Narrow Individual Variations in Serum T4 and T3 in Normal Subjects: A Clue to the Understanding of Subclinical Thyroid Disease

STIG ANDERSEN, KLAUS MICHAEL PEDERSEN, NIELS HENRIK BRUUN, AND PETER LAURBERG

Department of Endocrinology (S.A., N.H.B., P.L.) and Clinical Biochemistry (K.M.P.), Aalborg Hospital, Aalborg, Denmark

High individuality causes laboratory reference ranges to be insensitive to changes in test results that are significant for the individual. We undertook a longitudinal study of variation in thyroid function tests in 16 healthy men with monthly sampling for 12 months using standard procedures. We measured serum T4, T3, free T4 index, and TSH. All individuals had different variations of thyroid function tests (P < 0.001 for all variables) around individual mean values (set points) (P < 0.001 for all variables). The width of the individual 95% confidence intervals were approximately half that of the group for all variables. Accordingly, the index of individuality was low: T4; 0.58; T3; 0.54; free T4 index; 0.59; TSH; 0.49. One test result described the individual set point with a precision of plus or minus 25% for T4, T3, free T4 index, and plus or minus 50% for TSH. The differences required to be 95% confident of significant changes in repeated testing were (average, range); T4; 28, 11–62 nmol/liter; T3; 0.55, 0.3–0.9 nmol/liter; free T4 index; 33, 15–61 nmol/liter; TSH; 0.75, 0.2–1.6 mU/liter. Our data indicate that each individual had a unique thyroid function. The individual reference ranges for test results were narrow, compared with group reference ranges used to develop laboratory reference ranges. Accordingly, a test result within laboratory reference limits is not necessarily normal for an individual. Because serum TSH responds with logarithmically amplified variation to minor changes in serum T4 and T3, abnormal serum TSH may indicate that serum T4 and T3 are not normal for an individual. A condition with abnormal serum TSH but with serum T4 and T3 within laboratory reference ranges is labeled subclinical thyroid disease. Our data indicate that the distinction between subclinical and overt thyroid disease (abnormal serum TSH and abnormal T4 and/or T3) is somewhat arbitrary. For the same degree of thyroid function abnormality, the diagnosis depends to a considerable extent on the position of the patient’s normal set point for T4 and T3 within the laboratory reference range. (J Clin Endocrinol Metab 87: 1068–1072, 2002)

OVER ABNORMALITIES IN thyroid function are common endocrine disorders affecting 5–10% of individuals over a lifespan (1). Clinical symptoms and signs are often nonspecific, and the diagnosis and monitoring of therapy depends crucially on measurements of thyroid hormones and TSH in blood (2, 3).

Minor abnormalities in thyroid function with subclinical hypothyroidism or hyperthyroidism are even more common (1, 4, 5). Both subclinical hypothyroidism and hyperthyroidism are associated with an increase in risk of disease (4–6) as well as abnormalities in biochemical and physiologic measures that are often abnormal in patients with overt thyroid disease (6–9). Still it is debated to what extent subclinical thyroid disease should be treated (6, 10–12).

Subclinical thyroid disease is defined by high- or low-serum TSH with T4 and T3 within laboratory reference ranges. An important issue is whether serum concentrations of T4 and T3 are normal for the individual in subclinical thyroid disease.

Population-based reference ranges for serum T4 and T3 are quite wide because of large differences in thyroid function tests among normal subjects. These differences are caused by analytical and biological variation (13). Biological variation consists of between-individual and within-individual variation, the latter being characterized by rhythmic aberrations of multiple frequencies ranging from 30 min to 365 d (14, 15).

In general, population-based reference ranges are of limited value for interpretation of measurements in the individual if variation within individuals is small, compared with variation between individuals (16, 17). This may cause test results that are abnormal for the individual to be unnoticed within the wide group reference range.

Previous studies of the biological variation in thyroid function tests were performed under standardized conditions over shorter periods of time. Hence, the clinical importance of variation in thyroid function tests remains to be clarified (18–20).

We estimated biological [i.e. within (intra-) and between (inter-) individual] variation in thyroid function tests over a 12-month period in a group of healthy men in a routine laboratory setting. This was used to assess the utility of population-based reference ranges for serum TSH, total T3, total T4, and free T4 index (FTI). In addition, we calculated the number of tests needed to estimate the individual set point with certain levels of confidence and the difference required between two hormone values for significance in repeated testing of an individual.

Subjects and Methods

Subjects and protocol

Sixteen healthy Caucasian men participated. Median age was 38 yr, range 24–52 yr. Five were nonsmokers and 11 were smokers (5–25 cigarettes/d). They had an average body mass index of 25.4 kg/m2, range 21.3–30.9. Participant number 7 turned out to be subclinical hy-
perthyroid with permanently suppressed serum TSH and normal T₄, T₃, and FTI in serum. He was excluded from calculations. Of the remaining 15, none had clinical goiter or known previous or present thyroid disorders. None took any medication. They lived in Jutland, Denmark, where the iodine intake is moderately low (21), and their median urinary iodine excretion in spot urine samples, collected monthly for 12 months, was 50 μg/liter (22). We made no restrictions to their daily or yearly routines. Approval by the Regional Ethics Committee was obtained before the commencement of this study.

Blood samples were collected monthly for 12 months. Venepuncture was done between 0900 and 1200 h using standard procedures. Whole blood was allowed to clot and serum was separated and stored at −20°C until analysis.

**Assays**

Serum TSH was determined using immunochromiluminometric technology and a third-generation assay (LUMTest, Brahms, Berlin, Germany) with an intraassy coefficient of variation (CV) in our laboratory of 2.3% (TSH measured in duplicate in sera from 87 healthy subjects, level 1.5 mU/liter). Serum total T₃ (CV 2.2% in duplicate measurements of 25 sera from healthy subjects, level 2.3 nmol/liter) and serum total T₄ (CV 2.4% in duplicate measurements of 108 sera at 109 nmol/liter) were determined by radioimmunoassays (Amerlex-M T₃ RIA kit and Amerlex-M T₄ RIA kit, Johnson & Johnson, Cardiff, UK). T₃ uptake for calculation of FTI (CV 4.0% in 91 normal samples) was measured using an assay from Farmos Diagnostica (Oulunsalo, Finland). Single determinations were performed, and all samples from a subject were analyzed in random order in the same assay to eliminate interassay variation. Analytical error was determined for intraassay variation as described. Also, no attempt was made to detect or exclude outliers. Procedures were designed to picture the standard procedures at our laboratory.

**Statistical analysis**

Serum TSH, T₃, T₄, and FTI were normally distributed in all individuals (in the individual participants P = 0.3–1.0 for TSH; 0.2–1.0 for T₃, and 0.6–1.0 for FTI) and mild to moderately positively skewed in the group (P for normally distributed = 0.08, 0.24, 0.26, 0.02, respectively) as evaluated by the Shapiro-Wilk test and normality plots. As recommended, none of the variables were transformed (17). Samples were assumed independent and the results were compared using Bartlett’s test for homogeneity of variance and Kruskal-Wallis test. A P value of less than 0.05 was considered significant.

The individuality of reference ranges was estimated from the individuality index (16, 17). This index of individuality was calculated from SDanalytical + intraindividual/SDintraindividual (17) where the average SDanalytical + intraindividual was computed as the square root of the mean of the individual variances and the SDintraindividual as the estimated variance between means (17).

The number of tests needed to determine the homeostatic set point for the individual was estimated from n = (Z × CVanalytical + intraindividual/ D)² where n was the number of specimens obtained, Z was 1.96 (i.e. 95% confidence interval), and D was the percentage closeness to the homeostatic set point (17).

The difference required for significance in serial testing was calculated from 1.96 × √ n × SDanalytical + SDintraindividual².

Statistical analyses were performed using the statistical package for the social sciences version 10.0 and Corel QuattroPro8 (SPSS, Inc., Chicago, IL).

**Results**

Figure 1 illustrates within-individual and between-individual differences in serum TSH, serum T₃, serum T₄, and serum FTI over 1 yr. For each variable large variations were seen both in individuals and within individuals. Also, considerable differences in the variations within individuals were seen. This was independent of position of mean serum TSH, serum T₃, serum T₄, and serum FTI in the reference range. When testing for this, homogeneity of variance was not present between individuals (P < 0.001 for all variables) indicating an unpredictable difference in variation among individuals. Also, the participants had different set points (P < 0.001 for all variables). Hence, each subject’s thyroid function was unique, as evaluated from hormone concentrations in serum.

The relationship between serum T₃, T₄, and FTI levels and serum TSH was examined in each individual. The associations were marginally positive (mean Spearman’s rho with 95% confidence intervals: TSH vs. T₃ 0.14 (−0.43 to 0.72), TSH vs. T₄ 0.13 (−0.56 to 0.81), and TSH vs. FTI 0.18 (−0.47 to 0.83)).

Participant number 7 turned out to be subclinically hyperthyroid with permanently suppressed TSH. He was excluded from calculations. However, the measured values were included in Table 1.

Means and individual 95% reference ranges for each of the 16 participants are listed in Table 1. Also, means and individual reference ranges for an average participant as well as means and group reference ranges for participants 1 through 6 and 8 through 16, and the standard laboratory reference ranges are presented in Table 1. In addition, the individuality index is shown. The width of the individual participants reference ranges was about half of that of the group for all thyroid function tests, indicating that test results for an individual varied within 50% of the distribution for the group. Accordingly, the individuality index was below the critical limit of 0.6 (16, 23) for all thyroid function tests demonstrating that reference ranges based on the group are insensitive to significant changes in the individual.

Figure 2 illustrates the distribution of serum T₄ in one individual, compared with the distribution of measurements of T₄ in the group. In this individual serum T₄ may be substantially above the individual reference range and still lie well within the group reference range.

Table 2 shows the number of tests required to obtain a 95% confident estimate of the homeostatic set point for TSH, T₃, T₄, and FTI in an individual within plus or minus 5%, plus or minus 10%, and plus or minus 25% variation. Intraindividual variations caused a relatively high number of tests needed to describe the individual set points.

Table 3 shows the difference required in serum TSH, T₃, T₄, and FTI to be 95% and 99% confident of real changes in repeated testing. These changes are significantly larger than the normal spontaneous variation observed in the individual. Because of differences in variations among individuals, it is difficult to make general recommendations for evaluation of significant changes in repeated testing.

**Discussion**

Population-based reference ranges include both within- and between-individual variation. The value of reference ranges diminishes if between-individual variation exceeds within-individual variation. Harris (16, 23) explored this relationship computing the individuality index (16) and stated that population-based reference intervals are insensitive to real changes in individuals when the index of individuality for the test is less than 0.6. An index greater than 1.4 indicates...
that observed values can be evaluated usefully against reference values (16, 17, 23).

Our data indicate that within an individual thyroid hormone concentrations are maintained within relatively narrow limits. In addition, we found large variations among individuals and thus a low index of individuality for all thyroid function tests. This demonstrates that conventional population-based reference intervals for thyroid function tests may be unable to detect abnormal test results that are considerably outside the normal range for the individual being tested. These findings are consistent with previous observations (18–20). Browning et al. (19) studied 12 subjects over a period of 5 wk and found ratios of intra- to interindividual variance about 0.6 for all thyroid function tests. Rasmussen et al. (18) found ratios of intra- to interindividual variances quite variable (0.27–2.7) because of large differences in variance between individuals. Nagayama et al. (20) collected five specimens from each of 47 subjects and confirmed a marked degree of individuality over a period of 2 wk. These studies were performed over shorter periods of time using highly standardized preanalytical conditions. Less standardized preanalytical conditions and rhythmic aberrations of lower frequencies in thyroid function tests (14, 15) may have a different impact on within- vs. between-individual variance and thus change the ratios of intra- to interindividual variance, the individuality index. However, our data indicate that the index of individuality is low in routine laboratory testing of thyroid function over a 1-yr period also.

In our study each individual had a unique thyroid function. Individual set points could be determined by a single blood test within 25% variation for T4, T3, and FTI and within 50% variation for TSH. The difference in test results required for significant changes from the individual set point was relatively large. Previous studies have shown that the problem of unique thyroid function in the individual when testing against a laboratory reference range cannot be solved by stratification (24, 25) or by computation of multivariate reference regions (23, 26). Hence, separate reference ranges for serum TSH, T3, T4, and FTI, depending on, for example, sex and age or evaluation of TSH for different levels of serum T4, will not improve clinical interpretation.

The study was done in men. Variation in thyroid hormones may be different in menstruating women, although hormone levels are unchanged throughout the normal menstrual cycle (14, 27) and circadian variation showed no difference between men and women (28).

Participant number 7 had no clinical symptoms or signs of thyrotoxicosis, but he had a permanently low TSH, which is common in our area (29). Over 1 yr he had one measurement of elevated serum T3 compatible with overt hyperthyroidism.

![Fig. 1](image-url) Serum TSH, total T3, total T4, and FTI in 16 normal subjects taken monthly for 12 months. Each dot represents a monthly measurement and horizontal bars indicate individual parametric means. Participants are sorted by increasing mean. Laboratory reference ranges are: TSH, 0.3–5.0; T3, 1.2–2.7; T4, 60–140; and FTI, 70–140. Large differences were seen between individual set points, and unpredictable differences were seen in variations within individuals for all thyroid function tests.
TABLE 1. Hormone levels and individual interval of variation in 16 subjects (mean ± 2 SD)

<table>
<thead>
<tr>
<th>Participant</th>
<th>TSH, mU/liter, mean (± 2 SD)</th>
<th>Total T4 nmol/liter, mean (± 2 SD)</th>
<th>Total T3 nmol/liter, mean (± 2 SD)</th>
<th>Free T3 index, nmol/liter, mean (± 2 SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.71 (0.54–2.89)</td>
<td>1.65 (1.09–2.21)</td>
<td>108 (63–152)</td>
<td>107 (68–146)</td>
</tr>
<tr>
<td>2</td>
<td>1.54 (0.72–2.35)</td>
<td>1.67 (1.12–2.23)</td>
<td>98 (84–113)</td>
<td>87 (67–107)</td>
</tr>
<tr>
<td>3</td>
<td>1.61 (0.89–2.32)</td>
<td>1.58 (1.04–1.22)</td>
<td>130 (99–162)</td>
<td>123 (91–154)</td>
</tr>
<tr>
<td>4</td>
<td>0.95 (0.52–1.38)</td>
<td>1.57 (1.13–2.00)</td>
<td>88 (75–102)</td>
<td>87 (71–104)</td>
</tr>
<tr>
<td>5</td>
<td>1.49 (0.96–2.01)</td>
<td>1.34 (0.85–1.83)</td>
<td>95 (80–110)</td>
<td>93 (80–105)</td>
</tr>
<tr>
<td>6</td>
<td>0.88 (0.32–1.43)</td>
<td>1.64 (1.12–2.15)</td>
<td>99 (74–124)</td>
<td>84 (65–102)</td>
</tr>
<tr>
<td>7</td>
<td>0.00 (0.00–0.01)</td>
<td>2.36 (1.69–3.03)</td>
<td>129 (109–149)</td>
<td>122 (104–141)</td>
</tr>
<tr>
<td>8</td>
<td>1.68 (1.04–2.31)</td>
<td>1.03 (0.44–1.61)</td>
<td>85 (79–91)</td>
<td>83 (75–90)</td>
</tr>
<tr>
<td>9</td>
<td>1.11 (0.79–1.42)</td>
<td>2.01 (1.62–2.41)</td>
<td>110 (92–128)</td>
<td>109 (90–128)</td>
</tr>
<tr>
<td>10</td>
<td>0.67 (0.34–1.00)</td>
<td>2.10 (1.86–2.35)</td>
<td>137 (130–144)</td>
<td>135 (121–148)</td>
</tr>
<tr>
<td>11</td>
<td>2.42 (1.60–3.24)</td>
<td>1.72 (1.53–1.92)</td>
<td>81 (63–99)</td>
<td>84 (65–104)</td>
</tr>
<tr>
<td>12</td>
<td>0.48 (0.32–0.64)</td>
<td>2.05 (1.72–2.37)</td>
<td>136 (119–153)</td>
<td>127 (114–141)</td>
</tr>
<tr>
<td>13</td>
<td>1.07 (0.50–1.65)</td>
<td>1.72 (1.52–1.91)</td>
<td>122 (88–157)</td>
<td>116 (73–159)</td>
</tr>
<tr>
<td>14</td>
<td>1.12 (0.45–1.79)</td>
<td>1.54 (1.31–1.76)</td>
<td>108 (86–129)</td>
<td>105 (79–131)</td>
</tr>
<tr>
<td>15</td>
<td>1.15 (0.70–1.59)</td>
<td>1.51 (1.35–1.68)</td>
<td>89 (72–106)</td>
<td>86 (70–102)</td>
</tr>
<tr>
<td>16</td>
<td>1.25 (0.91–1.60)</td>
<td>1.47 (1.30–1.64)</td>
<td>109 (100–118)</td>
<td>111 (75–146)</td>
</tr>
<tr>
<td>Averageb</td>
<td>1.27 (0.66–1.89)</td>
<td>1.64 (1.24–2.04)</td>
<td>106 (87–126)</td>
<td>102 (80–125)</td>
</tr>
<tr>
<td>Overall groupc</td>
<td>1.27 (0.16–2.39)</td>
<td>1.64 (0.97–2.31)</td>
<td>106 (65–148)</td>
<td>102 (61–144)</td>
</tr>
<tr>
<td>Referencerangesd</td>
<td>(0.3–5.0)</td>
<td>(1.2–2.7)</td>
<td>(−60–140)</td>
<td>(−70–140)</td>
</tr>
<tr>
<td>Individuality indexd</td>
<td>0.49</td>
<td>0.58</td>
<td>0.54</td>
<td>0.59</td>
</tr>
</tbody>
</table>

a Participant number 7 was excluded from calculations because he was subclinically hyperthyroid.
b Calculated values for an average participant; reference intervals calculated as mean ± 2 SD; average SD is the square root of the average variance.
c Calculated values for the group, excluding participant 7.
d The conventional population-based reference limits used in our routine laboratory.

e Individuality index was calculated from $SD_{analytical}^2 + SD_{intraindividual}^2$. The individuality index describes the sensitivity of population-based reference ranges to abnormal test results in the individual (<0.6 insensitive; >1.4 sensitive).

TABLE 2. Number of tests required to describe the homeostatic set point in an individual

<table>
<thead>
<tr>
<th>Precision of set point</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSH</td>
<td>85</td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>T3</td>
<td>25</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>T4</td>
<td>25</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>FTI</td>
<td>25</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

Calculated from: $n = (Z \times CV_{analytical}^2 + CV_{intraindividual}^2)^{1/2}/D^2$, where: D is percent closeness to the homeostatic set point, Z is the number of standard deviations required for a confidence interval (i.e. 1.96 for 95%), n is the number of specimens.

but 11 measurements of T3 within the population-based reference range, indicating subclinical hyperthyroidism. If it is hypothesized that participant number 7 would normally have the same set point for serum T3 as participant number 8, then all 12 measurements of T3 in participant number 7 would be outside the 95% confidence interval for the individual, indicating overt hyperthyroidism. This illustrates how subclinical thyroid disease could be overt hypo- or hyperthyroidism hidden behind vague symptoms and insensitive reference ranges for serum T3 and T4.

The pituitary gland is sensitive to minor changes in serum TSH and T4, and serum TSH responds heavily to such changes. When thyroid function is abnormal, the association between serum TSH and both T3 and T4 is log linear (25, 31). This amplified response of serum TSH to changes in serum T3 and T4 may cause serum TSH to leave the population-based reference range when serum T3 and T4 are outside the individual reference range, even when they are still within the population-based reference range. This is labeled subclinical thyroid disease. The view that individuals with subclinical thyroid disease have abnormal thyroid function is supported by increasing amounts of data on the biological
importance of subclinical thyroid disease for a number of organs (4–9).

Our data indicate that the distinction between subclinical and overt thyroid disease is somewhat arbitrary because it depends to a considerable extent on the position of the patient’s normal set point for T₃ and T₄ within the laboratory reference range. For example, serum FTI had to decrease by 14 nmol/liter below the set point in participant 8 to be reported subnormal by the laboratory. In participant 10, on the other hand, FTI had to decrease by 66 nmol/liter below the individual set point. Hence, participant 10 had to be considerably more hypothyroid to have a diagnosis of overt hypothyroidism. The opposite relation existed to have a diagnosis of overt hyperthyroidism based on an increase in FTI. This would happen much earlier in participant 10 (increase of 6 nmol/liter) than in participant 8 (increase of 58 nmol/liter required). These differences emphasize the importance of abnormalities in serum TSH, more or less independent of serum T₃ and T₄.

In conclusion, we found that individual reference ranges for serum T₃ and T₄ are about half the width of population-based reference ranges. Hence, a test result within the laboratory reference limits is not necessarily normal for the individual. Serum TSH outside the population-based reference range indicates that serum T₃ and serum T₄ are not normal for the individual.

Acknowledgments

Received June 29, 2001. Accepted October 8, 2001.

Address all correspondence and requests for reprints to: Stig Andersen, M.D., Department of Endocrinology and Medicine, Aalborg Hospital, DK-9000 Aalborg, Denmark. E-mail: stiga@daldnet.dk.

References

12. Samuels MH 1998 Subclinical thyroid disease in the elderly. Thyroid 8:803–813