NON-GAUSSIAN DISTRIBUTION OF SERUM THYROXINE LEVELS FOR EUTHYROID POPULATION

Valerie A. Brookeman

University of Florida, College of Medicine, Gainesville, Florida

The purpose of this study is to analyze the distribution of T_4I values for euthyroid subjects and to establish the euthyroid T_4I range that results in fewest errors of classification. A single series of 1,578 T_4I by CPB values, determined by one technician using a commercial kit in a consistent manner, is assigned to eight categories. The distribution of 1,355 euthyroid T_4I values is unequivocally skewed and not Gaussian. Hence the mean ±2 s.d. limits do not define the 95% euthyroid range. Comparison of our euthyroid distribution with the large euthyroid T_4 by CPB series reported some years ago by another laboratory indicates that both distributions are equivalent. For the combined common non-Gaussian population the euthyroid range giving maximum discrimination between hypothyroid, euthyroid, and hyperthyroid subjects is 3.0–9.2 μg% T_4I (corrected for alcohol extraction efficiency). General T_4 limits of mean −2 s.d. to mean +2.7 s.d. are proposed for use by any laboratory to define a euthyroid T_4 range. By using the correct euthyroid range, the T_4 by CPB test gives 97% accuracy in diagnosing hypothyroid, euthyroid, and hyperthyroid subjects with no additional subject information, and 99% with knowledge of pregnancy, nephrosis, and administration of Dilantin, estrogens, and androgens.

The determination of total serum thyroxine by competitive protein binding analysis (T_4 by CPB) was first reported in 1964 (*1*) and a clinical evaluation using a large number of patients was first published in 1966 (*2*). When a commercial T_4 kit, Tetrasorb-125, using the CPB method was introduced in 1968 (*3*), the nuclear medicine laboratory at the University of Florida College of Medicine began to use it for serum thyroxine determinations, and in 1970 it replaced the PBI test as a routine screening test for evaluation of thyroid function. The purpose of this paper is to report an analysis of the results of T_4 tests performed by our laboratory between November 1969 and December 1971.

MATERIALS AND METHODS

The Tetrasorb-125 T_4 commercial kits were used according to the procedure recommended by the manufacturer (*3*). The same technician analyzed all samples in a consistent manner, thus reducing variations due to individual differences in technique. The standard curve supplied with each different Tetrasorb lot was checked in duplicate at four points when that particular batch of ^{125}I-thyroxine binding globulin (TBG) was first used. Each time after that ^{125}I-TBG solution was used, the percent resin sponge uptake was checked in duplicate at one point (10 μg% T_4). Samples of serum, obtained from patients with suspected thyroid function abnormality, were refrigerated for 1 day or frozen up to 4 days until used. Each sample was defrosted only once, just before use.

Extraction efficiency. Extraction of T_4 from serum was performed primarily with absolute alcohol. Ninety-five percent ethanol and methanol were also used to a lesser extent. It has been shown (*4*) that mean extraction efficiencies of the three types of alcohol are the same. The mean T_4 extraction efficiency using absolute alcohol was 81% with a standard deviation of 3% when measured 26 times on two occasions, according to the method recommended by the manufacturer (*3*). Previously we reported 81% T_4 extraction efficiency for absolute alcohol with standard deviation 6% (*5*). The mean values of the two series of extraction efficiency determinations are insignificantly different at 0.05

Received Oct. 18, 1972; revision accepted Mar. 15, 1973.

For reprints contact: Valerie A. Brookeman, Radiology Dept., University of Florida College of Medicine, Gainesville, Fla. 32601.

TABLE 1. WITHIN-BATCH REPRODUCIBILITY (T_4I μg%)

	1969	1971	Total
No of duplicate samples	65	52	117
Range of differences between duplicates	0–1.6	0–1.8	0–1.8
Mean difference	0.31	0.36	0.33
Variance of differences σ^2	0.086	0.115	0.099
Standard deviation σ	0.29	0.34	0.31

probability level. All T_4 results were corrected for alcohol extraction efficiency and then multiplied by 0.653 to give thyroxine iodine (T_4I).

Reproducibility. Within-batch reproducibility of the test had previously been determined by our laboratory (*5*) and was repeated 14 months later by analyzing 52 patient samples in duplicate. The present and previous results of duplicate T_4I determinations are summarized in Table 1. The variances and means of the two series are insignificantly different at 0.05 probability level, and thus there has been no significant change of within-batch reproducibility of the T_4 kit as used by our laboratory. The two series may therefore be combined into one sample of 117 duplicate T_4I determinations with net mean difference between duplicates 0.33 μg% and net standard deviation 0.31 μg% (see statistics).

Quality control. Versatol control serum was analyzed in duplicate with each batch of patient sera to determine between-batch reproducibility of the T_4 test. Two hundred ten T_4I determinations were made with five different lots of Versatol and the results are summarized in Table 2. The standard deviation of between-batch measurements has been maintained over the 25-month period at 0.6 μg%, which is about 0.08 of the mean T_4I value for each Versatol lot. The mean deviation of measured T_4I values from the assigned value was +0.1 μg% (s.d. = 0.6 μg%).

Clinical material. The medical records of approximately 2,000 patients were reviewed at least 6 months after the T_4I determination. The T_4I values were assigned into euthyroid, hypothyroid, and hyperthyroid categories plus additional categories representing euthyroid patients who had received drugs or had conditions known to increase (estrogens, antiovulatory medication, pregnancy) or decrease (androgens, anabolic steroids, Dilantin, nephrosis) the T_4I estimate. All cases diagnosed as hypothyroid had low ^{131}I thyroid uptakes and with thyroid hormone therapy showed appropriate clinical response and increased T_4I values. All cases diagnosed as hyperthyroid had high ^{131}I thyroid uptakes and with antithyroid treatment (antithyroid drugs, ^{131}I therapy, or thyroid surgery) showed appropriate clinical response and decreased T_4I values. The T_4I values for patients with previously diagnosed hyperthyroidism who were being followed in our thyroid clinic on antithyroid medication or after ^{131}I therapy or surgery and patients with previously diagnosed hypothyroidism who were being followed in our thyroid clinic while on exogenous thyroid medication were excluded from the analysis. Patients with morphological or radionuclide scan abnormalities, i.e., solitary cold nodule, multinodular goiter, simple goiter, hyperfunctioning adenoma, chronic thyroiditis, but who were clinically euthyroid and had euthyroid ^{131}I uptake values, were included in the euthyroid category.

RESULTS

The breakdown of 1,578 T_4I values into eight categories is given in Table 3 with the range, mean, and standard deviation in each category. Individual T_4I values are plotted in Fig. 1 for all categories except euthyroid for which the frequency distribution is shown in Fig. 2.

Comparison of T_4I values for males and females. The sample of 1,355 euthyroid T_4I values was comprised of 388 males and 967 females. Since the euthyroid values do not follow a normal or Gaussian distribution (see below), the T_4I values for males and females were compared using the Kolmogorov-Smirnov test (see statistics), which compares the cumulative frequency distributions. The maximum difference between the two distributions was 0.077, which was less than the critical difference 0.082 at 0.05 probability level. Thus since the differences

TABLE 2. QUALITY CONTROL WITH VERSATOL SERUM

Assigned value* (μg%)	No. of determinations	Range (T_4I μg%)	Mean (T_4I μg%)	s.d. (T_4I μg%)	s.d. as fraction of mean
6.8	10	6.2–8.5	7.4	0.7	0.10
7.1	11	5.5–7.7	6.9	0.7	0.10
6.8	76	5.1–8.6	6.9	0.6	0.09
7.7	73	6.5–9.2	7.8	0.6	0.08
8.2	40	6.8–9.8	8.4	0.6	0.07

* PBI, corrected for 90% recovery (5).

TABLE 3. DISTRIBUTION OF T_4I VALUES FOR 1,578 SUBJECTS

Category	No. of subjects	T_4I (μg%) Min-max. range	Mean	s.d.
Euthyroid	1,355	2.4–10.3	5.7	1.4
Hypothyroid	47	0.3– 2.6	1.3	0.6
Hyperthyroid	32	8.5–16.8	13.0	2.4
Pregnancy	27	4.8–13.0	8.3	2.0
Estrogens	83	3.8–15.1	7.6	2.1
Nephrosis	6	2.8– 3.9	3.3	0.4
Dilantin	23	2.5– 5.9	4.1	0.9
Androgens	5	2.5– 4.8	4.1	0.9

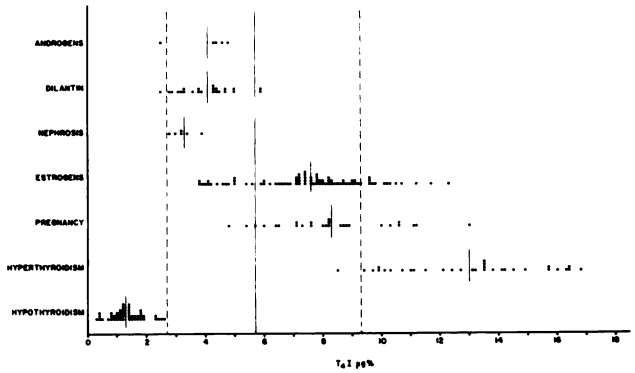

FIG. 1. Frequency distributions of T_4I values in seven categories. Mean T_4I value for each category is shown by short vertical line. Long vertical lines represent mean T_4I value for euthyroid subjects and "best" euthyroid T_4I limits (see text).

between them were not significant, the euthyroid male and female T_4I distributions were combined into a single population of 1,355 euthyroid T_4I values for further analysis.

Correlation with age. The age of the euthyroid subjects ranged from 1 week to 95 years. Determination of a linear correlation coefficient between age and the T_4I value assumes that the distributions are normal, which is not so (see below). Hence bivariate correlation analysis was performed by computing Spearman and Kendall rank-order correlation coefficients (*6*), which are nonparametric; that is, neither depends upon a normal distribution. Values of −0.07 and −0.05, respectively, were obtained, which are so small that, for practical considerations, negligible correlation may be assumed between T_4I value and age.

Serum T_4 may be elevated in the neonatal period, and mean T_4I values (corrected for extraction efficiency) of 11.9, 10.1, 9.5, 8.7, 7.4, and 8.7 μg% have been reported for ages 1–3 days, 1–2 weeks, 2–4 weeks, 1–4 months, 4–12 months, and for cord blood, respectively (*7*). The euthyroid sample reported in this paper included T_4I values for six subjects (0.4% of the total) less than 4 months old. The results of statistical analyses, as described below, of the euthyroid distribution were insignificantly different with or without these six values included in the sample.

Normality of euthyroid T_4I distribution. The frequency distribution of T_4I values for 1,355 euthyroid subjects (Fig. 2) was analyzed for normality by determining its skewness and kurtosis. These values and the mean, standard deviation, and (mean ±2 s.d.) limits of the distribution are given in Table 4. A normal or Gaussian distribution has skewness = 0 and kurtosis = 3. The positive skewness of +0.35 is significant, and hence the euthyroid T_4I values are not normally distributed. (A description of skewness and kurtosis and determination of their significance are outlined in statistics.) The logarithmic and square-root transformations of the euthyroid T_4I distribution are also given in Table 4. With logarithmic transformation the skewness has been reversed to −0.34, which is significant. The square-root transformation is insignificantly skewed but has kurtosis significantly different from 3. Thus neither of these simple transformations fit the euthyroid T_4I values to a normal distribution.

To eliminate any possibility that the nature of the distribution of euthyroid T_4I values and subsequent conclusions were due to any of the T_4I values outside the mean ±2 s.d. range being from other than confirmed euthyroid subjects, the medical records of the 37 and 12 euthyroid patients with T_4I values above 8.4 μg% and below 2.9 μg%, respectively, were reviewed again for confirmatory laboratory evidence that they were not borderline or mild

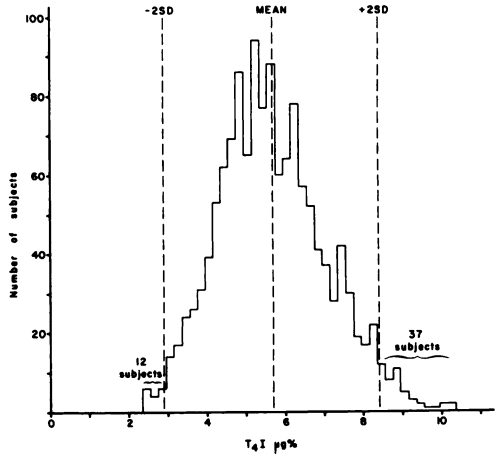

FIG. 2. Frequency distribution of T_4I values for 1,355 euthyroid subjects.

TABLE 4. DISTRIBUTION PARAMETERS OF 1,355 EUTHYROID T_4I VALUES AND OF TWO TRANSFORMATIONS

Distribution	Mean	s.d.	Mean ± 2 s.d. (μg% T_4I)	Skewness	Kurtosis
T_4I (μg%)	5.67	1.37	2.9–84	+0.35 (significant)*	2.96 (not significant)
$\log_{10}(T_4I)$	0.74	0.11	3.3–9.0	−0.34 (significant)	3.14 (not significant)
$\sqrt{T_4I}$	2.36	0.29	3.2–8.6	+0.03 (not significant)	2.60 (significant)

* At 0.05 probability level.

hyperthyroid or hypothyroid or did not have changes in TBG levels or conditions known to change TBG levels. Evidence accepted as confirming the euthyroid state included the following: (A) repeat T_4I value within the mean ±2 s.d. range of 2.9 — 8.4 μg%; (B) normal T_3 resin uptake test; (C) repeat normal T_4 value by column chromatography; (D) normal ^{131}I thyroid uptake and scan; and (E) normal TBG test. When this was done, 15 values above 8.4 μg% and 6 values below 2.9 μg% were retained. Those patients whose T_4I values were excluded because of the absence of any other thyroid function test result had signs or symptoms that either were explained by other diseases or were so few that the referring physician thought that further thyroid function tests were unnecessary. Complete statistical analyses were performed for this reduced sample of 1,327 values and even also for the euthyroid sample with *all* T_4I values >8.4 μg% and <2.9 μg% excluded. As expected, all statistical results and conclusions were negligibly affected. Skewness and/or kurtosis factors of the raw euthyroid distributions and of the logarithmic and square-root transformations were still significantly different from those of a Gaussian distribution, and mean ±2 s.d. ranges still proved inadequate for defining the best euthyroid range for discrimination between hypothyroid, euthyroid, and hyperthyroid subjects.

Euthyroid T_4I range. Since neither the raw euthyroid T_4I data nor the two simple transformations studied follow a normal distribution, the 95% range of the euthyroid population is *not* defined by two standard deviations from the mean. Inspection of T_4I values in Figs. 1 and 2 and Table 3 indicated that for our laboratory the euthyroid T_4I range giving the fewest misclassifications in hypothyroid, euthyroid, and hyperthyroid categories is 2.7–9.3 μg%. Only 16 euthyroids (1.2% of 1,355 euthyroids) lie outside this range (eight above and eight below); no hypothyroids and only one hyperthyroid (3.1% of 32 hyperthyroids) fall within this range. The lower limit, 2.7 μg%, is the mean −2.16 s.d., and the upper limit, 9.3 μg%, is the mean +2.65 s.d.; the asymmetry thereby allows for the positive skewness of the euthyroid T_4I distribution.

These chosen limits for the euthyroid T_4I range are shown in Fig. 1. Table 5 compares the numbers of T_4I values outside the range, mean ±2 s.d., and outside our optimum range, 2.7–9.3 μg%, for all categories except hypothyroid and hyperthyroid. The latter range reduces the number of misclassifications from 90 to 39. Table 6 compares the numbers of subjects outside both euthyroid T_4I ranges for all categories; the total number of misclassifications is reduced from 90 to 40 when the range 2.7–9.3 μg% is used. With no additional information regarding which subjects had conditions or diseases or were taking medications known to affect the T_4I value, the diagnostic accuracy of the T_4 by CPB test is 97.5% when the optimum euthyroid T_4I range, 2.7–9.3 μg%, is used, compared with 94.3% using the range, mean ±2 s.d. Knowledge of pregnancy or administration of estrogens, Dilantin and androgens reduces the number of misclassifications with the range, 2.7–9.3 μg%, to 17 (16 euthyroids and 1 hypothyroid—see Tables 5 and 6), which is 1.1% of the total number of subjects. Hence the T_4 by CPB test gives 98.9% accuracy in diagnosis of

TABLE 5. SUBJECTS IN SIX CATEGORIES WITH T_4I VALUES OUTSIDE TWO EUTHYROID T_4I RANGES

	2.9–8.4 μg% (mean ± 2 s.d.)		2.7–9.3 μg%	
Category*	No. below	No. above	No. below	No. above
Euthyroid	12	37	8	8
Pregnancy	0	11	0	7
Estrogens	0	26	0	14
Nephrosis	1	0	0	0
Dilantin	2	0	1	0
Androgens	1	0	1	0
Total number	16	74	10	29
	90		39	

* Excluding hypothyroid and hyperthyroid.

TABLE 6. NUMBER OF MISCLASSIFICATIONS ACCORDING TO TWO EUTHYROID T_4I RANGES

Category	Euthyroid T_4I range: 2.9–8.4 μg% (mean ± 2 s.d.)	2.7–9.3 μg%
Hypothyroid	0	0
Hyperthyroid	0	1
Other (from Table 5)	90	39
Total number misclassified	90	40
Total as % of all subjects	5.7%	2.5%
Diagnostic accuracy	94.3%	97.5%

TABLE 7. CORRECTED T_4I (μg%) VALUES BY CPB: EUTHYROID SUBJECTS

Author	No. of subjects	Min-max range	Mean	s.d.
Murphy, et al (2)	792	2.1–12.3	5.5	1.3
Nakajima, et al (11)	39	3.8– 7.8	5.7	1.1
Siersbaek-Nielsen (8)	49		5.9	1.5
Abbott Laboratories (3)	96		6.5	1.5
Sparagana, et al (10)	129		6.1	1.4
Kaihara, et al (12)	40	3.5– 8.4	5.9	1.2
Fitzgerald, et al (13)	221	2.9–12.7	6.9	1.9
Brookeman, et al (5)	115	2.9– 9.7	6.0	1.5
Braverman, et al (14)*	53	2.2– 7.4	5.0	1.0
This report	1,355	2.4–10.3	5.7	1.4

* Using Sephadex column; corrected for 97.9% extraction efficency (14).

hypothyroid and hyperthyroid subjects if the euthyroid range, mean ±2 s.d., is *not* used and a better range, accounting for the non-normality of the euthyroid T_4I distribution, is correctly chosen.

COMPARISON WITH LITERATURE

In agreement with our results, other laboratories have reported no statistical difference between T_4 values for euthyroid males and females (*8–10*) and no significant variation with age (*2,8*), except for the neonatal period.

Euthyroid T_4I distribution. Published single series of T_4I values for euthyroid subjects are given in Table 7, corrected for extraction efficiency. This series of 1,355 subjects is the largest to date. The series by Murphy, et al (*2*) of 792 euthyroid subjects is the second largest and the only other published series with sufficient T_4I values for comparable analysis with this euthyroid distribution. The publication of Murphy, et al (*2*) reports three consecutive series of T_4 by CPB determinations on a total of 1,146 euthyroid subjects. The largest single series containing 792 euthyroid subjects was compared with the sum of the three series using the Kolmogorov-Smirnov two-sample test (see statistics) to determine the validity of combining the three euthyroid samples into one large sample for statistical analysis. The maximum absolute difference between the two cumulative frequency distributions was 0.02 (at a T_4 value of 6.4 μg%), which was less than the critical difference, 0.06, at 0.05 significance level. Thus it was concluded that the combined sample of 1,146 T_4 values for euthyroid subjects was equivalent to the single sample of 792 values, and the skewness of the distribution and of the logarithmic and square-root transformations were determined using the published T_4 data (*2*). The results, given in Table 8, are compared with the skewness of this report's euthyroid T_4I distribution and both transformations and confirm that neither the T_4 by CPB values for euthyroid subjects nor the log transformation is normally distributed. Also given in Table 8 are the mean and standard deviation of the combined sample of 1,146 euthyroid T_4 values, corrected for the extraction efficiency (77%) of the ethanol used (*2*) and converted to T_4I. Finally, this sample of 1,355 euthyroid T_4I values was compared with the combined sample by Murphy, et al (*2*) of 1,146 euthyroid T_4I values using the Kolmogorov-Smirnov test. The cumulative frequency distributions are shown in Fig. 3. The maximum difference between them is 0.035, which is less than the critical difference, 0.055, at 0.05 level of significance. Hence it can be assumed that the two samples came from a common non-normal euthyroid population and are equivalent.

Best euthyroid T_4I range. Murphy, et al (*2*), recognizing that their euthyroid distribution was skewed, rejected a normal T_4 range of mean ±2 s.d. and arbitrarily adopted a normal uncorrected T_4 range of 4–11 μg%. Inspection of their published data (*2*) indicates that the uncorrected T_4 range giving fewest misclassifications of hypothyroid, euthyroid, and hyperthyroid subjects is 3.8–10.8 μg%, which is equivalent to 3.2–9.2 μg% T_4I, corrected for ethanol extraction efficiency. These limits are given in Table 8, expressed as the number of standard deviations below and above the mean T_4I value, and compared with equivalent results for this report's euthyroid T_4I data.

It is justifiable to consider a total sample of 2,501 euthyroid subjects (see Table 8) with mean T_4I value 5.6 μg%, standard deviation 1.3 μg%, and a best normal range of 3.0 to 9.2 μg%, and to propose that for maximum discrimination between hypothyroid, euthyroid, and hyperthyroid subjects, the normal T range (mean −2.0 s.d.) to (mean +2.7 s.d.) may be used by all laboratories.

TABLE 8. COMPARISON OF TWO EUTHYROID T_4I DISTRIBUTIONS

		Skewness			T_4I (μg%)		Best lower limit†		Best upper limit‡	
	No. of euthyroid subjects	T_4	$\log_{10} T_4$	$\sqrt{T_4}$	Mean	s.d.	T_4I (μg%)	No. of s.d. below mean	T_4I (μg%)	No. of s.d. above mean
This report	1,355 (1 series)	+0.35 (significant)*	−0.34 (significant)	+0.03 (not significant)	5.7	1.4	2.7	2.16	9.3	2.65
Murphy, et al (2)	1,146 (sum of three series)	+0.44 (significant)	−0.30 (significant)	+0.09 (not significant)	5.6	1.3	3.2	1.86	9.2	2.73
Average	2,501				5.6	1.3	3.0	2.0	9.2	2.7

* At 0.05 probability level.
† For maximum discrimination between euthyroid and hypothyroid subjects.
‡ For maximum discrimination between euthyroid and hyperthyroid subjects.

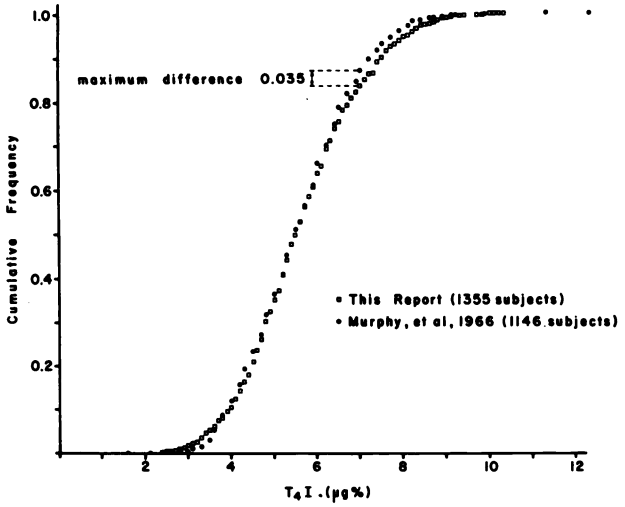

FIG. 3. Cumulative frequency distributions of two independent large series of T_4 by CPB measurements in euthyroid subjects. T_4 values published by Murphy, et al (2) were corrected for ethanol extraction efficiency (77%) and converted to T_4I.

Other categories. Tables 9 and 10 summarize those T_4I values by CPB values (corrected for extraction efficiency) published for hypothyroid, hyperthyroid, and pregnant subjects and for subjects with nephrosis or taking medications (estrogens or Dilantin) known to effect the T_4I value.

The 83 females in our estrogens category were either postmenopausal on Premarin (1.25–3.75 mg/day) or stilbestrol (0.25–0.50 mg/day) or premenopausal on oral contraceptives, the current commercial brands of which contain at least 50 μg estrogen. It has been shown (*18*) that 50 μg estrogen after 1 month of therapy results in definite elevation of the serum PBI value (by 1–2) and the T_4I-by-column value by about 1.4 μg% to 4.2–7.5 μg%. Our results are in agreement with those listed in Table 10 and with others (*2,9,19,20*); namely, that in pregnancy and with administration of estrogens, mean T_4 levels are increased because of increased TBG concentration.

It has been known since 1961 (*21*) that the administration of diphenylhydantoin (Dilantin) depressed the PBI concentration from a mean of 5.4 to 3.8 μg%. Our result of a low mean T_4I value for 23 subjects receiving Dilantin are in excellent agreement with those of Murphy, et al (*2*) (see Table 10). Six subjects with nephrosis had decreased T_4I values because of decreased thyroxine-binding proteins, in agreement with the results of Nakajima, et al (*11*) (see Table 10). Five patients on androgens had decreased T_4I values (see Table 3) because of decreased TBG concentration.

STATISTICS

All computations for statistical analyses of T_4I data were performed on an IBM 360 computer using a statistical software package (*22*).

Comparison of variances and means. For comparison of two sets of determinations of alcohol extraction efficiency or T_4I reproducibility, variances were first tested in pairs for significance using the F-distribution (*23*). Then the means were tested in pairs for significance using the t-test (student distribution). If the variances (σ_1^2 and σ_2^2) and means (x_1 and x_2) of the two samples are insignificantly different, as with the two sets of reproducibility measurements, then the two samples (sizes n_1 and n_2) come from a common population with mean $(n_1\bar{x}_1 + n_2\bar{x}_2)/n_1 + n_2)$.

The Kolmogorov-Smirnov two-sample test. This statistic (*24*) tests the null hypothesis H_0 that two

TABLE 9. CORRECTED T_4I (μg%) VALUES BY CPB: HYPOTHYROID AND HYPERTHYROID SUBJECTS

Author	Hypothyroid subjects				Hyperthyroid subjects			
	No. of subjects	Min-max range	Mean	s.d.	No. of subjects	Min-max range	Mean	s.d.
Murphy, et al (2)	67	0–3.4	1.8		35	8.3–>13.6	11.9	
Nakajima, et al (11)	7	1.4–3.4	2.3	0.7	16	8.4–17.1	10.9	2.3
Siersbaek-Nielsen (8)	24		1.6	0.8	100		13.1	3.5
Kennedy, et al (15)	12	0.1–3.0	1.2		19	6.6–17.0	11.3	
Cassidy, et al (16)	22	0–3.4	1.9		33	9.1–24.5	13.1	
Sparagana, et al (10)	9	0.5–3.3	1.8	0.8	14	10.3–33.4	17.7	6.2
Kaihara, et al (12)	15	0.5–3.8	2.5	0.9	32	7.4–23.8	10.5	2.9
Braverman, et al (14)*	14	0.2–1.5	0.6	0.5	21	7.6–17.2	12.0	3.0
This report	47	0.3–2.6	1.3	0.6	32	8.5–16.8	13.0	2.4

* Using Sephadex column; corrected for 97.9% extraction efficiency (14).

TABLE 10. CORRECTED T_4I (μg%) VALUES BY CPB: OTHER CATEGORIES

Author	Category	No. of subjects	Min-max range	Mean	s.d.
Nakajima, et al (11)	Pregnancy	16	5.2– 9.6	7.0	1.3
Ekins, et al (17)	Pregnancy	93		6.5	1.2
Braverman, et al (14)*	Pregnancy	10	4.8– 9.1	7.7	1.1
This report	Pregnancy	27	4.8–13.0	8.3	2.0
Ekins, et al (17)	Oral contraceptives	87		6.3	1.1
Braverman, et al (14)*	Estinyl	8	7.3– 9.7	8.6	0.9
This report	Estrogens	83	3.8–15.1	7.6	2.1
Murphy, et al (2)	Dilantin	22	2.2– 5.9	4.2	0.9
This report	Dilantin	23	2.5– 5.9	4.1	0.9
Nakajima, et al (11)	Nephrosis	4	1.3– 3.5	2.3	0.9
This report	Nephrosis	6	2.8– 3.9	3.3	0.4

* Using Sephadex column; corrected for 97.9% extraction efficiency (14).

independent samples were drawn from the same population. The two-tailed test uses the maximum absolute difference between the cumulative frequency distributions, regardless of direction, and is sensitive to any kind of difference in the distributions from which the two samples were drawn such as differences in location (central tendency), dispersion, and skewness. If the observed maximum difference is less than the critical difference $[(n_1 + n_2)/n_1n_2]^{1/2} \times 1.36$, for sample sizes n_1 and n_2, H_0 cannot be rejected at the 0.05 level of significance. Although this does not prove the identity of the two distributions, the differences between them are not significant.

Normality of a frequency distribution. To test for non-normality of a population frequency function, such as the euthyroid T_4I distribution, certain functions of the moments of the sample are calculated and the significance of their departure from the expected values for a normal population is examined. Departure of $\sqrt{b_1}$, $m_3/m_2^{3/2}$ from zero is an indication of skewness in the distribution, and departure of b_2, m_4/m_2^2, from 3 is an indication of kurtosis, where $m_r = \sum_{i=1}^{n} (x_i - \bar{x})^r/n$, $(r \geq 2)$ (25). A distribution, such as euthyroid T_4I values, with significant positive skewness is skewed to the right with high euthyroid T_4I values further from the mean than low euthyroid T_4I values. A distribution, such as the square root of euthyroid T_4I values, with kurtosis significantly less than 3, is platykurtic and has a flatter top and more abrupt tails than a normal distribution.

DISCUSSION AND CONCLUSIONS

The distributions of many physiological variables in normal subjects do not follow a Gaussian or normal distribution, but are skewed (26), and hence mean ± 2 s.d. limits do not define the central 95%. If the departure from normality is major, serious diagnostic misclassifications may arise from use of a normal range based on a falsely assumed Gaussian distribution.

Determination of serum T_4 by CPB has been established as the single in vitro screening test of

thyroid function that gives greatest diagnostic accuracy (*27*), and T_4 by CPB tests are performed routinely by many laboratories. The method of selection of a euthyroid T_4 range varies, and several laboratories have suggested various euthyroid T_4 limits, generally based on analyses of limited numbers of cases. Some laboratories take the mean ±2 s.d. as the limits of the central 95% range of an assumed Gaussian distribution of euthyroid T_4 values. Others, observing that their distribution of euthyroid T_4 by CPB values was skewed (*2,9,12,17*), arbitrarily selected by inspection those T_4 limits that could suitably distinguish between hypothyroid, euthyroid, and hyperthyroid subjects. However, such statistical and clinical analyses require large numbers of cases. This is not practical for the average nuclear medicine laboratory, and it is difficult to select a euthyroid T_4 range from the many published in the literature, particularly in view of the increasing numbers of commercial T_4 kits using different modifications to the basic T_4 by CPB procedure introduced in 1964 (*1*). Our laboratory formerly reported T_4 values based on a normal range of 2 s.d. from the mean. However it was apparent that some borderline values were being misclassified, and inspection of a small series of T_4I values (*5*) indicated a skewed euthyroid distribution.

To test the validity of assuming a normal distribution for T_4 by CPB values in euthyroid subjects, a large enough sample must be accumulated. To date, a full statistical analysis of a single series of T_4 values for a large number of euthyroid subjects has not been performed. An opportunity to carry out such a study was afforded our laboratory by the availability of T_4I values determined over a two-year period by one technician using one method in a consistent manner. We have performed a statistical analysis of our single series of 1,578 T_4 by CPB values including 1,355 euthyroids. The principal reason for the survey was the desire to establish diagnostic limits for the better interpretation of T_4 by CPB levels and to clarify the question of selection of a euthyroid T_4 range for routine clinical use by any laboratory.

Since the distribution of T_4I values for euthyroid subjects is unequivocally skewed, the 2 s.d. limits do not define the central 95% of the population and cannot properly be applied to determine a euthyroid T_4 range. Neither are the logarithmic nor square-root transformations normally distributed.

Comparison of our euthyroid distribution with a combined series of 1,146 euthyroid T_4 by CPB values (*2*), determined some years ago by a quite independent laboratory using nonidentical techniques, indicates that these two distributions are equivalent and independent of geographic location of the laboratory and method and technique of T_4 by CPB analysis. Hence a common non-normal euthyroid population is assumed. Although the means, standard deviations, and "best" euthyroid ranges of the two euthyroid T_4I distributions compared in Table 8 are almost identical, it is not suggested that the absolute average values given in Table 8 may be used by all laboratories determining T_4 by CPB since slight differences in method or technique could give rise to different values. However, the nature and shape of the distribution of T_4 values among euthyroid subjects should be the same for any laboratory, and a general formula for a euthyroid T_4 by CPB range is proposed (mean −2 s.d.) to (mean +2.7 s.d.), which accounts for the skewness of the euthyroid T_4 distribution, and may be used by any laboratory. It is suggested that each laboratory compile its own mean T_4 by CPB value and standard deviation for a group of euthyroid subjects and determine their euthyroid T_4 range from the above general formula.

For our laboratory the best euthyroid T_4I range by inspection of 1,578 T_4I values was 2.7–9.3 μg% and gave 97.5% diagnostic accuracy with no additional information (see Table 6). If the general formula (mean −2 s.d.) to (mean +2.7 s.d.) is applied to our mean T_4I value of 5.67 μg% and standard deviation of 1.37 μg%, a euthyroid T_4I range of 2.9–9.4 μg% is obtained. If this range were used, rather than 2.7–9.3 μg%, five additional subjects would be misclassified: two euthyroids, one hyperthyroid, one subject with nephrosis, and one subject on Dilantin (see Fig. 1 and Table 5). Now the total number of misclassifications, 45 out of 1,578 subjects, would represent an accuracy of diagnosis of hypothyroid and hyperthyroid subjects of 97.1%, with no additional information. However, knowledge of pregnancy and nephrosis and administration of estrogens, Dilantin, and androgens would reduce the number of misclassifications of euthyroid, hypothyroid, and hyperthyroid subjects to 21 and increase the diagnostic accuracy of the T_4 by CPB test as a single screening test of thyroid function to 98.7%.

The preceding general discussion about the selection and use of a normal T_4I range applies also to other in vitro thyroid function tests, such as the recently introduced effective thyroxine ratio (*28*) and normalized serum thyroxine level (*29*). The results of such tests should be similarly analyzed since the distribution of values for a euthyroid population may not be Gaussian, in which case the mean ±2 s.d. range does not define 95% of the euthyroid population, and an alternative appropriate normal range must be defined.

ACKNOWLEDGMENTS

The author is indebted to Clyde M. Williams, for reviewing the clinical records of the subjects included in this analysis, assigning T_4I values into the various categories, and for helpful discussions and critical review in the preparation of this manuscript. The author also thanks C. E. Davis, for advice and suggestions regarding the statistical analyses and Randy E. Roberts for performing all the T_4I determinations. This work was originally presented at the 19th Annual Meeting of the Society of Nuclear Medicine in July 1972.

REFERENCES

1. MURPHY BEP, PATTEE CJ: Determination of thyroxine utilizing the property of protein-binding. *J Clin Endocrinol Metab* 24: 187–196, 1964

2. MURPHY BEP, PATTEE CJ, GOLD A: Clinical evaluation of a new method for the determination of serum thyroxine. *J Clin Endocrinol Metab* 26: 247–256, 1966

3. *Tetrasorb™-125 Diagnostic Kit Instructional Manual.* Abbott Laboratories/Radiopharmaceuticals. North Chicago, Illinois, November, 1968

4. BROOKEMAN VA: T_4 extraction efficiencies of three alcohols. *J Nucl Med*: to be published

5. BROOKEMAN VA, WILLIAMS CM: Evaluation of the resin strip technique for determining serum T_3 binding capacity and serum thyroxine. *J Nucl Med* 12: 55–60, 1971

6. PEARSON ES, HARTLEY HO, eds: *Biometrika Tables for Statisticians,* vol 1, London, Cambridge University Press, 1966, pp 93–95

7. FISHER DA: Advances in the laboratory diagnosis of thyroid disease. Part 1. *J Pediatr* 82: 1–9, 1973

8. SIERSBAEK-NIELSEN K: Determination of serum thyroxine: the diagnostic value in thyroid diseases. *Acta Med Scand* 181: 327–333, 1967

9. HOWORTH PJN, MACLAGAN NF: Clinical application of serum-total thyroxine estimation, resin uptake, and free-thyroxine index. *Lancet* 1: 224–228, 1969

10. SPARAGANA M, PHILLIPS G, KUCERA L: Serum thyroxine by competitive protein binding analysis: clinical, statistical and comparative evaluation. *J Clin Endocrinol Metab* 29: 191–199, 1969

11. NAKAJIMA H, KURAMOCHI M, HORIGUCHI T, et al: A new and simple method for the determination of thyroxine in serum. *J Clin Endocrinol Metab* 26: 99–103, 1966

12. KAIHARA S, CARULLI N, WAGNER HN: Comparison of radioisotopic and column chromatographic assay of serum thyroxine. *J Nucl Med* 10: 281–283, 1969

13. FITZGERALD LT, BRUNO FP, GLASSMAN A, et al: Serum thyroxine determination by competitive protein binding analysis: the normal range. *J Nucl Med* 11: 669–673, 1970

14. BRAVERMAN LE, VAGENAKIS AG, FOSTER AE, et al: Evaluation of a simplified technique for the specific measurement of serum thyroxine concentration. *J Clin Endocrinol Metab* 32: 497–502, 1971

15. KENNEDY JA, ABELSON DM: Determination of serum thyroxine using a resin sponge technique. *J Clin Pathol* 20: 89–94, 1967

16. CASSIDY CE, BENOTTI J, PENO S: Clinical evaluation of the determination of thyroxine iodine. *J Clin Endocrinol Metab* 28: 420–421, 1968

17. EKINS RP, WILLIAMS ES, ELLIS S: The sensitive and precise measurement of serum thyroxine by saturation analysis (competitive protein binding assay). *Clin Biochem* 2: 253–288, 1969

18. MISHELL DR, COLODNY SZ, SWANSON LA: The effect of an oral contraceptive on tests of thyroid function. *Fertil Steril* 20: 335–339, 1969

19. GIMLETTE TMD, PIFFANELLI A: Assessment of thyroid status in pregnant women and in patients taking oral contraceptives by a free thyroxine index. *J Clin Pathol* 21: 767–770, 1968

20. GODWIN ID, SWOOPE HB: Comparison of a T_4 resin sponge uptake method with a protein-bound iodine procedure. *Am J Clin Pathol* 50: 194–197, 1968

21. OPPENHEIMER JH, FISHER LV, NELSON KM, et al: Depression of the serum protein-bound iodine level by diphenylhydantoin. *J Clin Endocrinol Metab* 21: 252–262, 1961

22. NIE N, BENT DH, HULL CH: *Statistical Package for the Social Sciences,* New York, McGraw-Hill, 1970, pp 102–109, 153–156

23. FISHER RA: On a distribution yielding the error function of several well-known statistics. *Proc Int Math Cong* (Toronto) 2: 805–813, 1924

24. SIEGEL S: *Nonparametric Statistics for the Behavioral Sciences.* New York, McGraw-Hill, 1956, pp 127–136

25. PEARSON ES, HARTLEY HO, eds: *Biometrika Tables for Statisticians,* vol 1, London, Cambridge University Press, 1966, pp 67–69

26. ELVEBACK LR, GUILLIER CL, KEATING FR: Health, normality and the ghost of Gauss. *JAMA* 211: 69–75, 1970

27. MURPHY BEP: In vitro tests of thyroid function. *Sem Nucl Med* 1: 301–315, 1971

28. MINCEY EK, THORSON SC, BROWN JL: A new in-vitro blood test for determining thyroid status—the effective thyroxine ratio. *Clin Biochem* 4: 286–291, 1971

29. ASHKAR FS, BEZJIAN AA: Use of normalized serum thyroxine (T_4N). *JAMA* 221: 1483–1485, 1972