Clinical Strategies in the Testing of Thyroid Function

Last Updated: June 1, 2011

Authors

• Jim Stockigt M.D., FRACP, FRCPA
  Monash University and Alfred and Epworth Hospitals Melbourne, Australia Tel: +613-9276-2460 Fax: +613-9276-3782

In the past decade, emphasis has shifted from testing of thyroid function in individuals who are likely to have a thyroid disorder to a broader population, an approach that identifies so-called subclinical thyroid dysfunction in up to 10% of women over fifty. The key assays that are used to detect thyroid dysfunction are serum thyroid stimulating hormone (TSH) and the main circulating thyroid hormones thyroxine (T4) and triiodothyronine (T3), either as total or estimated free concentrations. For these key parameters, it is preferable for results to be interpretable in relation to reference intervals that are independent of the particular method that is used. This requirement is satisfied for well standardized assays for serum TSH, total T4 and total T3. However, when free T4 is estimated, assay results often need to be evaluated in relation to method-specific reference intervals or “normal ranges”. This limitation is particularly cogent during pregnancy and in the face of critical illness (see 6.4.2 and 6.7.2 below).

The widespread search for thyroid dysfunction is influenced by two key questions. How damaging are the effects of subclinical dysfunction? Does treatment confer benefit? The answers are not uniform across populations. The diagnostic imperative is now radically different for the population at large and for women who are pregnant or about to become pregnant. Some key practice points that relate to the testing of thyroid function are summarized in box 1.

Key Practice Points

Many disorders are associated with increased prevalence of thyroid dysfunction; optimal testing strategy requires information on all co-existing conditions and medications.

The more widely thyroid function is tested, the greater the proportion of abnormal results that show borderline or “subclinical dysfunction”.

Testing of thyroid function is now widely advocated before and early in pregnancy, especially where fertility is impaired, assisted reproduction is used, or where pregnancy complications have occurred.

Except for “subclinical” hypothyroidism in early or impending pregnancy, intervention for subclinical thyroid dysfunction should only be considered after a sustained abnormality has been demonstrated over a minimum of three months.

There is increasing documentation of adverse effects from sustained or progressive subclinical
hypothyroidism, although evidence of benefit from treatment is less clear.

The combination of raised serum TSH and positive peroxidase antibody is predictive of likely long-term progression towards overt hypothyroidism.

The relationship between serum TSH and circulating thyroid hormones gives a better index of thyroid status than any single parameter. Six key assumptions underpin the diagnostic value of this relationship.

The “TSH alone” first line approach to thyroid function testing has important drawbacks and limitations.

Serum TSH is a cornerstone of thyroid diagnosis, but it is not possible to define its “normal” range or reference interval for all clinical circumstances. Age, pregnancy and fertility issues, diurnal variation and pulse secretion, and associated antibody status militate against fixed cut-off points.

Both TSH and T4 show spontaneous fluctuations that are much greater than analytical imprecision. A serial change can be inferred with confidence with about 50% alteration in serum TSH and 25% change in the free T4 estimate.

During thyroid hormone therapy, either replacement or suppressive, the optimal target TSH range may differ from the reference interval that is used to establish a new diagnosis.

Interpretation of anomalous test results should take into account the effects of all associated medications.

No estimate of circulating free thyroxine is impeccable. Especially in situations where assessment is difficult, eg late pregnancy and severe associated illness, there are strong arguments for re-establishing total T4 measurement as the preferred “gold standard”.

Identification of marked iodine excess, detected by urinary estimation, may identify important reversible thyroid abnormalities, eg iodine-induced exacerbation of primary hypothyroidism, atypical thyrotoxicosis with blocked isotope uptake, or resistance to standard doses of antithyroid drugs. Some food sources (eg soy, sea weed) and alternative health care products can be heavily iodine-contaminated.

1 Who should be tested for thyroid dysfunction?

It has long been recognized that the clinical manifestations of hyperthyroidism (thyrotoxicosis), or hypothyroidism are so diverse that diagnosis based on clinical features lacks sensitivity and specificity. Hence, reliance is placed on measurements of circulating thyroid hormones and thyroid stimulating hormone (TSH) to confirm or rule out thyroid dysfunction (TD).

After the publication of guidelines from the American College of Physicians in 1998 (1, 2) testing for detection of TD became widely applied, especially in women over 50, the group most likely to have either overt or subclinical thyroid dysfunction. Testing of this group is generally advocated at the time of presentation for medical care, i.e. a case finding strategy, rather than screening of a whole population.
A normal serum TSH value in ambulatory patients without associated disease or pituitary dysfunction has a high negative predictive value in ruling out both primary hypothyroidism and hyperthyroidism (1, 2), which has led to a short-cut approach in which free T4 may only be estimated if TSH is abnormal. However, this “TSH first” strategy of thyroid function testing has important limitations (see 5.1.1 below). If there is no suspicion of pituitary or thyroid disease, a normal TSH value need not be re-tested for about 5 years (3). To this broad indication has recently been added the still controversial recommendation that universal testing of thyroid function has a place before or as early as possible in pregnancy (4).

In some groups (table 1), known to be at increased risk of TD, there is a case for routine testing even in the absence of any suggestive clinical features.

<table>
<thead>
<tr>
<th>Table 1 Groups with an increased likelihood of thyroid dysfunction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous thyroid disease or surgery</td>
</tr>
<tr>
<td>Goiter</td>
</tr>
<tr>
<td>Associated autoimmune disease(s) (5, 5a)</td>
</tr>
<tr>
<td>Diabetes mellitus, type 1 (5,6)</td>
</tr>
<tr>
<td>Celiac disease (7,7a)</td>
</tr>
<tr>
<td>Scleroderma (8)</td>
</tr>
<tr>
<td>Chromosome 18q deletions (15)</td>
</tr>
<tr>
<td>Chronic renal failure (16)</td>
</tr>
<tr>
<td>Williams syndrome (17)</td>
</tr>
<tr>
<td>Fabry’s disease (18)</td>
</tr>
<tr>
<td>Irradiation of head and neck (19,20)</td>
</tr>
<tr>
<td>Radical laryngeal/pharyngeal surgery</td>
</tr>
<tr>
<td>Recovery from Cushing’s syndrome (21,22)</td>
</tr>
<tr>
<td>Gout (22a)</td>
</tr>
<tr>
<td>Environmental irradiation ? (23)</td>
</tr>
<tr>
<td>Thalassemia major (24)</td>
</tr>
<tr>
<td>Primary pulmonary hypertension (25)</td>
</tr>
<tr>
<td>Polycystic ovarian syndrome (26)</td>
</tr>
<tr>
<td>Endometriosis (26a)</td>
</tr>
<tr>
<td>Drug therapy</td>
</tr>
<tr>
<td>• Amiodarone (27)</td>
</tr>
<tr>
<td>• Lithium (28)</td>
</tr>
<tr>
<td>• Thalidomide (29)</td>
</tr>
<tr>
<td>• Chemotherapy for sarcoma (30)</td>
</tr>
<tr>
<td>• Stavudine, ? other potent retroviral agents (31)</td>
</tr>
<tr>
<td>• Tyrosine kinase inhibitors (32)</td>
</tr>
<tr>
<td>• Sunitinib, Imatinib, Motesanib, Sorafenib</td>
</tr>
<tr>
<td>Sjögren syndrome (9)</td>
</tr>
<tr>
<td>Morbid obesity ? (10)</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>Breast cancer (11)</td>
</tr>
<tr>
<td>Hepatitis C (pre-treatment) (12)</td>
</tr>
<tr>
<td>Down’s syndrome (13)</td>
</tr>
<tr>
<td>Turner’s syndrome (14)</td>
</tr>
<tr>
<td>Biological agents</td>
</tr>
<tr>
<td>• Interferon alpha (33)</td>
</tr>
<tr>
<td>• Ribavirin (34)</td>
</tr>
<tr>
<td>• Interferon beta (35)</td>
</tr>
<tr>
<td>• Interleukin-2 (36)</td>
</tr>
<tr>
<td>• Therapeutic use of antibodies (37)</td>
</tr>
<tr>
<td>• Growth hormone treatment (38)</td>
</tr>
<tr>
<td>• Denileukin (39)</td>
</tr>
<tr>
<td>Pituitary or cerebral irradiation (40)</td>
</tr>
<tr>
<td>Head trauma (41)</td>
</tr>
<tr>
<td>Very low birth weight premature infants (42,43)</td>
</tr>
</tbody>
</table>

Almost all developed countries now have routine neonatal screening programs for congenital hypothyroidism using heel prick filter paper blood spots (see chapter 15b). The value of such programs has long been clear (44), but neonatal screening is not yet routine in numerous developing countries where the prevalence of neonatal hypothyroidism may be high, often associated with iodine deficiency (45) (see chapter 20). In terms of benefit from allocation of health care resources in developing countries, the establishment of neonatal screening (46) probably takes precedence over routine testing of adults.

### 1.1 The basis for a case-finding strategy

Routine laboratory testing of particular population groups becomes well founded if a testing strategy satisfies the following criteria:

a. An abnormality cannot be identified in a reliable and timely way by standard clinical assessment.

b. Dysfunction is sufficiently common to justify routine testing, either by case finding or by population screening.

c. There are adverse consequences of failure to identify an abnormality, including the possibility of progression towards more severe disease.

d. The laboratory test method is cost-effective and sufficiently sensitive and specific to identify those at risk of adverse consequences.

e. There are no major adverse consequences of testing

f. Treatment is safe and effective and prevents some or all of the adverse consequences.

g. Abnormal findings can be adequately followed-up to ensure an appropriate clinical response. (An early detection program may have little value if this last requirement cannot be met.)

While the first five of the above criteria are reasonably established, the last two are less secure.
1.2 Sensitivity and accuracy of clinical assessment

Studies of unselected patients evaluated by primary care physicians show that clinical acumen alone lacks both sensitivity and specificity in detecting previously undiagnosed TD. In two Scandinavian studies of over 3000 unselected patients who were assessed by both clinical and laboratory criteria, a thyroid disorder was not suspected by primary care physicians in over 90% of those who tested positive, even when clinical features were apparent in retrospect (47,48). Furthermore, in up to one-third of patients evaluated for suspected thyroid dysfunction by specialists, laboratory results led to revision of the clinical assessment (49). Systematic comparison of the standard clinical features of hypothyroidism with laboratory tests (50) showed that clinical assessment identified only about 40% with overt hypothyroidism and that classical signs were present only in those severely affected. Both overt hyper- and hypothyroidism can have important consequences before the usual clinical features are obvious, and clinicians may fail to recognize diagnostic features even when they are present (47,48).

Boelert et al (50a) have recently confirmed that the typical multiple classical symptoms of hyperthyroidism becomes less prevalent with advancing age, with greater importance of weight loss, atrial fibrillation and shortness of breath as presenting features (50a). They propose a low threshold for assessment of thyroid function in patients older than 60 years who have any of these features.

Clinical evaluation remains of central importance to assess severity of TD, evaluate discordant results, establish the specific cause of TD and monitor the response to treatment. There is little doubt that repeated laboratory confirmation of normal thyroid function can be wasteful; strategies have been suggested to improve cost-effectiveness (51,52).

1.3 Prevalence

In considering the prevalence of thyroid dysfunction, a distinction needs to be made between so-called subclinical and overt abnormalities; paradoxically, this distinction is based on laboratory rather than clinical criteria. There is a trend to replace the term ‘subclinical hypothyroidism’ with the designation ‘mild thyroid failure’ (53).

In the progressive development of TD, abnormal values for serum TSH generally occur before there is a diagnostic abnormality of serum T4, because of the markedly amplified relationship between serum T4 and release of TSH from the anterior pituitary (see 4.1 below and chapter 4). For a two-fold change in serum T4, up or down from the set-point for that individual, the serum TSH will normally change up to 100-fold in the reverse direction (54,55). Thus, TSH becomes recognisably abnormal long before the serum concentrations of T4 or T3 fall outside the population reference interval.

The more widespread the testing of thyroid function in the absence of suggestive clinical features, the greater the proportion of abnormal results in which only TSH is abnormal. In evaluating serum TSH, typically defined with a normal reference interval of about 0.4-4.0 mU/l, it is important to note that normal values approximate to a logarithmic distribution, with mean and median values at 1.0-1.5 mU/l (56,57). While values of 2-4 mU/l lie within the reference range, the likelihood of eventual hypothyroidism increases progressively for values above 2 mU/l, especially if thyroid peroxidase antibody is positive (58).

A population study in Colorado (59), of over 25,000 individuals of mean age 56 years, 56% of whom were female, showed TSH excess in 9.5 %, with a 2.2 % prevalence of suppressed TSH; over half the group with suppressed TSH were taking thyroid medication. In women, the prevalence of TSH excess increased progressively from 4% at age 18-24 to 20% over age 74 (59).

The National Health and Nutrition Examination Survey (NHANES III) (60), found hypothyroidism in
4.6% of the US population (0.3% overt and 4.3% subclinical) and hyperthyroidism in 1.3% (0.5% overt), with increasing prevalence with age in both females and males (figure 1). Abnormalities were more common in females than males. The prevalence of positive thyroid peroxidase antibody was clearly associated with both hyper- and hypothyroidism, with important ethnic differences in antibody prevalence.

Figure 1 Percentage of the US population with abnormal serum TSH concentrations as a function of age. The disease-free population excludes those who reported thyroid disease, goiter or thyroid-related medications; the reference population excluded, in addition, those who had positive thyroid autoantibodies, or were taking medications that can influence thyroid function. Note the much higher prevalence of TSH abnormalities in the total population, than in the reference population. (from reference 60, NHANES, JCEM 2002; 87: p493)

Prevalence data from one region do not necessarily apply in other populations, because of differences such as ethnic predisposition or variations in iodine intake. For example, in Hong Kong, where iodine intake is marginally deficient, only 1.2% of Chinese women aged over 60 years had serum TSH values > 5mU/l, with a comparable prevalence of suppressed values indicating possible hyperthyroidism (61). Several European studies (62,63) have compared the effect of various levels of iodine intake on the prevalence of thyroid over- and under-function. Hypothyroidism is generally more common with abundant iodine intake, while goitre and subclinical hyperthyroidism are more common with low iodine intake (62,63). These regional differences may influence the choice of diagnostic test and target population. For example, in an iodine replete environment, emphasis could be placed on testing younger or pregnant women for subclinical hypothyroidism by measurement of TSH and peroxidase antibody, whereas in an iodine-deficient region there might be additional emphasis on early detection of thyroid autonomy and hyperthyroidism in older people, using a highly sensitive TSH assay.

2 Subclinical thyroid dysfunction

The proven or presumed importance of subclinical thyroid dysfunction will have a major effect on the
extent to which thyroid function testing is applied in any population. The term ‘subclinical’ is used when the serum concentration of TSH is persistently abnormal (however defined), while the concentrations of T4 and T3 remain within their reference intervals. Because results can fluctuate spontaneously, a new diagnosis of subclinical thyroid dysfunction is not warranted on the basis of a single laboratory sample. The following five criteria define endogenous subclinical thyroid dysfunction:

a TSH increased above, or decreased below designated limits (see 3.1)

b Normal free T4 concentration (and free T3 for hyperthyroidism)

c The abnormality is not due to medication (see 6.4)

d There is no concurrent critical illness or pituitary dysfunction.

e A sustained abnormality is demonstrated over 3-6 months.

Apart from the situation of impending or early pregnancy, where there is clear consensus that subclinal hypothyroidism should be promptly and fully treated, the approach to subclinical thyroid dysfunction remains uncertain. Various authorities express divergent views on the importance of detecting the mild TSH abnormalities that reflect subclinical thyroid dysfunction. Extremes of opinion can be summarized as follows. On the one hand, some take the position that subclinical thyroid dysfunction, both hypothyroidism and hyperthyroidism, are disorders that need to be treated in order to avert potential harm (64). To achieve optimal sensitivity, particularly for the diagnosis of hypothyroidism, some have advocated that the upper limit of the TSH reference interval should be lowered (65) because values in the range 2-4 mU/l, usually regarded as normal, are associated with an increased prevalence of future hypothyroidism (58). Active search for subclinical thyroid dysfunction is based on the view that treatment is usually justified, because of potential adverse outcome, even if proof of benefit is still lacking. At this end of the opinion spectrum, there is support for general community screening for thyroid dysfunction (64), in contrast to a case-finding strategy for women over 50 when they present for medical care (1).

Others have taken the view that while there is circumstantial evidence that subclinical thyroid dysfunction can have adverse long-term effects, there is currently a lack of strong evidence that treatment confers benefit (66,67).

A definitive position on this dilemma may emerge from long-term studies of outcome, but if differences are small, studies may be under-powered and the results may be indeterminate. Other factors to be taken into account in establishing an approach to widespread thyroid testing include ethnic or environmental predisposition to thyroid dysfunction in various communities, balance with other healthcare priorities that may be more compelling, cost of laboratory testing and the extent to which competent clinical assessment and therapeutic response may be overwhelmed by reliance on laboratory measurements. The special issues that need to be considered in establishing a strategy for the testing of thyroid function before, during and after pregnancy are considered in section 6.7 and chapter 14.

2.1 Adverse consequences of subclinical thyroid dysfunction

The issues that identify the clinical importance of subclinical TD are summarised in table 2; many of these adverse effects relate to the cardiovascular system (68,69).

There is still conflicting evidence on whether mild thyroid abnormalities influence cardiovascular mortality and, as yet, no convincing support for the proposition that treatment of subclinical thyroid dysfunction improves survival. In a 12 year follow-up study of women over 65, neither TSH > 5 mU/l, nor <0.5 mU/l were associated with any increase in all-cause or cardiovascular mortality, although a
previous history of hyperthyroidism had a minor adverse effect (70). In contrast, another survey of the relationship between serum TSH and all-cause and cardiovascular mortality over a 10 year period in individuals over 60, showed that the group with serum TSH below 0.5 mU/l had a significantly increased mortality, apparently due to cardiovascular disease (71), although an increased serum TSH was not associated with excess mortality (71).

More recent studies provide strengthening evidence for adverse effects from minor degrees of thyroid dysfunction. A Scottish study (72) of over 17,000 people followed for an average of 4.5 years, correlated fatal or non-fatal first episodes of cardiovascular disease, arrhythmias and osteoporotic fractures with ambulatory serum TSH values. With elevated TSH > 4 mU/l there was an increased incidence of dysrhythmia, cardiovascular ischemic episodes and fracture (hazard ratios 1.8 – 1.95) (72). With suppressed serum TSH <0.03 mU/l, all three endpoints were also increased in frequency (hazard ratio 1.4-2.0). It is important to note that these authors distinguished between clearly suppressed TSH (<0.03 mU/l) and subnormal-detectable TSH 0.04-0.4 mU/l, which was not associated with significant adverse effects. These findings appeared to be at odds with the initial evaluation of the Wickham data that showed no adverse cardiovascular effects of subclinical hypothyroidism. However, reanalysis of the Whickham findings (73) did show an adverse effect if the beneficial effect of thyroxine treatment was excluded.

Table 2 Reported effects of subclinical thyroid dysfunction

<table>
<thead>
<tr>
<th>Subclinical hyperthyroidism (suppressed TSH, normal free T4, normal free T3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Exposure to iodine may precipitate severe thyrotoxicosis (74)</td>
</tr>
<tr>
<td>• Threefold increased risk of atrial fibrillation after 10 years (75)</td>
</tr>
<tr>
<td>• Abnormalities of cardiac function (76,77)</td>
</tr>
<tr>
<td>• Osteoporosis risk increased (78,79)</td>
</tr>
<tr>
<td>• Progression to overt hyperthyroidism (80,81)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subclinical hypothyroidism or mild thyroid failure (increased TSH, normal free T4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Non-specific symptoms may improve with treatment (83)</td>
</tr>
<tr>
<td>• Progression to overt hypothyroidism (84)</td>
</tr>
<tr>
<td>• Independent risk factor for atherosclerosis? (85)</td>
</tr>
<tr>
<td>• Increased risk of coronary artery disease (86,87,88)</td>
</tr>
<tr>
<td>• Increased frequency of congestive heart failure (86,87,89)</td>
</tr>
<tr>
<td>• Adverse effects on vascular compliance (90,91,92)</td>
</tr>
<tr>
<td>• Abnormal cardiac function may improve with treatment (93)</td>
</tr>
<tr>
<td>• Beneficial effect of treatment on lipids (94,95)</td>
</tr>
<tr>
<td>• Increased prevalence of depressive illness? (96)</td>
</tr>
<tr>
<td>• Impaired fibrinolysis (97)</td>
</tr>
</tbody>
</table>

2.1.1 Hyperthyroidism

2.1.1.1 Natural history of subclinical hyperthyroidism

Follow up studies suggest that spontaneous progression to overt hyperthyroidism is uncommon and that subnormal-detectable levels of TSH in the range 0.05-0.4 mU/l frequently return to normal within one
year (81). Meyerovitch (82) reported the results of sequential tests of thyroid function over a 5-year period in a large community-based cohort. During the follow-up period, test results returned to normal in 27%, 62% and 51% respectively of untreated patients whose initial serum TSH values were >10mU/l, 5.5-10 mU/l and < 0.35 mU/l (82). While some subjects did show progression, the high chance of resolution towards normal suggests that retesting after follow-up, possibly after at least 6 months, is advisable before considering any intervention. However, quantitative conclusions from that retrospective study may be insecure, as many subjects with abnormal TSH were excluded from follow-up because they were treated after their initial result (82). If those treated were the more severely affected, whether based on symptoms, presence of goiter, or degree of TSH abnormality, the analysis may over-estimate the likelihood of resolution during follow-up. Spontaneous remission of subclinical hyperthyroidism may occur more frequently in Graves’ disease than in nodular thyroid disorders (98). The chance of spontaneous progression to overt hyperthyroidism appears to be no greater than 10% per year (81,98). However, it should be noted that the transition from immune subclinical to overt hyperthyroidism can occur more rapidly than is generally the case in immune hypothyroidism (98a).

2.1.1.2 Cardiovascular effects

From the Framingham study it was found that undetected subclinical hyperthyroidism, defined only by suppression of TSH, carried a three-fold increased risk of atrial fibrillation within 10 years (75). As yet, there is no study that shows that treatment given on the basis of low TSH alone, modifies this risk, although it is clear that survival is adversely affected by atrial fibrillation (99). Impaired left ventricular ejection fraction and reduced exercise capacity has been documented in subclinical hyperthyroidism due to high dosage thyroxine and may be alleviated by beta blockade (77).

2.1.1.3 Iodine-induced thyrotoxicosis

Undiagnosed subclinical hyperthyroidism due to autonomous nodular thyroid disease, a condition especially prevalent in iodine deficient regions, carries the risk of progression to severe overt thyrotoxicosis after iodine exposure (100). An Australian study from a region that is not known to be iodine deficient, was suggestive of recent iodine exposure, most often from radiologic contrast agents, in up to 25% of elderly thyrotoxic patients (74). Prior knowledge of subnormal serum TSH may identify a high risk group with thyroid autonomy in whom iodine exposure carries the risk of iatrogenic hyperthyroidism.

2.1.1.4 Osteoporosis

While longstanding overt hyperthyroidism is an important risk factor for osteoporosis, the association has until recently been less clear for subclinical hyperthyroidism. Overt hyperthyroidism is associated with reduced bone density, predominantly affecting cortical rather than trabecular bone, so that the femoral neck is more affected than the lumbar spine, associated with an increase in fracture rate (101). It has now been shown that women over 65 with suppressed TSH are at increased risk of hip and vertebral fractures (79). Notably, in a controlled trial of suppressive T4 treatment for multinodular goitre, TSH suppression without clear excess of serum T4 or T3 resulted in a mean 3.6% decrease in lumbar spine density within 2 years (78).
2.1.2 Hypothyroidism

2.1.2.1 Natural history of subclinical hypothyroidism

The benchmark study of thyroid epidemiology from Wickham, UK, (102), showed that the likelihood of overt hypothyroidism after 20 years was directly related to the initial serum TSH, even when that value was in the range 2-4 mU/l, within the upper reference interval. The study by Meyerovitch (82) showed that with the initial serum TSH in the range 5.5-10 mU/l, the chance of normalization after prolonged follow-up was greater than the chance of progression to TSH > 10 mU/l (see above). Antibody status was not assessed in that cohort. Huber et al (84) followed 82 Swiss women with subclinical hypothyroidism, with normal free T4 and serum TSH > 4 mU/l, for a mean of 9.2 years (figure 2). About half of their cohort had had previous ablative treatment for Graves’ disease. The cumulative incidence of overt hypothyroidism, defined as low free T4 with TSH > 20 mU/l, was directly related to the initial serum TSH, with 55% of women with initial serum TSH >6 mU/l progressing to overt hypothyroidism. Progression was not uniform, and over half of the cohort showed no deterioration of thyroid function, but positive microsomal antibodies increased the likelihood of progression. It is now clear that the opposite sequence may also occur, with spontaneous normalization of elevated TSH values (82,103).

Notably, a recent prospective study has shown that the progression of immune subclinical hypothyroidism tends to be slower than for subclinical hyperthyroidism (98a). Hence, the need for prolonged follow-up and patient education, where the decision to treat is deferred.
Figure 2 Kaplan Meier estimates of the cumulative incidence of overt hypothyroidism in women with subclinical hypothyroidism (initial serum TSH > 4mU/l) as a function of initial serum TSH, thyroid secretory reserve in response to oral TRH and detectable microsomal antibody. Serum TSH appears to be the strongest of these predictors. (from reference 84, Huber et al JCEM 2002 87 page 3223)
2.1.2.2 Atherosclerosis

Subclinical hypothyroidism, or mild thyroid failure, was shown to be an independent risk factor for both myocardial infarction and radiologically-visible aortic atherosclerosis in a study of Dutch women over 55 years of age (85). This effect was independent of body mass index, total and HDL cholesterol, blood pressure and smoking status. The attributable risk for subclinical hypothyroidism was comparable to that for the other major risk factors, hypercholesterolemia, hypertension, smoking and diabetes mellitus. The association was slightly stronger when subclinical hypothyroidism was associated with positive peroxidase antibody, but thyroid autoimmunity itself was not an independent risk factor. Thyroxine replacement for 18 months has been reported to improve blood pressure, lipids and carotid intimal thickness in women with subclinical hypothyroidism (104).

2.1.2.3 Vascular compliance

The finding of impaired flow-mediated, endothelium-dependent vasodilatation in subjects with borderline hypothyroidism or high-normal serum TSH values (90) was at first unexpected. Baseline artery diameter and forearm flow were comparable, but flow mediated vasodilatation during the period of reactive hyperemia was significantly impaired even in the group with serum TSH of 2-4 mU/l, compared with the group with serum TSH 0.4-2 mU/l (90). The difference could not be attributed to a difference in maximal nitrate-induced vasodilatation, age, sex, hypertension, diabetes, smoking, serum cholesterol, or levels of total T3 and T4 (90). This finding suggests that even minor deviation from an individual’s pituitary-thyroid set point may be associated with alteration in vasodilatory response. There is no known direct action of TSH that would account for this effect. A Japanese placebo-controlled study of women aged 60-70 with subclinical hypothyroidism with mean pre-treatment serum TSH of 7.3 mU/l showed improvement in pulse wave velocity, an index of vascular stiffness, in response to TSH normalization for 2 months by progressive low dose T4 replacement only up to 37.5 ug/day (91).

2.1.2.4 Cardiac function

From echocardiographic studies, there is evidence that mild thyroid failure can significantly increase systemic vascular resistance and impair cardiac systolic and diastolic function, as demonstrated by decreased flow velocity across the aortic and mitral valves (69). These changes, which were associated with reduced cardiorespiratory work capacity during maximal exercise, were reversed by T4 treatment sufficient to normalize serum TSH (69). Impairment of both diastolic and systolic function was demonstrable by echocardiography in a subclinically hypothyroid group of patients with TSH in the range 4-12 mU/l (93). Thyroxine treatment sufficient to normalise TSH to a mean of 1.3 mU/l for 6 months was associated with improvement in myocardial contractility (93). It remains to be established how these reversible abnormalities relate to cardiovascular prognosis.

2.1.2.5 Lipids

Overt hypothyroidism is associated with increase in the serum cholesterol concentration and correction of overt hypothyroidism results in a decrease in total and LDL cholesterol, apolipoprotein A1, apo B and apo E, while serum triglyceride concentrations may also decrease (105). A defect in receptor-mediated LDL catabolism, similar to that seen in familial hypercholesterolaemia, has been described in severe overt hypothyroidism (106), but there is no evidence to support such an abnormality in mild thyroid failure. At the other extreme, several large population studies report a positive correlation between serum lipids and serum TSH across its normal range (107), a correlation also associated with increasing blood pressure (108). The clinical impact of these associations remains unknown.
The Colorado study of over 25,000 subjects showed a continuous graded increase in serum cholesterol over a range of serum TSH values from <0.3 to >60 mU/l (59). However, there is still no consensus that mild thyroid failure has an adverse effect on plasma lipids, or that T4 treatment sufficient to normalize isolated TSH excess has a beneficial effect. A meta-analysis suggests that T4 treatment of subjects with mild thyroid failure does lower the mean total and LDL cholesterol, and is without effect on HDL cholesterol or triglyceride (94). In a prospective double-blind, placebo-controlled trial of thyroxine in subclinical hypothyroidism in which the response was carefully monitored with TSH, Meier et al (95) reported that the decrease in LDL cholesterol was more pronounced with higher initial TSH levels >12 mU/l or with elevated baseline LDL concentration.

It remains uncertain whether the serum concentration of the highly atherogenic Lp(a) particle is increased in overt hypothyroidism and whether T4 treatment sufficient to normalize TSH has a favourable influence. Serum concentrations of Lp(a) have been found to be increased in overt hypothyroidism with normalization after treatment in some studies (109,110) while others fail to confirm this finding (111,112).

2.1.2.6 Neurobehavioural effects

It is well known that overt TD may present with psychological or psychiatric symptomatology. A small retrospective study has shown a 2-3 fold increased frequency of previous depression in subjects with mild thyroid failure (96); T4 treatment has been reported to improve neuropsychological responses in this group (113). However, contrary to these findings, more recent controlled studies of unselected patients suggest that subclinical hypothyroidism is not associated with any consistent deficit in quality of life indices or improvement with treatment (114,115).

2.1.2.7 Insulin sensitivity and syndrome X

Several studies suggest that there may be a link between insulin sensitivity, serum lipids and thyroid function, whether assessed by serum TSH or circulating thyroid hormone levels. Bakker et al (116) noted that while serum TSH showed no overall correlation with insulin sensitivity or serum lipids, there was a complex interaction between these parameters such that the association between TSH and LDL-C was much stronger in insulin- resistant than in insulin- sensitive subjects. The same group showed that low free T4 levels within the normal range are dually associated with LDL-C and insulin resistance (117). Further results will show whether these findings account for the purported link between subclinical hypothyroidism and the metabolic syndrome (syndrome X) (118) and whether this link contributes to increased cardiovascular risk in a subgroup of patients with subclinical hypothyroidism. Notably, there is a positive correlation between serum TSH and BMI in euthyroid obese women (10,119). The value of routine thyroid testing in this group remains uncertain.

2.1.2.8 Studies in children

A recent study (119a) has shown that long-term idiopathic subclinical hypothyroidism did not appear to have an adverse effect on linear growth or intellectual development in children aged 4-18 years. Hence, it may be justifiable to follow this group without early recourse to lifelong replacement.

2.2 Safety and effectiveness of treatment

The benefits of early diagnosis and treatment are self-evident from the obvious decline in hospitalisation and mortality rates for severe TD over past decades. Before reliable tests of thyroid function became widely used, severe hyperthyroidism approaching thyroid storm and hypothyroidism
with impending myxoedema coma occurred quite regularly, but these presentations are now very uncommon. While the arguments for seeking out and treating mild TD in individuals are less compelling, there may be potential benefit for large numbers of people.

The points in favour of treating mild TD relate directly to the adverse consequences listed in table 2, but for many of these adverse outcomes there is still a lack of long-term studies that show benefit. On the basis of potential benefit from simple straightforward treatment and freedom from adverse effects, the argument for active treatment is generally stronger for mild thyroid failure than for subclinical hyperthyroidism. Conservative T4 therapy aimed at normalising TSH is simple, inexpensive and generally safe (120), although replacement may not be warranted in the extremely elderly. Where cardiovascular disease precludes full thyroid hormone replacement, detailed evaluation of the cardiac abnormality is appropriate. In contrast, treatment of subclinical hyperthyroidism needs to be evaluated in relation to adverse drug effects and potential for hypothyroidism.

3. Use of laboratory assays for case finding and screening

If the identification of abnormal thyroid function is to be based on laboratory testing, it is desirable that population reference intervals should not vary between methods. As recently emphasized, serum T3, T4 and serum TSH concentrations are among the many hormone parameters where between-assay standardization is crucial to ensure optimal assay specificity (120a). At present that aim is not satisfied for free T4 estimates, a problem that is particularly troublesome during pregnancy (see below). Analytically, serum TSH, total T4 and total T3 are well standardized, so that considerations of so-called normal ranges relate to the clinically relevant issues. By contrast, the diverse, ingenious manoeuvres involved in the estimation of free T4 and T3 lead to poor standardization between methods, so that method-specific reference intervals need to be used, both for individual clinical diagnosis and population studies. Between-method free T4 variations are especially troublesome in pregnancy and critical illness (see below).

3.1 TSH Reference Interval or “Normal Range” (See also chapter 6a)

For serum TSH, arguably the most important parameter for diagnosis of primary thyroid dysfunction, no firm consensus range has been agreed, despite reliable standardization between methods. Firstly, a reference interval for the diagnosis of hypothyroidism in the elderly may be far too broad and lenient to identify women whose fertility or pregnancy outcome might be improved by thyroid supplementation. Secondly, it has been shown that the reference TSH set-point for each individual can be defined with a narrow band of the broad population range (see below 4.1). Third, criteria for the new diagnosis of thyroid dysfunction may not be the same as those required for optimal adjustment of therapy. Fourth, for population studies, whether screening or case-finding, a reference interval with higher sensitivity will have lower specificity.

Thus, the controversies as to whether the standard TSH reference interval of about should be 0.4 – 4.0 mU/l should be narrowed, with lowering of the upper limit (65), or retained (121-123) do not address the needs of individuals. Population studies to define the upper limit of the “normal range” for serum TSH will be influenced by whether those with positive peroxidase antibody are excluded (see chapter 6a). However, even after exclusion of individuals with clinical, antibody or sonographic evidence of
any thyroid disorder, Hamilton et al supported an upper reference limit at about 4 mU/l (121), although quite different criteria may apply around pregnancy (see 6.7 below). Ethnic differences (124), differences and time of sampling in relation to diurnal variation are also important.

Terminology for abnormal TSH values has also become inconsistent, as for example in the use of the term *suppressed* to describe lower-than-normal TSH values. In some studies (125,126) any subnormal value is classified as suppressed, while others reserve this term for lower levels (60) that allow a distinction between undetectable (eg <0.03 mU/l) and a subnormal-detectable range below about 0.4 mU/l (72). These two categories may have different diagnostic and prognostic significance (72). Based on follow-up studies of the probability of progression to overt thyroid dysfunction, there is a strong case for regarding TSH values in the subnormal-detectable range, arbitrarily 0.05-0.4 mU/l, as distinct from the even lower levels that are typical of hyperthyroidism. It is a personal view that the term “suppressed” should be avoided in describing subnormal-detectable values, a semantic point that may affect up to 1% of the population. Since the gradation from normality to severe thyroid dysfunction is a continuum, studies of adverse outcomes or benefits from intervention will be critically dependent on uniform terminology.

### 3.2 Choice of initial test

The definitive diagnosis of TD should always be made using the typical relationships between trophic hormone and target gland secretion that define endocrine dysfunction. In contrast, case-finding studies of untreated subjects may begin with measurement of TSH alone (127), with T4 and T3 assays added only if TSH is abnormal, or if an abnormality of TSH secretion is suspected. It is self-evident that serum TSH loses its diagnostic value when pituitary function is abnormal (128).

In the absence of associated disease, a normal serum TSH concentration by a so-called third generation assay (a functional lower limit of sensitivity of about 0.03 mU/l), has high negative predictive value in ruling out primary hypothyroidism and hyperthyroidism. Such immunometric assays, which use two antibodies against different epitopes of the TSH molecule, give a wide separation between the lower limit of the normal reference range at about 0.4 mU/l and the typical TSH values found in hyperthyroidism. While some subjects with subnormal-detectable TSH values do progress to overt hyperthyroidism, values in this range may also revert to normal (see above).

In 1996, Danese et al (3) analysed the cost-effectiveness of TSH measurement, performed every 5 years from the age of 35 and concluded that for females, such a strategy compared favourably with other preventative medical practices (3). From age 35, the cost per quality-adjusted life year was about US$9,000 for females and $22,500 for males, but this cost dropped by half if testing was commenced from age 65. Cost effectiveness was highly dependent on the price of the TSH assay and was adversely affected by more frequent testing (3).

There are some clinical situations in which assessment of thyroid function will give a high prevalence of abnormalities that cannot be interpreted with certainty. Notably, glucocorticoids and dopaminergic agents have a potent effect to suppress TSH secretion (130), while TSH is also frequently subnormal in starvation or caloric deprivation (131). Transient increases to above normal can occur in euthyroid subjects during recovery from critical illness (132). The finding that 33% of serum TSH values fell more than 2 SD from the geometric mean in acutely hospitalized patients, with 17% of values more than 3 SD from the mean value, indicates that TSH value lacks diagnostic specificity in this setting (133). A serum free T4 estimate will generally follow from an abnormal TSH value, but during critical illness, free T4 estimates often show non-specific abnormalities (134) (see section 6.4.2 below). Lack of specificity was the basis for a recommendation against routine assessment of serum TSH and free T4 during acute critical illness in the absence of risk factors, or clinical features suggestive of a thyroid
Testing of thyroid function is appropriate in a wide range of psychiatric disorders, but diagnostic specificity is limited by a high prevalence of transient non-specific abnormalities at the time of acute psychiatric admission (136). Thus, it has been recommended that laboratory evaluation should be delayed for 2-3 weeks after acute presentation, unless there are specific risk factors for TD (137).

3.3 Potential adverse effects of testing

In terms of potential for adverse effects, there may be important differences between screening of unselected populations and the case-finding strategy that is now recommended for TD. There are no reports of a “labelling effect” (i.e. perception of chronic illness in previously asymptomatic subjects), described in hypertension screening programs (138), when testing for TD is done at the time of presentation for medical care. Nevertheless, further attention needs to be given to the potential for unwarranted treatment based on false positive results, as well as the cost of follow-up investigations for perceived abnormalities that may not warrant treatment at any stage. In particular, there is a need for clinical consensus as to how marginally abnormal results should be classified. Widespread laboratory testing will lead to an increase in the number of false positive results; the potential for a diagnostic method to give misleading variations from normal may not become known for some years until the full diversity of the non-diseased population is documented (139).

3.4 Follow-up of abnormal results

If serum TSH is used as a single initial test for case-finding, a value outside the reference interval should lead to estimation of serum free T4 on the same sample, if possible without recall of the patient. This requires an algorithm-based testing protocol, which should also include measurement of serum free T3 if TSH is suppressed, in order to identify T3 toxicosis. It may also be relevant to measure thyroid peroxidase antibody if TSH is increased, so as to define an autoimmune mechanism of hypothyroidism, particularly as a raised level of this antibody is associated with an increased likelihood of progression to overt hypothyroidism (58).

For screening or case-finding to be effective, patients with unsuspected overt TD should be actively traced because they will benefit most from treatment and have the most to lose if the abnormal finding is ignored. For mild TD, a practitioner who has continuing contact with the patient should evaluate the assay result in clinical context and initiate any necessary follow-up. However, there is currently no consensus as to how an appropriate clinical response to abnormal laboratory findings can be assured. A transition towards identification of TD by laboratory measurement, rather than on clinical criteria, modifies, but does not diminish the clinician’s role. The severity of thyroid dysfunction cannot be judged from the extent of the laboratory abnormality (140), which indicates neither the duration of exposure, nor individual susceptibility. Whether a decision is made to treat or to observe, patient education is crucial in establishing effective compliance and rational cost-effective long-term follow-up. Computer based programs can identify affected individuals, but do not replace direct involvement of both patient and clinician. There may be potential medicolegal consequences of failure to respond to abnormal results, if widespread laboratory testing is initiated without an established follow-up plan.
4 Interpretation of TSH and T4 assays

4.1 The TSH-T4 relationship

Definitive assessment of thyroid status requires both a sensitive serum TSH assay and a valid T4 estimate, with interpretation based on the relationship between these two values. It is generally assumed that inverse TSH responses to changes in free T4 are approximately logarithmic. While this concept is useful, recent data suggests that the relationship varies with thyroid status (141) with progressively greater amplification of the TSH response as thyroid function declines. Diurnal variation in serum TSH, with amplitude about 50% with higher levels between 2100 and 0600, is superimposed on rapid pulse secretion with amplitude about 10%. Basal secretion and pulsatility are both increased in primary hypothyroidism, with retention of diurnal variation (142).

Because of the feedback relationship between tropic and target gland hormones, typical disease-related changes lead to diagonal deviations from the normal T4-TSH relationship. (figure 3).

Figure 3 The relationship between serum TSH and free T4 concentration is shown for normal subjects (N) and in the typical abnormalities of thyroid function: A, primary hypothyroidism; B, central or pituitary-dependent hypothyroidism; C, hyperthyroidism due to autonomy or abnormal stimulation of the gland; D, TSH-dependent hyperthyroidism or thyroid hormone resistance. Note that linear changes in the concentration of T4 correspond approximately to logarithmic changes in serum TSH.
4.2 The T4-TSH setpoint

Small changes in serum T4 and T3 concentrations, within the normal range, alter serum TSH, indicating that the inverse feedback relationship between serum free T4 and TSH applies across their normal ranges as well as in disease states (143,144). Studies of normal subjects demonstrate significant individual variation, independent of sex and age, in the setpoint of the pituitary-thyroid axis (145,146), which suggests that the TSH set-point for a particular serum free T4 or free T3 concentration is an individual characteristic. It follows that the concept of “normality” for an individual may be narrower for an individual than it is for a population at-large. Studies of monozygotic and dizygotic twin pairs also suggest that genetic factors influence the serum concentrations of total and free T4 within the normal range (147). The recent demonstration (147a) that multiple genetic factors can influence inter-individual differences in the TSH response to circulating thyroid hormones may ultimately increase the precision and the complexity of interpreting thyroid function.

Andersen et al (146), in a study of normal subjects whose samples were taken monthly between 0900 and 1200 for a year, showed that individual references ranges for T4 and T3 were only about half the width of the population reference ranges, indicating that a test result within the population range is not necessarily normal for that individual. Serum TSH showed greater between-sample variation for each individual than serum T4 or T3. Based on the degree of individual variation, it was estimated that a normal serum TSH concentration needed to change by 0.2-1.6 mU/l to be confident of a serial change in thyroid status. Based on this analysis, it was estimated that a single morning sample defined serum T4 and T3 to within 25%, and serum TSH only to within 50%. Since TSH shows diurnal variation and pulsatile secretion in both normal and hypothyroid subjects (146), random samples are likely to show even greater variation.

It has been suggested that some healthy elderly subjects have normal serum TSH concentrations despite having low serum free T4 values, attributed to resetting of the threshold for TSH inhibition (148). In a large cohort of Danish patients with newly diagnosed hypothyroidism, the increase in serum TSH for a given degree of lowering of serum free T4 was less in the elderly, suggesting that equivalent TSH increases in the elderly may be accompanied by more severe thyroid hormone deficiency (149). It is notable that a higher level of serum free T4 was necessary to normalize serum TSH in children with congenital hypothyroidism than in adults with acquired immune hypothyroidism (150). It is uncertain whether this is a further reflection of age-related difference, or whether this difference reflects a perinatal shift in the central setpoint for regulation of TSH secretion.

4.3 The TSH-T4 relationship: diagnostic assumptions

The precise diagnosis of thyroid dysfunction can generally be established from a single serum sample from the relationship shown in figure 2, subject to six key assumptions (table 3). It should be noted that only the last three of these assumptions can be validated in the laboratory; the first three are best verified clinically.

<table>
<thead>
<tr>
<th>Table 3: Assumptions inherent to diagnostic use of the T4 -TSH relationship (Conditions that may breach these assumptions are shown in italics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Steady-state conditions (n.b. difference in half-lives of TSH and T4)</td>
</tr>
<tr>
<td>Early treatment with antithyroid drugs (151)</td>
</tr>
</tbody>
</table>
Early response to T4 therapy

Evolution of transient thyroid dysfunction

Recovery from severe illness (132)

2. Normal trophic-target hormone relationship

Alternative thyroid stimulators

Immunoglobulins (152)

Chorionic gonadotrophin (153)

Medications that influence TSH secretion

T3, triiodothyroacetic acid (154)

Other thyroid hormone analogues? (154a)

Glucocorticoids (155)

Dopamine (156)

Amiodarone (27)

Recent hyperthyroidism (151)

Recent longstanding hypothyroidism

Treated congenital hypothyroidism (157)

TSH receptor mutations (158, see also chapter 16a)

Variable individual setpoint (145-149)

3. Tissue responses proportional to hormone concentration

Hormone resistance syndromes (159, see also chapter 16b)

Slow onset/offset of thyroid hormone action

Drug effects?

Amiodarone (27)
Phenytoin (160?)

4. The assay measurement represents the active hormone

Unmeasured agonist in excess (e.g. T3, Triac, hCG)

TSH of altered biologic activity (161,162)

Spurious immunoassay results

TSH

Heterophilic antibodies (163)

Free T4

Abnormal serum binding proteins (164)

Autoantibodies (165)

Heterophilic antibodies (166)

Medications that inhibit protein binding (164, 167)

Heparin artefact (168,169)

5. The assay can reliably distinguish low from normal values

Lack of precision at the limit of detection (170, see chapter 6a)

6. Reference ranges are appropriate

Influence of age (171, 172)

Associated illness (see 6.4 below)

4.3.1 Steady-state conditions.

This first assumption should be questioned whenever anomalous results occur during associated illness, or with medications that perturb the pituitary-thyroid axis. The half lives of plasma TSH (approx. 1 hour) and plasma T4 (approx. 1 week) differ so widely that acute perturbation of the pituitary-thyroid axis will often result in transient non steady-state conditions. With its much shorter half-life, serum TSH deviates more rapidly from the steady state. Other common deviations from the steady state relate to short-term pulsatile or diurnal fluctuations in hormone secretion, responses to treatment and spontaneous evolution of disease, as can occur in subacute thyroiditis or postpartum thyroid dysfunction.
Figure 4 Measurement of serum T4, rather than serum TSH, is the more reliable single test of thyroid function when steady state conditions do not apply, as in the early phase of treatment for hyperthyroidism or hypothyroidism. (from reference 173)

### 4.3.2 Normal tropic-target hormone relationship.

During treatment of prolonged hyperthyroidism, TSH secretion may remain low for several months after serum free T4 becomes normal (151). Conversely, after severe prolonged hypothyroidism, or in some children treated for congenital hypothyroidism (150), TSH hypersecretion may persist despite normalization of serum T4. Serum TSH will then give an inaccurate indication of thyroid status, with the potential for over-treatment if this parameter alone is used to assess therapy.

### 4.3.3 Tissue responses proportional to the hormone concentration.

The active or free concentrations of T3 and T4 generally correlate well with clinical features, but in generalized thyroid hormone resistance, high serum free T3 and T4 concentrations are entirely appropriate to maintain the euthyroid state. The onset and offset of thyroid hormone action is slow, so that tissue responses may lag behind changes in serum concentrations of free T4 and T3. There is a notable lack of convenient, sensitive, specific, objective indices of thyroid hormone action (see chapter 6e), so that assessment remains predominantly clinical. Corroborative measurements that can be useful, especially in following the response of individuals to therapy, include measurement of oxygen consumption (174), sex hormone binding globulin (175), angiotensin converting enzyme (176) and serum ferritin (177), as well as several indices of cardiac contractility.

### 4.3.4 The assay measurement reflects the active hormone(s).

TSH and iodothyronine assays make comparative, rather than absolute, measurements of hormone
concentrations, based on the premise that samples and assay standards differ only in their concentration of analyte. This assumption breaks down if there is any other difference between a serum sample and assay standards that influences the measured parameter, as, for example, dissimilar protein binding of tracer (178), the presence of binding competitors (167,179), or possible nonspecific interference with enzymatic, fluorescent, or chemiluminescent detection systems. Circulating T3 and T4 autoantibodies may invalidate immunoassays by sequestering the assay tracer (180), while heterophile mouse or sheep (166) antibodies and rheumatoid factor can interfere with immunoglobulin aggregation, or with cross linking of the signal and capture antibodies (181,182).

If the biologic activity of circulating immunoreactive TSH is increased or decreased, the normal relationship between measured serum TSH and free T4 may be altered. Secreted immunoreactive TSH is heterogeneous, due to differences in its three oligosaccharide side chains. In hypothalamic hypothyroidism, the secreted TSH has decreased bioactivity (162, 183), whereas activity may be enhanced in thyroid hormone resistance, primary hypothyroidism and in some TSH-producing tumors (184,185).

4.3.5 The assay can reliably distinguish low from normal levels

Assay precision inevitably deteriorates as the limit of detection is approached; this characteristic is crucial in evaluating TSH assays that are used to distinguish hyperthyroidism from normal (54,173). The lower working limit of a TSH assay should be defined in terms of its between-assay reproducibility, defined as functional sensitivity, rather than by the analytical sensitivity of individual assay runs (54,173).

4.3.6 Reference ranges are appropriate

Reference ranges (table 4) show little change with advancing age, except for a possible decline in normal serum T3 with age (148). Serum T3 is significantly higher in children (172) and probably also in young adults. For statistical analysis, TSH reference ranges are generally evaluated after logarithmic transformation; the geometric mean is then used to define a realistic lower normal limit. There may be an age-related change in the central setpoint for TSH secretion, with progressive decline in the TSH response to decrease in serum T4 (149). Associated illness, nutritional changes, and medications frequently cause assay results to fall outside the normal reference ranges as defined in healthy subjects, so that it may be relevant to use wider than normal reference ranges in the face of associated illness (173).

<table>
<thead>
<tr>
<th>Hormone</th>
<th>Metric</th>
<th>SI Units</th>
<th>Variations unrelated to thyroid disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total T4</td>
<td>4-11 μg/dl</td>
<td>60-140 nM</td>
<td>Binding protein changes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Competitors for T4 binding</td>
</tr>
<tr>
<td>Free T4</td>
<td>0.7-2.1 ng/dl</td>
<td>10-25 pM</td>
<td>Methodologic factors</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Albumin changes</td>
</tr>
</tbody>
</table>

(Ranges should be determined for the particular methods used in each laboratory)
<table>
<thead>
<tr>
<th>Indicator</th>
<th>Range</th>
<th>Unit</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total T3</td>
<td>75-175 ng/dl</td>
<td>1.1-2.7 nM*</td>
<td>Binding protein changes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Competitors for T3 binding</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Age-related changes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Nutrition, illness, drugs</td>
</tr>
<tr>
<td>Free T3</td>
<td>0.2-0.5 ng/dl</td>
<td>3-8 pM*</td>
<td>Methodologic factors as for T4 and influences on total T3</td>
</tr>
<tr>
<td>Reverse T3</td>
<td>15-45 ng/dl</td>
<td>0.2-0.7 nM</td>
<td>Nutrition, illness, surgery, drugs</td>
</tr>
<tr>
<td>TSH Φ</td>
<td>0.3-4.0 mU/l</td>
<td>1-15 pM#</td>
<td>Diurnal variation, pulsatile secretion</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Medications</td>
</tr>
<tr>
<td>TSH alpha-subunit</td>
<td>&lt;0.2 ng/dl</td>
<td>&lt;100 pM</td>
<td>Postmenopausal women</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Primary hypogonadism</td>
</tr>
<tr>
<td>Thyroglobulin</td>
<td>1-20 μg/l+</td>
<td>1.5-30 pM</td>
<td>Variation in assay standards and antibodies</td>
</tr>
</tbody>
</table>

* Values in children may be higher (172)

Φ These ranges do not apply in pregnancy (see 6.7 below and chapter 14)

# Assumes biologic potency of 7-15 U/mg

Dependent on serum TSH and amount of thyroid tissue; variation between methods

---

5 Application and interpretation of individual serum assays

5.1 Serum TSH as the initial test of thyroid function

Either serum TSH or free T4 can be used as the initial test of thyroid function, with TSH previously assessed as giving better first-line discrimination at slightly higher cost (127). A normal serum TSH concentration has high negative predictive value in ruling out primary thyroid disease, and this assay has become increasingly used as the single initial test of thyroid function, with further assays done routinely if serum TSH is normal. If serum TSH is increased, free T4 is measured on the same sample to distinguish between overt and subclinical hypothyroidism ( figure 5 ). If serum TSH is suppressed, i.e. < 0.05 mU/l, both free T4 and free T3 should be assessed to distinguish between overt hyperthyroidism, T3-thyrotoxicosis and subclinical hyperthyroidism. The interpretation of subnormal TSH values is influenced by the functional sensitivity of the particular assay ( see Chapter 6a ).
Figure 5: An algorithm for the initial assessment of thyroid function, based on initial assay of serum TSH. The limitations of this strategy are summarised in table 5.

5.1.1 Limitations of the “TSH first” testing strategy

The rationale for using TSH alone as a first-line test of thyroid function rests on the assumptions that thyroprivic, or primary hypothyroidism is much more common than central or secondary hypothyroidism and on the fact that serum TSH deviates outside the reference range early in the natural history of thyroid over-function or progressive thyroid failure. The major weaknesses of this approach are the likelihood of missing secondary hypothyroidism (in which the concentration of immunoreactive serum TSH is frequently normal rather than low), and the frequency of abnormal TSH concentrations in the absence of thyroid dysfunction, especially in patients with an associated illness.

Beckett and Toft (129) have pointed out the adverse consequences of missing secondary hypothyroidism due to pituitary failure by relying on normal serum TSH to rule out thyroid dysfunction. The diagnosis of this disorder is often difficult, with diverse presentations to a wide range of practitioners who may not be attuned to key clinical features. He estimated that as many as 1500 cases per year of central hypothyroidism could be missed in the UK if the “TSH first” policy were followed inflexibly. This potentially serious diagnostic problem was shown by Wardle et al (128) who reported an analysis of 56,000 requests for thyroid function testing over 12 months from a population of 471,000. Serum TSH was normal in association with subnormal total and free T4 in 15 patients who, on further investigation, had probable hypopituitarism that would have been missed by assessment of serum TSH alone. Cost benefit analyses would therefore need to balance the savings from not measuring serum T4 routinely, against the cost of further investigation and medical care for the missed diagnoses, in addition to a component for burden of suffering and potential litigation. If advances in technology can reduce the unit cost of measuring T4, the “TSH first” approach may be abandoned.

<table>
<thead>
<tr>
<th>Table 5: Situations in which serum TSH alone can give a false or uncertain indication of thyroid status.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Condition</strong></td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Primary abnormality of TSH secretion

<table>
<thead>
<tr>
<th>Condition</th>
<th>L- N</th>
<th>L</th>
<th>H- L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pituitary-hypothalamic abnormality</td>
<td>L- N</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Extremely premature infants</td>
<td>L- N</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Central TSH excess</td>
<td>N- H</td>
<td>H</td>
<td>H</td>
</tr>
</tbody>
</table>

Hyperthyroidism

<table>
<thead>
<tr>
<th>Condition</th>
<th>U</th>
<th>N</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>T3 toxicosis</td>
<td>U</td>
<td>N</td>
<td>H</td>
</tr>
<tr>
<td>Subclinical</td>
<td>U</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Early Treatment</td>
<td>U</td>
<td>H-N-L</td>
<td>H-N-L</td>
</tr>
<tr>
<td>TSH assay artefact</td>
<td>L- N</td>
<td>H</td>
<td>H</td>
</tr>
</tbody>
</table>

Hypothyroidism

<table>
<thead>
<tr>
<th>Condition</th>
<th>H</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subclinical</td>
<td>H</td>
<td>N</td>
</tr>
<tr>
<td>Early Treatment</td>
<td>H</td>
<td>L-N</td>
</tr>
</tbody>
</table>

Thyroid hormone resistance

<table>
<thead>
<tr>
<th>Condition</th>
<th>N- H</th>
<th>H</th>
<th>H</th>
</tr>
</thead>
</table>

Medications

<table>
<thead>
<tr>
<th>Medication</th>
<th>L</th>
<th>N</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dopamine</td>
<td>L</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Glucocorticoids</td>
<td>L</td>
<td>N</td>
<td>L-N</td>
</tr>
<tr>
<td>Amiodarone (acute)</td>
<td>H</td>
<td>N-H</td>
<td>L</td>
</tr>
</tbody>
</table>

N: normal; L: low; H: high; U: undetectable.

5.2 Serum free and total T4

Estimation of free T4, or total T4 linked to a measurement of the thyroid hormone binding ratio, in association with serum TSH, has now become the standard method of assessing thyroid function, except in settings where the TSH-first strategy persists. Assays for total T4 and T3 in unextracted serum include a reagent such as 8-anilinonaphthalene sulfonic acid that blocks T4 and T3 binding to serum proteins, so that total hormone is available for competition with the assay antibody. Assays for free T4 or T3 omit this blocking reagent and use a wide variety of manoeuvres to isolate a moiety that reflects the free hormone concentration. The theoretical basis, practical utility, and validity of the many different approaches to the estimation of serum free T4 and T3, have been considered in detail (186) (see chapter 6a).

Two key assumptions in any method of free T4 estimation are (i) that the dissociation of bound hormone with sample dilution is similar in samples and standards, and (ii) that samples and standards show identical protein binding of the assay tracer. If either of these conditions is breached, the assay is likely to give inaccurate results. Serum free T4 and free T3 can be estimated either by two-step methods that separate a fraction of the free hormone pool from the binding proteins before the assay incubation, or by one-step methods in which the free hormone concentration is measured in the presence of binding proteins (186). Many of the one-step methods become invalid when the sample and standard differ in their binding of assay tracer, but the two-step methods are less prone to non-specific artefacts. Almost all techniques of estimating free T4 give a useful correction for moderate variations in serum TBG concentration, but no method can yet accommodate the effect of circulating competitors for T4 binding (167,187).
5.3 Indications for measurement of serum T3

Measurement of serum T3 is indicated, in addition to serum T4, as follows:

a. In suspected hyperthyroidism with suppressed TSH and normal serum T4, to identify T3-thyrotoxicosis and distinguish this entity from subclinical thyrotoxicosis.

b. During antithyroid drug therapy to identify persistent T3 excess, despite normal or low serum T4 values (188).

c. For diagnosis of amiodarone-induced hyperthyroidism, which should not be based on T4 excess alone because of the frequency of euthyroid hyperthyroxinemia during amiodarone treatment (189).

d. To assess the extent of T3 excess in individuals treated with thyroid extracts of animal origin, to assess a potentially damaging hormone excess that is not reflected by the level of T4.

e. To identify T3-predominant thyrotoxicosis, an entity that is less likely to achieve remission (see below). In some studies this entity is been reported to respond poorly to radioiodine (189a), a phenomenon that may relate to rapid iodine turnover within the gland (189b).

Serum T3 measurements may also be useful:

f. For estimation of the serum T3-T4 ratio. A high ratio (>0.024 on a molar basis or >20 calculated as ng/µg) that persists during antithyroid drug treatment suggests that patients with hyperthyroid Graves’ disease are unlikely to achieve remission (190). This ratio usually is lower in iodide-induced hyperthyroidism (191) or hyperthyroidism caused by thyroiditis (192) than in Graves’ disease.

g. To detect early recurrence of hyperthyroidism after cessation of antithyroid drug therapy.

h. To establish the extent of T3 excess during high-dose replacement or suppressive therapy with T4, or after an accidental or intentional T4 overdose.

Low serum T3 concentrations have little specificity or sensitivity for the diagnosis of hypothyroidism. Many patients with nonthyroidal illness have low values, and the serum T3 concentration can remain in the reference range until hypothyroidism is severe. Serum T3 values are usually interpreted together with the levels of T4 (table 6).

<table>
<thead>
<tr>
<th>Table 6 Relationship between serum T4 and T3 in various disorders*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Serum T4</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Severe hypothyroidism.</td>
</tr>
<tr>
<td>T4 treatment.</td>
</tr>
</tbody>
</table>
Hypothyroidism.  
Hyperthyroidism with nonthyroidal illness.  
T4 binding autoantibodies.

| High | Iodine deficiency | T3 toxicosis | Hyperthyroidism.  
T3 treatment | T3 binding autoantibodies | Excess T4 ingestion | Hormone resistance.  
Antithyroid drugs | TBG excess # |

* Excludes short term changes related to initiation or cessation of therapy

# Effect on total hormone concentration; free hormone concentration remains normal.

### 5.4 Indications for TRH Testing

The development of high-sensitivity TSH assays has almost eliminated the need for TRH testing in clinical practice. With intact hypothalamic-pituitary function, there is a close correlation between the TSH response to TRH and the basal level of serum TSH, when measured by a highly sensitive assay (193). Hence, TRH testing now offers little diagnostic advantage over accurate measurement of the basal serum TSH concentration in the detection of hyperthyroidism (194). However, measurement of serum TSH 20 to 30 minutes after intravenous injection of 200 to 500 µg TRH remains useful for several purposes:

a. To assess patients whose basal serum TSH values are out of context, in order to identify assay artefacts. For example, a detectable serum TSH concentration that is unresponsive to TRH in a thyrotoxic patient suggests either non-specific assay interference (194), or the presence of a TSH-secreting pituitary tumor where serum TSH is often unresponsive to TRH (185).

b. To distinguish between cases of thyroid hormone resistance and pituitary-dependent hyperthyroidism. In most instances of hyperthyroidism caused by TSH-secreting pituitary tumors there is no increase in serum TSH in response to TRH (185), while an increased serum TSH response to TRH is seen in the thyroid hormone resistance (159).

### 5.5 Serum Thyroglobulin

This iodinated 660 kDa dimeric glycoprotein, which is the backbone for thyroid hormone synthesis, is normally detectable in serum and is released in excess in a wide variety of situations where thyroid tissue is overactive, inflamed or proliferating. Undetectable serum levels suggest absence or suppression of thyroid tissue. In the past decade functional assay sensitivity has improved almost ten-fold from about 1 ng/ml to around 0.1 ng/ml. As a consequence, Tg is now detectable without TSH stimulation in a greater proportion of patients and has greater diagnostic sensitivity during continuing T4 therapy (195). Nevertheless, undetectable serum Tg in the presence of high TSH remains the benchmark for proof of successful ablation after treatment of differentiated thyroid cancer.

Standard indications for measurement of serum thyroglobulin are as follows:
a. Follow-up of treatment for differentiated thyroid cancer to identify or rule out the presence of residual thyroid tissue, whether normal or metastatic. An undetectable serum thyroglobulin concentration in the presence of high TSH, whether achieved by thyroxine withdrawal, or injection of recombinant TSH, is presumptive evidence that differentiated thyroid tissue has been ablated (196). In contrast, the level of serum thyroglobulin that persists when serum TSH is suppressed, can give a useful index of tumor burden (196). In a 16-year follow-up study in differentiated carcinoma, post-ablative Tg concentration was a strong predictor of disease recurrence (196a). Concurrent measurement of serum TSH and thyroglobulin antibody is always required for interpretation of serum thyroglobulin values.

b. Investigation of atypical hyperthyroidism, where there is a suspicion of thyrotoxicosis factitia in which serum thyroglobulin is generally very low or undetectable (197).

c. Assay of thyroglobulin in the needle washes after biopsy of extrathyroidal neck masses is useful in identifying metastatic thyroid tissue (198). This technique has higher specificity and sensitivity than cytology from suspect lymph nodes (199).

d. Serum thyroglobulin may serve as a sensitive, if non-specific marker of iodine deficiency or ineffective synthesis of thyroid hormone (200)

Because elevation of serum thyroglobulin concentration occurs in a wide range of thyroid disorders, interpretation of results always requires a knowledge of (i) the clinical context, (ii) the serum TSH concentration and (iii) whether thyroglobulin antibodies are present.

Interactions between serum thyroglobulin and circulating antibodies have the potential to cause false-negative assay results, as discussed in chapter 6a.

5.6 Thyroglobulin antibody (TgAb)

Assessment of TgAb in serum is indicated:

a. As an essential component of the interpretation of assays for serum thyroglobulin, to establish whether endogenous antibodies could be responsible for spuriously low thyroglobulin values by interfering with assay separation, particularly in immunometric assays. It is cause for concern that current assays for TgAb may not give congruent results when cross-matched in the assessment Tg in antibody-positive samples (See chapter 6b).

b. As secondary follow-up criterion in differentiated thyroid cancer, where a progressive decline in antibody concentration may follow successful ablation.

5.7 Thyroid peroxidase antibody (TPOAb)

This entity, previously termed antimicrosomal antibody, is closely linked with autoimmune thyroid disease, in particular Hashimoto’s thyroiditis. Measurement of thyroid peroxidase antibody is indicated as follows:

a. To identify an autoimmune cause for primary hypothyroidism.

b. In individuals with the TSH excess of mild thyroid failure, in whom a positive TPOAb indicates an approximately two-fold increase in risk of progression to overt hypothyroidism (58). A strongly positive result in the presence of mild thyroid failure may influence the decision to commence replacement.

c. In screening for immune thyroid dysfunction or potential hypothyroidism before or early in
pregnancy, especially in high-risk groups and as a predictor of the potential postpartum thyroid dysfunction. (see below and chapter 6b).

d. In order to establish whether there is an immune basis for euthyroid goitre, and to evaluate the risk of future hypothyroidism at a stage before there is any increase in serum TSH.

### 5.8 TSH Receptor Antibody (TRAb) (See chapters 6b and 7)

Antibodies that interact with the TSH receptor can be measured either by bioassay, or competitive binding techniques. Bioassays that use thyroid cells of human or animal origin, or cells with transfected TSH receptor, generally depend on tissue production of cAMP for quantitation. These assays allow a distinction to be made between antibodies with thyroid stimulating (TSAb) and blocking (TBAb) activity. Radioreceptor (thyroid binding inhibitor immunoglobulin, TBII) assays that measure competition by circulating immunoglobulin for specific binding of labelled TSH, are more widely available, but do not distinguish between blocking and stimulating activity.

A competitive binding assay using a recombinant human TSH receptor, showed 98-99% positivity in active Graves’ disease, with positive results in <1% of subjects with non autoimmune thyroid diseases (201). Impressive sensitivity and specificity has been reported with an assay that uses solubilised native porcine TSH receptor with a biotinylated anti-porcine TSH receptor monoclonal capture antibody (202).

Indications for the measurement of TRAb vary in different practice environments, but are clearly applicable to the following situations:

a. During pregnancy in women with active or previous autoimmune thyroid disease, to assess the risk of neonatal thyroid dysfunction due to transplacental passage of TRAb

Guidelines from the European Thyroid Association (203) suggest that TRAb measurements should be made early in pregnancy in women who have received previous ablative treatment, and in the last trimester in women receiving drug treatment for active Graves’ disease. Antibody measurement during pregnancy was not regarded as necessary for Graves’ disease in remission (203).

b. In the differential diagnosis of atypical hyperthyroidism that may be due to Graves’ disease.

c. In atypical eye disease that may be due to thyroid ophthalmopathy.

d. To assess the chance of achieving a drug-induced remission of Graves’ disease and to predict whether an apparent remission of Graves’ disease is likely to be sustained, or whether relapse should be anticipated. (See chapters 6b, 7 and 11)

### 5.9 TSH alpha subunit

Assay of the TSH alpha subunit is indicated where hyperthyroidism appears to be the result of central autonomous TSH excess to distinguish this entity from thyroid hormone resistance (185). Most TSH producing pituitary adenomas show an increase in alpha subunit (185), and this assay may serve as a useful tumor marker after treatment. By contrast, levels generally remain normal in thyroid hormone resistance (159). Values are also high in postmenopausal women, in men with hypogonadism and in gonadotrophin-producing pituitary tumors; both thyrotrophs and gonadotrophs secrete this subunit.
5.10 Serum reverse T3

This assay does not have a diagnostic role. Previous suggestions that it might be useful in distinguishing true hypothyroidism from the hypothyroxinemia of severe illness have not been confirmed (204).

5.11 Urinary iodine estimation

Both iodine deficiency and excess can lead to important abnormalities of thyroid function. In contrast to the wide-reaching effects of iodine deficiency, especially in relation to pregnancy and the neonate (see Chapters 14 and 15b), possible iodine excess should be considered as a precipitating cause of hyperthyroidism in the following situations (205):

(i) Atypical hyperthyroidism with blocked isotope thyroid uptake, with features similar to those of excess thyroid hormone ingestion, or subacute thyroiditis without the expected systemic features.

(ii) High dose requirement for antithyroid drug, or failure to respond to these medications. (In some instances, iodine-induced thyrotoxicosis may be extremely resistant to standard therapy, sufficient to require emergency thyroidectomy (206)).

(iii) Rapid progression from subclinical to overt hyperthyroidism, especially in patients with autonomous multinodular goitre.

Thyrotoxicosis due to iodine excess, which may ultimately be self-limiting, may result from iodine-containing radiographic contrast media or medications, alternative or “natural” health care products and contaminated food products, as observed with soy milk preparations (207). In contrast to subtleties related to urine volume and concentration that are important in assessing urinary iodine excretion in population studies of iodine deficiency, the possibility of marked iodine excess can be assessed from a single urine sample (208). The urinary iodine concentration that can be regarded as excessive is ill-defined and may vary in different populations. Urinary iodine concentrations >1000 ug/l are certainly abnormal, but iodine-induced thyrotoxicosis may also occur at lower levels.

6 Thyroid testing strategies in special situations

6.1 Subclinical dysfunction

Follow-up studies suggest that it is inappropriate to commence treatment for subclinical thyroid dysfunction on the basis of a single laboratory result, or even a confirmed assay result within a short period of time (81,82) (see above). Because both free T4 and TSH show spontaneous fluctuation, it has been suggested that frequent testing may increase the likelihood of an apparent change (209).

6.2 Hyperthyroidism

6.2.1 Initial diagnosis

The initial diagnosis of hyperthyroidism is securely established when excess of free T4 and free T3 is associated with an undetectable serum TSH concentration in an assay of appropriate sensitivity. However, the condition can be present without any one of these three criteria. T3-thyrotoxicosis, in which serum T4 remains normal, is more prevalent in iodine deficient regions (210) and can be a
premonitory stage of typical hyperthyroidism (211). When hyperthyroidism coexists with another severe illness, serum T3 may be transiently normal or even subnormal (212). Detectable or increased serum TSH concentrations in the face of hyperthyroidism can be due to laboratory artefacts (194), or to autonomous over-secretion of TSH (185). The increase in free T4 and free T3 estimates is usually more marked than the increase in total hormone concentration. Progressive increases in serum total T4, will approach and eventually exceed the limited T4 binding capacity of TBG (about 300 nmol/l or 24 ug/dl), leading to disproportionate increases in the free serum concentrations of T4 (213) and T3 (214), relative to their total concentrations.

6.1.2 Treated hyperthyroidism

In the early drug treatment of hyperthyroidism, measurements of serum T4 and T3 are required for dose adjustment because suppression of TSH may persist for months after correction of longstanding hyperthyroidism (151). Hence, failure to decrease thionamide dosage while TSH suppression persists, can result in serious over-treatment during the early phase of therapy. During thionamide therapy, hyperthyroidism may persist due solely to T3 excess (188); assessment of therapy based on serum T4 alone can then result in under-treatment. A daily dose of 15 mg methimazole can result in hypothyroxinemia in about 10 % of previously thyrotoxic subjects within 4 weeks (215); hence a reassessment of both serum free T3 and free T4 is timely after about 3 weeks to allow appropriate dose-adjustment. In contrast to these discrepancies during early treatment, serum TSH generally gives a reliable index of therapy during the longterm drug treatment of hyperthyroidism. The rate of change in serum free T4 and free T3 during treatment of hyperthyroidism with thionamide gives a valuable guide to dose adjustment, even while levels remain increased and serum TSH is still undetectable. For example, if serum T4 decreases by half in the first four weeks of treatment, it is appropriate to decrease initial thionamide dosage by about half to minimize the possibility of over-treatment.

Davies et al (216) assessed the prognostic significance of various serum TSH levels in a large cohort of patients treated with radioiodine 2-35 years previously, who were receiving no other treatment. After a further two years follow-up, 83% of those with normal TSH had not changed their diagnostic category, although there was a trend for TSH to increase. An increased TSH was associated with a 14.5 % incidence of hypothyroidism after one year. Notably, spontaneous normalization of subnormal or undetectable TSH values during the follow-up period was more common than recurrence of overt hyperthyroidism (216). Hence, while an increased TSH value might be a pointer towards T4 treatment, there appears to be no basis for further radioiodine therapy solely because TSH remains suppressed.

6.3 Hypothyroidism

6.3.1 Initial diagnosis

The initial diagnosis of overt primary hypothyroidism is established by increase in serum TSH with subnormal free T4; serum free T3 may remain normal except in severe cases. Subclinical hypothyroidism or mild thyroid failure is defined by persistent elevation of serum TSH, while free T4 remains normal. The precise upper limit of the reference range for TSH is difficult to define, because of the approximately logarithmic distribution of this parameter. The upper “tail” of normal TSH in the range 2-4 mU/l is thin and subjects with TSH values in this range have an increased likelihood of future hypothyroidism, especially if peroxidase antibody is positive (58).

Serum TSH is the key to the distinction between primary and secondary hypothyroidism. A low serum free T4 in the absence of TSH elevation should always raise the possibility of a pituitary or
hypothalamic abnormality, although this combination of findings is also frequent in a number of other situations (table 7), especially during critical illness (156). It should be noted also that immunoreactive serum TSH is often detectable in secondary or central thyroid deficiency, a phenomenon that appears to result from dissociation between biological and immunological TSH activity (162). Hence, normal or even elevated serum TSH should not be interpreted as conclusive evidence against hypopituitarism.

<table>
<thead>
<tr>
<th>Table 7 Causes of subnormal free T4 without TSH excess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secondary or Central Hypothyroidism</td>
</tr>
<tr>
<td>Impaired biological activity of TSH</td>
</tr>
<tr>
<td>Critical Illness</td>
</tr>
<tr>
<td>Falsely low free T4 estimate (method-dependent)</td>
</tr>
<tr>
<td>Dilution-dependent artefact (see 6.4.2 below)</td>
</tr>
<tr>
<td>Effect of medications that compete for T4 binding</td>
</tr>
<tr>
<td>Impaired TSH response to hypothyroxinemia</td>
</tr>
<tr>
<td>Effect of severe illness</td>
</tr>
<tr>
<td>Medications eg dopamine, glucocorticoids</td>
</tr>
</tbody>
</table>

6.3.2 Response to treatment

A serum TSH value in the low-normal range between 0.5 and 1.5 mU/L, close to the geometric mean, is probably the best single indicator of appropriate thyroxine dosage during standard replacement therapy. In a study of ambulatory patients attending a thyroid clinic, hypothyroid patients taking T4 replacement seldom needed a serum free T4 measurement if the serum TSH was greater than 0.05 mU/L, although at lower TSH values, the magnitude of T4 excess did influence management (217). Numerous studies show that patients taking exogenous thyroxine show higher levels of serum total and free T4 for equivalent levels of serum TSH and T3 when compared with untreated euthyroid control subjects (218,219).

It should be noted that in some situations, for example patients with both ischemic heart disease and hypothyroidism, or in the extreme elderly, the appropriate dose of T4 should be influenced by clinical judgment as well as laboratory findings. The assessment of optimal dosage in patients with secondary or central hypothyroidism remains a challenge, because serum TSH does not serve as a reliable marker of under-replacement. Serum TSH may diminish during thyroxine replacement in hypopituitarism (219); while not of primary diagnostic value, the level may serve as an index of individual progressive response (220).

Since there is no readily available alternative parameter of thyroid hormone action, it is generally appropriate to assess replacement clinically and on the basis of both serum T4 and T3. Measured levels of free and total T4 are influenced by the interval between tablet ingestion and blood sampling. In athyreotic subjects who took 0.15-0.2 mg T4 orally, the serum free and total T4 concentrations were increased by about 20% one to four hours later, with return to baseline about nine hours after T4
ingestion; serum TSH and T3 levels showed no time-dependent variation in relation to timing of thyroxine dosage (221).

### 6.3.3 Suppressive treatment with T4

In line with recent guidelines for the stratification of risk in patients with differentiated thyroid cancer (222), there is no single TSH target during long-term postoperative thyroxine therapy (see chapter 18). In high risk patients there may be benefit from adjusting T4 dosage to suppress TSH to undetectable levels (e.g. <0.03 mU/l). However, in many situations it may be appropriate to aim for a TSH target in the lower normal range in order to minimize adverse effects on bone and cardiovascular system. When the aim of T4 suppressive therapy is regression of benign thyroid tissue, it may be adequate to give sufficient T4 to reduce serum TSH to 0.1 to 0.3 mU/L (223).

### 6.3.4 Treatment with T3

If currently available T3 formulations are used for replacement, it is difficult to monitor the effectiveness of treatment. Owing to its short plasma half-life, the serum concentration is highly dependent on the interval between dosage and sampling (224). There is also doubt as to whether TSH serves as an accurate index of thyroid hormone action during continuing T3 therapy, with the suggestion that doses of T3 required to normalise TSH could produce tissue hyperthyroidism (225). The difficulty of monitoring T3 replacement by currently available techniques is one of the arguments against the routine use of T3 or thyroid extract (226).

If the suggestion is confirmed that some genetically-identifiable individuals within the hypothyroid population are better served by combined T4-T3 therapy on the basis of polymorphism of type 2 deiodinase (227,228), the use of T4-T3 combination therapy will increase. That will require re-evaluation and refinement of monitoring strategies for combined replacement. The development of sustained-release T3 formulations will be advantage, so that the interval between dosage and sampling can be minimized as a variable.

### 6.3.5 Replacement with uncertain indication

The recent demonstration that fertility and pregnancy outcome may be adversely affected by minor degrees of thyroid dysfunction (see 6.7 below and chapter 15) has led to commencement of thyroxine treatment in women of reproductive age, who may lack conclusive criteria for lifelong therapy. It may later be questioned whether continuing replacement is required. Positive peroxidase antibody may be a basis for continuing therapy, but in the absence of this marker, measurement of serum TSH after 3-4 weeks withdrawal of treatment is likely to be definitive. The alternative of following the TSH response during partial replacement with T4, 50 ug daily, may be preferred. Similar approaches apply where no documented indication exists. Because of the difficulty in establishing the need for treatment during optimal replacement, it is an advantage for hypothyroid patients to retain a permanent record of pre-treatment documentation.

### 6.4 Assessment of thyroid function during non-thyroidal illness (NTI) (See chapter 5b)

Several distinct issues need to be considered when assessing thyroid function during critical illness. First, there is the possibility that an underlying thyroid abnormality might be missed, and second, prolonged severe illness per se may be associated with an abnormality of thyroid hormone secretion or
action that may benefit from treatment (229,230). Third, some of the observed abnormalities will reflect the effect of medications, or may be methodological artefacts.

It is difficult to rule out previously unrecognized or recent thyroid dysfunction by clinical assessment in critically ill patients; current laboratory tests often do little to resolve the problem. During severe illness, one or more of the assumptions that underpin the diagnostic use of the TSH-T4 relationship (see 4.3 above) may not be valid. For example, acute fluctuations from the steady-state can lead to an anomalous T4-TSH relationship simply because of the marked difference in the plasma half lives of T4 and TSH.

In general, the same sample assayed for TSH by various methods will give similar results during critical illness, while various estimates of free T4 may give widely divergent results (see below). When TSH and T4 changes are considered together, the abnormal results rarely correspond to standard criteria for diagnosis of primary hypothyroidism or hyperthyroidism. In contrast, persistent hypothyroxinemia without the corresponding anticipated rise in serum TSH is a common finding that suggests secondary or central hypothyroidism. These changes appear to be part of a broad neuroendocrine response that also involves the pituitary-adrenal axis, the pituitary-gonadal axis and the IGF binding proteins (230,231). Further study of these responses may eventually lead to therapy that could extend beyond thyroid hormone replacement, for example to substitution of hypothalamic releasing hormones (231, 232).

6.4.1 Effects of medications

Interpretation of thyroid function tests during critical illness can be influenced by multiple medications (table 8), in particular dopaminergics and glucocorticoids, which inhibit TSH secretion, and a wide range of inhibitors of T4 and T3 binding to circulating TBG.

<table>
<thead>
<tr>
<th>Table 8 Effects of medications that influence thyroid function</th>
</tr>
</thead>
<tbody>
<tr>
<td>(see also chapter 5a)</td>
</tr>
<tr>
<td><strong>Inhibit pituitary TSH secretion</strong></td>
</tr>
<tr>
<td>Dopamine, dobutamine, glucocorticoids, octreotide,</td>
</tr>
<tr>
<td>Bexarotine (233), metformin ? (234)</td>
</tr>
<tr>
<td><strong>Iodine load modifies hormone synthesis and release</strong></td>
</tr>
<tr>
<td>Contrast agents, amiodarone, topical iodine preparations</td>
</tr>
<tr>
<td><strong>Inhibit thyroid hormone synthesis or release</strong></td>
</tr>
<tr>
<td>Lithium, glucocorticoids, aminogluthethimide ?</td>
</tr>
<tr>
<td><strong>Inhibit T4-T3 5' Deiodination</strong></td>
</tr>
<tr>
<td>Amiodarone, glucocorticoids, beta blockers *</td>
</tr>
<tr>
<td>Hepatic Contrast agents, e.g. iopanoic acid, ipodate</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td><strong>Modify immune function</strong></td>
</tr>
<tr>
<td>Interleukin 1, interferon alpha, interferon beta</td>
</tr>
<tr>
<td>Monoclonal antibody therapy</td>
</tr>
<tr>
<td><strong>Modify binding of T4, T3 to plasma proteins</strong></td>
</tr>
<tr>
<td><strong>a. Alter concentration of T4 binding globulin</strong></td>
</tr>
<tr>
<td><strong>(i) Increase</strong> Estrogen, heroin, methadone Clofibrate, 5-fluouracil, perphenazine, tamoxifen</td>
</tr>
<tr>
<td><strong>(ii) Decrease</strong> Glucocorticoids, androgens, l-asparaginase</td>
</tr>
<tr>
<td><strong>b. Displace T4 and T3 from binding proteins (167)</strong> Furosemide, salicylates, henytoin, carbamazepine, meclofenamante, Heparin #</td>
</tr>
<tr>
<td><strong>c Displace T4 from tissue pool</strong></td>
</tr>
<tr>
<td>Alkylating agents, oral cholecystographic agents</td>
</tr>
<tr>
<td><strong>Modify thyroid hormone action</strong></td>
</tr>
<tr>
<td>? Amiodarone, ? phenytoin</td>
</tr>
<tr>
<td><strong>Increase clearance of T4,T3</strong></td>
</tr>
<tr>
<td>Barbiturates, phenytoin, carbamazepine</td>
</tr>
<tr>
<td>Rifampicin</td>
</tr>
<tr>
<td>Sertraline?, fluoxetine?, dothiepin?</td>
</tr>
<tr>
<td>Tyrosine kinase inhibitors (32)</td>
</tr>
<tr>
<td>Bexarotene (233)</td>
</tr>
<tr>
<td><strong>Impair absorption of ingested T4 (235,236)</strong></td>
</tr>
<tr>
<td>Impaired gastric acidity (237)</td>
</tr>
<tr>
<td>Aluminium hydroxide, ferrous sulfate, cholestyramine, coffee (238), Colestipol, sucralfate, soya preparations,</td>
</tr>
</tbody>
</table>
6.4.2 Methodological discrepancies in non-thyroidal illness

6.4.2.1 The heparin artefact and free T4

The effect of heparin to increase serum free T4 is an important in vitro phenomenon that can lead to spuriously high estimates of circulating free T4 (168). In the presence of a normal serum albumin concentration, non-esterified fatty acid (NEFA) concentrations >3 mmol/l are required to increase free T4 by displacement from TBG, but these concentrations are uncommon in vivo. However, in samples from heparin-treated patients, serum NEFA may increase to these levels during in vitro sample storage or incubation as a result of heparin-induced lipase activity (168) (figure 6). This effect is accentuated by incubation of serum at 37 C and by increased serum triglyceride or low serum albumin concentrations, particularly if the sample is pre-diluted (164). If heparin is given in vivo and the sample is then incubated at 37C, doses of heparin as low as 10 units may result in NEFA-induced increases in the apparent concentration of serum free T4 (167,169). The assay result is analytically correct, but does not reflect the in vivo concentration of free T4. Low molecular weight heparin preparations have a similar effect (240).
Figure 6: Heparin-induced release of lipase in vivo can lead to in vitro generation of non-esterified fatty acids during sample incubation or storage. An increase in serum NEFA concentrations to > 3 mmol/l is sufficient to displace T4 from TBG, but such values are uncommon in vivo. This artefact is accentuated by high triglyceride or by low albumin concentrations.

6.4.2.2 Competitors for plasma protein binding

The accuracy of virtually all methods of free T4 estimation is compromised by medications that displace T4 and T3 from TBG. Current methods tend to underestimate the concentration of free T4 in the presence of binding competitors because of dilution-related artefacts. Binding competitors are usually less protein-bound than T4 itself so that progressive sample dilution leads to a fall in the free concentration of competitor before the free T4 concentration alters (167). (For a hormone such as T4, with a free fraction in serum of about 1:4000, progressive dissociation will sustain the free T4 concentration up to at least 1:100 dilution. In contrast, 1:10 dilution of serum will result in a marked decrease in the free concentration of a drug that is 98% bound, i.e. has a free fraction in serum of 1:50). Because displacement depends on the relative free concentrations of primary ligand and competitor, the underestimate of free T4 will be greatest in assays with the highest sample dilution. This important dilution-dependent difference between various free T4 methods was shown by the relative ability of three commercial free T4 assays to detect the T4-displacing effect of therapeutic concentrations of furosemide (241) (figure 7).

Similarly, therapeutic concentrations of phenytoin and carbamazepine increased the free concentration
of T4 by 40-50% using ultrafiltration of serum that had not been diluted, while the free hormone estimate was spuriously low using a commercial single-step free T4 assay after 1:5 serum dilution (179).

Figure 7: Influence of increasing serum concentrations of added furosemide on estimates of serum free T4 using three commercial free T4 methods that involve varying degrees of sample dilution. The effect of the competitor is progressively obscured with increasing sample dilution. (Redrawn from 241).

It is possible that methodologic artefacts have influenced previous descriptions of free T4 changes during critical illness. On the one hand, an apparent increase in free T4 may arise from heparin-induced in vitro generation of free fatty acids during sample incubation (168). On the other hand, estimates of free T4 may be spuriously low in assays that use diluted serum (167,179).

6.4.2.3 Divergent estimates of free T4

That estimates of free T4 may show opposite discrepancies by different methods was shown by Sapin et al (242) in a prospective study of bone marrow transplant recipients. Twenty previously euthyroid subjects were studied on the seventh day after bone marrow transplantation using six commercial free T4 kits, during multiple drug therapy, including heparin and glucocorticoids (figure 8). Free T4 methods that involved sample incubation at 37 °C showed supranormal free T4 values in 20-40% of these subjects (see heparin effect above), while analog tracer methods that are influenced by tracer binding to albumin gave subnormal estimates of free T4 in 20-30%. By contrast, total T4 was normal in 19 of these 20 subjects. Serum TSH was <0.1 mU/l in half the subjects, independent of the method that was used. Thus, there was the possibility that an erroneous diagnosis of either hyperthyroidism or secondary or central hypothyroidism could be considered, solely as a result of variations in free T4 methodology.
Figure 8: Free T4 estimated by six different kit methods in 20 previously euthyroid patients on the seventh day after bone marrow transplantation. There was a high proportion of abnormal values, either increased or decreased, depending on the type of free T4 method used (see text). Therapy included heparin and glucocorticoids at the time of sampling. The mean for each method has been normalized to 100%, with the limits of the range shown by the box. Serum total T4 remained normal in 19 of the 20 study subjects, while serum TSH was subnormal in 11, independent of assay method. (Redrawn from 242)

6.4.3 Serum TSH in critical illness

Serum TSH assessment in severe NTI depends on the sensitivity of the particular method. The “third generation” assays with a functional sensitivity below 0.01 mU/L are generally sufficiently sensitive to distinguish the very low values in most thyrotoxic patients from the subnormal but somewhat higher TSH levels of NTI (243,244). Among a group of patients with low serum TSH values (<0.1 mU/L), almost all thyrotoxic patients had values less than 0.01 mU/L, when assessed with a highly sensitive assay, whereas most critically ill euthyroid patients had values between 0.01 and 0.1 mU/L (243). However, in another study, about 4% of patients with NTI had values below the functional sensitivity of a “third generation” assay, indicating that an absolute distinction cannot be made on the basis of TSH alone (244).

6.5 Differentiated Thyroid Cancer

There is now consensus that not all patients with differentiated thyroid cancer require T4 treatment in doses that achieve complete TSH suppression. Thus, based on a general assessment of risk (222), a TSH target should be determined as a guide to T4 dosage. In the follow-up of differentiated thyroid cancer the interpretation of TSH and serum thyroglobulin is inter-dependent. For studies done after
initial near-total thyroidectomy, following withdrawal or temporary reduction of suppressive therapy (245), or with the use of recombinant human TSH (246), serum TSH levels in excess of 30 mU/l appear to achieve adequate stimulation of potential thyroglobulin production (247). The failure of endogenous TSH to increase into this range suggests too short a period of T4 withdrawal, a compliance problem, or the presence of a substantial amount of active thyroid tissue.

Serum thyroglobulin measurement and whole body radioiodine scanning have generally been used in a complementary fashion, but there is now good evidence that current assays for thyroglobulin have greater sensitivity than follow-up whole body scanning with 2mCi 131I. There is a clear trend to place greater emphasis on measurements of thyroglobulin and to move away from repeated low dose diagnostic whole body radioiodine scanning, a procedure that has limited sensitivity (248,249). Management of low risk patients now tends to be based on assessment of serum thyroglobulin after recombinant TSH stimulation, without the need for diagnostic scanning (249). Undetectable serum thyroglobulin in the years after ablative treatment has been shown to be a reliable index in ruling out persistent or recurrent disease that requires further evaluation and treatment. Detectable serum thyroglobulin, together with detailed ultrasonography for localization has become the mainstay of further investigation (250).

Assay of thyroglobulin on needle washes of suspect lymph nodes has a high degree of sensitivity and specificity, apparently superior to cytological examination (198,199). During long-term follow-up, it is crucial that clinicians are kept informed of changes in serum thyroglobulin methodology that may give a false impression of remission or recurrence (see chapter 6a).

6.6 Psychiatric Illness

An unusual variety of euthyroid hyperthyroxinemia occurs in some patients hospitalized with acute psychiatric illness (251). Serum T4 is increased, but serum T3 is less frequently elevated; serum TSH is generally normal or slightly high (252). These abnormalities, presumed to be due to central activation of the hypothalamic-pituitary-thyroid axis, often resolve in several weeks (251).

6.7 Assessment of thyroid function before, during and after pregnancy (see also chapter 14)

Recent advances in the understanding of the importance of optimal maternal thyroid function for fetal brain development in early pregnancy, as well as the influence of thyroid immunity, or thyroid status, on numerous pregnancy outcomes, has greatly increased the frequency of thyroid testing before, during and after pregnancy.

Clinical guidelines from The Endocrine Society (253) suggest consensus, but significant controversies remain (254). The reproductive effects of abnormal thyroid function and its immune associations are considered in chapter 14. This section will consider the issues from the point of view of testing strategy and assay methodology. From studies of pregnancy outcome, evidence is becoming persuasive that miscarriage rate (256,257) and prematurity (255,257) are favourably influenced by maternal thyroid hormone replacement, even when hypothyroidism would be regarded as “subclinical” by standard criteria.

A key consideration is whether thyroid function should be universally tested before or early in pregnancy, in view of the fact that women at risk of adverse outcome cannot be identified reliably from their clinical history, even if this is assessed in full detail. (256)
6.7.1 Methodological issues

Some methodological issues are relevant before the application of tests is considered. There has been almost universal acceptance of free T4 estimates in preference to total T4 measurement in pregnancy because of the well-known estrogen effect to increase TBG and hence total T4. It is clear that normal pregnancy is associated with a marked increase of about 40% in total T4 concentration, but free T4 is maintained close to normal non-pregnant levels (258). Previous suggestions that free T4 actually declines significantly in late pregnancy (259) are uncertain because of wide method-dependent variations, with strong negative bias in some methods (260,261). In a comparative study of free T4 by seven commercial methods in 23 euthyroid women at term, Roti et al found that albumin-dependent methods show marked negative bias, with up to 50% subnormal values, while other methods gave values above their non-pregnant reference interval (260). Notably, particularly during pregnancy, methods of free T4 estimation fall far short of desirable assay standardization criteria that aim to minimize between-assay variation (120a).

A recent study by Lee et al confirmed marked negative bias in free T4 estimates during pregnancy, and questioned the basis for continuing to rely on free thyroxine estimates during pregnancy (262). In contrast to two kit free T4 methods, total serum T4 and its derivative, the free thyroxine index, showed the anticipated inverse relationship with serum TSH, with historically consistent results in numerous reports (262). Thus, because of consistency between methods, total T4 measurement may be superior to free T4 estimation as a guide to therapy during pregnancy, provided that reference values take account of the normal estrogen-induced increase in TBG during pregnancy.

Serum TSH is generally analytically reliable during pregnancy, but the reference interval is lower at the end of the first trimester, associated with the effect of HCG as a surrogate thyroid stimulator (263). Hence, the general advice that tests of thyroid function during pregnancy should relate to trimester-specific reference intervals (263). In the case of free T4, ranges also need to be method-specific (260-262). The magnitude of method-dependent variations between non-pregnant and third trimester free T4 estimates is shown in Fig 9.
Figure 9. Free T4 estimates in the third trimester of pregnancy show large method-dependent differences from the reference intervals for non-pregnant subjects. Reproduced from ref 261 and 261a which identify each method studied. The left panel shows values measured by equilibrium dialysis. The solid bars indicate third trimester free T4 estimates, the grey bars, non-pregnant results.

6.7.2 Before pregnancy or when pregnancy is confirmed.

Because of the adverse effects of maternal thyroid deficiency, there is agreement that pre-pregnancy testing is now mandatory to optimize thyroxine replacement and to assess thyroid status in women who have an increased risk of thyroid dysfunction (253), for example a history of recurrent miscarriage, impaired fertility requiring assisted reproductive technology, type 1 diabetes or any other factors known to be associated with increased risk of thyroid dysfunction (Table 1). Notably, in one key study, about a third of women with subclinical or overt hypothyroidism in the first trimester would have been missed if testing had been confined to those assessed as being at higher risk, based on detailed questioning about potential risk factors, in particular a personal or family history of thyroid or other autoimmune disorder, or previous thyroid-related treatment (256). A further study confirms that about half of the women at risk for pregnancy-related thyroid dysfunction would be missed by testing confined to the group assessed to be at high risk (256a). Hence, the view that routine testing may now be warranted (4, 254, 264). Because the major influence of thyroid hormone on fetal brain development is early in pregnancy and the onset of thyroid hormone action is slow, effective testing needs to be done earlier than would follow the first routine obstetric visit at 10-12 weeks.

While there is now good evidence that any level of TSH elevation that suggests maternal hypothyroidism should be treated, there is currently no consensus about responses to other test abnormalities. An important finding is that thyroxine treatment of peroxidase antibody-positive euthyroid women from about the end of the first trimester may improve obstetrical outcome, as judged by rates of miscarriage and premature delivery (255). A recent meta-analysis confirms that thyroid autoantibodies are strongly associated with miscarriage and preterm delivery and suggests that
thyroxine treatment diminishes these risks (255a). Pre-partum assessment of antibody status may also provide information that improves prediction of post-partum thyroid dysfunction (255b).

A recent study showed that almost 50% of thyroxine treated women were not optimally replaced in early pregnancy (TSH <0.4 or >4.0 mU/l) with the potential for significant adverse effect on pregnancy outcome (265). In thyroxine-treated women it should now be standard practice to assess optimal replacement prior to pregnancy and to re-evaluate this when pregnancy is confirmed, with a prompt 25-30% increase in replacement dosage (266), with two extra daily doses per week as an approximation. The required increase is larger in women who have had thyroid ablation than in those with spontaneous hypothyroidism (266).

A special situation arises in thyroxine-treated women who have in vitro fertilization. Ovulation-induction cycles, even in the absence of ongoing pregnancy, can be associated with 2-10 fold increases in serum TSH (267), indicating that previously adequate replacement may not accommodate estrogen-induced thyroxine requirements. This suggests that women who lack endogenous thyroxine should receive increased dosage from the time of ovulation induction, to accommodate increased thyroxine requirement, whether associated with a subsequent pregnancy, or not. The need to routinely test thyroid function before in vitro fertilization seems self-evident.

6.7.3 Assessment during pregnancy

In women treated for hypothyroidism, some guidelines still recommend that replacement thyroxine dosage should be progressively adjusted in pregnancy in response to increasing serum TSH, a response that is likely to be too late to assure optimal first trimester maternal T4 levels for optimal fetal brain development. Hence, the advice to increase dosage by two daily doses per week as soon as pregnancy is confirmed, followed by a check of serum TSH about 4 weeks later (253,266).

In women treated for Graves’ disease with antithyroid drug, dose reduction should be the aim, keeping TSH in the lower normal range. In contrast to the adverse effects of even the mildest grades of hypothyroidism, an extensive study has shown no adverse effect of mild overactivity during pregnancy (268). It is often possible to cease antithyroid drug late in pregnancy, with prospective follow-up to detect post-partum recurrence. In contrast to immune thyroid disorders, nodular thyroid disorders do not appear to show marked fluctuation during and after pregnancy.

In the event of uncontrolled or severe maternal hyperthyroidism, careful assessment should be made of fetal growth and thyroid size followed by fetal blood sampling in some cases as a guide to optimal therapy (269). Guidelines for the assessment of TSH receptor antibody to gauge the likelihood of neonatal thyrotoxicosis have been discussed above (157).

6.7.4 Postpartum Assessment

In prospective studies there is a high prevalence of thyroid dysfunction during the post-partum period (253). In an Australian study, 11.5% of women showed at least one laboratory abnormality of thyroid function when tested 6 months after delivery (270). TSH values were increased in 6%, of whom almost 90% showed positive peroxidase antibody, indicating an autoimmune abnormality. In about half the women with initial TSH elevation who remained untreated, TSH was still increased 30 months after delivery (270), as in other studies that show a high prevalence of continuing hypothyroidism after postpartum TSH elevation (271). Hence, women with postpartum hypothyroidism should be treated on the basis of confirmed TSH elevation even if the indication for lifelong replacement has not been firmly established. This strategy is especially important because untreated maternal hypothyroidism has critical implications for fetal development in subsequent pregnancies, By contrast, almost all women
who had suppressed or subnormal TSH values 6 months postpartum, showed normal TSH 12 and 18 months later, indicating that postpartum hyperthyroidism is likely to be temporary phenomenon (253,270).

The diagnosis of postpartum thyroid dysfunction is elusive (255b, 271a). When women were classified on the basis of screening early in pregnancy as having a high, or low risk of post-partum thyroid dysfunction, the chance of an abnormality was about 6-fold greater in the high risk group, but about half the affected subjects came from the low-risk group (255b, 271a). Notably, 54% of women with postpartum hypothyroidism were still hypothyroid 12 months post-partum (271a).

In women with established hypothyroidism, replacement dosage of thyroxine can generally be decreased to the pre-pregnancy dose with testing deferred for several months to review dosage. In Graves’ hyperthyroidism, where antithyroid drug has been ceased during pregnancy, resumption of treatment is generally deferred, based on symptoms and the results of testing at two-three monthly intervals. Where TPO antibody testing early in pregnancy has been positive in the absence of frank thyroid dysfunction, prospective testing may be warranted in the 12 months post-partum.

Knowledge of maternal thyroid status should be taken into account in assessing the significance of neonatal screening for congenital hypothyroidism, because of reports of temporary effects of maternal blocking antibody in the infant (see chapters 14 and 15b).

6.7.5 Iodine status (See also chapter 20)

The critical importance of adequate iodine nutrition during pregnancy is widely acknowledged. The increased pool of thyroxine that follows from the increase in serum TBG, as well as increase in urinary iodine excretion that will aggravate any degree of iodine deficiency, with potential for adverse effects on fetal development.

While assessment of iodine status by measurement of urinary iodine excretion in individual women does not have wide support, data on iodine nutrition within each geographical area should be an important public health issue, with appropriate supplementation to achieve estimated urinary iodine excretion in the range 150-300 microgram daily in pregnant women.

7 The physician-laboratory interface

Because of the diverse clinical presentations of thyroid dysfunction, initial requests for assessment of thyroid function are often made by clinicians who, while alert to the possibility of thyroid dysfunction, may not be familiar with the limitations of current assays, or with medication effects. Clinical decisions can be assisted by comments from the laboratory, based on detailed knowledge of current immunoassay limitations (272-5). The quality of this assistance will depend on two key components: the training and experience of the reporter and the available clinical information.

7.1 Historical

The competitive binding assays that are used for thyroid diagnosis were initially developed 30-35 years ago by clinical investigators who used ‘in house’ assay reagents and were often closely involved in patient care. This nexus between laboratory and clinical investigation allowed deficiencies in early assays to be readily appreciated, but advanced diagnostic technology was available to only a few practitioners, with results sometimes available only after long delay. In recent decades there has been a strong trend away from ‘in house’ reagents which have been replaced by kits that incorporate highly
sophisticated standardised reagents and automated instrumentation (e.g. solid phase antibodies, magnetic separation systems, chemiluminescent detection systems). Assay turnaround is much faster and up-to-date techniques are widely available to almost all practitioners, most of whom are inexperienced in endocrinology or laboratory methodology. Non-specialist users of endocrine assays are most likely to benefit from laboratory assistance in the interpretation of results, but as assay automation has increased, laboratory professionals have become more distant from the bedside. As clinicians receive less assistance, they tend to provide less relevant information and vice versa. Laboratory personnel in turn, see results that are uninterpretable or ambiguous without the relevant clinical background. Potential assay imperfections may be ignored simply because clinical correlation is not possible (see below).

7.2Diagnostic approach to discordant or apparently anomalous results

The relationship between laboratory results and clinical findings may be either concordant or discordant. With discordant results, a distinction needs to be made between a previously unsuspected diagnosis, subclinical disease and anomalous assay results. If the assay result is confirmed, the fundamental assumptions that underpin the diagnostic use of tropic-target hormone relationships should be considered (see above).

The following steps may be helpful in evaluating anomalous assay results:

a. Re-evaluation of the clinical context, with particular attention to pre-diagnostic probability, long-term features suggestive of thyroid disease and medication history.

b. Assessment of whether serum TSH is markedly suppressed (<0.05 mU/l) or simply in the subnormal-detectable range.

c. Further sampling to establish whether the anomalous result is persistent or transient.

d. Estimation of serum free T4 by an alternative method, preferably a two-step technique that removes binding proteins from the assay system before quantitation.

e. If hyperthyroidism is suspected, measurement of the serum free T3, or total T3 with appropriate binding correction.

f. Measurement of serum total T4 to establish whether the serum free T4 is disproportionately high or low, possibly due to a preanalytical or method-dependent artefact.

g. Possible evaluation of propositus and family members for evidence of unusual binding abnormalities or hormone resistance.

In some circumstances, the cause of an apparently anomalous or unusual laboratory result may become obvious from the clinical context (table 9). By taking account of drug therapy, or acute perturbation of the pituitary-thyroid axis, assay validity can be affirmed and unnecessary further investigation may be avoided.

| Table 9 Situations in which interpretation of unusual laboratory results depends on relevant clinical information |
|-------------------------------------------------|---------------|---------------|
| **Clinical background** | **Assay Results** |
| Pregnancy | fT4 | fT3 | TSH |
| | L* | L,N | N |
7.3 Clinical feedback, quality assurance and cost effectiveness

Clinical feedback will remain a key aspect of quality assurance in laboratory testing. While assay precision or reproducibility can be evaluated solely in the laboratory, diagnostic accuracy requires clinical correlation. As with any diagnostic test, the non-specificity of a procedure may not become apparent until the full diversity of the non-diseased population is appreciated (139). Premarketing evaluation of assays may fail to include the critical samples that probe the diagnostic accuracy of an assay, so that non-specific interference is not appreciated until procedures have been in use for some time.

In studies of “cost effectiveness” it is hard to evaluate the human and financial costs that result from needless duplication, unnecessary testing and inappropriate management. Diagnostic inaccuracy of immunoassays may account for substantial unnecessary expenditure on laboratory resources (274,275). Dilution effects and binding protein abnormalities that affect free T4 estimation, the multiple effects of medications in non-thyroidal illness and the problems associated with thyroglobulin autoantibodies will continue to challenge both clinical chemists and clinicians; further technological development will not substitute for collaboration between these two groups.

Clinical issues may influence the selection of thyroid tests that will best serve a particular population. For example, the effects of severe illness on free T4 estimates may be of minor importance for a laboratory that serves predominantly ambulatory patients. A different assay profile, with emphasis on measurement of total hormone as the “gold standard” for T4 assessment may be required in a laboratory that evaluates thyroid function during critical illness (231,232) or in obstetric practice (253, 262).

References


3 Danese MD, Powe NR, Sawin CT, et al. Screening for mild thyroid failure at the periodic health examination. JAMA 1996; 276:285-92.


19. Sinard RJ, Tobin EJ, Mazzaferri EL, et al. Hypothyroidism after treatment for nonthyroid head and


44. LaFranchi S. Congenital hypothyroidism: Etiologies, diagnosis, and management. Thyroid 1999; 9:735-40.


53. McDermott MT, Ridgway EC, Subclinical hypothyroidism is mild thyroid failure and should be treated. J Clin Endocrinol Metab 2001; 86:4585-90.


66. Surks MI, Ortiz E, Daniels GH et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. JAMA 2004; 291: 228-38.


73. Razvi S, Weaver JU, Vanderpump MP et al. The incidence of ischemic heart disease and mortality in people with subclinical hypothyroidism: reanalysis of the Wickham survey cohort. J Clin Endocrinol


80. Stott DJ, McLellan AR, Findlayson J, et al. Elderly patients with suppressed serum TSH but normal free thyroid hormone levels usually have mild thyroid overactivity and are at increased risk of developing overt hyperthyroidism. Q J Med 1991; 78:77-84.


98. Woeber KA, Observations concerning the natural history of subclinical hyperthyroidism. Thyroid 2005; 15:687-91.


105. Franklyn JA. Metabolic changes in hypothyroidism. In: Braverman LE and Utiger RG, eds.


121. Hamilton TE, Davis S, Onstad L et al. Thyrotropin levels in a population with no clinical, antibody or ultrasonographic evidence of thyroid disease: implications for the diagnosis of subclinical hypothyroidism J Clin Endocrinol Metab 2008; 93: 1124-30.


139. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of
142. Roelfsema F, Periera AM, Adriaanse Ret al. Thyrotropin secretion in mild and severe primary hypothyroidism is distinguished by amplified burst mass and basal secretion with increased spikiness and approximate entropy. J Clin Endocrinol Metab 2010; 95:928-34.


166. Monchamp T, Chopra IJ, Wah DT et al. Falsely elevated thyroid hormone levels due to anti-sheep antibody interference in an automated electrochemiluminescence immunoassay. Thyroid 2007; 17:271-5.


221. Ain KB, Pucino F, Shiver TM, Banks SM. Thyroid hormone levels affected by time of blood sampling in thyroxine-treated patients. Thyroid 1993; 3:81-5.


256. Vaida B, Anthony S, Bilous M et al. Detection of thyroid dysfunction in early pregnancy:


