Computerized Method for Validating Laboratory Reference Ranges for Triiodothyronine and Thyroxin Immunoassays

Brian Luttrell and Sally Watters

We used a computer-based method to help validate the reference ranges of assays for triiodothyronine (T₃) and thyroxin (T₄). A retrospective search of a database of laboratory results for the previous six months identified all patients with apparent euthyroid status, as defined by methods independent of the immunoassay under review. A computer-generated reference group (CGR Group) of 2001 records had a gaussian distribution of T₄ values and a reference range (mean ± 2 SD) of 56–161 nmol/L, compared with the supplier's suggested range for euthyroid subjects (58–148 nmol/L) and an in-house range of 60–144 nmol/L for a group of 97 normal subjects. A similar CGR Group of 1902 records gave a reference range for T₃ of 0.7–2.1 nmol/L (manufacturer's range 0.8–2.8; normal subjects 0.8–2.2). An attempt to devise a reference range for thyrotropin failed when we found that its concentration in the population of patients with normal values for thyroid hormones was distributed differently from that in the normal population. The method is intended to be used in addition to conventionally derived ranges based on results for healthy subjects. It allows the laboratory to conveniently verify the reference ranges for T₃ and T₄ assays at regular intervals by using very large samples with appropriate age, sex, and weight distribution, drawn from the population of patients' samples submitted for analysis.

Additional Keyphrases: thyrotropin • variation, source of

The establishment and maintenance of reference ranges for hormone assays is a continual problem for diagnostic laboratories (1). Statistical considerations dictate that a substantial sample size of normal healthy subjects is required and that the age and sex distribution should be similar to that of the patient population. Ideally, the reference ranges ought to be verified at regular intervals (e.g., each six months) to take account of any drift with time and should be repeated for each analyte measured.

To partly overcome these difficulties, we have attempted to develop a computer method to help verify reference ranges for some analytes. This effort was based on the premise that a proportion of the patients' samples analyzed by any diagnostic laboratory will be from normal subjects. These samples will have been submitted for investigations that return negative results. Such a sub-group would be well suited for the determination of laboratory reference ranges if these individual patients' records could be isolated from the database of all test results. The major advantages of this approach would include the large sample sizes achievable and the fact that the age, weight, and sex distribution of the subjects chosen would be totally representative of the patient population.

Previous attempts to derive normal reference ranges from patients' data (2, 3) relied on statistical manipulation of data from the actual assay under review. Several of these have been criticized by Amador and Hsi (4). For the procedure to be valid, the means of identification of suitable patients for the reference ranges must be independent of the analytical method under consideration.

This condition can be met for the thyroid hormones triiodothyronine (T₃) and thyroxin (T₄), and for thyrotropin (TSH), which are related physiologically, through the hypothalamic–pituitary–thyroid axis, but are determined separately and independently in the laboratory. Thus the measurement of serum TSH alone can provide one measure of thyroid status because it is capable of discriminating between hypothyroid, euthyroid, and hyperthyroid states. Although the strategy has not been widely adopted, several groups of authors have suggested that the new generation of sensitive TSH immunoassays, used alone, may provide an adequate initial thyroid screen for primary investigation (5, 6). Conversely, thyroid status can also be assessed by using only thyroid hormone estimations, with an appropriate correction for binding protein effects. This approach was in general use before immunoassay of TSH became available. Although reliance on thyroid hormones alone will result in some euthyroid subjects being misclassified as abnormal, particularly in patients with nonthyroidal illness, as demonstrated by Beaman and Woodhead (7), the exclusion of these subjects does not adversely affect our selection of a reference group. Thus, two separate and independent assessments of thyroid status can be based on routine laboratory tests of a single serum specimen.

In this center, a routine laboratory thyroid investigation includes estimations of total T₃, total T₄, TSH, and thyroxin-binding globulin (TBG), with the TBG concentrations being used to correct the T₃ and T₄ estimates for binding protein effect. We wrote a program to search the computer database of laboratory results for all patients' records, within a given time interval, to determine which would be suitable for inclusion in reference

---

1 Nonstandard abbreviations: T₄, thyroxin; T₃, triiodothyronine; TBG, thyroxin-binding globulin; TSH, thyrotropin; and CGR, computer-generated reference (group).
ranges for the analytes T₃, T₄, and TSH. Thus, for the T₄ computer-generated reference group (CGR Group), we included all records for which the values for the remaining two variables, TSH concentration and T₃/TBG ratio, fell within the currently accepted laboratory reference ranges.

A similar method (based on TSH concentrations and T₄/TBG ratios) was used to obtain a separate and independent CGR Group for T₃, and a TSH CGR Group was generated from all records with T₃/TBG and T₄/TBG ratios within the laboratory reference ranges. The T₄/TBG ratio is a well-recognized method for correcting for the effects of protein binding on thyroid hormone concentrations except at very low concentrations of TBG. It parallels free T₄ concentrations during pregnancy (8), is unaffected by free fatty acid interactions (9), and shows negligible influence from variations of concentrations in serum of albumin, thyroxin-binding prealbumin, and TBG (10) in nonthyroidal illness.

Materials and Methods

Computer Methods

All laboratory results were stored in a customized database utilizing Foxbase+ (Fox Software Inc., Pelham, OH), running on a Token-Ring network of IBM PS/2 computers. A computer search was carried out to identify all records in a particular period (the previous six months) for which the complete set of T₄, T₃, TBG, and TSH estimates were available. This search established a pool of 3436 records.

A computer-generated reference group for T₄ was isolated by deleting all records with TSH values or T₄/TBG ratios outside the current laboratory reference ranges. The 2001 records that remained were designated the T₄ CGR Group.

A copy of the file containing the initial pool was processed separately to obtain a T₃ CGR Group (1902 records with normal TSH and T₄/TBG ratio). The TSH CGR Group (1963 records) was acquired by deleting all records with T₃/TBG ratio or T₄/TBG ratio outside the laboratory reference ranges (3.14–6.55 and 0.045–0.096 μmol/g, respectively). The data files containing the CGR Groups were converted into ASCII text files and then imported into the statistics program Minitab (Minitab Inc., State College, PA). The data were then evaluated for mean, standard deviation, estimates of skewness, etc. Histograms were printed to provide visual output of the distribution of analytical values (Figure 1). The running time for the program to accumulate the patients’ records from a laboratory database of 70 000 records and generate the ASCII files containing the CGR Groups was ~30 min.

Details of the program used can be obtained from the corresponding author.

Immunoassays

The following immunoassay kits were used to generate the CGR Groups: Total T₄ was measured with Clinical Assays Gammacoat RIA kits (Baxter Healthcare Corp., Cambridge, MA). The lower limit of detection was 13 nmol/L and the interassay coefficient of variation (CV) was 19, 44, and 100 nmol/L was 12.2%, 10.2%, and 9.0%, respectively. The current laboratory reference range is 60–144 nmol/L.

For total T₃, we used an Amerlex-M RIA kit (Amer- sham, Australia, North Ryde, N.S.W.). It had a detection limit of 0.5 nmol/L and CVs of 11.2%, 8.7%, and 7.1% at analyte concentrations of 0.9, 2.0, and 3.3 nmol/L, respectively. The laboratory’s current reference range is 0.8–2.2 nmol/L.

The TBG assays were performed with Immophase kits (Corning Medical, Medfield, MA). The detection limit was 10 mg/L, and the CVs at 14, 22, and 30 mg/L were 9.3%, 10.9%, and 9.8%, respectively. The laboratory’s
current reference range for subjects who are neither pregnant nor receiving estrogens is 11.5–32 mg/L.

TSH was assayed with RIABEAD II solid-phase immunoradiometric assay kits (Abbott Diagnostics Division, Australia, North Ryde, N.S.W.). This assay had a detection limit of 0.1 milli-int. unit/L, with CVs of 6.5%, 4.9%, and 5.2% at analyte concentrations of 2.8, 17, and 40 milli-int. units/L, respectively. The manufacturer’s recommended range of 0.34–3.5 milli-int. units/L was used as the laboratory reference range at the time of this study.

The T₄/TBG and T₃/TBG ratios were obtained by dividing total T₄ and T₃, respectively, by TBG and expressing the results as µmol/g. The laboratory reference ranges were 3.14–6.55 and 0.045–0.096 µmol/g, respectively.

Manufacturers’ Reference Ranges

Each of the manufacturers of the kits listed above quoted suggested reference ranges in their pamphlets accompanying the kits. All made clear, however, that these are provided as a guide only, and that individual users should develop their own ranges.

For the T₃ assay, Amersham Australia supplied the results of studies on 182 euthyroid subjects, and Abbott supplied results for similar studies of 505 euthyroid subjects as the basis for their respective reference ranges. Neither manufacturer detailed the method of calculating the reference range from these data.

With the T₄ assay, Baxter refers the user to the 1969 report of Ekins et al. (11). However, those workers used not the Baxter kit but rather a competitive protein-binding assay involving transcortin. They obtained a reference range for T₄ of 45–115 µg/L (58–148 nmol/L) for 152 euthyroid adults living in England. The booklet accompanying the Baxter assay indicates that this range has also been confirmed for the Boston area.

Subjects

Patients. The laboratory database contained results of tests on all patients’ specimens analyzed in the Endocrinology Laboratories of this hospital. The majority were from ambulatory patients attending outpatient clinics, with a smaller number coming from ward patients and from external centers. Because the Hospital is a university teaching center, many of the patients were secondary referral cases who had been previously investigated elsewhere.

Healthy subjects. Blood samples were drawn from a group of 97 subjects ("Normal Group"), who were not known to have any endocrine abnormalities. The group comprised 69 women and 28 men, ages 18 to 80 years. The sera were analyzed for T₃, T₄, TBG, and TSH.

Results

Histograms of the T₃, T₄ and TSH CGR Groups are presented in Figure 1. The T₃ and T₄ groups appear to have a gaussian distribution, which was verified by deriving normal scores for the data. Comparison with the T₃ and T₄ distributions of the Normal Group yielded very similar means and standard deviations (Table 1). For T₄, the reference range derived from the Normal Group (mean ± 2 SD) agrees more closely with that from the CGR Group than with the range recommended by the manufacturer. For T₃, all three estimates of the lower limit of the reference agreed well, but the CGR Group provided a slightly higher upper limit (161 nmol/L) than either the Normal Group or the manufacturer’s range (144 and 148 nmol/L, respectively).

By contrast, the distribution of the TSH CGR Group was broadened and skewed, with 9.5% of data points exceeding the upper limit of the manufacturer’s recommended normal range and 13% falling below the lower limit. The range of values was much broader than was observed for the Normal Group, which, although also skewed, had the central 95 percentiles (0.3–2.4 milli-int. units/L) within the limits of the manufacturer’s recommended range (0.3–3.5 milli-int. units/L). Because of the divergence between the computer-generated distribution of TSH values and that for the Normal Group, we abandoned attempts to obtain a TSH reference range from the CGR Group.

The stability of the reference ranges with time was tested by separately generating CGR Groups for each of nine sequential two-month intervals (Figure 2). In the examples shown, the CVs of the mean T₃ and T₄ values for the CGR Groups were 4.0% and 2.2%, respectively, indicating that the performance of the assays was stable over this period.

Discussion

Thyroid Status and Reference Ranges

The computer approach to the determination of reference ranges described in this study requires that a "normal" group of subjects must be identified by a means independent of the analytical method under consideration.

With the increased sensitivity of immunoradiometric assays (and equivalent methods with nonisotopic labels) for TSH, one can now clearly distinguish between hyperthyroid, euthyroid, and hypothyroid subjects. Thus, CGR Groups of apparently normal subjects can be iso-

**Table 1. Reference Ranges for T₃, T₄, and TSH Assays Compared by Subject Groups**

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>Range*</th>
</tr>
</thead>
<tbody>
<tr>
<td>T₄ CGR Group</td>
<td>2001</td>
<td>108.5</td>
<td>26.2</td>
<td>56–161</td>
</tr>
<tr>
<td>T₄ normals</td>
<td>97</td>
<td>102.0</td>
<td>21.1</td>
<td>60–144</td>
</tr>
<tr>
<td>Manufacturer’s range</td>
<td>152</td>
<td></td>
<td></td>
<td>58–148</td>
</tr>
<tr>
<td>T₃ CGR Group</td>
<td>1902</td>
<td>1.40</td>
<td>0.35</td>
<td>0.7–2.1</td>
</tr>
<tr>
<td>T₃ normals</td>
<td>97</td>
<td>1.50</td>
<td>0.34</td>
<td>0.8–2.2</td>
</tr>
<tr>
<td>Manufacturer’s range</td>
<td>182</td>
<td></td>
<td></td>
<td>0.8–2.7</td>
</tr>
<tr>
<td>TSH CGR Group</td>
<td>1953</td>
<td>1.7</td>
<td>2.3</td>
<td>&lt;0.1–7.5</td>
</tr>
<tr>
<td>TSH normals</td>
<td>97</td>
<td>1.4</td>
<td>1.0</td>
<td>0.3–2.4</td>
</tr>
<tr>
<td>Manufacturer’s range</td>
<td>505</td>
<td>1.32</td>
<td>0.34</td>
<td>0.34–3.5</td>
</tr>
</tbody>
</table>

*For the T₃ and T₄ assays, reference ranges were determined as mean ± 2 SD (units: milli-int. units/L). For the TSH range the central 95 percentiles are quoted (units: milli-int. units/L) because the distribution was skewed.
lated from the patient database and appear suitable for generating reference ranges for the $T_3$ and $T_4$ assays.

However, a substantial proportion of patients having $T_3$/TBG and $T_4$/TBG ratios within their respective reference ranges had apparently inappropriate concentrations of immunoreactive TSH. Because the pattern of distribution of values in the TSH CGR Group did not conform to that of the Normal Group, or to the manufacturer's reference range, we did not use the computer method to generate a laboratory reference range for TSH.

The dissociation between the TSH distributions for normal subjects and apparently euthyroid patients is curious and is currently the subject of a further investigation. A similar observation was made by Spencer et al. (12), who studied 1580 hospitalized patients and found that a substantial proportion with above-normal TSH did not have thyroid disease. They attributed their findings to several causes, including glucocorticosteroid therapy and nonthyroidal illness.

Our findings also appear to lend some support to the use of TSH as a "front-line" screening test of thyroid function. Thus patients with TSH values within the normal range could be predicted to also have normal values of thyroid hormones. Under this approach, follow-up testing, including thyroid hormone estimations, would be reserved for cases involving abnormal concentrations of TSH. However, these observations are based on data from the patient population of one hospital laboratory only and may not reflect the situation in other centers.

The method may also be applicable to other analytes. Presumably, reference ranges for free $T_3$, free $T_4$, and free thyroxin index might be derived by a process analogous to the method used above. For example, for a laboratory using the combination of TSH, free $T_3$, and free $T_4$ assays, a pool of patients with normal free $T_3$ and TSH values may provide a suitable CGR Group for the derivation of a free $T_4$ reference range. Similarly, a cortisol reference range might be derived from patients with normal corticotropin values, although complicating factors such as the circadian rhythms of corticotropin and cortisol may result in reference ranges that vary throughout the day.

Further, although the successful use of the method described here generated gaussian-distributed CGR Groups for $T_3$ and $T_4$, the method probably would also be applicable to populations that are not gaussianly distributed. Studies by Shultz et al. (13), Hyltoft-Petersen et al. (14), Linnet (15), and others have described non-parametric methods for assigning reference intervals to such data. Note that our failure to obtain a reference range for TSH was not due to the skewness of the data, but rather to the apparent divergence of the TSH CGR Group and the "normal" healthy populations.

Strategy for Routine Use

The method for computer generation of reference ranges is not intended to displace conventional approaches, based on the assay of specimens from selected healthy subjects. These approaches should still be carried out at regular intervals. Initially, the computer method would be run at the same time as studies of healthy subjects, as in the present study, to verify that appropriate results are obtained for a given assay. It can then be repeated at intervals (e.g., every two to six months).

If a significant drift in reference ranges is detected in this way, the laboratory will be alerted to correct any faults that may have developed with the analytical method, or to assemble a new group of normal healthy subjects so that the reference range can be reset. The practice of continually resetting the reference ranges based on data from previous computer studies must be avoided, because a progressive systematic error may be introduced.

A better approach is to perform each successive computer analysis with use of constant values for the exclusion limits by which the CGR Groups are selected. These limits should be derived by noncomputer methods (e.g., manufacturer's recommended ranges or the most recent ranges obtained with healthy subjects).

In conclusion, we do not suggest that a diagnostic laboratory should rely solely on computer-generated reference ranges. The values obtained by assaying sam-
False