up.

an, and the circles indicate extreme outliers. As indicated, the distribution of K-HSS scores in the control group was almost the same at baseline and at follow-up. However, the distribution for the patient group differed between baseline and posttreatment. Fig. 2 shows the distribution of TSH and K-HSS scores in patients who were subclinically thyrotoxic or euthyroid posttreatment and in the control group at follow-up, and demonstrates that the distribution of TSH in patients who became euthyroid posttreatment was similar to that in the control group, but the K-HSS scores showed a wider distribution compared with that of the control group.

### Changes in each item score and reliability of K-HSS

Comparison of the changes in K-HSS scores between baseline and follow-up by item showed that the scores for nine of the 10 items were significantly decreased at follow-up in the patient group, but no significant changes were observed in any of the items in the control group (Table 4). The item that did not change significantly in the patient group were item 9, related to appetite and weight loss. The item with the greatest difference in scores was item 7 related to heart rate. Again, this analysis included the scores only of those subjects whose free T4 was improved to within the reference range. For the control group, some items could not be analyzed because they scored 0 points at both baseline and follow-up. For the whole 10 items of the K-HSS, Cronbach's  $\alpha$  coefficient for internal consistency was 0.860.

### Clinical feasibility

There was no difficulty in evaluating patients using K-HSS during the outpatient clinic. It took less than 5 minutes to evaluate the items through history taking and physical examination.

**Table 4.** Comparison of Scores at Baseline and Posttreatment/Follow-up for Each Item Included in the K-HSS

	Baseline	Posttreatment/Follow-up	P value
Patients group ( <i>n</i> =25)			
Total	12.0 (7.8–18.0)	3.0 (1.0–7.5)	<0.001
K-HSS, #1 (nervousness)	1.0 (0.0–2.0)	0.0 (0.0–1.0)	0.002
K-HSS, #2 (diaphoresis)	1.0 (0.0–2.0)	0.0 (0.0–2.0)	0.017
K-HSS, #3 (heat intolerance)	1.0 (1.0–2.5)	1.0 (0.0–1.0)	0.001
K-HSS, #4 (motor activity)	1.0 (0.0–2.0)	0.0 (0.0–1.0)	0.047
K-HSS, #5 (tremor)	2.0 (0.0–2.0)	0.0 (0.0–0.5)	0.002
K-HSS, #6 (weakness)	1.0 (1.0–2.5)	0.0 (0.0–0.5)	<0.001
K-HSS, #7 (hyperdynamic precordium)	2.0 (1.0–3.0)	0.0 (0.0–2.0)	0.001
K-HSS, #8 (diarrhea)	1.0 (0.0–1.0)	0.0 (0.0–0.0)	0.003
K-HSS, #9 (weight loss/appetite)	1.0 (0.0–2.5)	0.0 (0.0–1.0)	0.050
K-HSS, #10 (overall function)	1.0 (0.0–2.0)	0.0 (0.0–1.0)	0.002
Control group ( <i>n</i> =10)			
Total	0.5 (0.0–2.8)	0.0 (0.0–2.0)	0.496
K-HSS, #1 (nervousness)	0.0 (0.0–1.0)	0.0 (0.0–1.0)	1.000
K-HSS, #2 (diaphoresis)	0.0 (0.0–0.3)	0.0 (0.0–0.0)	0.157
K-HSS, #3 (heat intolerance)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	1.000
K-HSS, #4 (motor activity)	0.0 (0.0–0.3)	0.0 (0.0–0.3)	1.000
K-HSS, #5 (tremor)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	1.000
K-HSS, #6 (weakness)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	1.000
K-HSS, #7 (hyperdynamic precordium)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	1.000
K-HSS, #8 (diarrhea)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.317
K-HSS, #9 (weight loss/appetite)	0.0 (0.0–0.3)	0.0 (0.0–0.0)	0.317
K-HSS, #10 (overall function)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.317

Values are expressed as median (interquartile range).

K-HSS, Korean version of the hyperthyroidism symptom scale.

## DISCUSSION

In this study, K-HSS scores were positively associated with the baseline serum free T4 levels for all subjects, but not for thyrotoxic patients. For the patient group, the K-HSS scores decreased significantly after 3 months of observation or treatment with ATDs. In contrast, there were no significant changes in the scores for the control group even when compared item by item. The reliability of the K-HSS as assessed by Cronbach's  $\alpha$  was also confirmed to be high.

This study showed a significant positive correlation between serum free T4 levels and K-HSS scores in all subjects at their first visit. However, there was no significant correlation between them in patients group. There is little information available about the correlation between the degree of clinical thyro-

toxicosis and the results of TFT. One study that evaluated the severity of clinical thyrotoxicosis using the original HSS reported that in 25 patients with untreated GD the HSS was not associated with thyroid function assessed by free T4 index or serum T3 measured by radioimmunoassay [13]. Kolawole et al. [14] also reported that there was no significant association between laboratory T3, T4, and TSH values and HSS in 15 patients with newly diagnosed and untreated thyrotoxicosis. Therefore, the results for patient group from this study were consistent with previous reports which suggested the severity of peripheral endocrine symptoms, as measured by the K-HSS (HSS) was not linearly related to thyroid hormone levels such as T3 and free T4. According to the previous report, this might be due to potential variables such as cellular responsiveness to thyroid hormones or catecholamines on account of the difference in the de-

gree of expression of thyroid hormone receptor or the density of  $\beta$ -adrenergic receptor [13]. In addition, given that initial free T4 and K-HSS (HSS) scores were inevitably similar to each other in untreated patients with thyrotoxicosis, it could be one of the reasons for non-significant correlation between free T4 and K-HSS (HSS) that they did not include enough number of subjects to confirm the correlation between them. Although not shown in the results of this study, the correlation between free T4 and K-HSS before and after the treatment in thyrotoxic patients was significant ( $B=3.245$ , 95% confidence interval, 2.161 to 4.330;  $P<0.001$ ) (lineal model of generalized estimating equation analysis). In 28 patients in our study, thyroid function gradually improved with time or administration of ATDs. After approximately 3 months of observation or treatment, six patients were euthyroid and 19 patients were subclinically thyrotoxic. We excluded three patients who remained overtly thyrotoxic from the analysis because the purpose of this study was to analyze the changes in K-HSS scores accompanying an improvement in thyroid function induced with medication. The K-HSS scores decreased significantly with improved thyroid function. These results adequately validate the K-HSS and suggest that it could be useful for predicting the response to treatment in thyrotoxic patients. However, although there was no significant difference in free T4 between the posttreatment patients group and the control group at follow-up, K-HSS scores differed significantly between these groups. Considering that the majority of the posttreatment patient group were subclinically thyrotoxic, this could be interpreted to indicate that the K-HSS scores are related not only to free T4 levels but also to overall thyroid function. To evaluate this possibility, we divided the posttreatment patient group into two subgroups according to their thyroid function (subclinically thyrotoxic or euthyroid), and compared K-HSS scores between the subclinically thyrotoxic, euthyroid posttreatment patients and the controls at follow-up. The three patients who remained overtly thyrotoxic were excluded. There was no significant difference in the K-HSS scores of subclinically thyrotoxic and euthyroid subjects from the posttreatment patient group. As indicated by the box plots presented in Fig. 2, the euthyroid posttreatment patient group was relatively small ( $n=6$ ) and showed a broad distribution of K-HSS scores. Therefore, further evaluation is needed in a larger number of patients to clarify whether the K-HSS scores reflect free T4 as well as overall thyroid function.

The K-HSS is composed of 10 items, as is the original HSS, and most of these could be answered by patients after self-evaluation of their condition. Most of the items showed significant

differences between scores at baseline and follow-up in thyrotoxic patients, and only item 9 showed marginal significant difference ( $P=0.050$ ). This item concerns the degree of appetite and body weight, and could have been confusing because these variables are fairly subjective and the criteria for scoring were not clear. To confirm the internal consistency of K-HSS, we calculated Cronbach's  $\alpha$  coefficient [15]. There are different definitions of the acceptable values of  $\alpha$ , ranging from 0.70 to 0.95 [16]: a high value of  $\alpha$  ( $>0.90$ ) may suggest redundancies and indicate that the test should be shortened [17]. In the current study,  $\alpha$  was 0.86, high enough to confirm the reliability of the K-HSS.

To the best of our knowledge, this is the first report to validate a Korean-language scoring scale that evaluates the clinical status of thyrotoxicosis. Moreover, this study was a prospective study that followed up the changes in K-HSS scores after observation or administration of ATDs. Although a remarkable downgrading of the importance of clinical aspects of hyperthyroidism has occurred as implementation of TFT becomes more common [18], if a simple assessment of K-HSS scores could provide a clue to predict thyroid function it could be useful information for many aspects of the treatment of patients suspected of or diagnosed with thyrotoxicosis. For example, in patients with GD, their K-HSS scores could be used to estimate the response to ATDs and to prompt those patients who relapse after discontinuation of ATDs to visit the hospital as soon as possible. Furthermore, it could be used to estimate whether the dose of thyroid hormone was appropriate in patients with thyroid cancer undergoing TSH-suppressive therapy, who develop an iatrogenic subclinical thyrotoxicosis, to prevent the recurrence or progression of thyroid cancer with further studies on the thyroid cancer patients [19]. In addition to these clinical aspects, an index that could objectively assess the symptoms of thyrotoxicosis might be helpful for research.

This study had some limitations. The first was the failure to analyze changes after administration of  $\beta$ -blockers, which are commonly administered during thyrotoxicosis. This was because, according to the study protocol, the follow-up blood tests and K-HSS were performed 1 to 2 months after patients took the  $\beta$ -blocker, and the patients were trained to stop or reduce the dosage of the medication with symptom improvement. Next, the K-HSS was validated only in a single group recruited from a single center and assessed by one endocrinologist. As previously mentioned, there are some subjective items that could be given different scores by different endocrinologists, so analysis of multicenter patient groups may be required to confirm the re-

sults.

In conclusion, the K-HSS could be an easy-to-use tool for assessing the severity of thyrotoxicosis and the response to treatments such as ATDs. Based on the results from our study, the K-HSS could be used in research and clinical settings to enhance clinical understanding and decision-making, but not as the sole basis for making a clinical diagnosis.

## CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

## ACKNOWLEDGMENTS

This research was supported by a research grant (02-2016-035) from the Seoul National University Bundang Hospital to Jae Hoon Moon.

## AUTHOR CONTRIBUTIONS

Conception or design: J.H.M. Acquisition, analysis, or interpretation of data: J.H.M., J.E.L., K.M.K., T.J.O. Drafting the work or revising: J.H.M., J.E.L., D.H.L., S.H.C., S.L., Y.J.P., D.J.P., H.C.J. Final approval of the manuscript: J.H.M., J.E.L.

## ORCID

Jie-Eun Lee <https://orcid.org/0000-0002-1039-5769>  
 Tae Jung Oh <https://orcid.org/0000-0002-5078-6123>  
 Kyoung Min Kim <https://orcid.org/0000-0002-6515-9544>  
 Sung Hee Choi <https://orcid.org/0000-0003-0740-8116>  
 Soo Lim <https://orcid.org/0000-0002-4137-1671>  
 Young Joo Park <https://orcid.org/0000-0002-3671-6364>  
 Do Joon Park <https://orcid.org/0000-0001-9630-3839>  
 Hak Chul Jang <https://orcid.org/0000-0002-4188-6536>  
 Jae Hoon Moon <https://orcid.org/0000-0001-6327-0575>

## REFERENCES

- Devereaux D, Tewelde SZ. Hyperthyroidism and thyrotoxicosis. *Emerg Med Clin North Am* 2014;32:277-92.
- De Leo S, Lee SY, Braverman LE. Hyperthyroidism. *Lancet* 2016;388:906-18.
- Smith TJ, Hegedus L. Graves' disease. *N Engl J Med* 2016; 375:1552-65.
- Boelaert K, Torlinska B, Holder RL, Franklyn JA. Older subjects with hyperthyroidism present with a paucity of symptoms and signs: a large cross-sectional study. *J Clin Endocrinol Metab* 2010;95:2715-26.
- Goichot B, Caron P, Landron F, Bouee S. Clinical presentation of hyperthyroidism in a large representative sample of outpatients in France: relationships with age, aetiology and hormonal parameters. *Clin Endocrinol (Oxf)* 2016;84:445-51.
- Larsen PR. Thyroid-pituitary interaction: feedback regulation of thyrotropin secretion by thyroid hormones. *N Engl J Med* 1982;306:23-32.
- Crooks J, Murray IP, Wayne EJ. Statistical methods applied to the clinical diagnosis of thyrotoxicosis. *Q J Med* 1959;28: 211-34.
- Moura S, Almeida F, Nascimento F. Wayne's clinical index in the diagnosis of hyperthyroidism. *Arq Bras Endocrinol Metabol* 1965;14:41-6.
- Kalra S, Khandelwal SK, Goyal A. Clinical scoring scales in thyroidology: a compendium. *Indian J Endocrinol Metab* 2011;15(Suppl 2):S89-94.
- Klein I, Trzepacz PT, Roberts M, Levey GS. Symptom rating scale for assessing hyperthyroidism. *Arch Intern Med* 1988;148:387-90.
- Ladenson PW, Singer PA, Ain KB, Bagchi N, Bigos ST, Levy EG, et al. American Thyroid Association guidelines for detection of thyroid dysfunction. *Arch Intern Med* 2000; 160:1573-5.
- Geffner DL. Symptom rating scale for assessing hyperthyroidism. *Arch Intern Med* 1989;149:1466-7.
- Trzepacz PT, Klein I, Roberts M, Greenhouse J, Levey GS. Graves' disease: an analysis of thyroid hormone levels and hyperthyroid signs and symptoms. *Am J Med* 1989;87:558-61.
- Kolawole BA, Ikem RT, Lawal OO. Relationship between thyroid hormone levels and hyperthyroid signs and symptoms. *Niger J Clin Pract* 2002;5:29-31.
- Cortina JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol* 1993;78:98-104.
- Bland JM, Altman DG. Cronbach's alpha. *BMJ* 1997;314: 572.
- Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011;2:53-5.
- O'Reilly DS. Thyroid function tests-time for a reassessment. *BMJ* 2000;320:1332-4.
- Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel



SJ, Nikiforov YE, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid

Association guidelines Task Force on thyroid nodules and differentiated thyroid cancer. *Thyroid* 2016;26:1-133.