Biologic Variation is Important for Interpretation of Thyroid Function Tests

Stig Andersen, Niels Henrik Bruun, Klaus Michael Pedersen, and Peter Laurberg

Large variations exist in thyrotropin (TSH) and thyroid hormones in serum. The components of variation include preanalytical, analytical, and biologic variation. This is divided into between- and within-individual variation. The latter consists of circadian and seasonal differences although there are indicators of a genetically determined starting point. The ratio of within- to between-individual variation describes the reliability of population-based reference ranges. This ratio is low for serum TSH, thyroxine (T₄) and triiodothyronine (T₃) indicating that laboratory reference ranges are relatively insensitive to aberrations from normality in the individual. Solutions are considered but reducing the analytical variation below the calculated analytical goals of 7%, 5% and 12% for serum T₃, T₄, and TSH does not improve diagnostic performance. Neither does determination of the individual set-point and reference range. In practice this means that population-based reference ranges are necessary but that it is important to recognize their limitations for use in individuals. Serum TSH responds with amplification to minor alterations in T₄ and T₃. A consistently abnormal TSH probably indicates that T₄ and T₃ are not normal for the individual even when inside the laboratory reference range. This underlines the importance of TSH in diagnosis and monitoring of thyroid dysfunctions. Also, it implies that subclinical thyroid disease may be defined in purely biochemical terms. Under critical circumstances such as pregnancy where normal thyroid function is of importance for fetal brain development, subclinical thyroid disease should be treated. Even TSH within the reference range may be associated with slightly abnormal thyroid function of the individual. The clinical importance of such small abnormalities in thyroid function in small children and pregnant women for brain development remains to be elucidated.

Introduction

Symptoms and signs of hyperthyroidism and hypothyroidism are often nonspecific and vague if present (1–3) and measurement of thyrotropin (TSH), thyroxine (T₄), and triiodothyronine (T₃) in serum is important for diagnosis of overt and subclinical thyroid dysfunction (4–8).

Usually, a test result in an individual is compared to a laboratory reference range. Such a reference range is often defined by a probability distribution including 95% of test results in healthy individuals in a population. Hence, the variation in the population determines the width of the population-based reference range.

The overall variation consists of analytical and biologic variation (9,10). Biologic variation is divided into two components: variation in the individual and variation between individuals (9,10). The former is characterized by rhythmic aberrations of multiple frequencies (11–14) and the latter is caused by different set-points around which each individual varies (9).

This paper aims to provide insight into the components of variation in TSH, T₄, and T₃ in serum, the determinants of biologic variation, and the impact of biologic variation on interpretation of test results, specifically, the impact of variation on reference ranges and on alternative methods for detecting abnormal test results. Also, the largest acceptable analytical variation in measurements of TSH, T₄, and T₃ in serum is estimated. Finally, the impact on delineating the entity subclinical thyroid disease and the importance for individuals with TSH at the extremes of laboratory normal ranges are estimated with consideration of the importance for brain development.

Components of Variation in Thyroid Function Tests

Variation in thyroid function tests is described by a number of factors summarized in Table 1. A measured value, x, is determined by an overall group mean, μgroup. Deviation from this grand mean value is caused by the sum of components of variation. Minimizing each of these components improves the precision of the measured value, x.
Preanalytical variation, $\sigma_{\text{preanalytical}}$

Recommendations are that preanalytical conditions should be standardized to minimize preanalytical variation in the study of biologic variation (10,15): phlebotomy undertaken at the same hour, by the same individual, without tourniquet, in a fasting, rested individual. Such conditions were applied in two studies (16,17), inconsistently in one study (18), while information was limited in some studies (19–22).

Highly standardized preanalytical conditions may hamper the external validity. Hence, one study of variation was performed to illustrate everyday testing of thyroid function with blood samples taken during laboratory opening hours using standard procedures (23). The larger variations found in this study probably reflect the impact of preanalytical conditions.

Analytical variation, $\sigma_{\text{analytical}}$

Analytical variation adds imprecision to biologic variation. Conversely, if biologic variation is high then analytical variation becomes relatively less important and the requirement for analytical precision diminishes (10,24). The goals set are that variance as a result of analytical error should not exceed 20% of the variation (24). It can be shown that this is acceptable when $\text{CV}_{\text{analytical}} \leq \frac{1}{2} \text{CV}_{\text{intraindividual}}$ (10,24) or $\leq \frac{1}{2} \text{CV}_{\text{biologic}}$ (17).

Table 2 shows the analytical goals calculated from variation in a study using highly standardized preanalytical conditions (17), in a routine laboratory setting (23), and the recommended analytical coefficients of variation (CVs; 25–27). The analytical variation is less important for TSH because the biologic variation is quite high (27,28). The differences in CVs reflect different preanalytical conditions.

### Table 1. The Components Included in a Measurement of Serum TSH, $T_3$ or $T_4$ Can Be Described by This Equation

\[
x = \mu_{\text{group}} + \sigma_{\text{preanalytical}} + \sigma_{\text{analytical}} + \sigma_{\text{between-individual}} + \sigma_{\text{within-individual}} + \sigma_{\text{other}}
\]

- $x$ is a measured value of TSH, $T_3$, or $T_4$.
- $\mu_{\text{group}}$ is the overall group mean value.
- $\sigma_{\text{preanalytical}}$ is variance due to preanalytical conditions.
- $\sigma_{\text{analytical}}$ is variance due to analytical errors.
- $\sigma_{\text{between-individual}}$ is interindividual variance.
- $\sigma_{\text{within-individual}}$ is intraindividual variance.
- $\sigma_{\text{other}}$ is random variation not accounted for.

Biologic variation consists of $\sigma_{\text{between-individual}} + \sigma_{\text{within-individual}}$.

TSH, thyrotropin; $T_3$, triiodothyronine; $T_4$, thyroxine.

### Table 2. The Analytical Goals (CV%) Calculated From Data on Biologic Variation in a Highly Standardized Setting (17), and in a Routine Laboratory Setting (Calculated From Andersen 2002 [23]), and the Recommended Maximum CV% (27)

<table>
<thead>
<tr>
<th>Component</th>
<th>Standardized</th>
<th>Routine</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CV% $\leq^a$</td>
<td>CV% $\leq^a$</td>
<td>CV% $\leq^b$</td>
</tr>
<tr>
<td>$T_3$</td>
<td>5.2</td>
<td>6.8</td>
<td>10.2</td>
</tr>
<tr>
<td>$T_4$</td>
<td>2.5</td>
<td>5.1</td>
<td>9.7</td>
</tr>
<tr>
<td>FT$_4$I</td>
<td>4.7</td>
<td>5.8</td>
<td>10.2</td>
</tr>
<tr>
<td>TSH</td>
<td>8.1</td>
<td>11.6</td>
<td>21.9</td>
</tr>
</tbody>
</table>

$^a$Calculated from $\text{CV}_{\text{analytical}} \leq \frac{1}{2} \text{CV}_{\text{intraindividual}}$ (10,24).

$^b$Calculated from $\text{CV}_{\text{analytical}} \leq \frac{1}{2} \text{CV}_{\text{intraindividual}}$ + interindividual.

T$_3$, total triiodothyronine; FT$_4$I, total thyroxine; TSH, thyrotropin.

### Figure 1

**A:** Individual mean in each of 15 healthy subjects of estimated 24-hour urinary iodine excretion, thyrotropin (TSH), total triiodothyronine ($T_3$), and total thyroxine ($T_4$). Each dot represents the mean value of 12 consecutive monthly measurements in 1 individual. Individual mean values are indexed around the group mean value (index 100: urinary iodine 49.1 mg per 24 hours, serum TSH 1.27 mU/L, $T_3$ 1.64 nmol/L, $T_4$ 106 nmol/L). (From Andersen et al. [23,29]).

**B:** Individual coefficients of variation (CV%) for each variable. Each dot represents the CV% in one individual. (From Andersen et al. [23,29]).
tion tests in 15 healthy men (23,29). This illustrates that each individual had his own mean value around which the person varied. The difference between individual mean values was highly significant (Kruskal-Wallis test; \( p < 0.001 \) for all variables) (23). This is consistent with previous findings of large variance between individuals (16–18,30).

**Within-individual variation, \( \sigma_{\text{within-individual}} \)**

Repeated measurements of thyroid function tests in one individual scatter around an individual mean value. This variation differed widely between subjects (Fig. 1B, Bartlett’s test, \( p < 0.001 \) for all variables) in our study (23) as found by others (16–18).

The variation around the individual mean value is described by a probability distribution. The larger the difference between a measured value and the mean value, the less likely is this caused by random variation. Thus, a 95% confidence interval in the individual and the distance required between two measurements for significance can be calculated from data on variation.

**What Determines Biological Variation in Thyroid Function Tests**

**Circadian variation**

A circadian variation in serum TSH in healthy subjects is well documented (Table 3) (11–13, 31–50). Low levels of serum TSH (nadir) are found during the daytime with a characteristic nocturnal rise of more than 100% that peaks just after midnight. No increment of serum \( T_3 \) or \( T_4 \) seems to follow the nocturnal surge in TSH (32–35).

Lucke et al. (36) stated that “fluctuations do not exceed the normal ranges.” This is the key point: these fluctuations contribute to the width of the reference ranges.

The circadian rhythm of TSH is influenced by environmental factors. Sleep (13,37,38) decreased TSH pulse amplitude but not pulse frequency (13), as did fasting (39) and \( T_3 \) and \( T_4 \) infusion (40), while nocturnal physical activity may phase-delay the rhythms (41). Interestingly, the circadian pulsatile secretion pattern of serum TSH was found to be remarkably reproducible in the individual but with considerable and unaltered differences between individuals (13,35). This favors a genetic component in determining the pattern of TSH secretion.

**Seasonal variation**

Table 4 lists some longitudinal studies of seasonal variation in thyroid function tests in healthy adults (14,29,30,51–67). In general, serum \( T_3 \) was higher during winter while seasonal changes in \( T_4 \) and TSH were less consistent. Seasonal variation has been suggested to be caused by ambient temperature and luminosity (51,52) and an increase in \( T_3 \) with cold has been demonstrated (68,69). Seasonal variation in iodine intake has been demonstrated in some populations (70–72) though not in all (73,74). Because iodine in-

### Table 3. Some Studies of Circadian Variation of Thyroid Function Tests in Healthy Adult Subjects

<table>
<thead>
<tr>
<th>Participants (n)</th>
<th>Method for data evaluation</th>
<th>Increase from nadir to peak</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>TSH</td>
<td>( T_3 )</td>
</tr>
<tr>
<td>7</td>
<td>Cosinor</td>
<td>44%</td>
<td>16%</td>
</tr>
<tr>
<td>26</td>
<td>Pulse</td>
<td>180%</td>
<td>—</td>
</tr>
<tr>
<td>8</td>
<td>Pulse</td>
<td>170%</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>Pulse</td>
<td>110%</td>
<td>14%</td>
</tr>
<tr>
<td>6</td>
<td>Modified Wilcoxon</td>
<td>170%</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>Pulse</td>
<td>150%</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>ANOVA</td>
<td>320%</td>
<td>—</td>
</tr>
<tr>
<td>7</td>
<td>ANOVA</td>
<td>250%</td>
<td>—</td>
</tr>
<tr>
<td>24</td>
<td>Cosinor</td>
<td>87%</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>—</td>
<td>120%</td>
<td>22%</td>
</tr>
<tr>
<td>10</td>
<td>Power-spectra</td>
<td>170%</td>
<td>35%</td>
</tr>
<tr>
<td>6</td>
<td>t test</td>
<td>97%</td>
<td>20%</td>
</tr>
<tr>
<td>5</td>
<td>F test, t test</td>
<td>14%</td>
<td>22%</td>
</tr>
<tr>
<td>8</td>
<td>Covariance regression, t test</td>
<td>96%</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>t test</td>
<td>105%</td>
<td>—</td>
</tr>
<tr>
<td>8</td>
<td>ANOVA</td>
<td>—</td>
<td>15%</td>
</tr>
<tr>
<td>5</td>
<td>—</td>
<td>—</td>
<td>54%</td>
</tr>
<tr>
<td>13</td>
<td>t test</td>
<td>115%</td>
<td>—</td>
</tr>
<tr>
<td>6</td>
<td>ANOVA</td>
<td>88%</td>
<td>—</td>
</tr>
</tbody>
</table>

—, Data not included or specified.
\( a_p < 0.05 \).
\( b_p > 0.05 \).
\( c \) Modified power spectrum analysis (42).
\( d \) Significance level not specified or unclear.
\( ep < 0.001 \).
\( f p < 0.01 \).
\( sp < 0.02 \).
TSH, thyrotropin; \( T_3 \), triiodothyronine; \( T_4 \), thyroxine; ANOVA, analysis of variance.
fluences the secretion of thyroid hormones (75) some seasonal variation in thyroid function may be attributable to differences in iodine intake.

Statistically significant seasonal variations were not detected in our study (Fig. 2) although the trends found (29) were similar to those statistically significant in studies with a lower overall variation (14,30). Our study was performed in an area with mild to moderate iodine deficiency and part of the individual variation in serum TSH could be caused by variation in iodine intake. Thus, we found a statistically significant trend in the average correlation between iodine excretion and TSH (Fig. 3) with a negative correlation between urinary iodine and serum TSH in subjects with the lowest average annual iodine intake (29). Probably this indicates a variable lack of the substrate iodine for thyroid hormone synthesis in these subjects and thus different effects of variation in iodine on thyroid function.

Other factors that may influence thyroid hormones and TSH in serum

Gender differences was not found either in the pattern (11–13,45,64) or in the variance (16–18) of TSH secretion. Protein concentrations in serum or their affinity to thyroid hormones may vary without changes in thyroid status (76). Prolonged venipuncture and posture during phlebotomy (12) may alter serum concentration of T3 and T4 slightly because their concentrations relate to protein binding. This contributes to the difference in variance between studies (23) and (16,17) but the effect is small. The level of TSH and thyroid hormones in serum may be influenced by drugs (77,78) smoking (79–81) and pregnancy (82,83) but little is known about the effect on variation in thyroid function.

Impact of Biologic Variation on Interpretation of Tests of Thyroid Function

Reliability of reference ranges

The ability of laboratory reference ranges to detect an abnormal test result in an individual depends on the balance between contributors to the biologic variation. If the major part of the overall variation is the result of variation within individuals, while differences between individual set-points are small, then a population-based reference range matches variations in each of any individual. Conversely, if the overall variation is mainly caused by narrow individual variation around dispersed individual set-points then a population-based reference range is unlikely to detect minor deviations from the individual set-point.

This can be described by a ratio of within- to between-individual variation, the individuality index (10,84,85). When the ratio is less than 0.6, the population-based reference range is an insensitive measure in the large majority of individuals. When the ratio is greater than 1.4, the reference range works as intended (10,84).

Table 4. Some Longitudinal Studies of Seasonal Variation in Thyroid Function Tests in Healthy Adult Subjects

<table>
<thead>
<tr>
<th>Participants (n)</th>
<th>Method for data evaluation</th>
<th>TSH</th>
<th>T3</th>
<th>T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>ANOVA</td>
<td>8%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>20</td>
<td>ANOVA</td>
<td>31%&lt;sup&gt;b&lt;/sup&gt;</td>
<td>13%&lt;sup&gt;b&lt;/sup&gt;</td>
<td>—&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>13</td>
<td>Cosinor</td>
<td>25%&lt;sup&gt;c&lt;/sup&gt;</td>
<td>6%&lt;sup&gt;c&lt;/sup&gt;</td>
<td>&lt;1%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>43</td>
<td>ANOVA</td>
<td>—</td>
<td>77%&lt;sup&gt;c&lt;/sup&gt;</td>
<td>35%&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>166</td>
<td>Cosinor</td>
<td>21%&lt;sup&gt;f&lt;/sup&gt;</td>
<td>18%&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>13</td>
<td>Friedman’s</td>
<td>16%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>6</td>
<td>t test</td>
<td>—</td>
<td>17%&lt;sup&gt;b&lt;/sup&gt;</td>
<td>22%&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
<tr>
<td>32</td>
<td>Cosinor</td>
<td>25%&lt;sup&gt;b&lt;/sup&gt;</td>
<td>18%&lt;sup&gt;e&lt;/sup&gt;</td>
<td>4%&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
<tr>
<td>4</td>
<td>ANOVA</td>
<td>—</td>
<td>12%&lt;sup&gt;b&lt;/sup&gt;</td>
<td>—&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>15</td>
<td>t test</td>
<td>39%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7%&lt;sup&gt;d&lt;/sup&gt;</td>
<td>8%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>45</td>
<td>t test</td>
<td>15%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
</tr>
<tr>
<td>53</td>
<td>t test</td>
<td>11%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>27%&lt;sup&gt;c&lt;/sup&gt;</td>
<td>6%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>5</td>
<td>Cosinor</td>
<td>—</td>
<td>—</td>
<td>49%&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
<tr>
<td>13</td>
<td>Friedman’s</td>
<td>—</td>
<td>9%&lt;sup&gt;d&lt;/sup&gt;</td>
<td>14%&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>8</td>
<td>t test</td>
<td>53%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>22%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>14%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>12</td>
<td>—</td>
<td>—</td>
<td>14%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>26%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>9</td>
<td>ANOVA</td>
<td>—</td>
<td>—</td>
<td>18%&lt;sup&gt;a&lt;/sup&gt;,&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td>8</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>48%&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td>18</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>18%&lt;sup&gt;f&lt;/sup&gt;,&lt;sup&gt;b&lt;/sup&gt;,&lt;sup&gt;i&lt;/sup&gt;/87%&lt;sup&gt;f&lt;/sup&gt;,&lt;sup&gt;i&lt;/sup&gt;</td>
</tr>
<tr>
<td>4</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>68%&lt;sup&gt;f&lt;/sup&gt;,&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

—, Data not included or specified.
<sup>a</sup>p > 0.05.
<sup>b</sup>p < 0.05.
<sup>c</sup>—p < 0.001.
<sup>d</sup>p < 0.02.
<sup>e</sup>p < 0.1.
<sup>f</sup>Protein-bound iodine in plasma.
<sup>g</sup>Significance level not specified or unclear.
<sup>h</sup>Working indoor (nurse students).
<sup>i</sup>Working outdoor (postmen).
TSH, thyrotropin; T<sub>3</sub>, triiodothyronine; T<sub>4</sub>, thyroxine; ANOVA, analysis of variance.
Using highly standardized preanalytical conditions, this ratio was below 0.6 for all tests of thyroid function (17). Using preanalytical conditions similar to those used in everyday laboratory practice the average individual variation was only approximately half of that of the group (23) (Fig. 4). Accordingly, the ratio was still below 0.6 for all measures of thyroid function (23). Consequently, an individual test result may be far outside the individual reference range while still lie well within the laboratory reference range. This indicates a low sensitivity of the population-based reference ranges and causes uncertainty in the diagnosis of overt, and in particular subclinical thyroid disease.

**Individual reference ranges**

It has been suggested to use the individual as his own reference for detection of abnormal test results (18,86). To determine the individual set-point and reference range it is necessary to analyze a number of specimens from that individual. The required number of measurements can be estimated if we decide on two things: the percentage closeness to the true homeostatic set-point and the confidence interval. The number of tests required can then be calculated from

\[ n = \left( \frac{Z}{CV_{\text{analytical}} + CV_{\text{intraindividual}}} \right)^2 \]

where \( n \) is the number of specimens, \( Z \) is 1.96 for the 95% confidence interval, and \( D \) is the percentage closeness to the homeostatic set point (10).

Table 5 shows the average number of specimens required to describe the homeostatic set-point in an individual in a routine laboratory setting (23). The number of tests is large, and because of the large differences in individual variance (Fig. 1B) the number of tests required vary widely: from 30 to 183 tests of TSH, and from 2 to 126 measurements of \( T_4 \) and \( T_3 \). Thus, this approach is both impractical and unreliable.
Significant difference in repeated testing: the Δ check

A different way of evaluating a test result is to determine whether this is significantly different from previous ones: the delta check (Δ check). This may be calculated from $\Delta = Z \times \sqrt{n} \times \sqrt{SD_{analytical}^2 + SD_{intraindividual}^2 + \text{preanalytical}^2}$, where $n$ is the number of specimens obtained and $Z$ the confidence interval (10,87).

A problem may occur if a mean intraindividual variance is used for these calculations because of the large differences in variance between individuals ($p < 0.001$ in all variables) (23). Thus, an upper value was derived from the estimated distribution of intraindividual variances rather than a mean variance. Because only one type of deviation from mean variance was meaningful, a one-sided Z-value was used (0 for 50%, 1.64 for 95%, etc.). Hence, a true Δ check was computed, and the differences required to be 50% to 99.9% confident of an abnormal value compared to a previous test result in any of the individuals can be read from Figure 5: As a rule of thumb, a difference of approximately 40 nmol/L in T4, 0.7 nmol/L in T3, and 1.0 mU/L in TSH indicates 90% certainty of a true change.

![Image](40x591 to 258x717)

**FIG. 4.** The distribution of 12 consecutive monthly measurements of serum total thyroxine (T4) in 15 healthy subjects ($n = 180$) and in 1 individual ($n = 12$). The width of the distribution in one individual is approximately half of that of the group. (From Andersen et al. [23]).

**Table 5. Number of Tests Required to Describe the Homeostatic Set Point in an Individual (Adapted from Andersen 2002 [23]).**

<table>
<thead>
<tr>
<th>Precision of set point</th>
<th>5%</th>
<th>10%</th>
<th>20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSH</td>
<td>85</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>TT3</td>
<td>25</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>TT4</td>
<td>25</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>FT4I</td>
<td>25</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

Calculated from: $n = \left( Z^2(CV_{analytical}^2 + CV_{intraindividual}^2)^{1/2}/D \right)^2$

where:
- $D$ is % closeness to the homeostatic set point,
- $Z$ is the number of standard deviations required for a confidence level (1.96 for 95%; 1.64 for 90%; 1.28 for 80%),
- $n$ is the number of specimens.

TSH, thyrotropin; TT3, triiodothyronine; TT4, total thyroxine; FT4I, free thyroid index.

![Image](38x454 to 43x454)

**FIG. 5.** The change in thyrotropin (TSH), total triiodothyronine (T3), total thyroxine (T4) and free T4 index required for different levels of significance in repeated testing (the Δ check). The degree of confidence of a statistically significant difference from a previous measurement is read on the x-axis with the corresponding difference required in TSH, TT3, TT4, and free T4 index on the y-axis. This was calculated using $\Delta = Z \times \sqrt{n} \times \sqrt{SD_{analytical}^2 + SD_{intraindividual}^2 + \text{preanalytical}^2}$, where $n$ is the number of specimens obtained and $Z$ the two-sided confidence interval (i.e., 0.67 for 50%, 1.64 for 90%, 1.96 for 95%). Derived from Andersen et al. (23).
This Δ check, or critical difference in serial testing, could be perceived as a confidence interval around any individual test result for each of the variables (compare Fig. 2).

Clinical Implications

Relation between TSH and thyroid hormones: a clue to understanding subclinical thyroid disease

Serum TSH responds with amplification to minor changes in serum \(T_3\) and \(T_4\) in our data \((r = 0.30\) for log-linear relationship; \(r = 0.17\) for linear relationship) as described by others (88–91), and complete suppression of TSH could be found when \(T_3\) was still within the normal range (88).

The position of the individual set point of \(T_4\) and \(T_3\) within the wide laboratory reference ranges is important for the diagnosis of minor changes in thyroid hormones. If the set point of an individual is in the upper part of the laboratory reference range then only a slight increase in thyroid hormone secretion will cause these to leave the reference ranges, as will the amplified response in TSH. This individual will have diagnosed overt hyperthyroidism. If, however, the set point of an individual is in the lower part of the laboratory reference range (as in Fig. 4), then the same slight increase in thyroid hormone secretion will not cause these to leave the reference ranges while TSH will. This individual will have diagnosed subclinical hyperthyroidism. Thus, narrow individual variation combined with amplified response of TSH to minor alterations in thyroid hormones in serum provides a biochemical explanation for the entity subclinical thyroid disease and indicates that this is mild thyroid failure.

Before week 12 of gestation the maturing brain is critically dependent on the circulation maternal \(T_4\) and overt maternal hypothyroidism is associated with severely impaired neurologic development of the offspring. More importantly, even mild abnormalities in thyroid function may adversely affect brain development in the fetus (92). Impaired intellectual development can be found even with \(T_4\) within the reference range during early gestation (93). This emphasizes the importance of detection of minor abnormalities in thyroid function.

Abnormalities with both thyroid hormones and TSH within reference ranges

The ratio of within- to between-individual variation was low for serum TSH in all studies (16–18,23). This implies that some individuals with TSH within the population based reference range have a TSH outside the individual reference range (i.e., an abnormal serum TSH).

This is illustrated in cross-sectional population studies in which participants with serum TSH close to the outer limits of the laboratory reference range have a higher frequency of thyroid abnormalities. Thus, individuals with a high-normal serum TSH have a higher occurrence of autoimmunity (94,95) and a higher risk of developing thyroid disease (96). Also, individuals with a low-normal serum TSH have more nodules in the thyroid gland (97,98).

Thus, the individuals in the upper and lower parts of the population-based reference ranges for TSH seem to consist of two groups: those with normal thyroid function who just happen to have a high or a low set-point, and those who have small abnormalities in thyroid function. Whether this influences brain development is an issue for future investigations.

Impact of season and hour

Seasonal variation in serum TSH and thyroid hormones can be found in Table 4. Changes in \(T_3\) vary with latitude but on average serum total \(T_3\) increases approximately 0.2 nmol/L during winter. This is associated with an insignificant increase in serum TSH of approximately 0.1 mU/L.

Circadian variation in TSH is more pronounced (Table 3). However, the main increase occurs in the evening and the average decrease from 09.00 to 15.00 hours and may be estimated to be approximately 0.3 mU/L.

Conclusion and Recommendations

Individual variation in serum \(T_4\), \(T_3\), and TSH in healthy subjects is narrow compared to laboratory reference ranges. Consequently, a test result within the laboratory reference range does not necessarily indicate a normal thyroid function in the individual. No mathematical trick may overcome this problem because an impractically large number of tests are required to determine the individual set-point. The distances between two measurements of thyroid function required for statistical significance can be seen in Figure 5 but are quite large. However, serum TSH responds heavily to minor changes in thyroid hormone concentrations in serum. Hence, subclinical thyroid disease with abnormal TSH but \(T_4\) and \(T_3\) within laboratory reference ranges is probably always a sign that \(T_4\) and \(T_3\) are outside the individual reference range and thus an indicator for abnormal thyroid function in the individual. This emphasizes the importance of serum TSH relative to \(T_3\) and \(T_4\) this being total or estimated free hormone concentrations in serum. If there are clinical signs, or if other conditions such as pregnancy requires normal thyroid function to ensure normal fetal brain development, then there is a need for treatment.

The individual reference range for serum TSH is approximately half the width of laboratory reference range. Thus, some individuals in the upper and lower parts of the normal range will have TSH outside their individual reference range. It remains to be determined whether such small abnormalities in thyroid function are important (i.e., for the developing brain).

References

1998 Observer variation in the clinical and laboratory evaluation of patients with thyroid dysfunction and goiter. *Thyroid* 8:393–398.


83. Knudsen N, Laurberg P, Perrild H, Bulow I, Ovesen L, Jør-

Address reprint requests to:
Stig Andersen, M.D., Ph.D.
Department of Endocrinology and Medicine
University Hospital Aalborg
Reberbansgade
Box 561
DK-9100 Aalborg
Denmark
E-mail: stiga@dadlnet.dk
This article has been cited by:


7. Duk Jae Kim, Young Ho Khang, Jung-Min Koh, Young Kee Shong, Ghi Su Kim. 2006. Low normal TSH levels are associated with low bone mineral density in healthy postmenopausal women. *Clinical Endocrinology* 64:1, 86. [CrossRef]


12. 2004. Letters to the Editor. *Thyroid* 14:11, 983-986. [Citation] [PDF] [PDF Plus]