

# Evaluated Reference Intervals for Serum Free Thyroxine and Thyrotropin Using the Conventional Outlier Rejection Test without Regard to Presence of Thyroid Antibodies and Prevalence of Thyroid Dysfunction in Japanese Subjects

KYOKO TAKEDA\*, MASAOKI MISHIBA\*\*, HIDEKO SUGIURA\*, ATSUKO NAKAJIMA\*, MIKA KOHAMA\* AND SONOE HIRAMATSU\*\*\*

\*Clinical Laboratory Department, St. Luke's International Hospital, Akashi-cho 9-1, Chuo-ku, Tokyo 104-8560, Japan

\*\*Mitsubishi Chemical Medicine Co., 4-2-8, Shibaura, Minato-ku, Tokyo 105-0023, Japan

\*\*\*Center for Preventive Medicine, St. Luke's International Hospital, Akashi-cho 9-1, Chuo-ku, Tokyo 104-8560, Japan

**Abstract.** The determination of the reference intervals for serum free thyroxine (FT4) and thyrotropin (TSH) is usually based on central 95 percentile intervals using subjects without detectable antibodies against thyroid peroxidase (TPO) or thyroglobulin (Tg). However, some subjects with extreme data over reference intervals are generally included. The study objective was to evaluate the reference intervals for FT4 and TSH using different outlier tests. 1,007 Japanese subjects screened based on the National Academy of Clinical Biochemistry criteria in the United States participated in this study. Serum FT4, TSH, and TPO and Tg antibodies were measured in all subjects. To make appropriate reference intervals, the Smirnov-Grubbs' outlier test was taken for antibody-free subjects (Ab[-] S-G), and the conventional outlier rejection method rejecting the value out of  $\pm 3$  standard deviation was taken for antibody-free subjects (Ab[-] STD) and all subjects (AL STD), respectively. 12.8% of all subjects had either TPO or Tg antibodies in their serum. The 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles of reference intervals of serum FT4 (ng/dL) and TSH (mU/L) were 1.03~1.66 and 0.51~5.14 in (Ab[-] S-G), 1.03~1.65 and 0.51~4.57 in (Ab[-] STD) and 1.03~1.66 and 0.51~4.67 in (AL STD), respectively. FT4 in males were significantly and negatively correlated with age, and TSH was significantly and positively correlated with age ( $P < 0.00001$  and  $P < 0.00001$ , respectively). There was a significant difference between the sexes in FT4 ( $P < 0.00001$ ) but not in TSH. The prevalence of hypothyroidism, subclinical hypothyroidism, subclinical hyperthyroidism and hyperthyroidism were 0.2, 3.1, 2.3 and 0.3% (Ab[-] S-G), 0.3, 4.7, 2.3 and 0.3% (Ab[-] STD), and 0.3, 4.3, 2.3 and 0.3% (AL STD), respectively. This finding indicates that the conventional outlier rejection method is both convenient and appropriate to provide reference intervals for serum FT4 and TSH levels without regard to thyroid antibodies using large samples.

**Key words:** Free thyroxine, Thyrotropin, Smirnov-Grubbs' outlier test, Reference intervals

(Endocrine Journal 56: 1059-1066, 2009)

IT is important to use suitable reference intervals to assess thyroid function for patients with thyroid disease. The determination of the reference intervals has usually been based on central 95 percentile intervals [1], but other studies have introduced cut-off values to reduce the population without clinical or biochemi-

cal thyroid problems to a more normal sub-population [2]. The National Academy of Clinical Biochemistry (NACB) in the United States proposed that for the establishment of new TSH reference intervals, only euthyroid healthy volunteers be included, who should be free from detectable autoantibodies against thyroid peroxidase (TPOAb) or thyroglobulin (TgAb), any personal or family history of thyroid dysfunction and furthermore goiter, and be free of medications apart from estrogens [3]. There is a report making reference intervals for TSH and thyroid hormones to exclude subclinical thyroid disease by ultrasound [4] ad-

Received Apr. 27, 2009; Accepted Aug. 18, 2009 as K09E-123

Released online in J-STAGE as advance publication Sep. 9, 2009

Correspondence to: Kyoko TAKEDA, M.D., Ph.D., Clinical Laboratory Department, St. Luke's International Hospital, Akashi-cho 9-1, Chuo-ku, Tokyo 104-8560, Japan. E-mail: takedaky@luke.or.jp

ditionally according to NACB guidelines in the United States. However, the ultrasound procedure and/or the assay measuring anti-TPO and Tg immuno-reactivity was still not sensitive enough to detect occult thyroid dysfunction. Spencer also reported that TSH upper reference limits may be skewed by TPOAb-negative individuals with occult autoimmune thyroid dysfunction [5]. We explored whether there were a method to determine the reference intervals with the introduction of appropriate cut-off values excluding occult thyroid dysfunction.

Using serum without any thyroid antibodies we determined new reference intervals by the Smirnov-Grubbus' outlier test [6] and the conventional outlier rejection method [7]. Furthermore, we developed other reference intervals using serum without regard to thyroid antibodies to compare with the former reference intervals, examining whether the reference intervals made by this conventional method were applicable to evaluate thyroid function. Furthermore, we evaluated the prevalence of thyroid dysfunction in Japanese subjects using each method.

## Methods

One thousand and seven Japanese (487 males and 520 females), who had visited our health care center in 2005, were approved to participate in this study. Their average age was  $46.4 \pm 12.0$  years (male  $24 \sim 86$ ,  $48.5 \pm 12.6$  y. and female  $20 \sim 83$ ,  $44.4 \pm 11.1$  y.). Their count of white and red blood cells, hemoglobin, platelets and glycosylated hemoglobin were all within the reference range. Furthermore, serum level of total protein, albumin, BUN, creatinine, uric acid, total bilirubin, total cholesterol, low density cholesterol, triglyceride, aspartate aminotransferase, alanine aminotransferase,  $\gamma$ -glutamyl transpeptidase, alkaline phosphatase, lactate dehydrogenase, sodium, potassium, chloride, calcium, phosphorus, and glucose were also all within the reference range. They had neither goiter nor were taking thyroid medication, and no present or past history of thyroid disease based on NACB criteria in the United States. Furthermore, they had no present or past history of glomerular nephritis, nephrotic syndrome, renal failure, collagen disease, chronic hepatitis, liver cirrosis, broncheal asthma, brain tumor and depression in which the assay of FT4 and/or TSH might be influenced because of their disease

itself or pharmacological medication for it [8, 9].

Blood samples were drawn between 8:30 am  $\sim$  10:30 am. Immediately the serum FT4 and TSH concentrations were measured by Elecsys 2010 (Roche Diagnostics GmbH, Penzberg, Germany). The remaining serum samples were kept at  $-20^{\circ}\text{C}$  and titers of TPOAb and TgAb were measured by Elecsys A-Ag, A-TPO (Roche Diagnostics GmbH, Penzberg, Germany). Intra and interassay CVs ( $n=15$ ) of FT4 were 3.1% at 1.07 ng/dL and 3.2% at 2.0 ng/dL. Intra and interassay CVs ( $n=15$ ) of TSH were 2.2% at 0.51 mIU/L and 1.2% at 4.92 mIU/L. The functional sensitivity of FT4 and TSH assay were 0.03 ng/dL and 0.01 mIU/L [10], respectively. The value of FT4 in Elecsys was evaluated by that in Enzygum-Test FT4 which had been evaluated by free T4 by dialysis [11, 12]. Intra and interassay CVs ( $n=60$ ) of TPOAb were 24.4% at 12.4 IU/mL and those of TgAb were 6.3% at 49.8 IU/mL, respectively. Evaluations were deemed negative when the titer of serum TPO antibody was  $<16$  IU/mL or the titer of serum Tg antibody was  $<28$  IU/mL based on the comparison analysis of serum and tissues in patients with thyroid tumor [13]. In these cut-off values the sensitivity of TPOAb and TgAb were 92.3% and 96.6%, respectively [13].

The reference intervals were defined as within 95 percentiles. The median, 2.5 and 97.5 percentiles were calculated to generate possible reference limits for FT4 and TSH from the reference population [1, 14, 15]. For outlier rejection, we compared the Smirnov-Grubbus' outlier test ( $p<0.01$ ) and the conventional outlier rejection method in which sample values exceeding 3 standard deviations from the mean are rejected. The latter method was taken for subjects without any TPOAb and TgAb and all subjects without regard to TPOAb and TgAb. Furthermore, we evaluated the prevalence of thyroid dysfunction in all 1,007 subjects, in which the blood tests were within the reference range, and medical history was unremarkable.

Statistical analysis : Laboratory values are reported as mean  $\pm$  SD, or minimum, maximum, median and percentiles (2.5<sup>th</sup>, 97.5<sup>th</sup>), where appropriate. The data for FT4 exhibited an approximately normal distribution, while TSH values were skewed such that their logarithms were approximately normally distributed. Regressions were performed on log (TSH) vs. FT4, log(TSH) vs. age, and FT4 vs. age. Statistical significance was determined using Student's *t* test, F test and  $\chi^2$  test, with differences considered as significant at  $p<0.05$ .

The procedure of this study was approved by our hospital's ethics committee (2005-025). All participants gave their written informed consent.

## Results

### *Thyroid autoantibodies*

No serum TPO or Tg antibody was detected in 451 males and 427 females (Table 1). 12.8 % of all healthy subjects had detectable TPO or Tg antibody.

### *Reference intervals*

We compared the Smirnov-Grubbus' outlier test and the conventional  $\pm 3$  SD test. In the Smirnov-Grubbus' test, we used the 876 subjects who were antibody free; this group is denoted Ab[-] S-G. (See Table 1; subjects # 1 and 2 were excluded.) For the conventional  $\pm 3$  SD test, we used two different subgroups. The first included those who were antibody free, consisting of 857 subjects. (See Table 1; subjects #1-21 were excluded.) This subgroup is denoted Ab[-] STD. The second subgroup included all 1,007 subjects without regard to TPOAb and TgAb; 27 subjects were excluded (all except #5 and 21 in Table 2). This subgroup of 980 individuals is denoted AL STD.

The reference intervals (2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles) of serum FT4 (ng/dL) levels were 1.03 ~1.66 ( $n=876$ , Ab[-] S-G), 1.03 ~1.65 ( $n=857$ , Ab[-] STD) and 1.03 ~1.66 ( $n=980$ , AL STD), respectively (Table 3). The minimum value of serum FT4 levels after each outlier rejection was the same among the three methods. The maximum values of serum FT4 (ng/dL) levels among each outlier rejection method were similar: 1.85 (Ab[-] S-G), 1.77 (Ab[-] STD) and 1.80 (AL STD), respectively. The reference intervals (2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles) of serum TSH (mU/L) levels were 0.51~5.14 ( $n=876$ , Ab[-] S-G), 0.51~4.57 ( $n=857$ , Ab[-] STD) and 0.51~4.67 ( $n=980$ , AL STD), respectively (Table 4). The minimum values of serum TSH levels were the same among the three outlier rejection methods. However, the maximum values of serum TSH (mU/L) levels among the methods were different: 10.03 (Ab[-] S-G), 5.85 (Ab[-] STD) and 6.08 (AL STD), respectively. There was no significant difference of mean and SD between (AL STD) and (Ab[-] STD) by F test and Student's *t* test ( $P=0.382$  and  $P=0.680$ , respectively).

**Table 1.** Presence of TPOAb or TgAb in Japanese subjects.

	Male				Female		
	TgAb				TgAb		
	+	-	Total		+	-	Total
TPOAb	+	11	8	19	42	7	49
	-	17	451	468	44	427	471
	Total	28	459	487	86	434	520

**Table 2.** The list of outlier subjects.

No.	Age	Sex	FT4 (ng/dL)	TSH (mU/L)	TgAb (IU/mL)	TPO Ab (IU/mL)
1	38	M	2.12	0.02	15	<5.0
2	77	M	1.15	13.05	<10.0	<5.0
3	48	M	1.85	1.94	<10.0	5.7
4	34	F	1.81	1.29	<10.0	<5.0
5	30	M	1.80	0.56	<10.0	<5.0
6	50	M	1.47	10.03	<10.0	<5.0
7	66	F	1.24	8.99	<10.0	<5.0
8	54	M	1.15	8.39	<10.0	<5.0
9	57	M	1.40	8.36	<10.0	<5.0
10	28	M	1.27	7.51	<10.0	<5.0
11	40	F	1.41	7.04	<10.0	<5.0
12	62	M	1.63	6.93	<10.0	<5.0
13	44	F	1.15	6.91	<10.0	<5.0
14	51	M	1.25	6.56	<10.0	<5.0
15	30	F	1.39	6.45	<10.0	<5.0
16	51	F	1.19	6.41	<10.0	<5.0
17	36	F	1.41	6.34	<10.0	<5.0
18	38	F	1.22	6.33	<10.0	<5.0
19	53	F	1.18	6.30	<10.0	<5.0
20	38	F	1.26	6.25	10.2	<5.0
21	62	F	1.25	5.90	<10.0	<5.0
22	44	F	2.27	0.04	524.7	201.3
23	42	F	1.81	0.01	387.6	41.5
24	66	M	0.90	14.09	604.7	>600
25	37	F	0.86	8.20	571.6	5.6
26	68	M	1.25	6.73	107.5	<5.0
27	41	M	1.43	6.45	34.4	8.2
28	34	M	1.43	6.24	340.7	41.5
29	71	F	1.13	6.23	392.6	<5.0

**Table 3.** Serum FT4 (ng/dL) (2.5<sup>th</sup>, 50<sup>th</sup> and 97.5<sup>th</sup> percentiles, geometric mean and SD, and minimum and maximum) in the different groups for the total and the antibody-free, by three different methods (Ab[-] S-G, Ab[-]STD and ALSTD).

	n	2.5 <sup>th</sup> percentiles	median	97.5 <sup>th</sup> percentiles	Mean	SD	Minimum	Maximum
Total subjects	1007	1.00	1.32	1.64	1.32	0.16	0.85	2.27
Antibody free subjects	878	1.01	1.32	1.63	1.32	0.16	0.85	2.12
(Ab[-] S-G) group	876	1.03	1.3	1.66	1.32	0.16	0.85	1.85
(Ab[-] STD) group	857	1.03	1.3	1.65	1.32	0.15	0.85	1.77
(AL STD) group	980	1.03	1.3	1.66	1.32	0.16	0.85	1.80

Note: (Ab[-]S-G) Smirnov-Grubbs' outlier test for subjects without thyroid antibodies, (Ab[-]STD) standard method for subjects without antibodies and (AL STD) standard method for total subjects.

**Table 4.** Serum TSH (mU/L)(2.5<sup>th</sup>, 50<sup>th</sup> and 97.5<sup>th</sup> percentiles, geometric mean and SD, and minimum and maximum) in the different groups for the total and the antibody-free by three different methods (Ab[-] S-G, Ab[-] STD and AL STD).

	n	2.5 <sup>th</sup> percentiles	median	97.5 <sup>th</sup> percentiles	Mean	SD	Minimum	Maximum
Total subjects	1007	0.39	1.77	4.99	2.02	1.38	0.01	14.09
Antibody free subjects	878	0.46	1.68	5.15	1.97	1.30	0.02	13.05
(Ab[-] S-G) group	876	0.51	1.66	5.14	1.96	1.25	0.07	10.03
(Ab[-] STD) group	857	0.51	1.65	4.57	1.87	1.02	0.07	5.85
(AL STD) group	980	0.51	1.66	4.67	1.89	1.05	0.07	6.08

Note: (Ab[-] S-G) Smirnov-Grubbs' outlier test for subjects without thyroid antibodies, (Ab[-] STD) standard method for subjects without antibodies and (AL STD) standard method for total subjects.

### FT4 and TSH related age and sex

Serum FT4 and TSH concentrations ( $n=1,007$ ) were  $1.32 \pm 0.16$  ng/dL and  $2.02 \pm 1.38$  mU/L, respectively. The serum FT4 concentration in males was negatively correlated to age ( $y=1.51 - 0.0030 x$ ,  $R=-0.25$ ,  $P<0.000001$ ,  $n=449$ ) in (Ab[-] S-G) as well as (Ab[-] STD) and (AL STD) ( $P<0.000001$ ,  $n=441$  and  $P<0.000001$ ,  $n=474$ , respectively). The serum TSH concentration was significantly positively correlated to age ( $y=0.0070 x + 0.23$ ,  $R=0.16$ ,  $P<0.00001$ ,  $n=876$ ) in (Ab[-] S-G) as well as (Ab[-] STD) and (AL STD) ( $P<0.00001$ ,  $n=857$  and  $P<0.000001$ ,  $n=980$ , respectively). The serum TSH concentration separately in males and females was significantly positively correlated to age ( $y=0.0099 x + 0.10$ ,  $R=0.18$ ,  $P<0.0001$ ,  $n=449$  and  $y=0.0088 x + 0.20$ ,  $R=0.14$ ,  $P<0.01$ ,  $n=427$ , respectively) in (Ab[-] S-G) as well as (Ab[-] STD) ( $P<0.0001$ ,  $n=441$  and  $P<0.01$ ,  $n=416$ , respectively) and (AL STD) ( $P<0.0001$ ,  $n=474$  and  $P<0.01$ ,

$n=506$ , respectively). Mean serum FT4 concentration in males was significantly higher than in females ( $P<0.00001$ ) with (Ab[-] S-G) as well as (Ab[-] STD) and (AL STD). There was no significant difference in the serum TSH concentration between males and females in each method.

### Prevalence of thyroid dysfunction in all subjects

The prevalence of hypothyroidism and subclinical hypothyroidism was 0.2% and 3.1% (Ab[-] S-G), 0.3% and 4.7% (Ab[-] STD), and 0.3% and 4.3% (AL STD), respectively. The prevalence of subclinical hyperthyroidism and hyperthyroidism was 2.3% and 0.3% in each method, respectively. There was no significant difference of the prevalence of hypothyroidism and subclinical hypothyroidism and subclinical hyperthyroidism and hyperthyroidism among three methods ( $P=0.959$ ). Subjects with hypothyroidism in method (Ab[-] S-G) had all positive serum anti TPO or Tg an-

**Table 5.** Age and sex matched reference intervals for serum FT4 (ng/dL) and serum TSH (mU/L) (2.5, 50 and 97.5 percentiles, geometric mean and SD, and minimum and maximum).

	Age	n	2.5 <sup>th</sup> percentiles	50 <sup>th</sup> percentiles	97.5 <sup>th</sup> percentiles	Mean	SD	Minimum	Maximum
FT4	20s and 30s	129	1.11	1.39	1.70	1.40	0.15	1.05	1.77
	40s	125	1.16	1.36	1.66	1.38	0.13	1.04	1.75
	Male 50s	97	1.03	1.33	1.69	1.33	0.16	0.94	1.75
	60s	67	1.07	1.29	1.63	1.31	0.13	0.96	1.68
	Over 70s	23	1.00	1.29	1.53	1.27	0.12	0.99	1.54
	20s and 30s	178	1.02	1.26	1.59	1.28	0.15	1.00	1.68
	40s	125	0.99	1.25	1.53	1.26	0.14	0.85	1.58
	Female 50s	63	1.05	1.25	1.59	1.26	0.13	1.00	1.60
	60s	37	1.01	1.22	1.65	1.26	0.16	0.97	1.68
	Over 70s	13	1.05	1.24	1.46	1.24	0.10	1.05	1.46
TSH	20s and 30s	307	0.53	1.52	4.16	1.72	0.88	0.31	4.87
	40s	250	0.43	1.58	4.54	1.81	1.04	0.07	5.72
	50s	160	0.69	1.67	4.85	1.88	0.99	0.51	5.27
	60s	104	0.57	2.01	4.75	2.25	1.21	0.30	5.85
	Over 70s	36	0.75	2.17	5.37	2.38	1.19	0.65	5.63

tibodies. However, neither anti TPO or Tg antibodies were detectable in 0.1% of all cases with hypothyroidism in both method (Ab[-] STD) and (AL STD). Furthermore, neither anti TPO or Tg antibodies were detectable in 2.3% (Ab[-] S-G), 3.7% (Ab[-] STD) and 3.3% (AL STD) of all cases with subclinical hypothyroidism, respectively. In every method, 2.1% of all cases with subclinical hyperthyroidism and 0.1% of all cases with hyperthyroidism had neither anti TPO or Tg antibodies.

#### *Sex and age matched reference intervals*

As there was a significant difference between the sexes in the serum FT4 concentration and both the serum FT4 in males and the TSH concentration were significantly related to age, we computed other reference intervals regarding sex for FT4 and age for both FT4 and TSH after the conventional outlier rejection method ( $n=857$ ). The male and female FT4 reference intervals and TSH reference intervals in their 20s and 30s, their 40s, their 50s, their 60s and in the over 70s were shown on Table 5.

Using sex and age matched reference intervals, the prevalence of hypothyroidism, subclinical hypothyroidism, subclinical hyperthyroidism and hyperthyroidism were 0.2%, 4.4%, 2.3% and 0.3%, re-

spectively. There was no significant difference of the prevalence of hypothyroidism, subclinical hypothyroidism, subclinical hyperthyroidism and hyperthyroidism using sex and age matched reference intervals compared with (AL STD) ( $P=0.976$ ). Subjects with hypothyroidism all had positive serum anti TPO or Tg antibodies. However, neither anti TPO nor Tg antibodies were detectable in 3.6% of all cases with subclinical hypothyroidism, 2.0% of all cases with subclinical hyperthyroidism and 0.1% of all cases with hyperthyroidism.

## **Discussion**

There have been several population studies on FT4 and TSH reference intervals [2, 4, 16, 17, 18] although some large studies report thyroid hormone abnormalities [1, 19, 20]. There were some arguments in support of a narrower and optimal TSH reference range although the establishment of a more precise and true normal range for TSH was recognized to be important [21, 22, 23]. The mean TSH level and 95% reference intervals in our Japanese subjects was clearly higher than that in whites, blacks, Mexican Americans in the HANES III study in the United States [1] and Australians in the Busselton thyroid study [2] and



Germans [4, 16]. Additionally, the 2.5<sup>th</sup> percentile and 97.5<sup>th</sup> percentile of our TSH reference intervals were 0.1 and 1.0 mU/L higher than those of Kratzsch [4] using the same Roche analyser, respectively. This difference was more than the intra- and interassay variation [8]. Our upper limit of 4.57~5.14 mU/L is considerably higher than the currently used upper reference limit of 4.2 mU/L proposed by Roche Diagnostics [24] and the 97.5<sup>th</sup> percentile of 4.1 mU/L from the NHANES III study in the United States [1]. Thus, the differences in TSH reference intervals may be caused by the difference of sample number, age distribution, ethnicity and iodine intake and furthermore the exclusion criteria. We considered that it would be recommended to make the reference intervals for Japanese population when newly different equipment and reagents are available.

The most recommended strategy to make the reference intervals is to measure hormones of serum samples with negative thyroid antibodies [3]. We analyzed serum FT4 and TSH level in Japanese normal subjects with negative thyroid antibodies according to NACB guidelines in the United States. The Smirnov-Grubbus' outlier test is a common procedure to determine the cut-off values used to exclude the most typical exceeded results. The maximum value of the serum TSH level after the Smirnov-Grubbus' outlier rejection (Ab[-] S-G) was 10.03 mU/L, that was beyond 6 standard deviation and still a high evaluation for a healthy subject. As the Smirnov-Grubbus' outlier test is known to be useful to reject one sample among subjects, we used the conventional outlier test in which sample values out of  $\pm 3$  standard deviation were rejected. The maximum value of the serum TSH level after this outlier rejection was 5.85 mU/L that may be considered to be in more normal subjects. However, it is time consuming and potentially expensive to measure the serum thyroid antibodies in all samples each time to make reference intervals. We therefore tried to use the conventional outlier test without regard to thyroid autoantibodies (AL STD). The minimum value of reference intervals of FT4 and TSH were completely the same as in the other methods. The only difference was the maximum value of FT4 and TSH levels of reference intervals. There was no significant difference of mean and SD between (AL STD) and (Ab[-] STD) indicating that either (AL STD) or (Ab[-] STD) contains the same scattering diagram. So that, the maximum value of the serum TSH level after the conventional outlier rejection (AL STD) of 6.08 mU/L may be also acceptable because of locat-

ing between +3.5SD and +4SD as well as 5.85 mU/L in (Ab[-] STD).

Hypothyroidism was found in 3.3~5.0 % of Japanese subjects ( 0.2~0.3% clinical and 3.1~4.7% subclinical ) and hyperthyroidism in 2.6% (0.3% clinical and 2.3% subclinical). The prevalence of hypothyroidism in Japanese subjects was almost the same as in other reports. While the prevalence of hyperthyroidism was lower found compared with other reports [1, 17] although the lower reference interval of serum TSH in our study was higher than that in other reports. We may consider that the ethnic background and iodine intake [25] in Japanese that could suppress thyroid function may be a causative factor in the difference. The 0~3.8 % of subjects with hypothyroidism and 2.2% of those with hyperthyroidism did not have any thyroid antibodies in our study. Our data supported the previous report by Baloch that an increased odd-ratio for hypothyroidism was even seen in antibody-negative subjects [3]. It also advocated that the occult thyroid disease could not be completely excluded to make the reference intervals using antibody free subjects. Furthermore, there was no significant difference in the prevalence of thyroid dysfunction among (Ab[-] S-G), (Ab[-] STD) and (AL STD), suggesting the conventional outlier rejection (AL STD) could be enough appropriate to evaluate thyroid dysfunction.

We observed that the serum TSH level had a significantly positive relation with age and furthermore the serum FT4 level in males had a significantly negative relation to age as reported previously [1, 4]. We evaluated the reference intervals regarding the sex and age difference. Using sex and age matched reference intervals, the prevalence of hypothyroidism, subclinical hypothyroidism, subclinical hyperthyroidism and hyperthyroidism were not significantly differed from that using (AL STD). Subsequently other additional reference intervals regarding the sex and age difference were considered not to be more excelled in the evaluation of thyroid disease compared to the prevalent reference intervals without regard to the sex and age.

It was suggested that the method using the conventional outlier test without regard to thyroid autoantibodies (AL STD) would be a convenient and permissible procedure in making reference intervals when such large normal subjects as this study are available to be used. This method may be conducive to reduce the cost of medical procedures to measure thyroid autoantibodies.

In conclusion, we reported the investigation of Japanese subjects for thyroid antibodies and thyroid dysfunction. We proposed a strategy to provide new reference intervals using the conventional outlier rejection test without regard to thyroid antibodies for large data. This procedure may contribute to make the acceptable reference intervals even if the equipment and reagents used are different.

## Acknowledgements

The authors wish to thank Drs. K. Ichihara, Y. Tokuda, S. Takahashi and J. Butler for helpful comments and Mr. S. Chang, H. Umei, J. Nakayama, H. Satou and Y. Otsuka for their assistance on statistical analysis.

## References

- Hollowell JG, Staehling NW, Flanders WD, Hannon WH, Gunter EW, Spencer CA, Braverman LE. (2002) Serum TSH, T4, and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). *J Clin Endocrinol Metab* 87: 489-499.
- O'Leary PC, Feddema PH, Michelangelit VP, Leedman PJ, Chew GT, Knuiman M, Kaye J, Walsh JP. (2006) Investigations of thyroid hormones and antibodies based on a community health survey: the Busselton thyroid study. *Clin Endocrinol* 64: 97-104.
- Baloch Z, Carayon P, Conte-Devoix B, Demers LM, Feldt-Rasmussen U, Henry JF, LiVolsi VA, Niccoli-Sire P, Jhon R, Ruf J, Smyth PPA, Spencer CA, Stockigt JR. (2003) Laboratory support for the diagnosis and monitoring of thyroid disease. *Thyroid* 13: 3-126.
- Kratzsch J, Fiedler GM, Leichele A, Brugel M, Buchbinder S, Otto L, Sabri O, Matthes G, Thiery J. (2005) New reference intervals for thyrotropin and thyroid hormones based on National Academy of Clinical Biochemistry Criteria and regular ultrasonography of the thyroid. *Clin Chem* 51: 1480-1486.
- Spencer CA, Hollowell JG, Karzarosyan M, Braverman LE. (2007) National Health and Nutrition Examination Survey III thyroid - stimulating hormone (TSH) - thyroperoxidase antibody relationships demonstrate that TSH upper reference limits may be skewed by occult thyroid dysfunction. *J Clin Endocrinol Metab* 92: 4236-4240.
- De Muth JE, Demuth J. (2006) Basic statistics and pharmaceutical statistical applications. *CRC Press*, Boca Raton, Florida, p 638.
- Davies L, Gather U. (1993) The identification of multiple outliers. *J Am Stat Assoc* 88: 782-792.
- Thyroid disease manager. (2007) Endocrine education. <http://www.thyroidmanager.org> (accessed January 2005)
- Okabayashi T, Takeda K, Kawada M, Kubo Y, Nakamura S, Chikamori K, Terao N, Hashimoto K. (1996) Free thyroxine concentrations in serum measured by equilibrium dialysis in chronic renal failure. *Clin Chem* 42: 1616-1620.
- Neumeier D, Chan DW eds. 1998 Multicentre evaluation of the Boehringer Mannheim Elecsys<sup>R</sup> (2010) immunoassay system. *The Mid Euro J of Med Suppl* 3: 27-40.
- Ekens RP. (1990) Measurement of free hormones in blood. *Endocr Rev* 11: 5.
- Ekens RP, Ellis SM. (1975) The radioimmunoassay of free thyroid hormones in serum. In: Robbins J, Braverman LE(eds). *Thyroid research, Proceedings of the seventh international thyroid conference*. Boston. Amsterdam, *Excerpta Medica* 597.
- Morita S, Nishihara E, Kubo S, Hukada S, Hirokawa M, Amino N, Miyauchi A, Kuma S. (2006) Evaluation of anti thyroglobulin antibody kit "Elecsys Anti-Tg" and anti thyroid peroxidase antibody kit "Elecsys Anti-TPO"- comparison with cut off value using ROC curve and present method. *Jpn J Med Pharm Sci* 55: 775-782.
- PetitClerc C, Solberg HE. (1987) Approved recommendation (1987) on the theory of reference values; Part 2. Selection of individuals for the production of reference values. *J Clin Chem Biochem* 25: 639-644.
- Solberg HE. (1983) The theory of reference values. Part 5. Statistical treatment of collected reference values; determination of reference limits. *J Clin Chem Biochem* 21: 749-760.
- Hubl W, Schmieder J, Gladrow E, Demant T. (2002) Reference intervals for thyroid hormones on the Architect analyzer. *Clin Chem Lab Med* 40: 165-166.
- Pedersen IB, Laurberg P, Kundsén N, Jorgensen T, Perrild H, Ovesen L, Rasmussen LB. (2006) A population study of the association between thyroid autoantibodies in serum and abnormalities in thyroid function and structure. *Clin Endocrinol* 62: 713-720.
- Atzmon G, Barzilai N, Hollowell JG, Surks MI, Gabrieli I. (2009) Extreme longevity is associated with increased serum thyrotropin. *J Clin Endocrinol Metab* 94: 1251-1254.
- Kundsén N, Bulow I, Jorgensen T, Laurberg P, Ovesen L, Perrild H. (2000) Comparative study of thyroid

- function and types of thyroid dysfunction in two areas in Denmark with slightly different iodine status. *Euro J Endocrinol* 143: 485-491.
20. Aoki Y, Belin RM, Clickner R, Jeffries R, Phillips L, Mahaffey KR. (2007) Serum TSH and total T4 in the United States population and their association with participant characteristics: National Health and Nutrition Examination Survey (NHANES 1999-2002). *Thyroid* 17: 1211-1223.
21. Brabant G, Beck-Peccoz P, Jarzab B, Laurberg P, Origiazzi J, Szabolcs I, Weetman AP, Wiersinga WM. (2006) Is there a need to redefine the upper normal limit of TSH? *Eur J Endocrinol* 154: 633-637.
22. Surks MI, Goswami G, Daniels GH. (2005) Controversy in clinical endocrinology. The thyrotropin reference range should remain unchanged. *J Clin Endocrinol Metab* 90: 5489-5496.
23. Wartofsky L, Dickey RA. (2005) The evidence for a narrower thyrotropin reference range is compelling. *J Clin Endocrinol Metab* 90: 5483-5488.
24. RocheDiagnostics. (2004) Reference intervals for children and adults. In Elecsys thyroid tests. Mannheim, Germany: *Roche Diagnostics*, 5.
25. Tajiri J, Higashi K, Morita M, Umeda T, Sato T. (1986) Studies of hypothyroidism in patients with high iodine intake. *J Clin Endocrinol Metab* 63: 412-417.