Simple Modifications of Three Routine in Vitro Tests of Thyroid Function

Roy W. Pain

Semi-automation of equipment and simple modifications of technique reduced the work load without loss of diagnostic accuracy for three commonly used in vitro tests of thyroid function (total thyroxine, thyrobinding index, and free thyroxine index). Major innovations were the use of serum standards for all tests and having each duplicate for tests performed by a different technician. Attention is drawn to the false-positive and false-negative errors that occur when the 95% euthyroid limits is the sole reference range used.

Additional Keyphrases: total thyroxine • thyroxine binding index • free thyroxine index • diagnostic aid • thyroid status • normal values

Measurement of the total thyroxine concentration in serum and of the number of unoccupied thyrobinding sites on the serum proteins are both well established in in vitro tests of thyroid function. However, each test is influenced markedly by abnormalities of the thyrobinding proteins, which can be caused by drugs, disease, and genetic factors. This is a major problem; 31% of one large series of sera from euthyroid patients had evidence of thyrobinding abnormalities (1). To prevent the misdiagnosis of thyroid status in patients with abnormalities of thyrobinding proteins, the two measurements mentioned above are frequently combined to produce a mathematical index that is directly proportional to the concentration of the unbound or "free" thyroxine. This "free thyroxine index" gives a very good assessment of thyroid function (2-5).

Below are reported several modifications of equipment and technique for the above tests, which reduced the laboratory workload and ensured the accuracy and reproducibility of results. Attention is drawn to the diagnostic limitations of conventional 95% reference ranges.

Methods and Materials

Total Thyroxine (total T₄)

This was measured by the competitive protein-binding method of Murphy and Jachan (6) with the following modifications.

A multiple channelled aluminum evaporator block, 25 cm × 23 cm × 2.4 cm, was constructed with 90 nozzles to which were fitted 21-gauge needles cut to 1.25 cm length. This evaporator was similar to that used by Murphy and Jachan (6) except that it was larger and stood on legs 6 cm high so that it fitted over a metal rack containing up to 90 sample tubes, which were then evaporated simultaneously in a 45 °C water bath. Anion-exchange resin (Dowex 2 × 8, 20–50 mesh, Cl⁻ form), partly dried by exposure to room temperature (21 °C) for 48 h, was added by a Perspex (methyl methacrylate polymer) multiple resin dispenser similar in principle to, but larger than, one previously described (7). This consisted basically of two mounted and slotted plates of heavy Perspex (methyl methacrylate polymer; 28 × 25 × 1.3 cm and 27 × 23 × 1.2 cm), each containing 90 holes, 0.37 cm in diameter and of 0.54 ml capacity. The bottom plate was movable over a distance of 1.25 cm so that the two sets of holes in the plates could be either in or out of line. With the holes out of line, the holes in the upper plate were filled with resin from a wide-stemmed filter funnel with the stem cut short. A plastic scraper was then run over the upper Perspex plate, to wipe off any excess resin. When the metal rack holding up to 90 tubes was placed beneath the resin adder, resin was added simultaneously to all tubes by sliding the top plate along so as to oppose the two sets of holes in the two plates.

A mechanical shaker, to which the metal tray holding up to 90 tubes could be quickly clamped, was used to shake the tubes (750 oscillations per minute) at various stages. The shaker was fitted with a counter weight and top and bottom ball races to prevent ejection of the contents of the tubes. The designs of the evaporator block, resin dispenser, and mechanical shaker were suggested to me in 1968 after seeing similar equipment in use in the laboratory of Professor D. Curnow at the Royal Perth Hospital, Western Australia. Murphy has also described a mechanical shaker (6).

An Auto-Spenser diluter-dispenser (AO Instr. Co., American Optical Corp., Richmond, Calif.) was used in the final step to remove 1.5-ml aliquots of liquid sample into counting tubes, washing out with water between samples to prevent carryover of isotope.

At the introduction of this method serum standards were run, as well as standards of L-thyroxine (Sigma Chemical Co., St. Louis, Mo.), dissolved in alkaline ethanol. The original serum standards were made from a pool of serum consisting of 4.0 ml from each of the 100 male blood donors on whom the reference range was determined. The value of this pool when run against the thyroxine standards proved to be the midpoint of the reference range. The pool was dispensed into 10-ml
 aliquots, which were stored at −20 °C and used before the third day after thawing. This pooled serum was used to construct standard curves in the following manner. In the basic technique (6), after ethanol precipitation, 0.50 ml of ethanolic extract is used for the total T₄ assay. By using 0.25, 0.50, and 0.75 ml of extract, each in triplicate, from the standard serum, three points were obtained to construct a standard curve. Because the total thyroxine value of 0.50 ml of extract was 100% of the total thyroxine concentration of the serum, the values of 0.25 and 0.75 ml of extract were therefore 50% and 150% of this concentration, respectively. For reasons elaborated on in the Discussion section the serum standards proved superior to those of 1-thyroxine, and so the latter were eventually discontinued. New standard serum pools were made three-monthly from 4.0-ml aliquots from each of 100 male blood donors. Thirty aliquots of every new standard serum pool were checked to ensure calibration against the preceding pool.

Male blood donors were chosen to establish reference ranges for all tests. Nonpregnant women not taking estrogen preparations gave identical reference ranges to males, and so the latter were used to avoid estrogen effects. Euthyroid patients suffering from nonthyroidal disease states were not chosen because they might have been receiving medications or suffering from disorders that cause abnormalities of thyrobinding proteins.

Because the borderline areas at the lower and upper end of the reference range are critical for diagnostic discrimination, quality control sera with upper and lower borderline values for total thyroxine were prepared, the first by diluting pooled euthyroid serum with physiological saline so that the total T₄ value was in the lower borderline area, the second by rapidly freezing and slowly thawing serum from the same pool and then removing the diluted supernate. This entire procedure was repeated once. The resultant concentrated serum had a value for total T₄ in the upper borderline range. Each routine batch contained 38 specimens composed of 25 unknowns, three standards in triplicate, and the two quality controls in duplicate.

**Thyrobinding Index (TBI)**

I measured the concentration of free thyrobinding sites on the serum proteins by a kit technique (¹²³I-Tri-Ionex; Curtis Nuclear Corp., St. Louis, Mo.), using only 0.20 ml of serum sample instead of the 1.0 ml recommended by the manufacturer. It had been previously established in this laboratory that such a modification for this and a similar kit procedure (8) significantly improved diagnostic discrimination, perhaps because the greater dilution of serum diminished binding to albumin (9), thus giving a purer measure of thyrobinding globulin (TBG), prealbumin binding having also been inhibited by the buffer used. A thyrobinding index (TBI) was obtained by comparing the amount of ¹²³I-labeled triiodothyronine taken up by the test serum with that taken up by the standard serum. TBI, in contrast to the T₃ resin uptake, is increased in hypothyroidism and decreased in hyperthyroidism.

An Auto-Spenser was used to dispense 3.0 ml of buffer reagent and 0.20 ml of serum sample, in one procedure, into the resin vials. An elongated rotator, 56 cm long with a 15-cm square cross section and a central spindle through the long axis, rotated up to 200 specimens simultaneously at six revolutions per minute. The Auto-Spenser was then used to remove 1.5-ml aliquots for counting as for the total T₄ assay. Six samples of standard serum and two quality-control sera in duplicate were run with each batch of 25 patient sera. The standard and quality-control sera were the same as used in the total T₄ assay. By noting the average time in minutes for the six standard sera to yield 10 000 counts and then resetting the radioactive counter to count all specimens for this length of time, the number of counts for each specimen was 10 000 times its TBI value. This simple maneuver obviated the laborious calculations of the ratio of the result of each test to that of the standard.

The reference range for the TBI assay was established on the same 100 male blood donors used to establish the reference range for the total thyroxine estimation. The original pooled donor serum was used as the standard and was assigned an arbitrary value of 1.00. Each new pool of donor serum was run against the previous serum standard to establish its value.

**Free Thyroxine Index (FTI)**

Because of the equilibrium free thyroxine + unoccupied carrier protein = bound thyroxine, [free thyroxine] = k[bound thyroxine]/[unoccupied carrier protein], where k is the equilibrium constant. Bound thyroxine accounts for 99.95% of the total circulating thyroxine, and the TBI is directly proportional to the concentration of free thyrobinding sites. Therefore: free thyroxine index = total thyroxine/thyrobinding index = total T₄/TBI.² Low results are seen hypothyroidism and abnormally high results in hyperthyroidism.

**Reproducibility of Results**

Every specimen was tested by two technicians, each performing single estimations of total T₄ and TBI in independent batches. This manner of duplication therefore allowed for both between-technician and between-batch variation. Each individual technician's within-batch and between-batch coefficients of variation of total T₄ and TBI were calculated at monthly intervals from the results of the quality-control sera. All results for the quality-control sera duplicates were included, even those from batches that had to be repeated because of unsatisfactory quality-control results from any cause. The computer was programmed to perform a preliminary analysis and then to reanalyze the data after excluding outlying results more than three standard deviations from the mean.

**Diagnostic Accuracy**

The modified procedures were assessed on 89 patients posing diagnostic problems, who passed consecutively through a thyroid unit (10), six of whom were known to be taking drugs that affect thyrobinding proteins. The definitive clinical diagnosis was established in each instance by a team consisting of an endocrinologist, two physicians in nuclear medicine, and a medically qualified clinical chemist. The information used to make the diagnosis included clinical history, clinical examination, thyroidal uptakes and scans using both ⁹⁹ᵐTc-pertechnetate and ¹³¹I, serology, and biochemistry. In 16 instances the response to further tests (eight thyroidropin stimulation, six T₃-suppression, and two perchlorate discharge tests) were necessary for a definitive diagnosis.

**Results**

**Reference Ranges**

The distribution of the 100 results for each of the three tests was gaussian in each instance. With standards of L-thyroxine, the mean total T₄ was 92 µg/liter (SD, 14.0 µg/liter; 95% range, 64–120 µg/liter). The mean TBI was 1.00, the SD 0.075, and the 95% range 0.85–1.15. The mean FTI was 9.3, the SD 1.50, and the 95% range 6.3–12.3.

²Note that the value for the FTI depends on the units in which total thyroxine is expressed. It will be 10-fold greater for SI units, µg/liter, than for µg/dl for example, the "traditional" units.
Table 1. Diagnostic Accuracy of Three Thyroid Tests, with Use of 95% Reference Limits

<table>
<thead>
<tr>
<th>Definitive clinical diagnosis (and no. of cases)</th>
<th>Total T4</th>
<th>TBI</th>
<th>FTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothyroid (7)</td>
<td>5</td>
<td>71</td>
<td>4</td>
</tr>
<tr>
<td>Euthyroid (52)</td>
<td>49</td>
<td>94</td>
<td>39</td>
</tr>
<tr>
<td>Hyperthyroid (24)</td>
<td>22</td>
<td>92</td>
<td>18</td>
</tr>
<tr>
<td>Total (83)</td>
<td>76</td>
<td>92</td>
<td>61</td>
</tr>
<tr>
<td>Euthyroid* (58)</td>
<td>49</td>
<td>84</td>
<td>40</td>
</tr>
<tr>
<td>Total* (89)</td>
<td>78</td>
<td>88</td>
<td>59</td>
</tr>
</tbody>
</table>

* Including six euthyroid patients taking drugs that can affect thyrobinding proteins.

Reproducibility of Results

The mean and individual between-batch coefficients of variation (CV) for five technicians working in the thyroid laboratory for periods of three to five months are shown in Figures 1 and 2. During the first month of a technician's stay in the thyroid laboratory, his mean within-batch CV was on the average 0.8% lower than the mean between-batch CV. However, with time the difference between the within-batch and between-batch means regularly disappeared.

Diagnostic Accuracy

Of the 83 consecutive patients not on drugs affecting thyrobinding proteins assessed by the thyroid unit, 52 were considered euthyroid, 24 hyperthyroid, and seven hypothyroid. Using the derived 95% normal reference ranges, the total T4, TBI, and FTI agreed with the definitive clinical diagnoses, as shown in Table 1. Results that include the six euthyroid patients taking drugs that can affect thyrobinding proteins are also shown in this table.

Discussion

Semi-automation of the addition of resin in the total T4 method helped simplify and standardize that method. In this procedure serum standards were preferable to L-thyroxine standards on theoretical grounds, because they went through every step of the procedure including ethanolic deproteinization and extraction, and in practice thyroxine standards occasionally produced results for quality-control and test sera that were too high. Murphy (11) has commented that the occasional lack of correspondence between thyroxine standards and test samples seemed to be related to drying the standards in the absence of serum extract. Others have shown that ethanol extracts significant amounts of thyrobinding globulin from test sera and thus serum standards should be used (12).

Reading the TBI result directly off the printout from the counter saved time and avoided calculating errors. We reported the thyrobinding index rather than the T3-resin uptake because the concept of the TBI was easier for the clinician to grasp than that of the T3-resin uptake. (TBI is directly, and the latter is inversely, proportional to the serum concentration of unoccupied thyrobinding sites.)

Figures 1 and 2 show that the CV varied with the concentration of substance being tested, the technician, and the length of time the technician had spent performing that technique. Thus quoting a single figure for the CV of a method gives an incomplete picture of the reproducibility obtained. The CVs obtained compared favorably with those for other published similar methods (13-17) despite including in the...
analysis every result obtained for the quality-control sera duplicates, even those from batches that had to be repeated because of unsatisfactory quality-control results from any cause. Because each individual result for any technician was always matched with the corresponding result of a second technician, obtained in an independent batch, a new technician's proficiency could be easily assessed before he was permitted to test routine batches. Some technicians achieved good reproducibility almost immediately; others, although their reproducibility was acceptable, took two or even three months to achieve target reproducibility and very occasionally a technician had to be removed to another laboratory because of inability to obtain acceptable reproducibility. Most showed continuous improvement with time, but occasional technicians later relapsed because of boredom. Thus it was inadvisable to rotate technicians either too frequently or too infrequently through this laboratory.

Because we soon found that a common cause of unacceptable results was failure of adequate maintenance of equipment, especially the cleaning and lubrication of dispensers and dilutors, we made it a standing rule to routinely service all equipment at the beginning of each month. False results caused by isotope contamination were prevented by establishing three distinct areas in the thyroid laboratory, each with its own drainage sink. One area was reserved for centrifuging, preparing, and storing sera; a second for the actual testing procedures; and a third for disposal of isotope wastes. The laboratory was so arranged that fresh sera or reagents were never brought into the third area.

The 95% reference range for total T₄ obtained on blood donors and corrected for 80% extraction efficiency (64-120 µg/liter) was narrower than that "arbitrarily accepted" by Murphy et al. (18) for euthyroid patients, i.e., 52-143 µg/liter when corrected for 77% extraction efficiency. Other workers have obtained corrected reference ranges of 48-166 µg/liter (19) and 49-129 µg/liter (20) by using similar basic techniques. These few instances again show that each laboratory must establish its own reference range for every test.

The assessment of the diagnostic accuracy of any test procedure is likely to vary with the degree of selection and the precision of diagnosis of the abnormal subjects tested. The subjects included in this report were a consecutive series of patients, most of whom had only minor degrees of thyroid abnormality and some of whom required sensitive dynamic tests for definitive diagnosis. Even so, the diagnostic accuracy of the described triad of tests compared favorably with other series where a similar combination of tests was used (2, 21). The value of the FTI in the presence of abnormities of thyro-binding proteins was once again demonstrated. However, because 95% reference limits obtained on normal blood donors were used in this analysis, some euthyroid patients had results falling outside these limits (and therefore were misclassified as mildly abnormal, a positive error) while some patients with mild thyroid dysfunction had results within these limits (and therefore were misclassified as euthyroid, a negative error). A new diagnostic system was therefore devised which utilized borderline zones to allow for the overlap which always occurs between extreme euthyroid and mildly abnormal results. With these new diagnostic limits the total T₄, TBI, and FTI results were consistent with the definitive clinical diagnosis in 96%, 92%, and 100%, respectively, of the 83 patients assessed. (97%, 91%, and 100% when the 6 patients taking drugs which can affect thyro-binding proteins were included.) The diagnostic discrimination with these new limits was thus significantly better than that obtained by using conventional 95% limits. Details of these improved diagnostic limits are the subject of another publication (22).

I believe that the modifications reported here can be of value to other laboratories using similar methods, because they simplify the overall procedures, ensure their reproducibility, and do not reduce their diagnostic accuracy.

I gratefully acknowledge the valuable suggestions made by Dr. P. J. Phillips while I was writing this paper.

References

1. Pain, R. W., Thyroid function tests. The Royal College of Pathologists of Australia, Test and teach programme, 1972, p 33.